

# Echo Chambers \*

J. Anthony Cookson

*CU-Boulder*

Joseph E. Engelberg

*UC San Diego*

William Mullins

*UC San Diego*

November 22, 2021

## Abstract

We find evidence of selective exposure to confirmatory information among 400,000 users on the investor social network StockTwits. Self-described bulls are 5 times more likely to follow a user with a bullish view of the same stock than self-described bears. Consequently, bulls see 62 more bullish messages and 24 fewer bearish messages than bears over the same 50-day period. These “echo chambers” exist even among professional investors and are strongest for investors who trade on their beliefs. Finally, beliefs formed in echo chambers are associated with lower ex-post returns, more siloing of information and more trading volume.

---

\*Cookson ([[tony.cookson@colorado.edu](mailto:tony.cookson@colorado.edu)]), Leeds School of Business at University of Colorado-Boulder. Engelberg ([[jengelberg@ucsd.edu](mailto:jengelberg@ucsd.edu)]) and Mullins ([[wmullins@ucsd.edu](mailto:wmullins@ucsd.edu)]), Rady School of Management, UC San Diego. We are grateful for comments from Snehal Banerjee, Francesco D'Acunto, James Doran, Anastassia Fedyk, Laurent Fresard, John Garfinkel, Rawley Heimer, David Hirshleifer, Alberto Rossi and Paul Tetlock. In addition, this draft has benefited from presentations at Social Finance Day, the 2020 Australasian AFBC Conference, Finance in the Cloud III, Midwest Finance Association, Western Finance Association, SFS Cavalcade, Institute for Behavioral and Household Finance Symposium, Shanghai Financial Forefront Symposium, China International Conference in Finance, CQA Fall Conference, European Economic Association, Columbia, Baruch, Baylor, Carnegie Mellon, George Washington, Georgetown, Indiana, Northwestern, Norwegian School of Economics, UIC/Michigan State, University of Victoria, CU-Boulder, Texas Tech, Rice, and London Business School.

# 1 Introduction

Traditional models in finance – where investors have common priors, observe the same public signals, and update their beliefs according to Bayes’ rule – have difficulty explaining the high trading volume observed in financial markets. Difference of opinion models argue that high volume can emerge when investors disagree, either because they interpret information differently (e.g., Varian, 1985; Harris and Raviv, 1993; Kandel and Pearson, 1995) or see different information (e.g., Hong and Stein 1999). But these papers are often silent about *why* processing or information sets are persistently different across investors. This paper proposes and finds evidence for a mechanism that can sustain disagreement: selective exposure to confirmatory information. In other words, investors deliberately choose to consume information that aligns with their prior views, a phenomenon known as *echo chambers*.

Evidence for echo chambers has been found among Republicans and Democrats, churchgoers and non-churchgoers, and smokers and non-smokers (Gentzkow and Shapiro, 2011; Brock and Balloun, 1967). We study echo chambers in the context of financial markets where, for example, a Tesla bull chooses to consume more positive information about Tesla than a Tesla bear, leading to persistent disagreement between bulls and bears about Tesla’s prospects.

At first blush, it might seem odd that investors would behave this way. After all, there is a strong financial incentive to form correct beliefs about prices in a financial market setting. If Republicans watch Fox News and Democrats watch MSNBC, there is no immediate mechanism that causes them financial losses. However, if Tesla bulls purposefully ignore negative information about Tesla this could lead to significant financial harm. Traders have a financial incentive to seek out value-relevant information, regardless of whether it confirms their prior.

Despite this strong incentive, we find overwhelming evidence of selective exposure to confirmatory information when we examine 33 million posts and 14.3 million follower-connections by 400,000 users of StockTwits, one of the largest social networks for investors and traders. Because StockTwits users mark their posts as bullish (or bearish), and because we observe who they choose to follow, we can measure the extent to which users place themselves in echo chambers.

We find that self-described bulls are 5 times more likely to follow a user with a bullish view of the same stock than self-described bears. Moreover, this selective exposure generates significant differences in the newsfeeds of bulls and bears: over a 50-day period, a bull will see 62 more bullish messages and 24 fewer

bearish messages than a bear over the same period. We find a similar pattern with *likes*: bulls will like 38 more bullish messages and 10 fewer bearish messages than a bear over the same 50-day period.

The granularity of the data allow us to focus our analysis at the user-stock-day level. Therefore, we are able to include stock-day fixed effects in our regressions to account flexibly for stock-specific news or attention shocks, which are the focus of the financial attention literature (Tetlock, 2007; Da et al., 2011; García, 2013). In these specifications, we identify echo-chamber effects by comparing the behavior of self-declared bulls and bears for the same stock on the same day. In this case, the degree of selective exposure to information we find is large: declaring as a bull (rather than a bear) today increases the baseline rate of following a bull by 41 percent.

To understand the importance of echo chambers, we examine three sources of heterogeneity. First, we examine whether self-reported experience (Novice, Intermediate, Professional) and activity on StockTwits affect selective exposure. Though we find more pronounced effects for non-professionals and more active users, we also observe significant selective exposure among professionals and less active users. A similar pattern emerges when classifying users as sophisticated based on whether they support their posts with external links.

Second, we find that investors are more likely to seek confirmatory information when they have “skin in the game.” To do this, we identify bullish and bearish posts that also include a declaration of trade (e.g., “\$TSLA. Just added 100 shares.”). We find that users with trade declarations are two to three times more likely to follow same-sentiment users, relative to users who declare the same sentiment but do not trade.

Third, we examine echo chambers around the arrival of news. Surprisingly, we find that selective exposure to information is nearly twice as large around earnings announcements, when we would expect public news to cause convergence in beliefs. In other words, information events push people further into their echo chambers, which makes it more difficult for their beliefs to converge. In this way, we provide a complementary mechanism for the results in Kandel and Pearson (1995), which observes that analyst disagreement and trading volume increase after earnings announcements. Kandel and Pearson (1995) argue that analysts differentially interpret the same public signal, whereas our findings imply that investors choose to be exposed to more polarized information.

If investors selectively expose themselves to information, we would expect information to cluster by sentiment within receivers. For example, if 4 bearish messages and 4 bullish messages are sent out by StockTwits’ users, we would *not* expect most users to receive 50% bearish and 50% bullish messages in

their newsfeed. Instead, we would expect information to be siloed, with a disproportionate share of users receiving only bullish or only bearish signals. This is precisely what we find when we compare the expected number of all-the-same-sentiment messages per user under random assignment to the empirical frequency. For example, when we would expect a user to receive all-the-same-sentiment messages with probability 38%, we see this occur 51% of the time. Moreover, consistent with echo chambers, we find receivers are more likely to receive all bullish (bearish) signals if they have recently declared themselves a bull (bear).

Our final tests consider the implications of echo chambers for returns and trading volume. First, we document an inverse relationship between beliefs on StockTwits and future returns: bullish (bearish) declarations on StockTwits are associated with 1.9% lower (higher) future abnormal returns over the next 5 trading days. However, the size of this under-performance depends on whether the declaration was made inside an echo chamber. For example, for a declaration by a user who has no diversity in their newsfeed over the prior 30-days (i.e. all signals received were the same), the under-performance jumps to 2.3%. On the other hand, for a declaration by a user who has maximum diversity in their newsfeed over the proceeding month (i.e. half the signals were bullish and half were bearish) the under-performance is 1 percentage point smaller. This finding suggests a potentially large welfare cost to selective exposure behavior.

Second, we relate echo chambers to trading volume by constructing measures that capture how information is clustered in the social network. For each stock-day when messages are sent by StockTwits users, we calculate both the mean and standard deviation of the sentiment that they receive. We call the dispersion in the mean of receivers' signals "received disagreement," and the average standard deviation of receivers' signals "received uncertainty." For example, suppose there are 4 new messages about Tesla, 2 bearish and 2 bullish, and 10 StockTwits users see at least one of them. If all 10 users see all 4 messages then received disagreement is low (they all saw the same set of messages) and received uncertainty is high (each of them gets mixed sentiment messages about Tesla). However, if half of them see the 2 bullish messages and the other half see the 2 bearish messages, then received disagreement is high and received uncertainty is low (each of them gets 2 consistent messages about Tesla). In this case, we say information is "siloed," consistent with selective exposure.

When we examine information silos and trading, we find higher trading volume precisely when information silos are more pronounced, i.e., when received disagreement is high and received uncertainty is low. For a standard deviation increase in these information siloing measures, the increase in trading volume is similar to a standard deviation increase in *sender* disagreement. That is, the relationship between volume

and disagreement is related to *both* the dispersion in signals sent as well as the dispersion in signals received.

Our main contribution is to provide evidence of echo chambers in a financial market context. Echo chambers are related to two well-established concepts in the psychology literature: confirmation bias and selective exposure theory. Confirmation bias occurs when individuals systematically acquire or interpret information in support of prior beliefs (Nickerson, 1998). Selective exposure theory is the study of biased information acquisition, of central importance in the study of media and communication (Knobloch-Westerwick, 2014). Combining these concepts, an echo chamber emerges when individuals tilt their *information acquisition* toward sources that *confirm* their prior views.

By studying information acquisition, we introduce a novel perspective to the behavioral finance literature on confirmation bias. Despite a long-standing interest in confirmation bias,<sup>1</sup> the literature largely focuses on how individuals interpret information, both in theoretical (Rabin and Schrag, 1999; Camerer, 1999) and in empirical work (Pouget et al., 2017; Charness and Dave, 2017). Echo chambers are evidence that market participants engage in confirmation-biased information acquisition, which slows the arrival of new information that is inconsistent with an individual's prior. Given the importance of information arrival for belief updating, echo chambers provide a rationale for why beliefs diverge in the first place.

Our evidence contributes to the understanding of selective exposure, which dates back to the original theory of cognitive dissonance (Festinger, 1957).<sup>2</sup> Most of the evidence of selective exposure is derived from surveys and controlled experiments with low stakes which could fail to replicate in real-life decisions (Knobloch-Westerwick and Meng, 2009). Our research overcomes this limitation by showing strong selective exposure effects by individuals in financial markets, which have large economic stakes.

This paper also relates to the economics literature on information avoidance (Golman et al., 2017). However, our findings are distinct from the “optimism maintenance” or motivated beliefs channel, which posits that optimistic beliefs are intrinsically valued, giving rise to wishful thinking (Brunnermeier and Parker, 2005; Benabou, 2015; Banerjee et al., 2019). While motivated beliefs could explain why bulls follow other bulls, our evidence that bears subscribe to other bears implies that the selective exposure effects we observe are not entirely driven by the utility benefits of optimism.

---

<sup>1</sup>Thaler's preface to the *Journal of Economic Perspectives* series on anomalies (Thaler, 1987) argues that confirmation bias could be one explanation for the literature's strict adherence to a rational paradigm.

<sup>2</sup>The literature in psychology, communications and politics has identified other possible reasons for selective exposure. Selected information may be cognitively easier to process (Ziemke, 1980), may reflect judgments about information quality (Fischer et al., 2008), and may be affected by moods and emotions (Valentino et al., 2009). Despite the extensive literature on selective exposure theory and its underlying mechanisms, empirical evidence for the selective exposure hypothesis is mixed (e.g., see critiques in Frey, 1986 and Taber and Lodge, 2006).

Our evidence also relates to the literature on limited and selective attention in financial markets (Barber and Odean, 2008; Golman and Loewenstein, 2016). Most of the empirical literature has focused on attention at the stock-day level either by retail or institutional investors (Da et al., 2011; Ben-Rephael et al., 2017, 2021; Fedyk, 2019), but information on individual information choices is scarce. Several papers expose individuals to exogenous information shocks, and observe their ex post response (e.g., Heimer, 2014; D'Acunto, 2015; Heimer, 2016; Bail et al., 2018). By contrast, we examine agents' *endogenous* information choices after observing their ex ante beliefs. Prior work has also examined the discrete choice to access online account information, and how the timing of account logins relates to periods of market stress (Sicherman et al., 2016; Gargano and Rossi, 2018) or personal financial hardship (Olafsson and Pagel, 2017). Though the timing of accessing account information is a related phenomenon, the selective exposure of investors to information sources on StockTwits is conceptually different. In our setting, users already pay attention to financial information, but systematic cross-sectional differences in the information they choose to consume amplifies dispersion in their initial beliefs. In a related paper, Park et al. (2013) conduct a field experiment on a South Korean message board, finding evidence of confirmation bias in selecting and trading on information. We show that echo chambers are also tightly linked with *aggregate* trading in the U.S. stock market, suggesting that these effects extend beyond social media. In addition, we show that echo chambers lead investors to persistently disagree, leading to sustained trading over several days, and that echo chambers are associated with stronger subsequent beliefs.

This paper also contributes to the growing literature on social media and retail investment (e.g., Welch, 2020; Barber et al., 2021; Eaton et al., 2021) and the trading frenzies surrounding meme stocks (e.g., Pedersen, 2021; Van Wesep and Waters, 2021). The retail investor literature has mixed findings on the effect of attention on investor performance. For example, Gargano and Rossi (2018) provide evidence that account logins can improve performance, whereas Barber et al. (2021) show that the highest attention stocks from Robinhood exhibit negative abnormal returns. By showing that selective attention is associated with negative abnormal returns, we provide a behaviorally grounded mechanism for why high attention stocks exhibit worse investment performance.<sup>3</sup>

Finally, we contribute to the recent literature on sources of disagreement, and the implications of dis-

---

<sup>3</sup>This evidence is distinct from existing research on social media and return predictability. For example, Renault (2017), Broadstock and Zhang (2019), Cookson and Niessner (2020) and Dessaint et al. (2021) consider return predictability and forecasting quality at different horizons. Rather than examining predictability, we use variation in ex post returns as a way to evaluate whether declarations made inside echo chambers are misinformed, analogous to work using account-level data (e.g., Barber and Odean, 2000).

agreement for market outcomes (Banerjee and Kremer, 2010; Banerjee et al., 2018; Giannini et al., 2018; Cookson and Niessner, 2020). This literature has argued that disagreement can arise because of differential interpretation of information (i.e., different models), or from different information sets (e.g., Kandel and Pearson, 1995; Hong and Stein, 1999).<sup>4</sup> Our work most closely relates to this second strand of research, which has focused on gradual information diffusion as an explanation for disagreement and trading (Hong and Stein, 2007; Bailey et al., 2018).<sup>5</sup> However, without a friction that slows information transmission, gradual information diffusion is a puzzling phenomenon. Our contribution is to show that selective exposure to confirmatory information leads to persistent *cross-sectional* information differences, which in turn drives trading volume.<sup>6</sup>

The paper proceeds as follows. In Section 2, we describe the data on following behavior and messages on StockTwits. Section 3 provides our main results on how investors selectively expose themselves to information sources on StockTwits. Section 4 connects our evidence on selective exposure to stock returns and trading volume. Finally, we conclude in Section 5 with implications for future research.

## 2 StockTwits Data

We have message-level data and follower interactions from January 2013 through June 2020 from the investor social network, StockTwits. In this section, we provide background on StockTwits, as well as describe details of the message-level data and follower-network data.

### 2.1 Background on StockTwits

StockTwits is a popular social networking platform for investors to share opinions about securities. For users of the platform, the interface resembles Twitter, using “cashtags” that place a dollar sign before the

---

<sup>4</sup>Recent work at the intersection of politics and finance (e.g., Meeuwis et al., 2019; Cookson et al., 2020) has shown evidence of political model-based disagreement. In contrast, we provide evidence on individuals’ information acquisition choices, and how this affects market outcomes.

<sup>5</sup>For example, Chang et al. (2014) provides evidence of slower financial information diffusion in more linguistically-diverse areas of China. Bailey et al. (2018) shows that differential exposure to US housing price optimism through Facebook connections leads to dispersion of house price expectations. These papers show that different exposures to information affect financial market outcomes, but they are agnostic regarding how individuals *choose* which information to consume.

<sup>6</sup>To see how echo chambers can create trading volume via persistent disagreement, suppose users tend towards neutral sentiment, but bullish (bearish) messages change the sentiment of a user to be more bullish (bearish). If a day 0 bull chooses to follow another bullish user and a day 0 bear chooses to follow another bearish user, both users will drift towards neutrality in the absence of receiving new messages. However, their follows will create a newsfeed which consistently pulls them back to their initial bullish and bearish positions. Such predictable, persistent changes in sentiment can cause disagreement and, hence, trading volume.

security's unique symbol (e.g., \$AAPL or \$BTC for Apple or Bitcoin). Cashtags allow users to aggregate opinions about specific symbols, similar to the role of hashtags on Twitter.

Table 1 presents summary information on the composition of our sample. StockTwits users comprise a cross-section of market participants, ranging across categories of experience from Novice, Intermediate to Professional. Panel (a) of Table 1 shows that most StockTwits users do not select an experience classification, but of those who do identify their level of experience, nearly 20% (15,000) indicate that they are professionals. From a reading of profiles, most professionals on StockTwits work in finance or list professional financial certifications (e.g., CFA charterholders). We report examples of professional investor profiles in the Appendix (Figure A.1). Although StockTwits users are not a perfectly-representative sample of investors, the opinions expressed on StockTwits have been shown to have external reliability – e.g., both Cookson and Niessner (2020) and Giannini et al. (2018) show that different proxies for dispersion of sentiment sensibly relate to market-level trading volume.

Another useful feature of StockTwits from the standpoint of academic research is that the platform encourages users to self-classify their messages using a button that indicates whether a message's sentiment is bullish or bearish. Approximately 80% of sentiment-stamped messages are bullish (Panel (a), Table 1). Further, old messages cannot be deleted from StockTwits, which ensures that the data we extract from StockTwits reflects an unselected view of how users viewed the market at each date in our sample.

### **2.1.1 Message Sample**

We restrict attention to messages that are classified by users as either bullish or bearish, keep symbols with at least 2,000 messages, eliminate “robo users” (users that ever post over 1,000 messages in a single day), and eliminate messages about more than one symbol (so that sentiment that can be directly linked to a specific stock). Our final sample contains approximately 33.4 million messages by nearly 400,000 unique users regarding 1,208 unique symbols (stocks, indexes and other assets). Aggregating to the stock-user-day level, our analysis sample contains approximately 14.4 million observations.

For each message in the sample, we observe the timestamp of the StockTwits post, the user identifier for the individual who posted the message and the self-declared sentiment (bullish, bearish or unclassified). We focus on the user-classified sample, excluding unclassified messages, because we do not wish to take a stand on the sentiment of unclassified messages, and because the sentiment-stamp on StockTwits is a salient signal to potential followers.

### **2.1.2 Follower Sample**

The follower data contain each following decision (i.e., one user follows another), user identifiers of both users involved in the connection and the precise time-stamp of the decision to follow another user. The follower data also contain information on the messages that each user likes, the user identifier for who posted these messages, and the timing of the liking.

Conceptually, we take decisions to follow another user as an information choice. This is true because the followed user's subsequent messages automatically enter the follower's newsfeed. The liking decisions provide complementary information about whether the user interacts with the message in question, thereby giving us indirect insight into both information preference and consumption.

In our tests of selective exposure, we relate the decision to follow other users to recent sentiment declarations by both users. More concretely, we use the user-identifier and the timestamp of the decision to follow another users to link these follower decisions to the message sample. For example, if a user Gary posted a bullish message about \$TSLA on January 4th, thereby declaring himself as a \$TSLA bull, we identify the users that Gary subsequently follows, as well as their declarations about \$TSLA. To the extent that Gary's subsequent follows are disproportionately \$TSLA bulls versus \$TSLA bears, we conclude that Gary selectively exposes himself to information that confirms the views in his initial declaration.

## **2.2 Identifying Bullish versus Bearish Declarations in StockTwits**

For the majority of our tests, we work with the message and follower data at the user-symbol-day level. This aggregation choice alleviates the concern that our findings are driven by a few users who post many messages about the same symbol per day. For users who post multiple messages per symbol-day, we classify a user as bullish (bearish) about a particular stock on date  $t$  if at least 90% of the messages posted by that user for a stock-day express bullish (bearish) sentiment. Our conclusions are not sensitive to the threshold we use in classifying sentiment because users rarely have conflicting sentiment about the same symbol on the same day (this occurs in only 1.4% of observations).

Using this classification, we observe that declared bulls about a particular stock are much more likely than a random person on StockTwits to express bullish sentiment about that same stock over the 50 days after declaring as a bull (see Panel (a) of Figure 1). Symmetrically, in Panel (b) of Figure 1, we observe that an individual who declares as bearish about a stock is also much more likely to continue to express bearish

sentiment over the subsequent 50 days. The within-individual persistence of sentiment about a particular stock is useful because we take an individual's declaration of bullish sentiment about a stock as a statement of their identity as a bull or a bear.

Our analysis focuses on bullish versus bearish sentiment and information acquisition decisions at the symbol-day level. However, optimism could also be a fixed characteristic of an individual, irrespective of the symbol. To evaluate this possibility, we check whether a user's declared sentiment is the same across symbols on a given day. Specifically, in Panel (b) of Table 1, we restrict attention to three subsets of user-day observations in which users make sentiment declarations about multiple stocks on the same day: user-days with declarations about 2 stocks, 3 stocks, and 4 stocks. In each case, we compute the frequency of all-bullish, all-bearish and mixed sentiment declarations, and, as a comparison, the theoretical probability of each possibility given the overall fraction of bullish declarations. Regardless of the number of stocks users declare about on a given day, the empirical frequency of all-bullish and all-bearish is more common than would be expected if the distribution were at random, indicating that optimism is – to some degree – an individual characteristic. However, there is substantial variation in sentiment within-user but across symbols (i.e., days where users express mixed sentiment is certainly well above zero). For this reason, our analyses account for individual heterogeneity in optimism by including individual fixed effects.

In the timing of our tests of selective exposure, decisions about information sources are made at date  $t + k$  ( $k$  days after  $t$ , the day a user self-declares as bullish versus bearish about the stock). We classify StockTwits users who are followed by the original user at date  $t + k$  in the same manner we classified the original user. That is, we say the original user followed another bullish user at date  $t + k$  if at least 90% of the followed user's messages about the *same stock* on date  $t + k$  are bullish (and similarly for bearish sentiment). The intuition is that – because expressed sentiment is persistent – the choice to follow someone who declares as bullish about stock  $s$  on date  $t + k$  is a choice to be exposed to (mostly) bullish information about stock  $s$ .

### **2.3 Echo Chambers by Security**

Table 2 presents lists of the top 10 symbols by intensity of selective exposure to bearish information out of the top 100 symbols by message volume. To identify bearish echo chambers, we keep only user-symbol-day

observations in which the user is a declared bear on day  $t$ . Then, we estimate the specification:

$$\mathbb{1}Follow\ Bear_{i,s,t+1} = \eta_t + \lambda_i + \gamma_s + \varepsilon_{i,s,t} \quad (1)$$

in which the dependent variable  $\mathbb{1}Follow\ Bear_{i,s,t+1}$  is an indicator for whether user  $i$  followed more bearish than bullish users about symbol  $s$  on day  $t + 1$ . The regression includes day ( $\eta_t$ ), user ( $\lambda_i$ ) and symbol ( $\gamma_s$ ) fixed effects. The magnitude of the symbol fixed effects allow us to rank symbols by the intensity of selective exposure among only bearish investors. To identify bullish echo chambers, we estimate the analogous specification.

The bearish echo chambers include stocks and assets that had sustained bullish runs during most of our sample period (2013-2020), and also had vocal groups of users who remained bearish in the presence of the bull run. Consistent with this interpretation, the SPDR S&P500 index ETF – which had its longest bull market in our sample frame – is the top bearish echo chamber in our data set. Other notable stock-level echo chambers in our top 10 list include Beyond Meat, Tesla, Snap, and Bitcoin. The top 10 list of bullish echo chambers tend to be pure play stocks in very particular markets: six of the top ten bullish echo chambers are in stocks of pharmaceutical or medical technology firms (some with their main products in clinical trials).

### 3 Evidence on Echo Chambers

#### 3.1 Graphical Evidence

In this section, we present several pieces of graphical evidence that users who declare as bullish (bearish) about a particular stock selectively expose themselves to information that confirms their initial declaration. To be consistent with the regression analysis in the following section, we perform the graphical analysis at the user-stock-day level.

Figure 2 illustrates the connection between user declarations of sentiment about a particular stock and whether subsequent follows are of users declaring the same sentiment in that stock. On StockTwits, the choice of whom to follow implies future exposure to the followed user's posts because these posts show up in the user's newsfeed. Specifically, Panel (a) of Figure 2 shows how the net number of follows of bullish users per declaration evolves over the 50 days after a user declares as a bull (solid line) or declares as a bear (dashed line). Consistent with echo chambers in sentiment, users who declare as bullish follow significantly

more new users who are also bullish about the same stocks, and this tendency to follow bulls is much greater than for users that declare themselves bearish. The magnitude of this difference is substantial: net follows of bulls increases 0.37 follows per declaration of bullish sentiment at date  $t = 0$ , but net follows of bulls only increase by roughly 0.08 per bearish declaration.

Panel (b) of Figure 2 shows that the relationship between declared sentiment and the type of subsequent follows is symmetric and opposite for the growth of bearish follows. Relative to declared bulls, declared bears follow significantly more new users who are bearish in the same stocks. Although the magnitudes are smaller because there are fewer bearish individuals to follow on StockTwits, the relative ratio is similar. In the 50-day window after declaration, declared bears increase the number of bearish follows by 0.095 per declaration, compared with a 0.026 additional bearish follows per bullish declaration. Simply put, both bullish and bearish users tend to follow other users whose opinions are more similar to their own.

A potential issue with equating decisions of whom to follow (bulls versus bears) with decisions about information sources is that these follows may not manifest into differential exposure to bullish versus bearish information if the followed users do not post much or change their views after the initial declaration. In Figure 3, we address this possibility by relating declarations of bearish versus bullish sentiment to subsequent information in the user's newsfeed. The number of bullish messages in a user's newsfeed is substantially greater for users who declared as bullish on date  $t = 0$  than for users who declared as bearish on date  $t = 0$ . Specifically, over a 50-day period following the user's declaration of bullish versus bearish sentiment, this difference amounts to roughly 62 more bullish messages for a declared bull versus a declared bear. In addition, a declared bull can expect to see 24 fewer bearish messages than a declared bear over this 50 day period.

One concern with the raw messages result in Figure 3 is that it could be driven by a few users who post a disproportionately large number of messages. We address this by counting the number of user impressions or user-days instead of messages (i.e., one bullish post by a user about the stock on date  $t$  is counted as one bullish impression, as is 10 bullish posts by a user about the same stock on the same day). Figure 4 presents the results. Similar to the result for messages, the number of bullish user impressions is substantially greater for users who declared as bullish on date  $t = 0$  than for users who declared as bearish on date  $t = 0$ . Indeed, on a per-day basis, roughly 95% of user impressions are bullish in the newsfeed of a declared bull, whereas only 65% of user impressions are bullish in the newsfeed of a declared bear.

Figures 2 through 4 show that declared bulls and bears selectively follow other users with like-minded

views (Figure 2), thereby leading to more information in the user’s newsfeed that confirms the user’s initial view (Figure 3). However, it is possible that the user may not pay attention to the inflow of posts in their newsfeed. To evaluate this possibility, we examine whether an user is more likely to *like* bearish versus bullish posts after the initial declaration of sentiment: a like implies that an individual read or engaged with the post, as well as approved of its content. Consistent with users actively paying attention to the differential information in their newsfeeds, Figure 5 shows that likes exhibit the same patterns as follows of bulls versus bears and the eventual sentiment in their newsfeeds. In the 50-day window after declaring as a bull or a bear, declared bullish users like more than 41 bullish posts in comparison to 3 likes of bullish posts by declared bearish users.

Beyond the consistency across the figures, it is worth noting how rare it is to have data on information consumption in a finance context. Using a political analogy, observing follows (Figure 2) is like seeing whether an individual records Fox News versus MSNBC, the newsfeed analysis (Figure 4) is like observing whether Fox News has conservative content versus MSNBC’s liberal content, and the analysis of likes (Figure 5) is similar to observing whether individuals actually watch the recorded news programs (via commenting on particular stories or liking particular pieces of information). The level of detail we have in the StockTwits data set is like having individual Nielsen viewership data in the political news arena.

## 3.2 Regression Analysis

In this section, we subject the graphical patterns in the previous section to regression analyses that account for time-varying heterogeneity across securities and individuals.

### 3.2.1 Choosing Information Sources

We begin by considering the decision to add confirmatory information sources to one’s newsfeed (Panel (a) of Figure 2). Specifically, we link declarations of sentiment to subsequent following decisions using a linear probability model of the form:

$$\mathbb{1}Follow\ Bull_{i,s,(t+1 \rightarrow t+k)} = \beta_1 \mathbb{1}Declare\ Bull_{i,s,t} + \lambda_i + \gamma_{s,t} + \varepsilon_{i,s,t} \quad (2)$$

where the dependent variable  $\mathbb{1}Follow\ Bull_{i,s,(t+1 \rightarrow t+k)}$  is an indicator for whether user  $i$  followed more bullish than bearish users about stock  $s$  between dates  $t + 1$  and  $t + k$ . The explanatory variable of interest is

$\mathbb{1}Declare Bull_{i,s,t}$ , which is an indicator equal to 1 if user  $i$  declared bullish sentiment about stock  $s$  on day  $t$  (the day of declaration). The coefficient of interest  $\beta_1$  is the change in the probability of following more bulls than bears (between dates  $t + 1$  and  $t + k$ ) for users who declare as bullish (rather than bearish) about a stock on day  $t$ . This coefficient captures the degree of assortative matching (homophily) in newsfeeds: bears following bears and bulls following bulls.

To absorb time-invariant individual heterogeneity (e.g., differences in following behavior), we include a user fixed effect  $\lambda_i$  in some specifications and a user-symbol fixed effect  $\lambda_{i,s}$  in others. We also absorb all time-varying heterogeneity by security by including symbol-day fixed effects  $\gamma_{s,t}$ . This fixed effects structure improves on the graphical evidence in the figures by accounting for omitted variables at the firm-day level, such as earnings announcements, information releases, media attention, or news more generally. Thus, the coefficient of interest  $\beta_1$  is identified from the bullish declarations about the same firms on the same days by different users. In other words, we use cross-user, within firm-date variation, netting out time invariant user heterogeneity. To account for within-user correlation of errors, standard errors are clustered at the user level.

Table 3 presents the results from estimating equation (2). Column (1) reports the estimated additional likelihood of following more bulls than bears on day one, after declaring oneself a bull (rather than a bear) on day zero. The estimated  $\beta_1$  coefficient in column (1) corresponds to a 54 percent increase in this likelihood, with a similar magnitude (a 53 percent increase, or 2.41 percentage points) over the next 5 days in column (3). This identifies the time-invariant-effect of being bullish versus bearish; both are strongly statistically significant. Columns (2) and (4) repeat the analysis adding a user fixed effect, reducing the magnitudes to around 40 percent of the base rate of following.

Next, we consider how strength of opinion relates to a user's echo chamber behavior. To do this, we replace the  $\mathbb{1}Declare Bull$  indicator with a set of indicator variables capturing how many net bullish declarations a user makes on day  $t$  about a stock. Specifically, we include an indicator for whether the user is bearish about a stock (net bullish declarations  $\leq -1$  on day  $t$ ) and three indicators for net bullish declarations  $= 2$ ,  $= 3$  and  $\geq 4$ . The omitted category is users with one net bullish declaration about a stock. The estimates in column (5) show that the degree of echo chamber behavior is stronger for users with stronger sentiment, and this propensity is monotonic in the net number of bullish declarations. In the Appendix Table A.3, we report specifications with differing fixed effects, and find that this monotonic pattern is present regardless.

Next, we refine the specification by including user-symbol fixed effects in place of user fixed effects.

This change means that we identify selective exposure from the *same* user declaring their sentiment about the *same* symbol at different times. The estimate in column (6) implies that a declared bullish user is 1.08 percentage points more likely to follow more bulls than bears over the next 5 days than a declared bearish user, approximately halving the effect size. This change in magnitude suggests important within-user heterogeneity in the degree of selective exposure.

These point estimates are large relative to the low base rate of decisions to follow others: most users do not choose to follow new users on most days. Given this, columns (6) through (9) consider the subset of observations for which the focal user follows at least one new user after their sentiment declaration. Column (6) indicates that users are 12 percentage points more likely to follow other same-sentiment users. Columns (7) through (9) extend the time window out to fifty days to examine how the selective exposure effect decays over time. We find that the effect declines monotonically, and disappears after thirty days.

One potential concern is that the effects we observe are driven by small stocks. For example, in an attempt to manipulate prices users might post messages with sentiment that is not genuine (i.e., “pump and dump” schemes), which is more likely among smaller stocks. To address this, we restrict the sample to firms with above median market capitalization as of the end of 2019 (approximately \$3.24bn) and re-run the analysis in Appendix Table A.2. The effect sizes are slightly larger once we drop small stocks, implying that the echo chamber effects we identify are not a small-stock phenomenon.

Another potential concern is measurement error, because followed users could declare about many symbols, not just the focal one. To address this concern, we estimate the main specification (columns 1-4 and column 6) only among follows of focused users - those who declare about a single symbol on average, across their time on StockTwits. These users attract fewer follows on StockTwits, but they offer a cleaner measure of the content of the information source because they produce information about a single symbol. We report the estimated specifications in Appendix Table A.4, and find consistent results. Indeed, among this subsample of focused, followed users where measurement error is presumably smaller we find a slightly stronger effect size (62% vs. 54% of the mean in column 1).

Taken together, these regressions present evidence of substantial associative matching (echo chambers) in users’ selection of information sources, after absorbing a wide variety of potential confounding effects.

### 3.2.2 Evidence of Selective Exposure from Trade Declarations

Another possible concern about our setting is that individuals who post on StockTwits do not necessarily have a financial stake in their opinions. We address this concern by analyzing the text of the tweets for indications that the user bought or sold the security which was the subject of the post (e.g., “I just bought \$TSLA” or “I just closed my position in \$SPOT”).

We use these indicators to examine whether selective exposure is stronger or weaker for individuals who have skin in the game. Specifically, we link declarations of sentiment and trading to later decisions to follow other users using a linear probability model:

$$\begin{aligned} \mathbb{1}Follow\ Bull/Bear_{i,s,(t+1 \rightarrow t+k)} = & \beta_1 \mathbb{1}Declare\ Bull_{i,s,t} \\ & + \beta_2 \mathbb{1}Declare\ Bull_{i,s,t} \times \mathbb{1}Trade_{i,s,t} + \lambda_i + \gamma_{s,t} + \epsilon_{i,s,t} \end{aligned} \quad (3)$$

where the dependent variable  $\mathbb{1}Follow\ Bull/Bear_{i,s,(t+1 \rightarrow t+k)}$  is an indicator for whether user  $i$  followed more bullish (or bearish) users about stock  $s$  between dates  $t + 1$  and  $t + k$  (net of unfollows). Relative to the base specification in equation (2), this specification also includes an interaction between a bullish sentiment indicator and whether the user bought or sold the security on the focal day  $t$  ( $\mathbb{1}Declare\ Bull_{i,s,t} \times \mathbb{1}Trade_{i,s,t}$ ). In this specification, the coefficient  $\beta_1$  is the change in the probability of following more bullish users in the days after a user declares as bullish about a stock on day  $t$  *but does not declare a trade*. The coefficient on the interaction  $\beta_2$  captures the change in the baseline selective exposure rate if the user also declares that a trade was made.

Table 4 presents the results from estimating equation (3). The odd columns of the table present the baseline estimates without the interaction, while the even columns introduce the interaction with buying. The baseline estimate in column (1) shows a similar estimate of selective exposure to our main specifications: A declared bullish user is 1.72 percentage points more likely to follow other bullish users between days  $t + 1$  and  $t + 5$  than a declared bearish user. Relative to this benchmark, declared buyers exhibit significantly more selective exposure. The specification in column (2) implies that a declared bullish user is an additional 1.19 percentage points more likely to follow another declared bull between days  $t + 1$  and  $t + 5$  if they also declared a purchase on day  $t$ . That is, if we condition on the users who have declared trades, the degree of selective exposure is greater, not less.

Furthermore, columns (3) and (4) refine the identification by including user-symbol fixed effects. In this

specification, the interaction coefficient is essentially comparing two bullish declarations by a user about the same security – one with a declared purchase and the other without. Using within user-symbol variation, buyers display approximately double the baseline selective exposure effect.

In columns (5) through (8), we estimate the analogous specifications for bears' propensity to follow other bears. Though the magnitude of the change in the probability is smaller than for bulls (around 0.5 percentage points), the degree of selective exposure relative to the base rate is greater (28%-39% of the base rate of following bears). Moreover, we find that bears are around *three* times more likely to follow other bears if they say they have sold the security (relative to bears not declaring a trade), a stronger effect than for bulls.

### 3.2.3 Evidence on Information Flows

We next examine whether the decision to follow someone affects the subsequent information flow observed in the user's newsfeed, and whether the sentiment matches the user's initial declaration.

The specification follows a similar structure to the analysis of follows, except that the dependent variable indicates how much bullish (bearish) information actually is present in the newsfeed after the user declares as bullish (bearish). In the case of bullish information, we estimate:

$$\text{Bullish User Impressions}_{i,s,t+k} = \beta_1 \mathbb{1}\text{Declare Bull}_{i,s,t} + \lambda_i + \gamma_{s,t} + \varepsilon_{i,s,t} \quad (4)$$

where  $\text{Bullish User Impressions}_{i,s,t+k}$  is the number of bullish user impressions about security  $s$  in the newsfeed of user  $i$  (i.e., the distinct number of users who posted a sentiment-stamped message in user  $i$ 's newsfeed on day  $t$ ),  $k$  days after user  $i$  declares as bullish.  $\mathbb{1}\text{Declare Bull}_{i,s,t}$  is the same indicator of declared bullish sentiment we use in our follower specifications. The coefficient of interest  $\beta_1$  represents the expected increase in the number of bullish user impressions in the user's newsfeed on date  $t+k$  after declaring as bullish on date  $t$ . As in the follow regressions, we use symbol-day fixed effects to account for time-varying heterogeneity by security, and user fixed effects to account for individual heterogeneity. We also estimate the analogous specification for *Bearish User Impressions*.

Table 5 presents the results, which confirm that the inflow of information into a user's newsfeed matches the user's initial declaration about the stock. Specifically, in column (1), we estimate that users who declare as bullish about security  $s$  on date  $t$  can expect to see 0.36 more bullish user impressions one day after the

initial bullish declaration. This effect on information flow represents an increase of approximately 17% of the average daily inflow of bullish user impressions in their newsfeed. Columns (3), (5), and (7) consider the impact of a bullish declaration on the inflow of bullish impressions in the user’s newsfeed for time windows that extend to fifty days, with effect sizes between 22% and 23% of the baseline. Unlike the follows result in Table 3, which decreases to no effect after 30 days, the information environment is persistently different through at least fifty days. This illustrates that initial a user’s selective exposure decisions have persistent consequences for their information environment.

The even columns of Table 5 reflect a similar inflow of bearish messages in the days following a bearish declaration about a security. Specifically, we estimate that a user who declares as bearish about security  $s$  on date  $t$  can expect to see 0.12 more bearish user impressions about security  $s$  one day later. Though the expected number of user impressions is smaller for bears than for bulls, this effect is 40% of the average daily inflow of bearish user impressions about a security in their newsfeed. When we extend the time horizon in columns (4), (6) and (8), we continue to see a significant increase in the average number of bearish impressions in the user’s newsfeed, though the effect size is somewhat smaller for longer horizons (19% of the average number of bearish user impressions for days 31 through 50). That is, we observe significant and persistent differences in the information environment of declared bulls compared with declared bears, another indication that users are systematically displaying selective exposure to confirmatory information.<sup>7</sup>

### 3.2.4 Evidence on Information Consumption

The evidence thus far indicates that users select information to be placed into their newsfeeds which reflects their initial views. However, it does not show that this differential exposure to bullish versus bearish information is *received* by the user. Figure 5 addressed this concern by showing that likes of bearish versus bullish posts exhibit the same pattern as follows and information content, thus showing that users receive and interact with the information they see.

We now examine this relation in a regression, analogous to our other specifications:

$$\text{Likes of Bullish Messages}_{i,s,t+k} = \beta_1 \mathbb{1} \text{Declare Bull}_{i,s,t} + \lambda_i + \gamma_{s,t} + \varepsilon_{i,s,t} \quad (5)$$

where the dependent variable is the number of user  $i$ ’s likes of bullish messages about security  $s$  on date  $t + k$  (i.e.,  $k$  days after we observe user  $i$  declare as bullish about security  $s$ ), and  $\mathbb{1} \text{Declare Bull}_{i,s,t}$  is the

---

<sup>7</sup>In the appendix, we report estimates from an analogous specification with messages (rather than user impressions) as the dependent variable. The estimates from the messages specifications yield identical insights. See Table A.5 for details.

same indicator of declared bullish sentiment we use in our previous regressions. The coefficient of interest  $\beta_1$  represents the expected increase in likes of bullish messages on date  $t + k$  after declaring as bullish about security  $s$  on date  $t$ . We include symbol-day fixed effects to account for time-varying heterogeneity by firm, and user fixed effects to account for individual heterogeneity.

Table 6 presents the results. In column (1), we find that declared bulls about security  $s$  on date  $t$  can be expected to like 3.9 additional bullish messages per day about security  $s$  on day  $t + 1$ . This effect represents an increase of 62% of the average daily liked bullish messages for a particular security. Referencing columns (3), (5) and (7), the effect remains significant out to a time horizon of 50 days, but the effect size diminishes somewhat to 36% of the average number of liked messages between days 31 and 50.

Turning to the bearish information consumption specifications in the even columns, we estimate that a declared bear about security  $s$  on date  $t$  likes 2.96 more bearish messages on day  $t + 1$ . From columns (4), (6) and (8), a bearish user will like significantly more bearish messages for horizons out to 50 days after the initial declaration. That is, we observe significant differences in the sentiment of liked messages for declared bulls compared with declared bears, an indication that the selective exposure of the information environment is attended to by the user.

### **3.3 Heterogeneity and Mechanisms**

Next, we turn to evaluating three sources of heterogeneity. First, we examine whether the arrival of news – e.g., on earnings announcement days – leads to a reduction or amplification of the degree of selective exposure. Second, we evaluate whether selective exposure is related to investor characteristics such as experience and activity on StockTwits. Third, we consider whether the returns experienced by users serve to push them out of their echo chambers.

#### **3.3.1 Echo Chambers and the Arrival of News**

First, we examine heterogeneity in the choice of information sources around the announcement of public (earnings) news. This exercise is complementary to Kandel and Pearson (1995), which finds that analysts differentially interpret the public signal (i.e., they use different models) in providing updates around earnings announcements. In our setting, the choice to selectively expose oneself to confirmatory information sources would naturally slow the arrival of the public signal. During periods of information arrival, do users increase

or decrease their degree of selective exposure? We estimate the following:

$$\begin{aligned} \mathbb{1}Follow\ Bull_{i,s,(t+1 \rightarrow t+2)} &= \beta_1 \mathbb{1}Declare\ Bull_{i,s,t} + \lambda_i + \gamma_{s,t} \\ &+ \beta_2 \mathbb{1}Declare\ Bull_{i,s,t} \times \mathbb{1}EA\ day_{s,t+1} + \varepsilon_{i,s,t} \end{aligned} \quad (6)$$

where the specification is similar to the main specification of follows, but it also includes an interaction with an indicator for whether there is an earnings announcement on day  $t + 1$  ( $\mathbb{1}EA\ day_{s,t+1}$ ), that is, the day after the sentiment declaration day  $t$ . The dependent variable is defined for follows on days  $t + 1$  and  $t + 2$  together because earnings announcements on day  $t + 1$  can be released either before market open or after market close. If information sources become more polarized around the arrival of new information, we would expect  $\beta_2 \geq 0$ , but if information sources were to converge, we would expect the opposite ( $\beta_2 < 0$ ). In addition to the main specification, where the dependent variable is an indicator for whether the user follows more bulls than bears, we also estimate specifications that consider whether the individual follows bulls (columns 3 through 5), and bears (columns 6 through 8).

Table 7 presents the results. Column (1) presents the baseline specification as in Table 3, but restricting to the subset of symbols that have earnings announcements – i.e., excluding assets like Bitcoin and ETFs. In column (2), we find that selective exposure to confirmatory information is nearly twice as pronounced upon the arrival of earnings news (0.77 more relative to the baseline selective exposure effect of 1.03), a finding that provides a complementary mechanism to Kandel and Pearson (1995) for why disagreement spikes on earnings days. Column (3) splits the earnings announcement day (EA day) interaction into good and bad news. To do this, we create an indicator for whether the earnings news was positive (revealed by positive abnormal returns in a 3-day window around the earnings day, i.e., days  $t$  to  $t + 2$ ), and its complement. When we interact the declared bull indicator with the positive EA indicator we observe an even larger coefficient. In other words, echo chamber behavior is higher on EA days and even higher when the EA news *aligns* with declared sentiment.

When we split this main effect out separately for bull follows and bear follows, we observe that the increase in selective exposure is greater when earnings news arrives for both types of connections: when bulls follow bulls (column 5) and bears follow bears (column 8). Relative to the baseline effect, this increase in selective exposure is greater for bears (the interaction is nearly twice the main effect) than it is for bulls (the interaction is roughly 70% of the main effect). This result does not reflect an increase in StockTwits

activity on these days, nor an increase in optimism, because the specifications employ day by symbol fixed effects. Instead, we find that both bulls and bears are more likely to put themselves in echo chambers when earnings news arrives.

We also examine how the content of news affects selective exposure for bulls and bears separately. In a regression that only includes earnings days in column (6), we observe that most of the selective exposure effect on the bullish side is driven by days with positive earnings news. In other words, in the presence of positive earnings news, bulls double down on their selective exposure to bullish information. However, in the specification for bear follows in column (9), we find no change in selective exposure for bears on earnings days with positive information. This latter finding is consistent with bears being more likely to seek other bears on earnings days, regardless of the earnings news.

### 3.3.2 Heterogeneity by User Characteristics

Next, we consider two user characteristics that potentially affect echo chamber behavior: investor experience and engagement. First, we consider the role of investor experience, using self-classified user experience categories from StockTwits.<sup>8</sup> If selective exposure to information is a behavioral bias that is costly to the user displaying it, we expect that selective exposure should decline with experience. Thus, we interact experience classifications with the indicator for a user declaring as bullish about symbol  $s$  on date  $t$  in a specification analogous to equation (6).

Table 8 presents estimates from this regression. Consistent with the motivating intuition, we observe that professionals exhibit less selective exposure to information than less experienced users. Specifically, in column (2), novice and intermediate investors exhibit similar selective exposure to the baseline category (i.e., users who do not indicate their experience), whereas professional investors exhibit 0.54 percentage points less selective exposure, which is 29% of the main effect. This estimate is statistically significant at the one percent level.

Importantly, experience moderates but does not eliminate selective exposure. Professional investors still exhibit a significant degree of selective exposure to confirmatory information: a professional user who declares as bullish increases the likelihood of following another bullish user by 1.31%, or approximately one-third of the baseline rate of following more bullish than bearish users (4.54% between days  $t + 1$  to  $t + 5$ ). The fact that professionals on StockTwits exhibit significant echo chamber behavior suggests that

---

<sup>8</sup>Though StockTwits users self-report their experience, Cookson and Niessner (2020) validate the StockTwits experience data.

this phenomenon could have real financial market consequences, a question we address in the next section.

As an alternative to the user-provided experience classification, we construct a measure of investor sophistication based on the frequency a user's post includes external links to support the investment opinion. Specifically, we create a user-level indicator variable that equals one if the user averages at least 1 link per message. This measure has external validity: for users who declare their experience category, professionals are approximately 5 times more likely to share at least one link than are novices. The estimate in column (5) shows that users who post many links are significantly less likely to place themselves into echo chambers than are less sophisticated users who do not support their opinions with external links.

In columns (6) through (8), we estimate heterogeneity in selective exposure by user activity. This interaction allows us to focus on users who consistently use StockTwits for information consumption versus inactive users who may infrequently check on their newsfeeds. We find significantly more selective exposure among more active users. Specifically, we interact  $\mathbb{1}Declare\ Bull$  with an indicator for whether the user is highly active – proxied by following more than the median number of users (column 6), having more than the median number of followers (column 7), and posting more than the median number of messages (column 8). All three columns deliver a consistent message: more active users exhibit greater selective exposure, but the main effect is remains significant. These findings highlight that both inactive users and active users exhibit selective exposure, but active users exhibit more than twice the degree of selective exposure.

### 3.3.3 Heterogeneity by Users' Return Experience

The preceding analysis focused on characteristics that correlate with echo chambers, i.e. the day when sentiment declarations are made (Table 7) or user characteristics (Table 8). In this section, we take sentiment declarations as given, and ask whether events push people out of their echo chamber. To illustrate the empirical design, consider a user who declares herself a Tesla bull on day 0. We know from the preceding analysis that such a user will choose to follow others who are also bullish on Tesla. But what if the user observes poor returns for Tesla? Will she continue the construction of her echo chamber by following yet more Tesla bulls?

Table 9 answers this question. The first seven columns consider bullish day 0 declarations about stock  $i$ , while columns (8) to (14) examine bearish declarations. The dependent variable is an indicator for whether the user increases her net follows of bulls (or bears) on days 11-30. In columns (2) and (3) (and 9-10 for bears), the independent variable is an indicator for whether the return of the entire market is positive

on days 2-10, while in columns (4) and (5) (and 11-12 for bears) the independent variable is an indicator for whether the return on focal stock  $i$  is positive on days 2-10. The negative coefficients for bulls and the positive coefficient for bears on the interactions between declarations and return indicators suggest that counter-attitudinal return information pushes people out of their echo chamber. In other words, when a Tesla bull (bear) experiences negative (positive) returns for Tesla on days 2-10, she becomes *less* likely to follow a like-minded user on days 11-30. In columns (6) and (7) (and 13-14 for bears) we include both market and focal stock interactions together; the coefficients are essentially unchanged. This suggests that counter-attitudinal firm-specific and market-wide returns have independent effects on selective exposure.<sup>9</sup>

### 3.4 Echo chambers and subsequent beliefs

One potentially important consequence of echo chambers is that they may reinforce agents' initial beliefs. To explore this, we estimate:

$$\begin{aligned}
 (\# \text{ Bull Days} - \# \text{ Bear Days})_{i,s,(t+11 \rightarrow t+30)} &= \beta_1 \mathbb{1} \text{Declare Bull}_{i,s,t} & (7) \\
 &+ \beta_2 \mathbb{1} \text{Declare Bull}_{i,s,t} \times \mathbb{1} \text{Net Bull Follows}_{i,s,(t+1 \rightarrow t+10)} \\
 &+ \lambda_i + \gamma_{s,t} + \varepsilon_{i,s,t}
 \end{aligned}$$

In this specification, we examine how frequently a user declares as bullish (bearish) in the future, as a function of whether they had placed themselves in an echo chamber. To be specific about the timing, we observe a user's declaration of bullishness or bearishness on day  $t$ , we then observe their subsequent following behavior between days  $t + 1$  and  $t + 10$ , and finally their declaration behavior on days  $t + 11$  to  $t + 30$ . For example, if a user declares as a Tesla bear on day  $t$ , follows more bears than bulls between days  $t + 1$  and  $t + 10$ , and then makes several more bearish Tesla declarations on days 11 to 30, this would be indicative of an association between echo chamber behavior and subsequent beliefs.

$\beta_1$  captures the baseline effect of users' day  $t$  sentiment on their future declarations (days 11 to 30).  $\mathbb{1} \text{Net Bull Follows}_{i,s,(t+1 \rightarrow t+10)}$  is an indicator for users that followed more bulls than bears over the  $t + 1$  to  $t + 10$  window. Thus,  $\beta_2$  captures how much additional bullishness (bearishness) is expressed by users who constructed echo chambers on days 1 to 10, i.e., this captures how much echo chambers appear to

<sup>9</sup>In a field experiment, Levy (2021) finds that insertions of opposite party news on Facebook decrease political polarization. Our finding that counter-attitudinal returns decrease investment-level selective exposure is analogous, and suggests that as an intervention, insertions of posts from different perspectives into users' newsfeeds might reduce echo chamber behavior.

reinforce initial beliefs. We report the estimated coefficients in Appendix Table A.6. We find that bullish users who choose to place themselves in echo chambers also re-declare their initial sentiment on nearly twice as many days (an additional 1.68 days, relative to the baseline of 2.18 for *Declare Bull*) as those who did not. This near-doubling is also present for Bears. Columns (3) and (4) replace the dependent variable with an indicator for whether a user made more bullish than bearish declarations between days 11 and 30. The estimated coefficients show that users who choose to construct a bullish echo chamber in the first 10 days increase their probability of declaring as a bull, on net, by 15 percentage points over the following 20 days, an increase of nearly 60% of the baseline. Clearly, there is a correlation between echo chamber behavior and subsequent beliefs, either because users with greater conviction place themselves in echo chambers or because echo chambers strengthen beliefs.

## 4 Selective Exposure and Market Outcomes

In this section, we examine the relationship between selective exposure to information and market outcomes such as investor returns and trading volume in the broader market.

### 4.1 Returns

So far, the analysis has been silent on the question of whether being inside an echo chamber affects the quality of a user's declarations; i.e., does it contribute or detract from investment performance? In this section, we use ex post returns to inform that judgment, following a similar approach in the retail trading literature (e.g., Barber and Odean, 2000).<sup>10</sup> For example, if a user makes a bullish declaration from inside an echo chamber, it is likely to be less informed than one made from a newsfeed that is not selectively exposed. In other words, we expect a bullish user inside an echo chamber to receive lower subsequent returns. We estimate the relationship between sentiment declarations and ex post returns using the following specification:

---

<sup>10</sup>The retail trading literature uses the timing of buys and sells with respect to subsequent returns as a measure of the quality of decision-making by retail traders. Similarly, we make no claims about the equilibrium effects of echo chambers on returns; rather, we use the timing of bullish versus bearish declarations with respect to subsequent returns to make a similar judgment about the quality of sentiment declarations inside versus outside echo chambers.

$$\begin{aligned}
\text{Abnormal return}_{i,s,(t+2 \rightarrow t+\tau)} &= \beta_0 \mathbb{1} \text{Declare Bull}_{i,s,t} \times sd(\text{Received Signals})_{i,s,t} & (8) \\
&+ \beta_1 \mathbb{1} \text{Declare Bull}_{i,s,t} + \beta_2 sd(\text{Received Signals})_{i,s,t} \\
&+ X'_{s,t} \delta + \eta_t + \lambda_{i,s} + \gamma_{s,m(t)} + \varepsilon_{i,s,t}
\end{aligned}$$

The dependent variable *Abnormal Return* is percentage abnormal return for stock  $s$  (return minus the CRSP value-weighted market return) in the forward-looking window from 2 to  $\tau$  days after user  $i$  makes a bullish (or bearish) declaration about stock  $s$  on day  $t$ . The main tests employ two time windows: a five-day ( $t+2, t+6$ ) and a ten-day window ( $t+2, t+11$ ).<sup>11</sup> The variable  $\mathbb{1} \text{Bull}_{i,s,t}$  is an indicator that equals 1 if user  $i$  declares as bullish about stock  $s$  on day  $t$ , and the variable  $sd(\text{Received Signals})_{i,s,t}$  equals the standard deviation of sentiment declarations in user  $i$ 's newsfeed about stock  $s$  in the 30 days before day  $t$ . Users in an echo chamber will see a concentration of similar signals and thus have a low  $sd(\text{Received Signals})$ , while those outside an echo chamber will see a diversity of signals and have a high  $sd(\text{Received Signals})$ .<sup>12</sup>

This specification is akin to following users' individual trades, as is common in analyses of transaction-level data, except that we evaluate the performance of trades after user *declarations*. Viewed this way, the coefficient  $\beta_1$  captures the under-performance for an investor who buys on the user's bullish declaration (or sells on the user's bearish declaration). Relative to this baseline, the coefficient on the interaction term,  $\beta_0$ , is the coefficient of interest, capturing how the under-performance differs based on whether the declaration is made in an echo chamber versus not. If selective exposure worsens decision-making, then sentiment declarations made in echo chambers will be associated with weaker *ex post* return performance, and thus,  $\beta_0 < 0$ .

The specifications include day fixed effects ( $\eta_t$ ), user-stock fixed effects ( $\lambda_{i,s}$ ), and stock-month ( $\gamma_{s,m(t)}$ ) fixed effects. The vector of controls  $X_{s,t}$  includes abnormal returns for trading days  $t-30$  to  $t-6$ , abnormal returns for days  $t-5$  to  $t-1$ , the natural logarithm of abnormal Google search volume (*Log GoogleASVI*)

<sup>11</sup>We skip day  $t+1$  in our future return calculations because sentiment declarations can be made after the market close. For example if a sentiment declaration is on Tuesday, day  $t+2$  begins with Thursday's close-to-close return, (measured from Wednesday 4pm to Thursday 4pm).

<sup>12</sup>We calculate  $sd(\text{Received Signals})$  by assigning a value of 1 to bullish signals and  $-1$  to bearish signals as in Cookson and Niessner (2020) and computing the standard deviation. The resulting measure of signal diversity ranges from 0 to 1. For example, a user who saw 4 bullish signals about Tesla and 0 bearish signals about Tesla over the prior 30 days would have  $sd(\text{Received Signals}) = 0$ , while a user that saw 2 bullish signals and 2 bearish signals about Tesla would have a  $sd(\text{Received Signals}) = 1$ .

from the preceding five days, and the interaction between this measure of recent search attention and preceding five-day abnormal returns.<sup>13</sup> We cluster standard errors by user and by stock, which allows observations within a stock to have an arbitrary serial correlation. As noted in Hodrick (1992), this accounts for serial correlation due to overlapping return windows.<sup>14</sup>

The results from estimating equation (8) without including  $sd(Received\ Signals)_{i,s,t}$  are reported in column (1) of Table 10. We find an inverse relationship between beliefs on StockTwits and future returns: bullish (bearish) declarations on StockTwits are associated with 1.88 pp lower (higher) abnormal returns over the next 5 trading days. The magnitude is somewhat larger for the 10-day return window, which gives an estimated under-performance of 2.48 pp. The marginal impact of adding additional days does not increase the magnitude, nor does the return reaction revert (see Figure 6). These negative returns following sentiment declarations suggest opinions on StockTwits are misinformed on average.

In column (2), we add our measure of signal diversity and its interaction with  $\mathbb{1}Bull_{i,s,t}$  to the abnormal return specification (8). Recall that the coefficient of interest is on  $\mathbb{1}Bull_{i,s,t} \times sd(received\ signals)_{i,s,t}$ . In Table 10 we estimate a positive interaction coefficient, indicating that declarations made in echo chambers (i.e., less signal diversity) are associated with greater under-performance. For example, the estimated main effect on  $\mathbb{1}Bull_{i,s,t}$  in column (2) of Table 10 implies that a declaration by a user in a pure echo chamber (i.e.,  $sd(Received\ Signals) = 0$ ) is associated with 2.33 pp under-performance over the 5 day window following the sentiment declaration. By contrast, a declaration by a user with an even split of bearish and bullish signals over the 30 days before  $t$  (i.e. maximum diversity of signals,  $sd(Received\ Signals) = 1$ ) is associated with 0.97 pp less under-performance over the 5-day return window, reducing the under-performance by more than a third. In column (6), the analogous test for a 10-day window yields an estimated under-performance of 3.18 pp for sentiment declarations made in an echo chamber, and under-performance is reduced by 1.49 pp for those users who see maximum signal diversity.

One potential concern is that being in an echo chamber is a stand-in for lack of investor sophistication. For this reason, in columns (3) and (4), we include our set of investor experience indicators – novice, intermediate and professional – interacted with  $\mathbb{1}Bull_{i,s,t}$ . The baseline (omitted) category is users who do not specify their experience. We find a monotonic relationship between experience and under-performance, with

---

<sup>13</sup>This interaction accounts for the possibility that high-attention stocks exhibit short-term reversals in abnormal returns, an effect noted in the literature on retail attention and returns (Barber and Odean, 2008; Barber et al., 2021).

<sup>14</sup>Although Hodrick (1992) proposes an explicit correction for this serial correlation, our clustering by stock is a more conservative approach. Thompson (2011) shows that clustering standard errors by firm fully addresses the overlapping return windows problem.

professionals outperforming intermediates, who outperform novices.<sup>15</sup> The professional and intermediate categories both outperform the baseline (missing experience) and novice categories. Importantly, including these experience controls does not affect our conclusion regarding under-performance and echo chambers. The coefficient on  $\mathbb{1}Bull_{i,s,t} \times sd(Rec. Signals)_{i,s,t}$  changes from 0.97 pp (column 2) to 0.92 pp (column 4) and has a similarly minor effect on this coefficient in the 10-day regressions (columns 6 and 8).

Figure 6 illustrates the dynamics of under-performance in echo chambers. Rather than estimating abnormal returns accumulated over a window, the figure presents the daily abnormal return coefficients for event days 2 through 31. We use the same controls and fixed effects as in equation (8). The under-performance coming from echo-chambers is large in the days after the sentiment declaration, and it decays to approximately zero by event day 13. For those in a pure echo chamber, daily under-performance begins at -0.63 pp on day  $t + 2$  and declines to -0.07 pp on day  $t + 13$ . By contrast, for users with maximum diversity of signals, the under-performance begins at -0.41 pp on day  $t + 2$ , and declines to -0.06 pp by day  $t + 13$ .<sup>16</sup>

The results indicate that the average sentiment declaration on StockTwits is misinformed, which resolves over the following two weeks. Being in an echo chamber appears to exacerbate this phenomenon, suggesting potential welfare consequences to selective exposure behavior for individuals who choose to participate in markets.<sup>17</sup>

## 4.2 Trading volume

Echo chambers have a distinct prediction about the structure of information within and across different users' newsfeeds, which we call *information siloing*. To see how information filters through echo chambers, suppose first that individuals follow other users independently of their sentiment. In this case, we should expect each user's newsfeed to be, on average, representative of the overall distribution of sentiment. By contrast, if individuals place themselves into echo chambers, their received sentiment about a particular stocks will be clustered. Relative to a benchmark that randomly allocates messages to users, users in echo chambers are more likely to see newsfeeds with all the same sentiment, and these messages will be less

<sup>15</sup>When we test the equality of coefficients between  $\mathbb{1}Bull_{i,s,t} \times \mathbb{1}Novice_i$  and  $\mathbb{1}Bull_{i,s,t} \times \mathbb{1}Intermediate_i$  in column (3) we reject the null with a  $p$ -value of 0.023 (and a  $p$ -value of 0.001 for  $\mathbb{1}Bull_{i,s,t} \times \mathbb{1}Professional_i$ ).

<sup>16</sup>In Appendix Table A.7, we conduct a similar analysis of returns at the stock-day level, using the information siloing measures developed in Section 4.2.2. We find that stock-days with more information siloing have a stronger negative relationship between aggregate sentiment and future returns over the subsequent 5 days, 10 days and 30 days.

<sup>17</sup>We are silent on the net welfare effect of investor social media. It could be that such platforms increase household participation in financial markets. However, conditional on their choice to invest, they may make behavioral mistakes. Whether the benefits of participation outweigh the costs of these mistakes is unclear.

representative of the overall distribution.

#### 4.2.1 Information Siloing

To evaluate the degree of information siloing in the StockTwits data, we calculate the likelihood that all of the messages received at the user-stock-day level are the same sentiment, assuming random linkages across users for each combination of messages posted (bullish versus bearish) and number of messages received by a user on that day. For each realization in our data, we compare these theoretical likelihoods to the empirical likelihoods. For example, if the original distribution of signals were 4 bullish and 2 bearish about a stock, but the user only saw two signals, we calculate the theoretical likelihood of all-the-same sentiment (both messages bearish or both messages bullish) as  $\left[ \binom{4}{2} + \binom{2}{2} \right] / \left[ \binom{6}{2} \right] = 47\%$ . If, in the data, we observe that this combination of signals sent leads to newsfeeds of all-the-same sentiment 60% of the time, then this would indicate clustering or information silos.

Figure 7 presents a graphical comparison of the theoretical likelihoods of all-same-sentiment messages in comparison to the empirical likelihoods in our data in 5 percentage point bins. Across the entire distribution, we observe greater clustering than we would observe if information were not siloed. Table 11 presents regression evidence using a linear probability model for whether all received messages are the same sentiment, separately for all-bullish (columns 1-3) and all-bearish (columns 4-6). Holding constant the expected probability of receiving all bullish messages if randomly assigned, a declared bull is 6.72 to 8.02 percentage points more likely to observe all bullish messages. Similarly, declared bears are 2.85 to 3.04 percentage points more likely to observe all bearish messages, holding constant the theoretical likelihood of observing all bearish messages if randomly received. That is, we observe that echo chambers result in significant information siloing.

#### 4.2.2 Measuring Information Siloing

We now construct empirical measures of information siloing driven by echo chambers, and relate these measures to trading volume.

For stock  $s$  at date  $t$ , denote the sentiment of each message (bullish = 1, bearish = -1) in the newsfeed of user  $i$  as  $Sent_{i,s,t,j}$ , and let  $j$  index the messages posted on date  $t$  by individuals followed by user  $i$ . User  $i$  sees  $N_{i,s,t}$  messages at date  $t$ , so  $j \in \{2, \dots, N_{i,s,t}\}$ . With this notation, we can compute the mean and standard

deviation of the sentiment of the  $N_{i,s,t}$  messages:

$$\hat{\mu}_{i,s,t} = \frac{1}{N_{i,s,t}} \sum_{j=1}^{N_{i,s,t}} Sent_{i,s,t,j}$$

$$\hat{\sigma}_{i,s,t} = \sqrt{\frac{1}{N_{i,s,t} - 1} \sum_{j=1}^{N_{i,s,t}} (Sent_{i,s,t,j} - \hat{\mu}_{i,s,t})^2}$$

$\hat{\mu}_{i,s,t}$  and  $\hat{\sigma}_{i,s,t}$  are summary statistics for user  $i$ 's information environment about stock  $s$  on day  $t$ . The mean of the signals  $\hat{\mu}_{i,s,t}$  is user  $i$ 's measure of other users' sentiment about the stock  $s$ , while the standard deviation  $\hat{\sigma}_{i,s,t}$  reflects the dispersion of opinion visible in user  $i$ 's newsfeed about stock  $s$  on day  $t$ .

To measure the degree of selective exposure for a stock  $s$  at day  $t$ , we aggregate these user-level summary statistics to the stock-day level. In an extreme echo chamber, each user would observe no dispersion in opinion within newsfeed, i.e.,  $\hat{\sigma}_{i,s,t} = 0$ . By contrast, users whose information environment is not siloed will tend to see more dispersed opinions within their newsfeed, i.e.,  $\hat{\sigma}_{i,s,t} > 0$ . Thus, one measure of the extent of selective exposure to information is the sample mean across users of  $\hat{\sigma}_{i,s,t}$ . Specifically, if a stock  $s$ , shows up in  $K_{s,t}$  newsfeeds at date  $t$ , we calculate:

$$Received\ Uncertainty_{s,t} = \frac{1}{K_{s,t}} \sum_{i=1}^{K_{s,t}} \hat{\sigma}_{i,s,t}$$

$Received\ Uncertainty_{s,t}$  is the within-user dispersion of their newsfeeds, averaged across users. It is mechanically greater if there is more disagreement in the sent messages, but for a given level of this "sender" disagreement,  $Received\ Uncertainty_{s,t}$  is lower if there is greater selective exposure. Thus, our tests of selective exposure on volume must condition on  $Sender\ Disagreement_{s,t}$ , which is the standard deviation of opinion about stock  $s$  on day  $t$ , defined following the literature (Antweiler and Frank, 2004).

In addition to  $Received\ Uncertainty$ , a complementary measure of selective exposure to information is the cross-user dispersion in the mean of their signal about stock  $s$  at date  $t$ , a measure we call  $Received\ Disagreement_{s,t}$ . Intuitively, if users choose to follow like-minded individuals, there will be a marked difference between the distributions of sentiment signals that are sent ( $Sender\ Disagreement$ ), and those that are received, which

we calculate as:

$$Received\ Disagreement_{s,t} = \sqrt{\frac{1}{K_{s,t} - 1} \sum_{i=1}^{K_{s,t}} (\hat{\mu}_{i,s,t} - \hat{\mu}_{s,t})^2}.$$

As selective exposure to information increases, we expect the cross-user dispersion of received signals to increase. Similar to *Received Uncertainty*, greater *Sender Disagreement* mechanically implies that *Received Disagreement<sub>s,t</sub>* is higher. However, even controlling for the level of *Sender Disagreement*, *Received Disagreement* is increasing in selective exposure. This is because selective exposure implies that users construct their personal network (through which messages are distributed) to be more homogeneous in sentiment, which leads users to *receive* a distribution of messages that is systematically different from the *sent* message distribution, on average.

#### 4.2.3 Information Silos and Turnover

We can then estimate the effect of information siloing on abnormal log stock turnover as follows:

$$Y_{s,t} = \beta_1 Sender\ Disagreement_{s,t} + \beta_2 Received\ Disagreement_{s,t} + \beta_3 Received\ Uncertainty_{s,t} + X'_{s,t} \delta + \eta_t + \gamma_{s,m(t)} + message_{s,t} + \varepsilon_{s,t}, \quad (9)$$

where  $Y_{s,t}$  is abnormal log turnover of stock  $s$  on date  $t$ , and  $X_{s,t}$  are control variables following Cookson and Niessner (2020).<sup>18</sup> We also include day and stock-month fixed effects ( $\eta_t$  and  $\gamma_{s,m(t)}$ ), as well as fixed effects for eight bins ( $message_{s,t}$ ) capturing the number of messages about a given stock on that day. In this specification, the coefficients of interest are  $\beta_2$  and  $\beta_3$ .  $\beta_2$  measures how increasing received uncertainty (i.e., increasing the dispersion of messages that each user sees – less selective exposure) is associated with abnormal turnover. If echo chambers lead to information siloing that generates trading, we expect  $\beta_2 < 0$ .  $\beta_3$  measures how increasing received disagreement (i.e. more dispersion of received signals across users – more selective exposure) relates to turnover; we expect  $\beta_3 > 0$ . These two *Received* measures

<sup>18</sup>The control variables are: the previous day's *Abnormal Log Turnover*, an indicator variable for media attention at the stock-day level (whether the stock was mentioned in the Dow Jones Newswire, which includes the Wall Street Journal), recent volatility (last five days), and recent abnormal returns (last five, and previous 25 trading days). We also add the natural logarithm of abnormal Google search volume. This variable is calculated following Niessner (2016): we take the daily Google SVI data for each symbol and divide by its median SVI between days  $t - 56$  and  $t - 35$ . We take the natural logarithm of this data, and replace missing values (caused by a missing median) with zero. Note that the SVI data come from 200 day downloads with a day of overlap that we concatenate to ensure consistency across time.

capture different aspects of selective exposure behavior. Thus, our preferred specification includes both *Received Disagreement<sub>st</sub>* and *Received Uncertainty<sub>st</sub>*.

We now link daily abnormal stock turnover to the measures of disagreement at both the sender and receiver levels, and to the dispersion in the received signal (*Received Uncertainty<sub>s,t</sub>*). Table 12 reports the results from estimating the specifications in equation (9). To ease interpretation, we subtract the mean and divide by the standard deviation (both calculated over the whole sample period) for both disagreement measures, as well as for *Received Uncertainty<sub>s,t</sub>*. Column (1), which includes the *Sender Disagreement<sub>s,t</sub>* measure by itself, provides a somewhat smaller estimate (0.017) to the equivalent specification in Cookson and Niessner (2020), whose data end in 2014. In our sample, a one standard deviation increase in disagreement increases abnormal turnover by 4% of its mean.

Column (2) adds the *Received Disagreement<sub>s,t</sub>* measure as a regressor. Holding constant the amount of sender disagreement, greater dispersion in the signals users receive indicates greater dispersion in information sets. However the magnitude of the coefficient on *Received Disagreement<sub>s,t</sub>* is smaller than the coefficient on *Sender Disagreement<sub>s,t</sub>* (0.004 versus 0.015).

Column (3) includes the average within-newsfeed dispersion (*Received Uncertainty<sub>s,t</sub>*), and we estimate a negative and statistically significant coefficient. That is, on stock-days in which selective exposure to information reduces the dispersion of sentiment observed by users, we see greater turnover. The magnitude of the coefficient on *Received Uncertainty<sub>s,t</sub>* amounts to over one half of the main effect of sender disagreement.

Columns (4) and (5) consider both the sender and receiver measures in the same specification. The reported magnitudes in column (4) are similar to those in column (5), which omits control variables. Taken together, both dispersion of opinion (measured via sender disagreement) and information siloing (measured by received disagreement and received uncertainty) contribute similarly to stock turnover. A one standard deviation increase in *Sender Disagreement<sub>s,t</sub>* increases abnormal turnover by approximately 5.8% of the mean, which is similar to combined effect of the information siloing variables (i.e., reducing *Received Uncertainty<sub>s,t</sub>* and increasing *Received Disagreement<sub>s,t</sub>* each by one standard deviation).<sup>19</sup>

---

<sup>19</sup>Because many StockTwits users are retail investors, it is natural to directly investigate retail trading. To do this, we follow Boehmer et al. (2021) and identify retail trades in the TAQ database from their preferential pricing. Based on this classification, we construct abnormal logged retail turnover, and the number of retail trades, buys and sells, all at the stock-day level. In Appendix Table A.8, we confirm that abnormal trading volume by retail investors relates to disagreement and information siloing in the same way as aggregate stock-day trading. However, while we find an increase in the *level* of retail trading, in untabulated results, we find no evidence of an increase in the *percentage* of retail trades, perhaps because the presence of retail investors improves market conditions for institutions (e.g., Cookson et al. (2021)).

#### 4.2.4 Information Silos and the Persistence of Disagreement

A common puzzle in the trading volume literature is the persistence of disagreement (Kandel and Pearson, 1995; Hong and Stein, 2007). We interpret our results as evidence of echo chambers contributing to sustaining disagreement, leading to more trading volume. To evaluate this idea, we measure daily disagreement in financial markets using the *Sender Disagreement* variable, and we run a regression of *future Sender Disagreement* on day  $t$  measures of *Sender Disagreement* and our information silo measures (*Received Disagreement* and *Received Uncertainty*) at the stock-day level:

$$\begin{aligned} \text{Sender Disagreement}_{s,t+x} = & \beta_1 \text{Sender Disagreement}_{s,t} + \beta_2 \text{Received Disagreement}_{s,t} \quad (10) \\ & + \beta_3 \text{Received Uncertainty}_{s,t} + \eta_t + \alpha_s + \text{message}_{s,t} + \varepsilon_{s,t} \end{aligned}$$

where  $\text{Sender Disagreement}_{s,t+x}$  is the (sender) disagreement measure about stock  $s$  on day  $t+x$  and  $x$  represents the number of days after the focal day  $t$ . We also include day ( $\eta_t$ ), stock ( $\alpha_s$ ) and message bins ( $\text{message}_{s,t}$ ) fixed effects, separately estimating this for different days, ranging from 1 to 50 days after  $t$ . In each case, standard errors are double clustered at the stock and day level.

Figure 8 presents the estimated coefficients and 95% confidence intervals for each of the information siloing measures. In Appendix Table A.9, we confirm that this relationship is not sensitive to the particular fixed effects we employ. Echo chambers appear to lead to persistent disagreement: while received uncertainty is unrelated to future sender disagreement, received disagreement is positively and persistently related to disagreement over the next 50 days, and the relationship fades gradually over time. Because we also control for sender disagreement as of day  $t$ , this relationship is not due to auto-correlation with past disagreement.

In summary, these findings suggest that echo chambers, which lead people to see different information from one another, have meaningful consequences for both the persistence of disagreement.

## 5 Conclusion

Selective exposure to confirmatory information has been documented in a variety of settings, from politics to vehicle ownership. It appears that once people form a belief about immigration or Chevy trucks, they selectively choose information that supports their belief and avoid information which contradicts it. By all

accounts, selective exposure appears to be a broad phenomenon.

This paper shows the phenomenon extends to an unlikely setting, financial markets, where users have a strong incentive to get prices right. Nevertheless, we find users behave the same way humans do in other settings: by following users who share their beliefs, they build a personalized newsfeed that supports their original views. This behavior is not doing investors any favors: we find that beliefs formed in echo chambers are associated with poor ex-post returns.

Moreover, selective exposure is a natural candidate to explain some persistent disagreement in financial markets, and we provide evidence that it is positively related to trading volume. To the extent that selective exposure drives disagreement in financial markets, there are still many unanswered questions. For example, how is the rapidly changing technological and information environment affecting the tendency to selectively expose? Thirty years ago, investors could get financial information from only a handful of sources. Today, as our study demonstrates, they have thousands of choices. Does technological innovation liberate those who would want to selectively expose and lead to more disagreement? More generally, there are many other areas in financial markets where agents have initial views and then make choices about the information they collect: board members have views on managers and collect information for the purposes of monitoring, analysts have views on firms and then collect information to make recommendations, rating agencies have views on firms and then collect information to update their ratings, and so on. To what extent does selective exposure lead agents to have views which are “too sticky” in these other settings? We leave these questions and others for future research.

## References

- Antweiler, W. and M. Z. Frank (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance* 59(3), 1259–1294.
- Bail, C. A., L. P. Argyle, T. W. Brown, J. P. Bumpus, H. Chen, M. F. Hunzaker, J. Lee, M. Mann, F. Merhout, and A. Volfovsky (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences* 115(37), 9216–9221.
- Bailey, M., R. Cao, T. Kuchler, and J. Stroebel (2018, December). The economic effects of social networks: Evidence from the housing market. *Journal of Political Economy* 126(6), 2224–2276.
- Banerjee, S., J. Davis, and N. Gondhi (2018, June). When transparency improves, must prices reflect fundamentals better? *Review of Financial Studies* 31(6), 2377–2414.
- Banerjee, S., J. Davis, and N. Gondhi (2019). Choosing to disagree in financial markets. *Working Paper*.
- Banerjee, S. and I. Kremer (2010). Disagreement and learning: Dynamic patterns of trade. *The Journal of Finance* 65(4), 1269–1302.
- Barber, B. M., X. Huang, T. Odean, and C. Schwarz (2021). Attention Induced Trading and Returns: Evidence from Robinhood Users. *Working Paper*.
- Barber, B. M. and T. Odean (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance* 55(2), 773–806.
- Barber, B. M. and T. Odean (2008). All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *Review of Financial Studies* 21(2), 785–818.
- Ben-Rephael, A., B. I. Carlin, Z. Da, and R. D. Israelsen (2021). Information consumption and asset pricing. *The Journal of Finance* 76(1), 357–394.
- Ben-Rephael, A., Z. Da, and R. D. Israelsen (2017). It Depends on Where You Search: Institutional Investor Attention and Under-reaction to News. *Review of Financial Studies* 30(9), 3009–3047.
- Benabou, R. (2015). The Economics of Motivated Beliefs. *Revue d'Economie Politique* 125(5), 665–685.
- Boehmer, E., C. M. Jones, X. Zhang, and X. Zhang (2021). Tracking retail investor activity. *The Journal of Finance* 76(5), 2249–2305.
- Broadstock, D. C. and D. Zhang (2019). Social-media and intraday stock returns: The pricing power of sentiment. *Finance Research Letters* 30, 116–123.
- Brock, T. C. and J. L. Balloun (1967). Behavioral receptivity to dissonant information. *Journal of Personality and Social Psychology* 6(4p1), 413.
- Brunnermeier, M. K. and J. A. Parker (2005). Optimal Expectations. *American Economic Review* 95(4), 1092–1118.
- Camerer, C. (1999, September). Behavioral economics: Reunifying psychology and economics. *Proceedings of the National Academy of Sciences* 96, 10575–10577.
- Chang, Y.-C., H. G. Hong, L. Tiedens, N. Wang, and B. Zhao (2014). Does diversity lead to diverse opinions? Evidence from languages and stock markets. *Working Paper* (168), 13–16.

- Charness, G. and C. Dave (2017). Confirmation bias with motivated beliefs. *Games and Economic Behavior* 104(1), 1–23.
- Cookson, J. A., J. E. Engelberg, and W. Mullins (2020, December). Does partisanship shape investor beliefs? Evidence from the COVID-19 pandemic. *Review of Asset Pricing Studies* 10(4), 863–893.
- Cookson, J. A., V. Fos, and M. Niessner (2021). Does disagreement facilitate informed trading? evidence from activist investors: Evidence from activist investors. *SSRN Working Paper*(January 12, 2021).
- Cookson, J. A. and M. Niessner (2020, February). Why don't we agree? Evidence from a social network of investors. *Journal of Finance* 75(1), 173–228.
- Da, Z., J. E. Engelberg, and P. Gao (2011). In search of attention. *Journal of Finance* 66(5), 1461–1499.
- D'Acunto, F. (2015). Tear down this wall street: The effect of anti-market ideology on investment decisions. *Working Paper*.
- Dessaint, O., T. Foucault, and L. Frésard (2021). Does alternative data improve financial forecasting? the horizon effect. *CEPR Discussion Paper No. DP15786*.
- Eaton, G. W., T. C. Green, B. Roseman, and Y. Wu (2021). Zero-commission individual investors, high frequency traders, and stock market quality. *High Frequency Traders, and Stock Market Quality* (January 2021).
- Fedyk, A. (2019). Front-Page News: The Effect of News Positioning on Financial Markets. *Working Paper*.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Row, Peterson & Company.
- Fischer, P., S. Schulz-Hardt, and D. Frey (2008). Selective exposure and information quantity: How different information quantities moderate decision maker's preference for consistent and inconsistent information. *Journal of Personality and Social Psychology* 94, 231–244.
- Frey, D. (1986). Recent research on the selective exposure to information. *Advances in Experimental Social Psychology* 19(41-80).
- García, D. (2013). Sentiment during recessions. *The Journal of Finance* 68(3), 1267–1300.
- Gargano, A. and A. G. Rossi (2018). Does it pay to pay attention? *The Review of Financial Studies* 31(12), 4595–4649.
- Gentzkow, M. and J. M. Shapiro (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics* 126(4), 1799–1839.
- Giannini, R., P. Irvine, and T. Shu (2018). The convergence and divergence of investors' opinions around earnings news: Evidence from a social network. *Journal of Financial Markets*.
- Golman, R., D. Hagmann, and G. Loewenstein (2017). Information Avoidance. *Journal of Economic Literature* 55(1), 96–135.
- Golman, R. and G. Loewenstein (2016). Information Gaps: A Theory of Preferences Regarding the Presence and Absence of Information. *Decision* 5(3), 143–164.
- Harris, M. and A. Raviv (1993). Differences of opinion make a horse race. *Review of Financial Studies* 6(3), 473–506.

- Heimer, R. Z. (2014). Friends do let friends buy stocks actively. *Journal of Economic Behavior & Organization* 107B(11), 527–540.
- Heimer, R. Z. (2016). Peer pressure: Social interaction and the disposition effect. *Review of Financial Studies* 29(11), 3177–3209.
- Hodrick, R. J. (1992). Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Review of Financial Studies* 5(3), 357–386.
- Hong, H. and J. C. Stein (1999). A Unified Theory of Underraction, Momentum Trading, and Overaction in Asset Markets. *The Journal of Finance*.
- Hong, H. and J. C. Stein (2007). Disagreement and the stock market. *The Journal of Economic Perspectives*, 109–128.
- Kandel, E. and N. D. Pearson (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy* 103(4), 831–872.
- Knobloch-Westerwick, S. (2014). *Choice and preference in media use: Advances in selective exposure theory and research*. Routledge.
- Knobloch-Westerwick, S. and J. Meng (2009). Looking the Other Way: Selective Exposure to Attitude-Consistent and Counterattitudinal Political Information. *Communication Research* 36(3), 426–448.
- Levy, R. (2021). Social media, news consumption, and polarization: Evidence from a field experiment. *American economic review* 111(3), 831–70.
- Meeuwis, M., J. A. Parker, A. Schoar, and D. I. Simester (2019). Belief Disagreement and Portfolio Choice. *Working Paper*.
- Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology* 2(2), 175–220.
- Niessner, M. (2016). Strategic Disclosure Timing and Insider Trading. *Working Paper*.
- Olafsson, A. and M. Pagel (2017). The ostrich in us: Selective attention to financial accounts, income, spending and liquidity. *NBER Working Paper No. 23945*.
- Park, J., P. Konana, B. Gu, A. Kumar, and R. Raghunathan (2013). Information valuation and confirmation bias in virtual communities: Evidence from stock message boards. *Information Systems Research* 24(4), 1050–1067.
- Pedersen, L. H. (2021). Game on: Social networks and markets. *Working Paper*.
- Pouget, S., J. Sauvagnat, and S. Villeneuve (2017). A Mind Is a Terrible Thing to Change: Confirmatory Bias in Financial Markets. *Review of Financial Studies* 30(6), 2066–2019.
- Rabin, M. and J. L. Schrag (1999). First Impressions Matter: A Model of Confirmatory Bias. *Quarterly Journal of Economics* 114(1), 37–82.
- Renault, T. (2017). Intraday online investor sentiment and return patterns in the us stock market. *Journal of Banking & Finance* 84, 25–40.
- Sicherman, N., G. Loewenstein, D. J. Seppi, and S. P. Utkus (2016). Financial Attention. *Review of Financial Studies* 29(4), 863–897.

- Taber, C. S. and M. Lodge (2006). Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science* 50(3), 755–769.
- Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance* 62(3), 1139–1168.
- Thaler, R. H. (1987). Anomalies: The January Effect. *Journal of Economic Perspectives* 1(1), 197–201.
- Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics* 99(1), 1–10.
- Valentino, N. A., A. J. Banks, V. L. Hutchings, and A. K. Davis (2009). Selective Exposure in the Internet Age: The Interaction between Anxiety and Information Utility. *Political Psychology* 30, 591–613.
- Van Wesep, E. D. and B. Waters (2021). All-in investors and unstable asset prices. *Working Paper*.
- Varian, H. R. (1985). Divergence of Opinion in Complete Markets: A Note. *Journal of Finance*.
- Welch, I. (2020). The wisdom of the robinhood crowd. Technical report, National Bureau of Economic Research.
- Ziemke, D. A. (1980). Selective exposure in a presidential campaign contingent on certainty and salience. *Annals of the International Communication Association* 4(1), 491–511.

## 6 Figures

Figure 1: Is sentiment persistent?

Panel (a) presents the probability that a bullish user stays bullish on each of the subsequent 50 days after their declaration (solid line). The dotted line shows the unconditional frequency of bullishness in the data. Panel (b) presents the analogous table for bearish users. We identify individuals as bullish or bearish about a symbol on event day 0 if more than 90 percent of their messages are bearish or bullish.

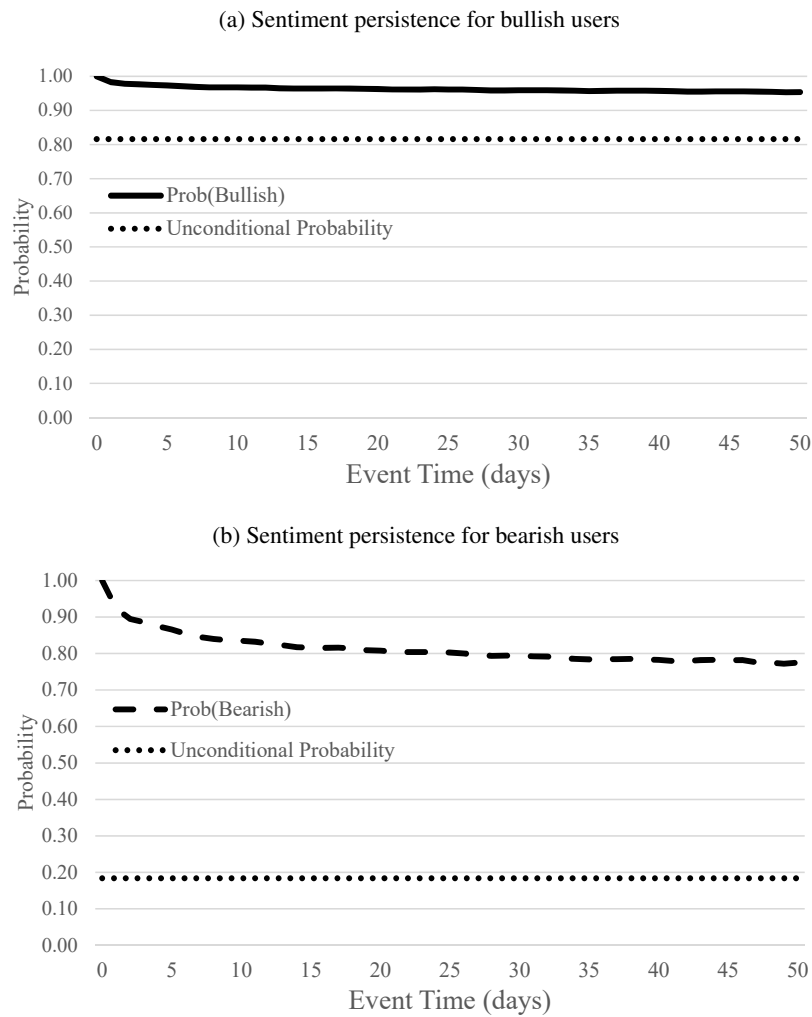


Figure 2: Who do users follow?

Panel (a) plots the cumulative number of net new follows of bullish users by an individual; Panel (b) is for net new follows of bearish users. We identify individuals as bullish or bearish about a symbol on event day 0 if more than 90 percent of their messages are bearish or bullish.

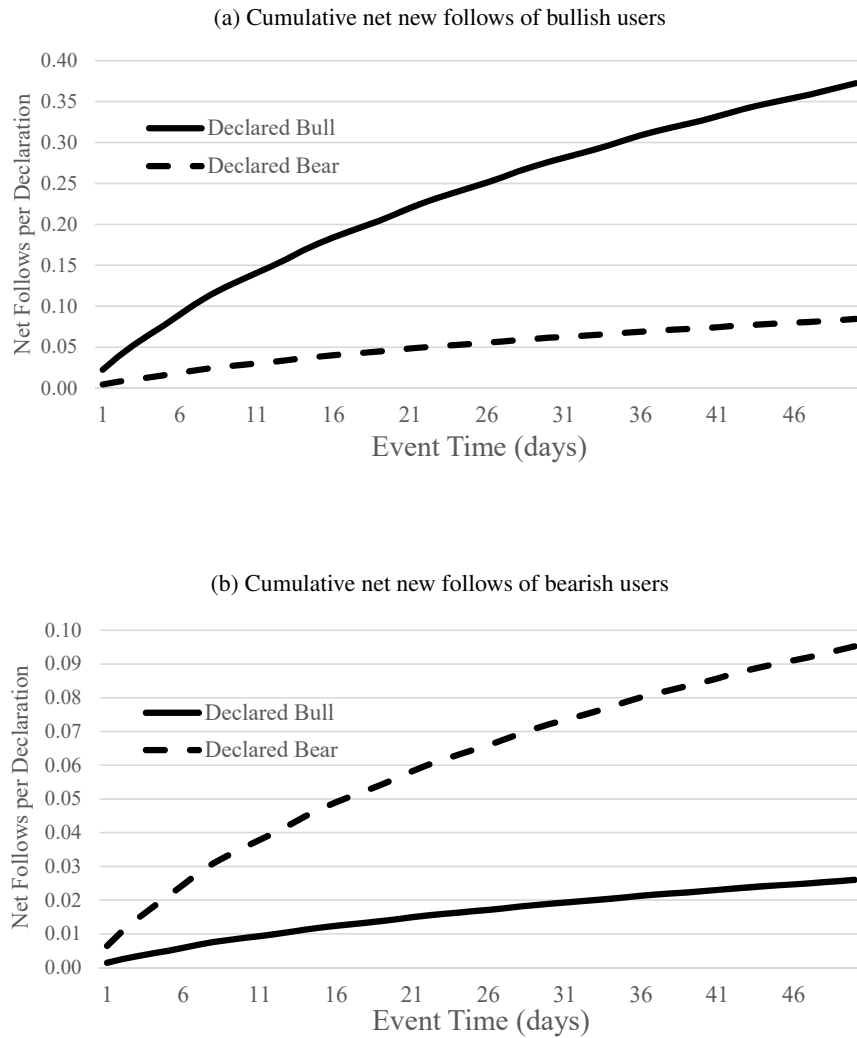


Figure 3: Do users' newsfeeds match their sentiment? Messages

Panel (a) plots the average number of bullish messages in a user's newsfeed following their bullish declaration about that symbol; Panel (b) plots the number of bearish messages. We identify individuals as bullish or bearish about a symbol on event day 0 if more than 90 percent of their messages are bearish or bullish.

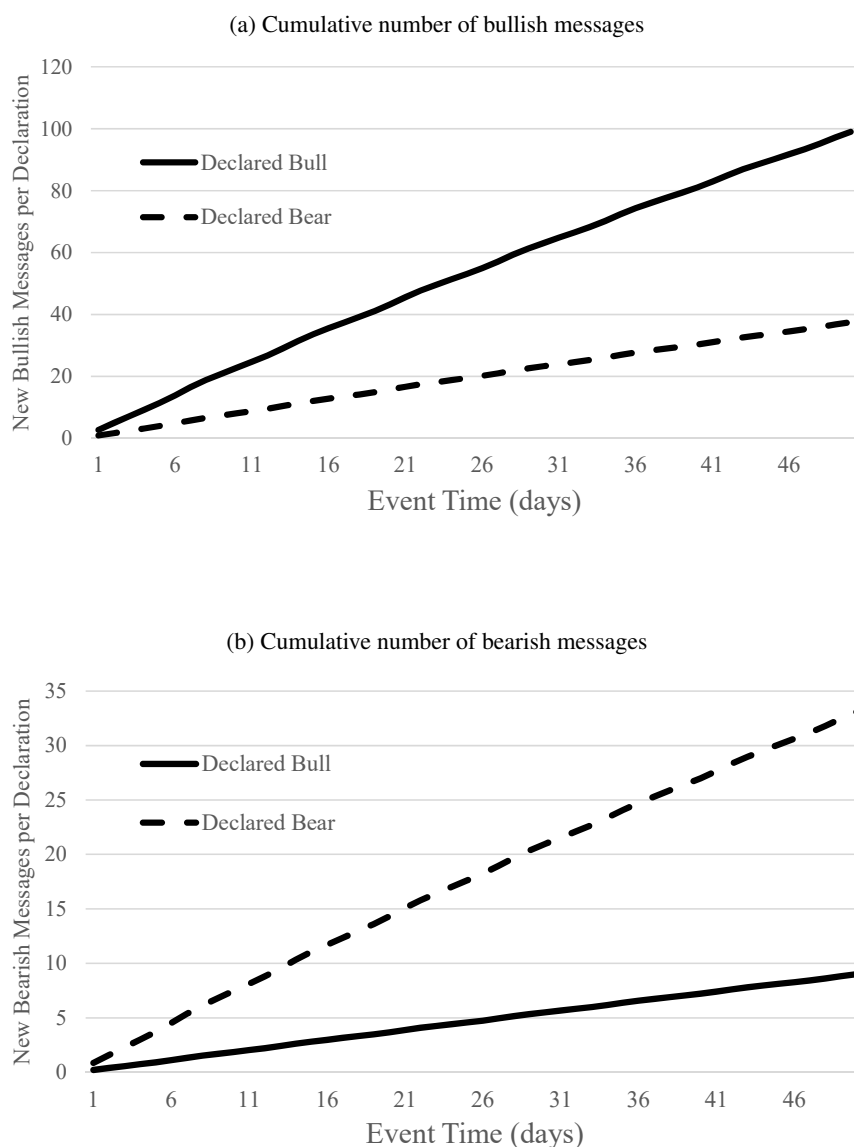


Figure 4: Do users' newsfeeds match their sentiment? User impressions

Panel (a) plots the number of bullish user impressions (i.e., the number of distinct users who post bullish messages in the focal user's newsfeed that day); Panel (b) plots the number of bearish user impressions. A bullish (bearish) user impression occurs on a security-day when an individual who is followed by the user posts at least one message with bullish (bearish) sentiment. We identify individuals as bullish or bearish about a symbol on event day 0 if more than 90 percent of their messages are bearish or bullish.

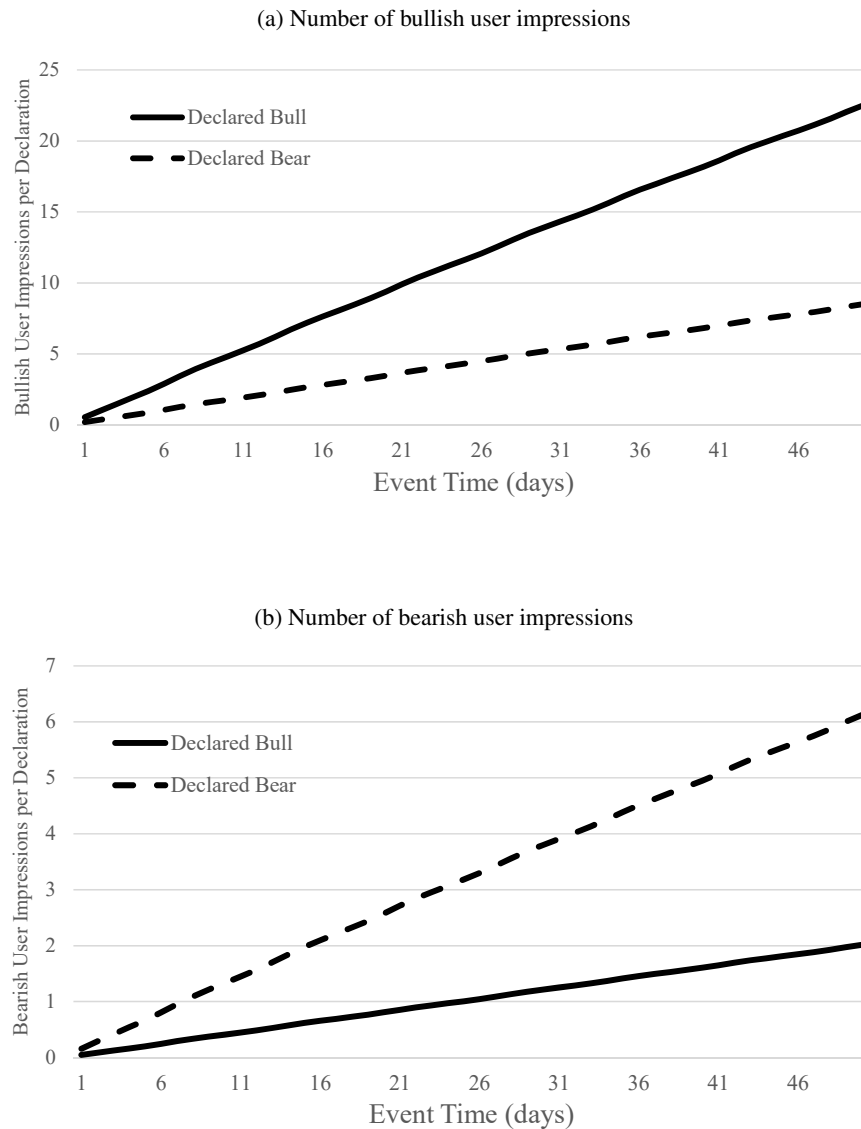


Figure 5: Do users' likes match their sentiment?

Panel (a) plots the number of bullish likes by the user; Panel (b) plots the number of bearish likes by the user. We identify individuals as bullish or bearish about a symbol on event day 0 if more than 90 percent of their messages are bearish or bullish.

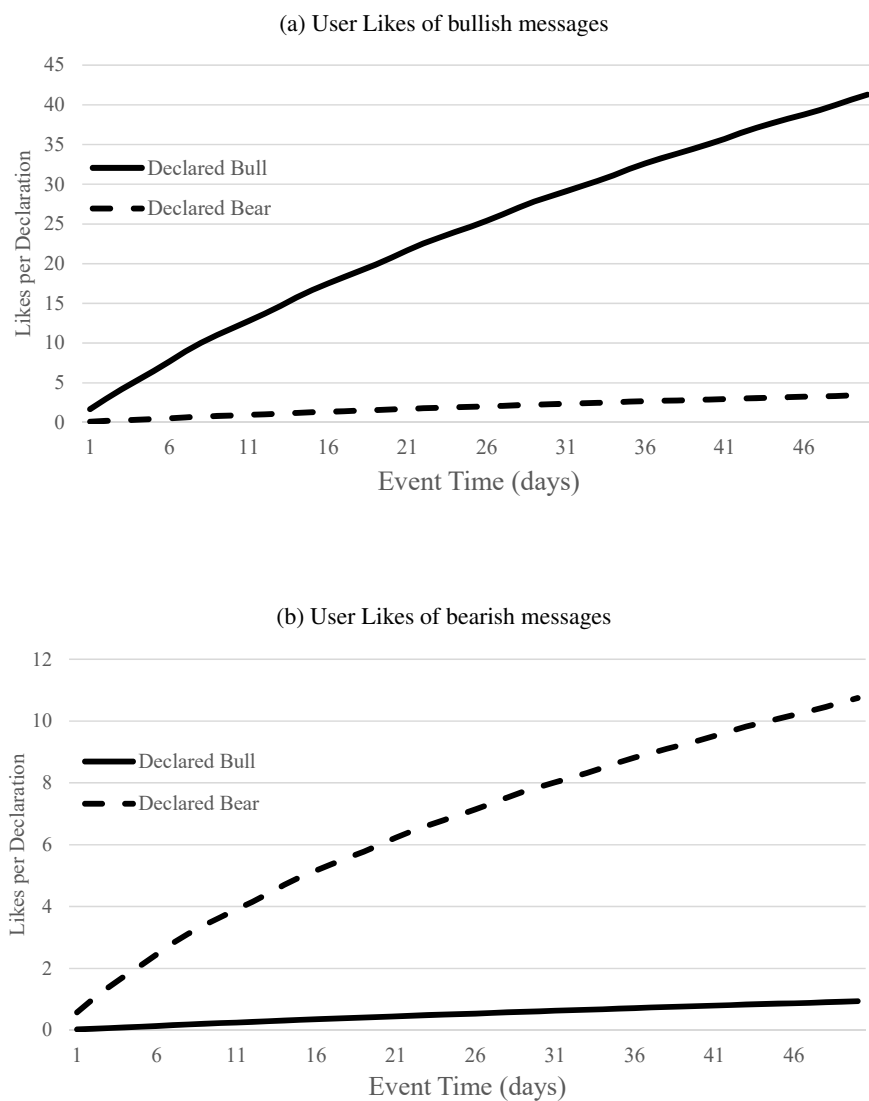


Figure 6: Comparing the return predictability of sentiment declarations made inside and outside echo chambers

This figure presents the coefficients from equation (8) associated with a user's bullish (bearish) declaration on day  $t = 0$ . The dependent variable is abnormal returns (in percentage points, 0-100) on days  $t + \tau$  (where  $\tau$  ranges from 2 through 31 days). We skip day  $t + 1$  because sentiment declarations on day  $t$  can be made after market close. The black diamonds illustrate the return under-performance that results from trading on users' sentiment declarations. The blue triangles denote the estimated coefficients on  $bull \times sd(received\ signals)_{i,s,t}$ , and illustrate the degree to which maximum received signal diversity over the thirty days preceding day  $t$  (i.e., when  $sd(received\ signals) = 1$ ) mitigates return under-performance on each day following  $t$ . The vertical bars represent 95% confidence intervals, clustering standard errors separately by user and by stock.

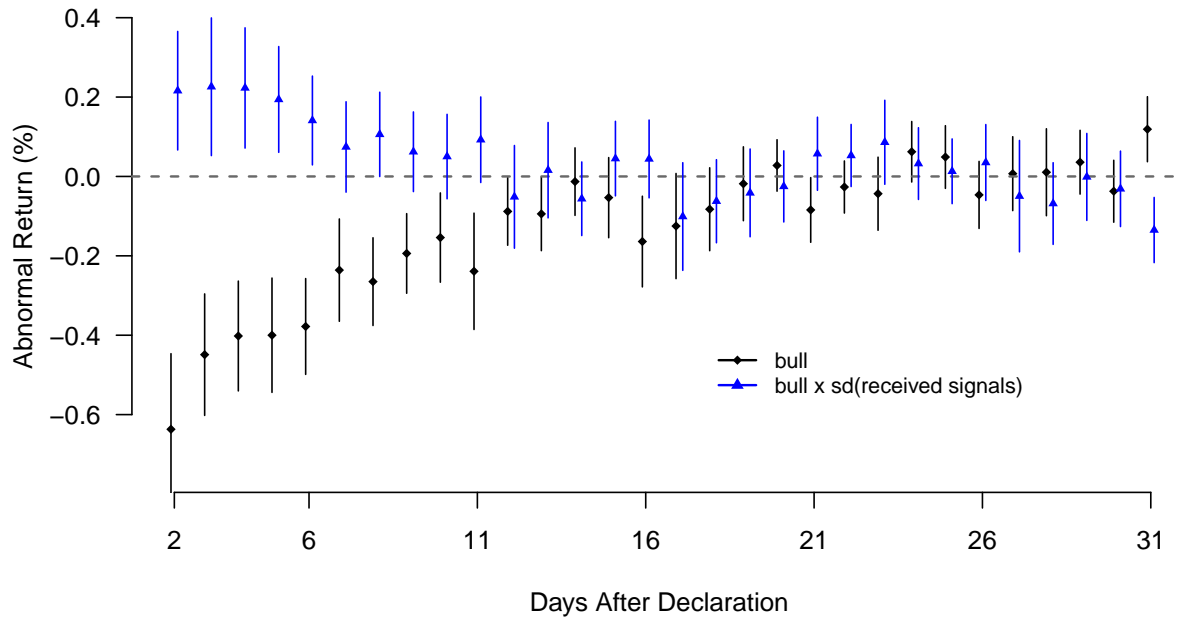


Figure 7: Do users receive *only messages with the same sentiment* more often than would be expected by chance?

The bars denote the empirical frequency that the sentiment-stamped messages received by a user are either all bullish or all bearish for bins five percentage points wide. The 45° line denotes the probability that a user receives only messages with the same sentiment, if messages were distributed at random.

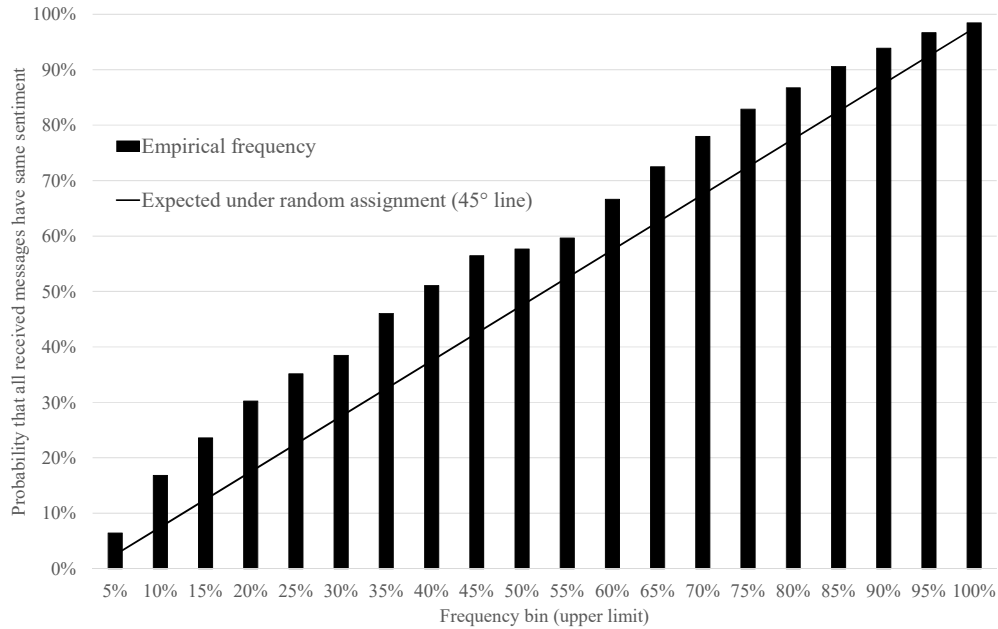
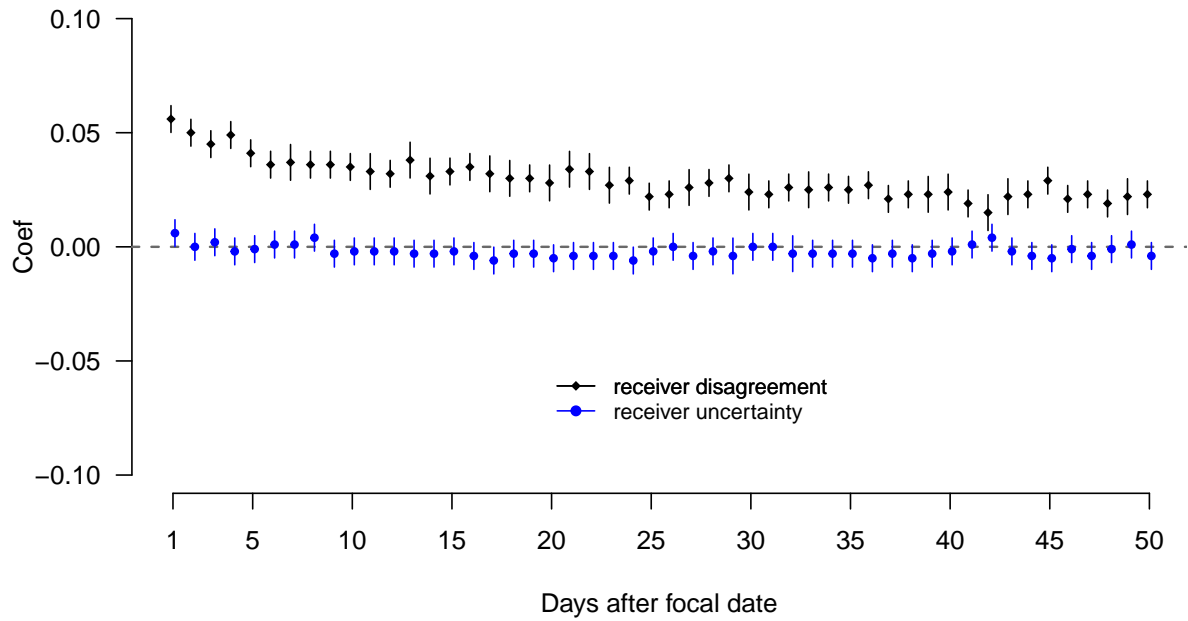


Figure 8: Do echo chambers lead to persistent disagreement?

This figure presents coefficients from regressions relating sender disagreement  $x$  days after the focal day ( $t$ ) to sender disagreement, receiver disagreement, and receiver uncertainty, all measured on day  $t$  (equation 10). The coefficients are from separate regressions, one per day, for days  $t + 1$  to  $t + 50$ . Regressions include message bin fixed effects, stock fixed effects, and day fixed effects. The vertical bars indicate 95% confidence intervals, double clustering the standard errors by day and by stock.



## 7 Tables

Table 1: Summary statistics

This table presents summary statistics. Panel (a) presents counts of the various units of observations that make up the dimension of our data – users, symbols, message sentiment and days. Restricting attention to user-days when a user posts multiple sentiment-stamped messages across different symbols, Panel (b) shows the empirical frequency of all-bullish, mixed-sentiment and all-bearish messages, and as a comparison, the theoretical probability assuming that messages are drawn independently from the overall mix of bullish versus bearish sentiment. Finally, Panel (c) presents statistics on the stock-day sample used in our regressions of abnormal log trading volume on our measures of disagreement and uncertainty (reported in Table 12).

(a) Dimensions of Data: Users, Symbols, Sentiment and Days

	Totals	Totals	
Users	395,474	Symbols	1,208
Novice	25,742	CRSP (e.g., Tesla)	1078
Intermediate	37,303	Non-CRSP (e.g., Bitcoin)	130
Professional	14,997		
Unclassified	317,432		
Sentiment Messages	33,386,587	Days	2,738
Bullish	27,090,113	Trading	1,887
Bearish	6,296,474	Non-Trading	851
User-Symbol-Sentiment Days	14,423,982		

(b) User Mixture of Sentiment Across Stocks on the Same Day

	2 Stocks		3 Stocks		4 Stocks	
	Theoretical Prob.	Empirical Freq.	Theoretical Prob.	Empirical Freq.	Theoretical Prob.	Empirical Freq.
All Bullish Sentiment	64.8%	70.8%	49.5%	62.7%	37.7%	57.6%
Mixed Sentiment	31.4%	19.5%	49.6%	30.1%	62.1%	36.5%
All Bearish Sentiment	3.8%	9.7%	0.9%	7.2%	0.2%	5.9%

(c) Summary Statistics on Stock-Day Sample for Trading Volume Evidence

	Mean	Median	Std. dev.	N obs.
<b>Main variables</b>				
Abnormal log volume <sub>s,t</sub>	0.426	0.209	1.135	421,915
Sender disagreement <sub>s,t</sub>	0.003	0.175	0.995	421,915
Received disagreement <sub>s,t</sub>	0.007	-0.691	1.001	421,915
Received uncertainty <sub>s,t</sub>	-0.002	-0.639	0.995	421,915
<b>Controls</b>				
Std dev. abnormal returns <sub>s,(t-5 to t-1)</sub>	0.046	0.029	0.077	421,915
Cum. abnormal returns <sub>s,(t-5 to t-1)</sub>	0.013	-0.001	0.222	421,915
Cum. abnormal returns <sub>s,(t-30 to t-6)</sub>	0.014	-0.019	0.416	421,915
Log Google ASVI <sub>s,t</sub>	0.532	0.666	0.448	421,915
1 if Media articles <sub>s,t</sub>	0.266	0.000	0.442	421,915
Num. of messages <sub>s,t</sub>	24.166	10.000	56.138	421,915

Table 2: Top 10 Bear and Bull stocks by selective exposure

These are the 10 StockTwits Symbols with the largest symbol fixed effects (estimated from Equation (1) with separate day, user and symbol fixed effects), out of the 100 symbols with the most messages in the sample. By conditioning on declared bulls in one estimation and declared bears in another estimation, we separately identify bearish echo chambers from bullish echo chambers.

**Bearish Echo Chambers**

Rank	Asset	Industry
1	SPDR S&P 500	<i>Index ETF</i>
2	Roku	<i>Technology - Consumer</i>
3	Beyond Meat	<i>Technology - Food</i>
4	Energous Corp	<i>Technology - Wireless</i>
5	Tesla	<i>Automobile</i>
6	Snap Inc.	<i>Technology - Mobile app</i>
7	Bitcoin USD	<i>Cryptocurrency</i>
8	AVEO Pharmaceuticals	<i>Pharmaceutical</i>
9	Advanced Micro Devices	<i>Computer processors</i>
10	SunEdison Inc	<i>Renewable energy</i>

**Bullish Echo Chambers**

Rank	Asset	Industry
1	Delcath Systems	<i>Technology - Medical</i>
2	CytRx Corporation	<i>Pharmaceutical</i>
3	Yangtze River Port & Logistics	<i>Real estate</i>
4	SunEdison Inc	<i>Renewable energy</i>
5	Tornier N.V.	<i>Technology - Medical</i>
6	MGT Capital Investments	<i>Cryptocurrency (Bitcoin mining)</i>
7	Workhorse Group	<i>Manufacturing</i>
8	Precipio	<i>Pharmaceutical</i>
9	TransEnterix Inc.	<i>Technology - Medical</i>
10	Neovasc Inc.	<i>Technology - Medical</i>

Table 3: Do users prefer to follow like-minded users?

This table examines whether bullish users predominantly choose to follow bullish posters (and vice versa). Observations are at the user-symbol-day ( $i, s, t$ ) level. We examine a user's choice of new follows after they declare themselves bullish about a symbol (on day  $t$ ), and classify a *poster* as bullish about a symbol if their posts on the day they were followed by the focal user were also bullish. The specification follows Equation (2), and the dependent variable is an indicator equal to one if net new follows (follows minus unfollows) of bulls strictly exceed net new follows of bears on day  $t + 1$  (cols 1 - 2), and days  $t + 1$  to  $t + 5$  (inclusive, cols 3 - 5). Note that when zero new net follows occur on a day (the modal case), this is coded as a zero. Because the dependent variable is binary, an identical coefficient results from a specification with bearish users following bearish posters. Columns (6) through (9) are run on the subsample of users that make at least one new net follow in the corresponding time window, which extends through  $t + 50$ . We multiply the dependent variable by 100 so coefficients are percentage points. Table 4 also shows the specification in column (4) separately for net bull and net bear follows. Standard errors (in brackets) clustered by user. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Dep. var.: $\mathbb{1} \times 100$ if new follows $_{i,s,t+x}$ are more Bull than Bear									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	t+1	+ User FE t+1	t+1 → t+5	+ User FE t+1 → t+5	with Bull bins t+1 → t+5	+ User-Symbol FE t+1 → t+5	Conditional on new follows t+1 → t+5	Conditional on new follows t+2 → t+10	Conditional on new follows t+11 → t+30	Conditional on new follows t+31 → t+50
$\mathbb{1}$ Declare Bull $_{i,s,t}$	0.88*** [0.02]	0.68*** [0.01]	2.41*** [0.04]	1.78*** [0.03]		1.08*** [0.04]	12.21*** [0.57]	7.70*** [0.47]	3.34*** [0.38]	0.27 [0.31]
$\mathbb{1}$ if net # Bull declarations $_{i,s,t} \leq -1$					-0.81*** [0.03]					
$\mathbb{1}$ if net # Bull declarations $_{i,s,t} = 2$					1.30*** [0.02]					
$\mathbb{1}$ if net # Bull declarations $_{i,s,t} = 3$					2.28*** [0.04]					
$\mathbb{1}$ if net # Bull declarations $_{i,s,t} \geq 4$					4.92*** [0.06]					
# observations	13,893,332	13,893,332	13,893,332	13,893,332	13,893,332	12,262,524	596,518	919,306	1,302,544	1,308,130
# clusters (users)	305,967	305,967	305,967	305,967	305,967	259,476	63,594	77,761	84,419	82,568
$R^2$	0.05	0.12	0.09	0.19	0.20	0.34	0.72	0.78	0.83	0.87
Mean of dependent var.(%)	1.64	1.64	4.54	4.54	4.54	4.92	77.03	70.13	60.39	40.85
Effect size (% of mean)	54	41	53	39		22	16	11	6	1
User FE	-	Y	-	Y	Y	-	-	-	-	-
User x symbol FE	-	-	-	-	-	Y	Y	Y	Y	Y
Day x symbol FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table 4: Does skin-in-the-game lead to more selective exposure?

Like Table 3, this table predicts the likelihood that users choose to follow like-minded users, but the specification follows Equation (3). Observations are at the user-symbol-day ( $i, s, t$ ) level. To examine on differential selective exposure behavior for bulls versus bears, columns (1) through (4) (respectively, columns (5) through (8)) have an indicator equal to one if net new follows of bulls (bears) exceed zero as the dependent variable. In addition to the main effect of selective exposure for declaring as bullish (or bearish), the even columns of the table include an interaction for whether the user also declares a trade at the same time (e.g., writes “just bought” or “just sold”), turning on the  $\mathbb{1}Buyer$  or  $\mathbb{1}Seller$  indicator variables. The coefficient on the interaction measures how selective exposure for declared bulls (bears) differs when they have also declared trading the security. We multiply the dependent variable by 100 so coefficients are percentage points. Standard errors (in brackets) clustered by user. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Dep. var.: $\mathbb{1} \times 100$ if net new Bull follows $_{i,s,(t+1 \rightarrow t+5)} > 0$				Dep. var.: $\mathbb{1} \times 100$ if net new Bear follows $_{i,s,(t+1 \rightarrow t+5)} > 0$			
	(1) Baseline	(2) + Buyer interaction	(3) Baseline + UserSym FE	(4) + Buyer interaction	(5) Baseline	(6) + Seller interaction	(7) Baseline + UserSym FE	(8) + Seller interaction
$\mathbb{1}Declare Bull_{i,s,t}$	1.72*** [0.03]	1.64*** [0.03]	1.02*** [0.04]	0.95*** [0.04]				
$\mathbb{1}Declare Bull_{i,s,t} \times \mathbb{1} Buyer_{i,s,t}$		1.19*** [0.04]		0.93*** [0.04]				
$\mathbb{1}Declare Bear_{i,s,t}$					0.50*** [0.02]	0.47*** [0.02]	0.38*** [0.03]	0.36*** [0.03]
$\mathbb{1}Declare Bear_{i,s,t} \times \mathbb{1} Seller_{i,s,t}$						1.08*** [0.09]		0.78*** [0.08]
# observations	13,892,913	13,892,913	12,262,096	12,262,096	13,892,913	13,892,913	12,262,096	12,262,096
# clusters (users)	305,966	305,966	259,474	259,474	305,966	305,966	259,474	259,474
$R^2$	0.27	0.27	0.40	0.40	0.47	0.47	0.53	0.53
Mean of dependent var.(%)	5.09	5.09	5.42	5.42	1.28	1.28	1.29	1.29
Main effect size (% of mean)	34	32	19	18	39	37	30	28
User FE	Y	Y	-	-	Y	Y	-	-
User x symbol FE	-	-	Y	Y	-	-	Y	Y
Day x symbol FE	Y	Y	Y	Y	Y	Y	Y	Y

Table 5: Do Bulls' newsfeeds reflect their bullish sentiment? (And vice versa)

This table examines whether declared bulls (bears) see more bullish (bearish) users on the days (t+1 to t+50) following their sentiment declaration about a specific symbol on day t. The specification follows Equation (4). Observations are at the user-symbol-day ( $i, s, t$ ) level, and the sample is conditional on seeing at least one sentiment-stamped message in the time period of the dependent variable. The dependent variable is the number of bullish (bearish) impressions over a given time period; that is, the number of other *users* that post with sentiment, and are followed by the focal user. Standard errors (in brackets) clustered by user. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	N impressions <sub>i,s,t+1</sub>		N impressions <sub>i,s,(t+2 → t+10)</sub>		N impressions <sub>i,s,(t+11 → t+30)</sub>		N impressions <sub>i,s,(t+31 → t+50)</sub>	
	(1) Bullish	(2) Bearish	(3) Bullish	(4) Bearish	(5) Bullish	(6) Bearish	(7) Bullish	(8) Bearish
1Declare Bull <sub>i,s,t</sub>	0.36*** [0.02]		2.02*** [0.09]		3.85*** [0.17]		3.67*** [0.18]	
1Declare Bear <sub>i,s,t</sub>		0.12*** [0.01]		0.38*** [0.04]		0.57*** [0.07]		0.46*** [0.08]
# obs.	3,175,857	3,175,857	5,721,704	5,721,704	6,553,985	6,553,985	6,127,827	6,127,827
# clusters (users)	122,923	122,923	163,209	163,209	170,133	170,133	157,791	157,791
R <sup>2</sup>	0.52	0.49	0.58	0.52	0.60	0.53	0.61	0.52
Mean of dependent var.(%)	2.14	0.29	9.31	1.25	17.00	2.33	16.57	2.35
Effect size (% of mean)	17	40	22	31	23	25	22	19
User FE	Y	Y	Y	Y	Y	Y	Y	Y
Day x symbol FE	Y	Y	Y	Y	Y	Y	Y	Y

Table 6: Do Bulls *like* more bullish posts than bearish posts? (And vice versa)

This table examines whether declared bulls (bears) like a greater number of bullish (bearish) posts than bearish (bullish) posts about that symbol on the days ( $t + 1$  to  $t + 50$ ) following their sentiment declaration. The specification follows Equation (5). Observations are at the user-symbol-day ( $i, s, t$ ) level, and the sample is conditional on liking at least one sentiment-stamped message in the time period of the dependent variable. The dependent variable is a count of the number of liked messages for a symbol-day combination. Standard errors (in brackets) clustered by user. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	N of liked msgs <sub>s,t+1</sub>		N of liked msgs <sub>s,(t+2 → t+10)</sub>		N of liked msgs <sub>s,(t+11 → t+30)</sub>		N of liked msgs <sub>s,(t+31 → t+50)</sub>	
	(1) Bullish	(2) Bearish	(3) Bullish	(4) Bearish	(5) Bullish	(6) Bearish	(7) Bullish	(8) Bearish
1Declare Bull <sub>s,t</sub>	3.86*** [0.10]		11.10*** [0.39]		15.91*** [0.77]		12.63*** [0.81]	
1Declare Bear <sub>s,t</sub>		2.96*** [0.08]		6.12*** [0.24]		6.41*** [0.40]		4.57*** [0.43]
# obs.	3,180,665	3,180,665	5,678,414	5,678,414	5,341,915	5,341,915	4,094,312	4,094,312
# clusters (users)	135,506	135,506	169,171	169,171	145,680	145,680	109,660	109,660
R <sup>2</sup>	0.30	0.28	0.38	0.30	0.41	0.33	0.45	0.38
Mean of dependent var.(%)	6.18	0.56	21.23	1.81	35.75	2.81	34.85	2.59
Effect size (% of mean)	62	531	52	339	44	228	36	176
User FE	Y	Y	Y	Y	Y	Y	Y	Y
Day x Symbol FE	Y	Y	Y	Y	Y	Y	Y	Y

Table 7: Does the preference for following like-minded users differ on earnings announcement days?

Like Table 3, this table predicts the likelihood that users choose to follow like-minded users. The specification follows Equation (6). Observations are at the user-stock-day ( $i, s, t$ ) level. Columns (1) to (3) have as the dependent variable an indicator equal to one if net new follows of bulls exceed net new follows of bears on days  $t+1$  and  $t+2$ . Earnings Announcements (EAs) on day  $t+1$  are released either before market open or after market close, so the information reaches prices on days  $t+1$  and  $t+2$  respectively. Thus, we define the follow period as days  $t+1$  and  $t+2$  for all dependent variables (recall the focal user's sentiment declaration always occurs on day  $t$ ). Columns (2), (5) and (8) add an interacted indicator for an EA day occurring on day  $t+1$ . Columns (4) to (6) (respectively, (7) to (9)) have an indicator equal to one if net new follows of bulls (bears) exceed zero as the dependent variable. We define an earnings announcement as providing positive news (denoted by  $EA\ day^+$  in the table) if the stock's cumulative excess returns over days  $t$  to  $t+2$  inclusive are positive. Columns (6) and (9) restrict the sample to EA days and add an interaction with  $EA\ day^+$ . We multiply the dependent variable by 100 so coefficients are percentage points. Standard errors (in brackets) clustered by user. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	$\mathbb{1}$ if follows $_{i,s,(t+1 \rightarrow t+2)}$ are more Bull than Bear			$\mathbb{1}$ if net new <i>Bull</i> follows $_{i,s,(t+1 \rightarrow t+2)} > 0$			$\mathbb{1}$ if net new <i>Bear</i> follows $_{i,s,(t+1 \rightarrow t+2)} > 0$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Baseline	+ EA day interaction	+ separate EA day	Baseline	+ EA day interaction	Only EA days	Baseline	+ EA day interaction	Only EA days
$\mathbb{1}Declare\ Bull_{i,s,t}$	1.04*** [0.02]	1.03*** [0.02]	1.03*** [0.02]	1.03*** [0.02]	1.02*** [0.02]	1.24*** [0.21]			
$\mathbb{1}Declare\ Bull_{i,s,t} \times \mathbb{1}$ if EA day		0.77*** [0.11]			0.69*** [0.11]				
$\mathbb{1}Declare\ Bull_{i,s,t} \times \mathbb{1}$ if EA day <sup>+</sup>			0.92*** [0.18]			0.57** [0.29]			
$\mathbb{1}Declare\ Bull_{i,s,t} \times \mathbb{1}$ if EA day <sup>-</sup>			0.66*** [0.13]						
$\mathbb{1}Declare\ Bear_{i,s,t}$							0.21*** [0.01]	0.20*** [0.01]	0.56*** [0.16]
$\mathbb{1}Declare\ Bear_{i,s,t} \times \mathbb{1}$ if EA day								0.36*** [0.09]	
$\mathbb{1}Declare\ Bear_{i,s,t} \times \mathbb{1}$ if EA day <sup>+</sup>									-0.01 [0.21]
# observations	11,108,268	11,108,268	11,108,268	11,108,268	11,108,268	126,911	11,108,268	11,108,268	126,911
# clusters (users)	277,040	277,040	277,040	277,040	277,040	29,613	277,040	277,040	29,613
$R^2$	0.15	0.15	0.15	0.21	0.21	0.39	0.45	0.45	0.41
Mean of dependent var.(%)	2.57	2.57	2.57	2.79	2.79	3.90	0.54	0.54	0.79
Main effect size (% of mean)	41	40	40	37	37	32	38	37	71
User FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Day x stock FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table 8: Do user characteristics affect their preference for following like-minded users?

Like Table 3, this table predicts the likelihood that users choose to follow like-minded users in the five days after declaring as a bull, but adds interactions with indicators for the user's self-declared investor experience category (novice, intermediate, professional, or missing). Observations are at the user-symbol-day ( $i, s, t$ ) level. Columns (1) and (2) include missing experience as the omitted category, whereas Columns (3) and (4) estimate the specification on the sample for which we have information on experience. Column (5) interacts an indicator for users that include one or more links per post, on average (links include charts). The next two columns add an interaction with an indicator for a user with an above median # of people they are following (column (6)), or being followed by (column (7)), as of the latest date each user appears in the sample. Column (8) interacts an indicator for users with above median activity (defined as the sum of all likes, follows and posts with sentiment for the entire sample period). We multiply the dependent variable by 100 so coefficients are percentage points. Standard errors (in brackets) clustered by user. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Dep. var: $\mathbb{1}$ x100 if new follows $_{i,s,(t+1 \rightarrow t+5)}$ are more Bull than Bear							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline	Omitted category: missing experience	Baseline with experience	Omitted category: intermediate	Many links	# Following > median	# Followers > median	Activity > median
$\mathbb{1}$ Declare Bull $_{i,s,t}$	1.78*** [0.03]	1.85*** [0.04]	1.69*** [0.06]	1.91*** [0.08]	1.78*** [0.03]	0.83*** [0.03]	0.84*** [0.04]	0.59*** [0.05]
$\mathbb{1}$ Declare Bull $_{i,s,t}$ x $\mathbb{1}$ Novice investor $_i$		0.04 [0.12]		0.02 [0.14]				
$\mathbb{1}$ Declare Bull $_{i,s,t}$ x $\mathbb{1}$ Intermediate investor $_i$		0.02 [0.08]						
$\mathbb{1}$ Declare Bull $_{i,s,t}$ x $\mathbb{1}$ Professional investor $_i$		-0.54*** [0.10]		-0.58*** [0.12]				
$\mathbb{1}$ Declare Bull $_{i,s,t}$ x $\mathbb{1}$ if (avg. # links per msg) $_i \geq 1$					-1.55*** [0.10]			
$\mathbb{1}$ Declare Bull $_{i,s,t}$ x $\mathbb{1}$ if # following $_i >$ median						1.52*** [0.06]		
$\mathbb{1}$ Declare Bull $_{i,s,t}$ x $\mathbb{1}$ if # followers $_i >$ median							1.10*** [0.05]	
$\mathbb{1}$ Declare Bull $_{i,s,t}$ x $\mathbb{1}$ if User activity $_i >$ median								1.26*** [0.06]
# observations	13,893,332	13,893,332	5,087,695	5,087,695	13,893,332	13,893,332	13,893,332	13,893,332
# clusters (users)	305,967	305,967	65,866	65,866	305,967	305,967	305,967	305,967
$R^2$	0.19	0.19	0.21	0.21	0.19	0.19	0.19	0.19
Mean of dependent var.(%)	4.54	4.54	4.19	4.19	4.54	4.54	4.54	4.54
Main effect size (% of mean)	39	41	40	45	39	18	18	13
User FE	Y	Y	Y	Y	Y	Y	Y	Y
Day x symbol FE	Y	Y	Y	Y	Y	Y	Y	Y

Table 9: Do users leave Echo chambers in response to experienced returns?

Like Table 4, this table predicts the likelihood that users choose to follow like-minded users. Observations are at the user-stock-day ( $i, s, t$ ) level. To examine differential selective exposure behavior for bulls versus bears, the dependent variable in columns (1) through (7) (respectively, columns (8) through (14)) is an indicator equal to one if net new follows of bulls (bears) exceed zero over a time window ( $t + 11 \rightarrow t + 30$ ), similar to Tables 4 and 7. In addition to the main effect of selective exposure for declaring as bullish (or bearish), most columns of the table include an interaction with an indicator for whether the user experienced below-median (p50) or negative returns in the  $t + 2 \rightarrow t + 10$  window, either for the whole market (CRSP value-weighted index return including distributions) or for the focal stock. The coefficient on the interaction measures how selective exposure for declared bulls (bears) differs over the subsequent 20 days when they have experienced poor returns over the  $t + 2 \rightarrow t + 10$  window. We skip day  $t + 1$  because sentiment declarations on day  $t$  can be made after market close. We multiply the dependent variable by 100 so coefficients are percentage points. Standard errors (in brackets) clustered by user. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Dep. var.: $\mathbb{1} \times 100$ if net new Bull follows $_{i,s,(t+11 \rightarrow t+30)} > 0$							$\mathbb{1} \times 100$ if net new Bear follows $_{i,s,(t+11 \rightarrow t+30)} > 0$						
	(1) Baseline	(2) $\times$ MktRet < p50	(3) $\times$ MktRet < 0	(4) $\times$ Ret $_s$ < p50	(5) $\times$ Ret $_s$ < 0	(6) $\times$ both < p50	(7) $\times$ both < 0	(8) Baseline	(9) $\times$ MktRet < p50	(10) $\times$ MktRet < 0	(11) $\times$ Ret $_s$ < p50	(12) $\times$ Ret $_s$ < 0	(13) $\times$ both < p50	(14) $\times$ both < 0
$\mathbb{1}$ Declare Bull $_{i,s,t}$	2.22*** [0.05]	2.34*** [0.06]	2.32*** [0.06]	2.44*** [0.06]	2.48*** [0.06]	2.54*** [0.07]	2.56*** [0.07]							
$\mathbb{1}$ Bull $_{i,s,t} \times \mathbb{1}$ MktRet $_{(t+2 \rightarrow 10)} < p50$ or 0		-0.25*** [0.05]	-0.30*** [0.06]			-0.22*** [0.05]	-0.27*** [0.06]							
$\mathbb{1}$ Bull $_{i,s,t} \times \mathbb{1}$ Ret $_{s,(t+2 \rightarrow 10)} < p50$ or 0				-0.42*** [0.05]	-0.47*** [0.06]	-0.40*** [0.05]	-0.45*** [0.05]							
$\mathbb{1}$ Declare Bear $_{i,s,t}$								0.18*** [0.03]	0.13*** [0.03]	0.14*** [0.03]	0.14*** [0.03]	0.14*** [0.04]	0.10** [0.04]	0.10*** [0.04]
$\mathbb{1}$ Bear $_{i,s,t} \times \mathbb{1}$ MktRet $_{(t+2 \rightarrow 10)} < p50$ or 0									0.10*** [0.04]	0.12*** [0.04]			0.10*** [0.04]	0.11*** [0.04]
$\mathbb{1}$ Bear $_{i,s,t} \times \mathbb{1}$ Ret $_{s,(t+2 \rightarrow 10)} < p50$ or 0											0.08** [0.04]	0.08** [0.04]	0.07* [0.04]	0.07** [0.04]
# observations	10,187,966	10,187,966	10,187,966	10,187,966	10,187,966	10,187,966	10,187,966	10,187,966	10,187,966	10,187,966	10,187,966	10,187,966	10,187,966	10,187,966
# clusters (users)	261,936	261,936	261,936	261,936	261,936	261,936	261,936	261,936	261,936	261,936	261,936	261,936	261,936	261,936
$R^2$	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.60	0.60	0.60	0.60	0.60	0.60	0.60
Mean of dependent var.(%)	8.23	8.23	8.23	8.23	8.23	8.23	8.23	2.15	2.15	2.15	2.15	2.15	2.15	2.15
User FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Day x stock FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table 10: Is being in an echo chamber associated with lower future abnormal returns?

This table examines stock returns following a sentiment declaration, and how this varies with the diversity of sentiment signals that a user receives from their network. Estimated coefficients are from equation 8. The dataset and the signals received are at the user-stock-day ( $i, s, t$ ) level and we drop observations for which the user has not seen at least five signals over the preceding thirty days, in order to calculate non-trivial standard deviations of received signals. We keep only StockTwits symbols with data in CRSP and sharecodes of 10, 11 or 12. The dependent variable is cumulative abnormal returns (returns minus the CRSP value-weighted index return) over days  $t + 2$  to  $t + 6$  (or  $t + 11$ ), where day  $t$  is the day of the user's sentiment declaration. When calculating future returns we skip day  $t + 1$  because sentiment declarations on day  $t$  can be made after market close. User experience is self-declared, and users who have not provided this information make up the omitted (base) category. We multiply the dependent variable by 100 so coefficients are percentage points. Standard errors (in brackets) are separately clustered by user and by stock. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Dep. var.: abnormal return $_{i,s,(t+2 \rightarrow t+6)}$				Dep. var.: abnormal return $_{i,s,(t+2 \rightarrow t+11)}$			
	(1) Baseline	(2) + sd(rec. signals)	(3) Baseline with experience	(4) + sd(rec. signals)	(5) Baseline	(6) + sd(rec. signals)	(7) Baseline with experience	(8) + sd(rec. signals)
$\mathbb{1}Bull_{i,s,t} \times sd(\text{rec. signals})_{i,s,t}$		0.97*** [0.20]		0.92*** [0.20]		1.49*** [0.27]		1.43*** [0.27]
$\mathbb{1}Declare Bull_{i,s,t}$	-1.88*** [0.33]	-2.33*** [0.31]	-2.21*** [0.36]	-2.61*** [0.34]	-2.48*** [0.36]	-3.18*** [0.36]	-2.86*** [0.39]	-3.48*** [0.39]
$sd(\text{received signals})_{i,s,t}$		0.06 [0.26]		0.11 [0.26]		0.11 [0.34]		0.16 [0.34]
$\mathbb{1}Bull_{i,s,t} \times \mathbb{1}Novice\ investor_i$			-0.00 [0.28]	-0.03 [0.27]			0.00 [0.32]	-0.05 [0.32]
$\mathbb{1}Bull_{i,s,t} \times \mathbb{1}Intermediate\ investor_i$			0.66*** [0.17]	0.61*** [0.17]			0.72*** [0.23]	0.64*** [0.23]
$\mathbb{1}Bull_{i,s,t} \times \mathbb{1}Professional\ investor_i$			1.12*** [0.26]	1.02*** [0.26]			1.30*** [0.25]	1.15*** [0.25]
Cum. ab. returns $_{s,(t-5\ to\ t-1)}$	-7.99*** [1.62]	-7.98*** [1.62]	-7.99*** [1.62]	-7.98*** [1.62]	-12.59*** [2.83]	-12.58*** [2.83]	-12.59*** [2.83]	-12.58*** [2.83]
Cum. ab. returns $_{s,(t-30\ to\ t-6)}$	-4.69*** [1.02]	-4.70*** [1.02]	-4.69*** [1.02]	-4.70*** [1.02]	-6.25*** [1.42]	-6.25*** [1.41]	-6.25*** [1.42]	-6.25*** [1.41]
Log GoogleASVI $_{s,(t-5\ to\ t-1)}$	-1.07*** [0.28]	-1.08*** [0.28]	-1.07*** [0.28]	-1.08*** [0.28]	-1.70*** [0.39]	-1.71*** [0.39]	-1.70*** [0.39]	-1.71*** [0.39]
(Log GoogleASVI $\times$ Cum. ab. ret.) $_{s,(t-5\ to\ t-1)}$	0.24 [0.15]	0.24 [0.15]	0.24 [0.15]	0.24 [0.15]	0.54** [0.22]	0.54** [0.22]	0.54** [0.22]	0.54** [0.22]
# obs.	2,670,753	2,670,753	2,670,753	2,670,753	2,640,068	2,640,068	2,640,068	2,640,068
# clusters (stock)	757	757	757	757	757	757	757	757
# clusters (users)	85,753	85,753	85,753	85,753	84,734	84,734	84,734	84,734
$R^2$	0.39	0.39	0.39	0.39	0.54	0.54	0.54	0.54
Day FE	Y	Y	Y	Y	Y	Y	Y	Y
User $\times$ stock FE	Y	Y	Y	Y	Y	Y	Y	Y
Stock $\times$ month FE	Y	Y	Y	Y	Y	Y	Y	Y

Table 11: Information Silos: are Bulls more likely to receive only bullish messages (and vice versa)?

This table predicts the likelihood that all messages received by a user on day  $t$  will have all-bullish sentiment (columns 1-3) or all-bearish sentiment (columns 4-6). As a control variable, we include the probability that all received messages will be bullish (columns 1-3) or all bearish (columns 4-6) under random assignment, conditional on the number of messages sent and received. Bull (Bear) is an indicator if the most recent sentiment declaration by the receiver is bullish (bearish) in the preceding week. Observations are at the user-symbol-day  $(i, s, t)$  level. We multiply the dependent variable by 100 so coefficients are percentage points. Standard errors (in brackets) clustered by user. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Dep. var.: $\mathbb{1} \times 100$ if all messages received on day $t$ have sentiment that is					
	(1) Bullish	(2) Bullish	(3) Bullish	(4) Bearish	(5) Bearish	(6) Bearish
Expected Pr(all Bull) if random $_{i,s,t}$	93.01*** [0.35]	78.53*** [0.32]	74.48*** [0.64]			
Expected Pr(all Bear) if random $_{i,s,t}$				86.58*** [0.61]	79.31*** [0.58]	75.61*** [0.75]
$\mathbb{1} \text{Declare Bull}_{i,s,t-7 \text{ to } t-1}$		8.02*** [0.20]	6.72*** [0.18]			
$\mathbb{1} \text{Declare Bear}_{i,s,t-7 \text{ to } t-1}$					2.85*** [0.12]	3.04*** [0.12]
# obs.	2,963,654	2,963,654	2,963,654	2,963,654	2,963,654	2,963,654
# clusters (users)	93,052	93,052	93,052	93,052	93,052	93,052
$R^2$	0.34	0.47	0.55	0.17	0.30	0.40
Mean of dependent var.(%)	67.0	67.0	67.0	4.3	4.3	4.3
Main effect size (% of mean)		12	10		67	71
User FE	-	Y	Y	-	Y	Y
Day x Symbol FE	-	-	Y	-	-	Y

Table 12: Does selective exposure behavior affect stock turnover?

This table examines how proxies for selective exposure behavior on StockTwits (received disagreement and received uncertainty), together with sender disagreement, relate to daily abnormal log turnover. Observations are at the stock-day ( $s,t$ ) level; we estimate equation 9. *Sender Disagreement* captures dispersion of sentiment among posts about a stock ( $s$ ) (i.e. the standard deviation of sentiment across posts). *Received Disagreement* captures how disagreement in posts is distributed across receivers (i.e. the mean across receivers of the standard deviation of received sentiment). *Received Uncertainty* captures the dispersion of sentiment about a stock across newsfeeds (i.e. the standard deviation across receivers of the mean of received sentiment). We standardize the disagreement measures by subtracting the mean and dividing by the standard deviation, both measured over the full sample period.  $AbLogTurnover_{s,t}$  is the difference between log turnover on day  $t$  and the average log turnover from  $t-140$  to  $t-20$  trading days (6-month period, skipping most recent month). Controls are *Abnormal log turnover* on day  $t-1$ ;  $\mathbb{1}Media\ Article_{s,t}$ , an indicator for days when stock  $s$  was mentioned in at least one article covered by Dow Jones Newswire data (including the Wall Street Journal); *Log GoogleASVI* $_{s,t}$ , a measure of abnormal google search volume for the ticker of stock  $s$ ; *Volatility*, measured as the standard deviation of abnormal returns over days  $t-5$  to  $t-1$ ; and *Cumulative abnormal returns*, measured over days  $t-30$  to  $t-6$  and  $t-5$  to  $t-1$ . Fixed effects for day, stock-month, and message number are included in all regressions. Message number fixed effects are defined for days with 0, 1, 2, 3, 4, 5-10, 11-30, and over 30 messages. Standard errors (in brackets) are separately clustered by stock and day. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Dep. var.: Abnormal Log Turnover $_{s,t}$				
	(1)	(2)	(3)	(4)	(5)
Sender Disagreement $_{s,t}$	0.017*** [0.001]	0.015*** [0.001]	0.027*** [0.002]	0.024*** [0.002]	0.025*** [0.002]
Received Disagreement $_{s,t}$		0.004** [0.002]		0.008*** [0.002]	0.011*** [0.002]
Received Uncertainty $_{s,t}$			-0.014*** [0.002]	-0.016*** [0.001]	-0.016*** [0.001]
Abnormal Log Turnover $_{s,t-1}$	0.188*** [0.004]	0.188*** [0.004]	0.189*** [0.004]	0.188*** [0.004]	0.205*** [0.005]
$\mathbb{1}Media\ Article_{s,t}$	0.128*** [0.006]	0.128*** [0.006]	0.128*** [0.006]	0.128*** [0.006]	
Log GoogleASVI $_{s,t}$	0.347*** [0.015]	0.346*** [0.015]	0.347*** [0.015]	0.346*** [0.015]	
Volatility $_{s,(t-5\ to\ t-1)}$	0.117*** [0.035]	0.116*** [0.035]	0.116*** [0.035]	0.115*** [0.035]	
Cum. Abnormal Returns $_{s,(t-5\ to\ t-1)}$	0.010 [0.012]	0.010 [0.012]	0.010 [0.012]	0.010 [0.012]	
Cum. Abnormal Returns $_{s,(t-30\ to\ t-6)}$	-0.051*** [0.009]	-0.051*** [0.009]	-0.051*** [0.009]	-0.051*** [0.009]	
# obs.	421,915	421,915	421,915	421,915	421,915
# clusters (stock)	1,075	1,075	1,075	1,075	1,075
# clusters (day)	1,886	1,886	1,886	1,886	1,886
$R^2$	0.83	0.83	0.83	0.83	0.82
Mean of dependent var.	0.43	0.43	0.43	0.43	0.43
Day FE	Y	Y	Y	Y	Y
Month x stock FE	Y	Y	Y	Y	Y
Message number FE	Y	Y	Y	Y	Y

## **8 Internet Appendix to: Echo Chambers**

**J. Anthony Cookson, Joseph E. Engelberg and William Mullins**

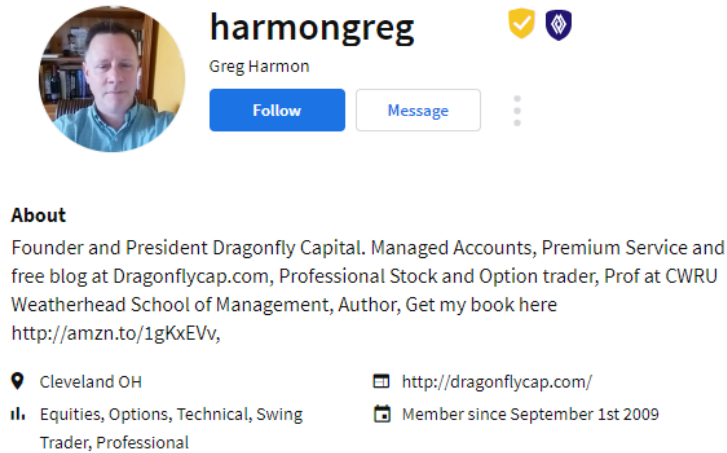
---

<sup>1</sup>Cookson (tony.cookson@colorado.edu), Leeds School of Business at University of Colorado-Boulder. Engelberg (jengelberg@ucsd.edu) and Mullins (wmullins@ucsd.edu), Rady School of Management, UC San Diego.

Figure A.1: Examples of StockTwits users

This figure presents screenshots of the user profile information for three prominent users on StockTwits. All three are verified professional traders and have public writing outside of StockTwits, as is indicated in the links in their profiles. These users also reflect diverse perspectives on investing. Greg Harmon is a prominent technical investor. Todd Sullivan is a long-term value investor. Aron Pinson is a long-term fundamental investor.

(a) Greg Harmon – Professional Technical Investor



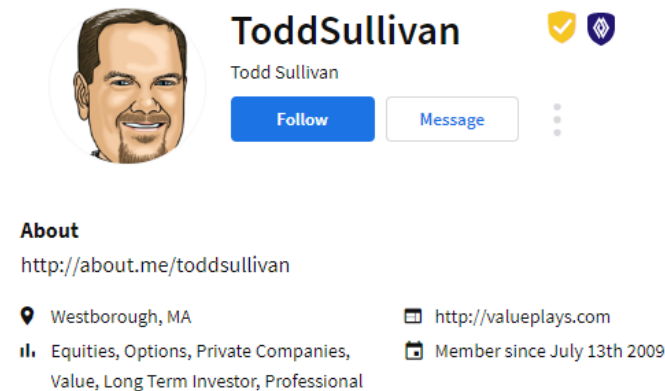
**harmongreg**  
Greg Harmon

Follow Message

**About**  
Founder and President Dragonfly Capital. Managed Accounts, Premium Service and free blog at [Dragonflycap.com](http://Dragonflycap.com), Professional Stock and Option trader, Prof at CWRU Weatherhead School of Management, Author, Get my book here <http://amzn.to/1gKxEVv>,

Cleveland OH <http://dragonflycap.com/>  
Equities, Options, Technical, Swing Trader, Professional Member since September 1st 2009

(b) Todd Sullivan – Professional Value Investor



**ToddSullivan**  
Todd Sullivan

Follow Message

**About**  
<http://about.me/toddsullivan>

Westborough, MA <http://valueplays.com>  
Equities, Options, Private Companies, Value, Long Term Investor, Professional Member since July 13th 2009

(c) Aron Pinson – Professional Fundamental Investor



**microfundy**  
Aron Pinson

Follow Message

**About**  
CIO at LPS Financial, CFA, and author of the MicroFundy.com blog. Tweets are not recommendations. Please read the full disclaimer at <http://microfundy.com/disclaimer/>

NYC <http://MicroFundy.com/>  
Equities, Options, Forex, Futures, Bonds, Private Companies, Fundamental, Long Term Investor, Professional Member since December 16th 2010

Table A.1: Example of Selective Exposure

As an illustrative example, this table presents information on the posts and newsfeeds of a declared bull to compare to a declared bear. Both users posted about Tesla on November 14th, 2018, but the declared bull – username: EVisthefuture – was bullish on Tesla, whereas the declared bear – username: DoctorBurry – was bearish on Tesla. On the next day, the bullish user’s newsfeed was 100% bullish (45 messages) and the bearish user’s newsfeed was 100% bearish (12 messages), providing an example of an information echo chamber. To illustrate the information content of the newsfeeds we report notable messages in each user’s newsfeed on November 15th, 2018.

**Declared Tesla Bull**

Nov 14, 2018: Bullish User (EVisthefuture) Message Posted About Tesla

---

Oil giant BP gets its first Tesla Powerpack project, says could lead to more

---

Nov 15, 2018: Notable Posts in EVisthefuture’s Newsfeed (45 Bullish, 0 Bearish)

---

Rishesh Singh: \$TSLA bout to rip <https://www.bloomberg.com/news/articles/2018-11-14/china-is-leading-the-world-to-an-electric-car-future>

Rishesh Singh: \$TSLA Musk says Tesla acquired trucking capacity to ensure Model 3 delivery by Dec 31

Tesla Long: \$TSLA Another frozen shut car for bears here... oh wait it’s not a Tesla so don’t mention it to people <https://www.youtube.com/watch?v=Dlc5Hmsm>

Tesla Long: \$TSLA Bears you gonna lose. The arguments by these CNBC bears are idiotic! Andrew Left the bear camp hah <https://youtu.be/RJppWHQc9p0>

Angry Panda: \$TSLA gonna be glorious tomorrow.... Powell was very optimistic about the economy.... reiterated twice.... I smell bear fear...

Dexter Wilson: \$TSLA Here is a great resource for Bulls, also maybe shorts can get a clue as to what they are in for! <https://twitter.com/nykchannel/status/1063128324711596038?s=21>

**Declared Tesla Bear**

Nov 14, 2018: Bearish User (DoctorBurry) Message Posted About Tesla

---

Lots of great companies with strong mgmt teams, profits and cashflows on sale. Why would anyone buy into this \$TSLA fraud

---

Nov 15, 2018: Notable Posts in DoctorBurry’s Newsfeed (0 Bullish, 12 Bearish)

---

posicaprinia: \$TSLA They really need this over \$360 in a hurry, and keep it up there. Musky will continue tweeting to try to get the price there. Scammer

posicaprinia: \$TSLA Heed caution folks. 20%+ correction coming soon? <https://twitter.com/EconguyRosie/status/1063159726324834306>

posicaprinia: \$TSLA Pray for the nasdaq tomorrow. NVDA down 14% AH

posicaprinia: \$TSLA not sure how they are going to get this to \$360 and keep it up there. Will take an intervention from the lord and savior (Elon Musk)

ThePatrickBateman1: \$TSLA only sold 20,000 vehicles last month but has one of greatest market caps of all autos. Total joke Big Short

ThePatrickBateman1: @HeyGuy @DoctorBurry superior LOL LOL LOL \$TSLA doors and windows don’t work in cold weather and spontaneously combust in hot weather

Table A.2: Table 3 restricted to larger firms

This table examines whether bullish users predominantly choose to follow bullish users, as in Table 3, but restricting the sample to firms with market capitalization above \$3.24bn (the median December 2019 value in Kenneth French’s data). Observations are at the user-symbol-day ( $i, s, t$ ) level. We examine a user’s new follows after they declare themselves bullish about a symbol (on day  $t$ ), and classify a user as bullish about a symbol if their posts on the day they were followed by the focal user were also bullish. The specification follows Equation (2), and the dependent variable is an indicator equal to one if net new follows (follows minus unfollows) of bulls strictly exceed net new follows of bears on day  $t + 1$  (cols 1 - 2), and days  $t + 1$  to  $t + 5$  (inclusive, cols 3 - 5). Note that when zero new net follows occur on a day (the modal case), this is coded as a zero. Because of the dependent variable is binary, an identical coefficient results from a specification with bearish users following bearish posters. Columns 6 through 9 are run on the subsample of users that make at least one new net follow in the relevant time window, which extends through  $t + 50$ . We multiply the dependent variable by 100 so coefficients are percentage points. Standard errors (in brackets) clustered by user. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Dep. var.: $\mathbb{1} \times 100$ if new follows $_{i,s,t+x}$ are more Bull than Bear								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	t+1	+ User FE t+1	t+1 → t+5	+ User FE t+1 → t+5	+ User-Symbol FE t+1 → t+5	Conditional on new follows t+1 → t+5	Conditional on new follows t+2 → t+10	Conditional on new follows t+11 → t+30	Conditional on new follows t+31 → t+50
$\mathbb{1} \text{Declare Bull}_{i,s,t}$	0.56*** [0.02]	0.41*** [0.02]	1.55*** [0.05]	1.04*** [0.04]	0.70*** [0.05]	12.62*** [1.18]	8.64*** [0.96]	4.74*** [0.71]	0.43 [0.55]
# observations	3,773,560	3,773,560	3,773,560	3,773,560	3,323,876	89,302	145,955	227,802	240,985
# clusters (users)	130,302	130,302	130,302	130,302	105,810	13,030	17,107	20,441	22,151
$R^2$	0.04	0.10	0.05	0.17	0.27	0.73	0.76	0.79	0.84
Mean of dependent var.(%)	0.84	0.84	2.42	2.42	2.66	68.44	63.79	56.91	39.77
Effect size (% of mean)	66	49	64	43	27	18	14	8	1
User FE	-	Y	-	Y	-	-	-	-	-
User x symbol FE	-	-	-	-	Y	Y	Y	Y	Y
Day x symbol FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table A.3: Table 3 with Bull bins

This table follows Table 3, but replaces the (single) declared bull indicator with multiple indicators capturing the distribution of the # of bull minus # of bear declarations on day  $t$ . Observations are at the user-symbol-day  $(i, s, t)$  level. Net bearish users (# bullish - # bearish declarations  $< 0$ ) make up approximately 20% of observations, and are captured by the indicator for negative one or below. There are no observations with zero declarations in the sample. The omitted category is users with only one net declaration about a symbol on day  $t$ , amounting to about half of observations. We multiply the dependent variable by 100 so coefficients are percentage points. Standard errors (in brackets) clustered by user. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Dep. var.: $\mathbb{1} \times 100$ if new follows $_{i,s,t+x}$ are more Bull than Bear				
	(1) t+1	(2) + User FE t+1	(3) t+1 $\rightarrow$ t+5	(4) + User FE t+1 $\rightarrow$ t+5	(5) + User-Symbol FE t+1 $\rightarrow$ t+5
$\mathbb{1}$ if net # Bull declarations $_{i,s,t} \leq -1$	-0.40*** [0.01]	-0.25*** [0.01]	-1.26*** [0.03]	-0.81*** [0.03]	-0.55*** [0.03]
$\mathbb{1}$ if net # Bull declarations $_{i,s,t} = 2$	0.56*** [0.01]	0.51*** [0.01]	1.51*** [0.03]	1.30*** [0.02]	0.68*** [0.02]
$\mathbb{1}$ if net # Bull declarations $_{i,s,t} = 3$	1.03*** [0.02]	0.96*** [0.02]	2.62*** [0.05]	2.28*** [0.04]	1.26*** [0.03]
$\mathbb{1}$ if net # Bull declarations $_{i,s,t} \geq 4$	2.43*** [0.05]	2.25*** [0.04]	5.58*** [0.10]	4.92*** [0.06]	2.91*** [0.05]
# observations	13,893,332	13,893,332	13,893,332	13,893,332	12,262,524
# clusters (users)	305,967	305,967	305,967	305,967	259,476
$R^2$	0.06	0.12	0.09	0.20	0.34
Mean of dependent var.(%)	1.64	1.64	4.54	4.54	4.92
User FE	-	Y	-	Y	-
User x symbol FE	-	-	-	-	Y
Day x symbol FE	Y	Y	Y	Y	Y

Table A.4: Table 3 only counting follows of focused users ( $\leq 1$  symbol on average)

This table follows Table 3, but changes the dependent variable to only count follows of focused users, i.e., users that on average declare about one or fewer symbols per month over their time on StockTwits. Observations are at the user-symbol-day  $(i, s, t)$  level. We multiply the dependent variable by 100 so coefficients are percentage points. Standard errors (in brackets) clustered by user. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Dep. var.: $\mathbb{1}$ x100 if new focused follows $_{i,s,t+x}$ are net Bull				
	(1)	(2)	(3)	(4)	(5)
	t+1	+ User FE t+1	t+1 $\rightarrow$ t+5	+ User FE t+1 $\rightarrow$ t+5	+ User-Symbol FE t+1 $\rightarrow$ t+5
$\mathbb{1}$ Declare Bull $_{i,s,t}$	0.12*** [0.00]	0.09*** [0.00]	0.41*** [0.01]	0.28*** [0.01]	0.14*** [0.01]
# observations	13,981,783	13,893,332	13,981,783	13,893,332	12,262,524
# clusters (users)	392,259	305,967	392,259	305,967	259,476
$R^2$	0.03	0.07	0.05	0.12	0.23
Mean of dependent var.(%)	0.20	0.20	0.66	0.66	0.74
Effect size (% of mean)	62	45	62	43	19
User FE	-	Y	-	Y	-
User x symbol FE	-	-	-	-	Y
Day x symbol FE	Y	Y	Y	Y	Y

Table A.5: Do Bulls' newsfeeds reflect their bullish sentiment? Messages instead of impressions

This table is a variant of Table 5, examining whether declared bulls (bears) see more bullish (bearish) posts about that symbol on the days (t+1 to t+50) following their sentiment declaration on day t. Observations are at the user-symbol-day ( $i, s, t$ ) level, and the sample is conditional on seeing at least one sentiment-stamped message in the time period of the dependent variable. The dependent variable is the number of bullish (bearish) messages seen by the focal user about that symbol. The specification follows Equation (4). Standard errors (in brackets) clustered by user. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	N messages <sub>i,s,t+1</sub>		N messages <sub>i,s,(t+2 → t+10)</sub>		N messages <sub>i,s,(t+11 → t+30)</sub>		N messages <sub>i,s,(t+31 → t+50)</sub>	
	(1) Bullish	(2) Bearish	(3) Bullish	(4) Bearish	(5) Bullish	(6) Bearish	(7) Bullish	(8) Bearish
1Declare Bull <sub>i,s,t</sub>	2.57*** [0.11]		12.69*** [0.57]		22.16*** [1.09]		20.60*** [1.11]	
1Declare Bear <sub>i,s,t</sub>		0.74*** [0.07]		2.71*** [0.31]		3.71*** [0.57]		2.81*** [0.57]
# obs.	3,206,087	3,206,087	5,732,782	5,732,782	6,562,290	6,562,290	6,135,101	6,135,101
# clusters (users)	123,718	123,718	163,448	163,448	170,302	170,302	157,950	157,950
R <sup>2</sup>	0.43	0.37	0.50	0.43	0.53	0.46	0.54	0.45
Mean of dependent var.(%)	10.63	1.46	43.86	6.32	75.49	11.47	70.13	11.17
Effect size (% of mean)	24	51	29	43	29	32	29	25
User FE	Y	Y	Y	Y	Y	Y	Y	Y
Day x symbol FE	Y	Y	Y	Y	Y	Y	Y	Y

Table A.6: Echo chambers and subsequent beliefs

This table provides evidence on whether users that place themselves in an echo chamber on days  $t + 1$  to  $t + 10$  are more likely to declare their original sentiment on days  $t + 11$  to  $t + 30$ . The specification follows equation 7:

$$(\# \text{ Bull Days} - \# \text{ Bear Days})_{i,s,(t+11 \rightarrow t+30)} = \beta_1 \mathbb{1} \text{Declare Bull}_{i,s,t} + \beta_2 (\mathbb{1} \text{Declare Bull}_{i,s,t} \times \mathbb{1} \text{Net Bull Follows}_{i,s,(t+1 \rightarrow t+10)}) + \lambda_i + \gamma_{s,t} + \epsilon_{i,s,t}$$

The dependent variable for columns (1) and (2) is the number of days in which a user ( $i$ ) declares themselves bullish about a stock ( $s$ ) in the  $t + 11$  to  $t + 30$  window, minus the number of days they declare themselves as bearish about  $s$  over the same period. For columns (3) and (4) the dependent variable is an indicator equal to one if the # of bull days exceeds the # of bear days for user  $i$  and stock  $s$  over that period.  $\mathbb{1} \text{Net Bull Follows}_{i,s,(t+1 \rightarrow t+10)}$  is an indicator for a user following more bulls than bears about a stock over the  $t + 1$  to  $t + 10$  window. Standard errors (in brackets) clustered by user. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Dep. var.: # net Bull days $_{s,(t+11 \rightarrow t+30)}$		$\mathbb{1}$ if Net Bull $_{s,(t+11 \rightarrow t+30)}$	
	(1)	(2)	(3)	(4)
$\mathbb{1} \text{Declare Bull}_{i,s,t}$	2.27*** [0.03]	2.18*** [0.03]	0.26*** [0.00]	0.26*** [0.00]
$\mathbb{1} \text{Bull}_{i,s,t} \times \mathbb{1} \text{Net Bull follows}_{i,s,(t+1 \rightarrow 10)}$		1.68*** [0.02]		0.15*** [0.00]
# observations	13,893,332	13,893,332	13,893,332	13,893,332
# clusters (users)	305,967	305,967	305,967	305,967
$R^2$	0.50	0.51	0.42	0.43
Mean of dependent var.(%)	1.88	1.88	0.42	0.42
User FE	Y	Y	Y	Y
Day x stock FE	Y	Y	Y	Y

Table A.7: Selective exposure and stock-day returns

This table examines how proxies for selective exposure behavior on StockTwits (received disagreement and received uncertainty), together with sender disagreement, relate to daily stock returns. Observations are at the stock-day ( $s,t$ ) level. The table presents estimates of equation 9, but with a different dependent variable (cumulative abnormal stock returns), and additional interactions with an indicator for stock-days with above median bullishness. *Sender Disagreement* captures dispersion of sentiment among posts about a stock ( $s$ ) (i.e. the standard deviation of sentiment across posts). *Received Disagreement* captures how disagreement in posts is distributed across receivers (i.e. the mean across receivers of the standard deviation of received sentiment). *Received Uncertainty* captures the dispersion of sentiment about a stock across newsfeeds (i.e. the standard deviation across receivers of the mean of received sentiment). We standardize the disagreement measures by subtracting the mean and dividing by the standard deviation, both measured over the full sample period. Controls follow those in Table 12, and include fixed effects for day, stock-month, and message number. Message number fixed effects are defined for days with 0, 1, 2, 3, 4, 5-10, 11-30, and over 30 messages. Standard errors (in brackets) are separately clustered by stock and day. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Cumulative abnormal returns $_{s,(t+1 \rightarrow \dots)}$			
	(1) to t+5	(2) to t+5	(3) to t+10	(4) to t+30
$\mathbb{1}$ if >median Bullishness $_{s,t}$	-0.69*** [0.09]	-0.51*** [0.13]	-0.80*** [0.21]	-0.70*** [0.17]
Sender Disagreement $_{s,t}$		0.30*** [0.11]	0.33*** [0.11]	0.17 [0.13]
$\mathbb{1}$ if >median Bullishness $_{s,t} \times$ Sender Disagreement $_{s,t}$		-0.47*** [0.16]	-0.36* [0.20]	0.17 [0.24]
Received Disagreement $_{s,t}$		0.18** [0.08]	0.13 [0.14]	0.33*** [0.05]
$\mathbb{1}$ if >median Bullishness $_{s,t} \times$ Received Disagreement $_{s,t}$		-0.43** [0.17]	-0.52* [0.29]	-1.17*** [0.25]
Received Uncertainty $_{s,t}$		-0.01 [0.05]	-0.02 [0.05]	-0.02 [0.04]
$\mathbb{1}$ if >median Bullishness $_{s,t} \times$ Received Uncertainty $_{s,t}$		0.21 [0.18]	0.29 [0.26]	0.73*** [0.27]
$\mathbb{1}$ Media Article $_{s,t}$	0.02 [0.11]	0.01 [0.11]	-0.14 [0.22]	0.11 [0.13]
Log GoogleASVI $_{s,t}$	-0.61** [0.28]	-0.62** [0.28]	-0.69 [0.75]	-1.13*** [0.41]
Volatility $_{s,(t-5 \text{ to } t-1)}$	-8.19*** [3.03]	-8.17*** [3.03]	-8.87** [4.41]	-5.07 [4.20]
Abnormal returns $_{s,t}$	-0.27*** [0.01]	-0.27*** [0.01]	-0.44*** [0.06]	-0.41*** [0.02]
Cum. Abnormal Returns $_{s,(t-5 \text{ to } t-1)}$	-0.21*** [0.02]	-0.21*** [0.02]	-0.35*** [0.07]	-0.31*** [0.03]
Cum. Abnormal Returns $_{s,(t-30 \text{ to } t-6)}$	-0.15*** [0.01]	-0.15*** [0.01]	-0.21*** [0.03]	-0.18*** [0.02]
# obs.	339,264	339,264	337,251	329,196
# clusters (stock)	760	760	760	760
# clusters (day)	1,880	1,880	1,875	1,855
$R^2$	0.32	0.32	0.48	0.77
Day FE	Y	Y	Y	Y
Month x stock FE	Y	Y	Y	Y
Message number FE	Y	Y	Y	Y

Table A.8: Retail turnover and trades

This table follows Table 12, changing only the dependent variable, and uses the procedure in Boehmer et al. (2021) to identify retail trades in the TAQ database from their preferential pricing. In columns (1) to (2) the dependent variable is abnormal log *retail* turnover, calculated as in Table 12 but with trading volume replaced by shares traded by retail investors. Columns (3) to (8) use the number of retail trades as the dependent variable. Standard errors (in brackets) clustered by day and by stock. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Dep var: Ab.In(retail turnover) <sub>s,t</sub>		# retail trades <sub>s,t</sub>		# buy retail trades <sub>s,t</sub>		# sell retail trades <sub>s,t</sub>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sender Disagreement <sub>s,t</sub>	0.03*** [0.00]	0.02*** [0.00]	64.24*** [9.12]	61.28*** [7.82]	30.24*** [5.26]	30.37*** [4.47]	34.06*** [4.03]	30.82*** [3.50]
Received Disagreement <sub>s,t</sub>	0.02*** [0.00]	0.01*** [0.00]	99.40*** [9.34]	34.38*** [8.06]	54.65*** [5.70]	18.68*** [4.84]	44.72*** [4.03]	16.55*** [3.62]
Received Uncertainty <sub>s,t</sub>	-0.02*** [0.00]	-0.02*** [0.00]	-40.62*** [8.39]	-45.52*** [7.87]	-19.62*** [4.85]	-23.06*** [4.47]	-21.01*** [3.67]	-22.33*** [3.51]
Dependent variable <sub>s,t-1</sub>		0.07*** [0.01]		0.42*** [0.02]		0.42*** [0.02]		0.40*** [0.02]
1Media Article <sub>s,t</sub>		0.13*** [0.01]		315.41*** [25.47]		169.08*** [14.61]		147.43*** [11.21]
Log GoogleASVI <sub>s,t</sub>		0.37*** [0.02]		2,051.42*** [203.80]		1,132.11*** [113.53]		928.81*** [92.09]
Abnormal Log Turnover <sub>s,t-1</sub>		0.14*** [0.01]		-316.15*** [24.63]		-181.07*** [14.32]		-120.22*** [11.21]
Volatility <sub>s,(t-5 to t-1)</sub>		0.18*** [0.04]		-2,088.59*** [296.97]		-1,101.58*** [167.99]		-967.05*** [130.28]
Cum. Abnormal Returns <sub>s,(t-5 to t-1)</sub>		-0.06*** [0.01]		-80.83 [125.92]		-54.41 [71.64]		-2.91 [56.47]
Cum. Abnormal Returns <sub>s,(t-30 to t-6)</sub>		-0.08*** [0.01]		-490.04*** [86.28]		-278.08*** [51.07]		-207.80*** [36.22]
# obs.	343,588	343,011	344,507	343,923	344,783	344,404	344,773	344,437
# clusters (stock)	859	859	862	862	862	862	862	862
# clusters (day)	1,876	1,875	1,876	1,875	1,876	1,875	1,876	1,875
R <sup>2</sup>	0.80	0.82	0.74	0.79	0.72	0.77	0.74	0.79
Mean of dependent var.	0.47	0.47	2,188.80	2,190.55	1,164.39	1,164.91	1,023.88	1,024.29
Day FE	Y	Y	Y	Y	Y	Y	Y	Y
Month x stock FE	Y	Y	Y	Y	Y	Y	Y	Y
Message number FE	Y	Y	Y	Y	Y	Y	Y	Y

Table A.9: Evidence on the persistence of disagreement

This figure presents coefficients from the following specification (equation 10):

$$Sender\ Disagreement_{s,t+x} = \beta_1 Sender\ Disagreement_{s,t} + \beta_2 Received\ Disagreement_{s,t} + \beta_3 Received\ Uncertainty_{s,t} + \varepsilon_{s,t}$$

Where the dependent variable is measured  $x$  days after the focal day ( $t$ ), and ranges from 1 (columns 1 and 5) to 50 (columns 4 and 8). The regressions in columns (5) to (8) also include message bin fixed effects, stock fixed effects, and day fixed effects. Standard errors (in brackets) clustered by day and by stock. \*\*\* 1%, \*\* 5%, \* 10% significance level.

	Dep. var.: Sender Disagreement <sub>s,t+x</sub>							
	(1) t+1	(2) t+10	(3) t+30	(4) t+50	(5) t+1	(6) t+10	(7) t+30	(8) t+50
Sender Disagreement <sub>s,t</sub>	0.417*** [0.007]	0.260*** [0.010]	0.216*** [0.010]	0.192*** [0.010]	0.313*** [0.005]	0.123*** [0.005]	0.072*** [0.005]	0.050*** [0.005]
Received Disagreement <sub>s,t</sub>	0.097*** [0.005]	0.085*** [0.006]	0.081*** [0.007]	0.081*** [0.007]	0.056*** [0.003]	0.035*** [0.003]	0.024*** [0.004]	0.023*** [0.003]
Received Uncertainty <sub>s,t</sub>	-0.000 [0.004]	-0.006 [0.005]	-0.009 [0.006]	-0.011* [0.006]	0.006** [0.003]	-0.002 [0.003]	-0.000 [0.003]	-0.004 [0.003]
# obs.	311,829	276,145	257,529	245,037	311,827	276,140	257,524	245,034
# clusters (stock)	1,075	1,072	1,069	1,068	1,075	1,070	1,067	1,068
# clusters (day)	1,886	1,877	1,857	1,837	1,884	1,874	1,854	1,834
$R^2$	0.24	0.11	0.08	0.06	0.32	0.24	0.23	0.23
Mean of dependent var.	0.05	0.05	0.06	0.07	0.05	0.05	0.06	0.07
Message number FE	-	-	-	-	Y	Y	Y	Y
Stock FE	-	-	-	-	Y	Y	Y	Y
Day FE	-	-	-	-	Y	Y	Y	Y