Echo Chambers

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

January 15, 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), 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. We are grateful to helpful suggestions from Snehal Banerjee, Francesco D'Acunto, James Doran, Anastassia Fedyk, John Garfinkel, Rawley Heimer, David Hirshleifer, Alberto Rossi and Paul Tetlock. In addition, this draft has benefited from workshop and conference presentations at Social Finance Day, the 2020 Australasian AFBC Conference, Finance in the Cloud III, Columbia University, Baruch, Baylor, Carnegie Mellon, George Washington, Georgetown, Indiana, Norwegian School of Economics, UC San Diego, University of Victoria, and University of Colorado-Boulder.

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 a difficult time explaining the high trading volume observed in financial markets. Difference of opinion models argue that high levels of 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."

Empirical 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 affects 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.

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 approximately 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), who observe 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 1.4% lower (higher) future abnormal returns over the next 5 trading days. However, the size of this underperformance depends on whether the declaration was made inside an echo chamber. For example, for a declaration by a user who has no diversity in his newsfeed over the prior 30-days (i.e. all signals received were the same), the underperformance jumps to 1.7%. On the other hand, for a declaration by a user who has maximum diversity in his newsfeed over the proceeding month (i.e. half the signals were bullish and half were bearish) the underperformance is 0.58 percentage points 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 each receiver's signal. 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 central contribution is to provide novel 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, which is 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,[1] the behavioral finance literature largely focuses on how individuals interpret information, which is a feature of models of confirmation bias (Rabin and Schrag, 1999; Camerer, 1999), as well as empirical evidence on confirmation-biased behaviors (Pouget et al., 2017; Charness and Dave, 2017). Our evidence of echo chambers is evidence of confirmation-biased information acquisition, which slows the arrival of new information that is inconsistent with the individual's prior. Given the importance of information arrival for the updating of beliefs, the emergence of echo chambers provides a rationale for why beliefs diverge in the first place.

Our findings also contribute to the broader literature on selective exposure, which dates back to the original theory of cognitive dissonance (Festinger, 1957).[2] Most of the evidence of selective

[1]The behavioral finance literature has long recognized that confirmation bias could manifest in financial contexts. In perhaps the earliest reference to the concept in behavioral economics, Thaler (1987)'s preface to the of *Journal Economic Perspectives* series on anomalies argues that confirmation bias could be one explanation for the literature's strict adherence to a rational paradigm.

[2]The literature in psychology, communications and politics has identified other possible reasons for selective exposure. For example, 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).

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.

Our findings on selective exposure also relate to the economics literature on information avoidance (Golman et al., 2017). Most notably, our findings are distinct from the "optimism maintenance" or motivated beliefs channel, which posits that optimistic beliefs are valued unto themselves, giving rise to wishful thinking (Brunnermeier and Parker, 2005; Benabou, 2015; Banerjee et al., 2019). While motivated beliefs could explain why bulls subscribe to 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.

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 on attention has focused on market aggregates of attention to particular stocks, either by retail investors or institutional investors (Da et al., 2011; Ben-Rephael et al., 2017, 2020; Fedyk, 2019), but information on individual information choices is scarce.[3] Prior work has 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 their cross-sectional selection of *which* information to consume serves to amplify dispersion in their initial beliefs.

Finally, our findings contribute to the recent literature on sources of disagreement, and the implications of disagreement 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., see seminal contributions by Kandel and Pearson, 1995, and Hong and

[3] 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; D'Acunto et al., 2020). By contrast, we examine agents' *endogenous* information choices after observing their ex ante beliefs.

Stein, 1999).[4] 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).[5] 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.[6]

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 taken 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

[4]Recent empirical work has shown that modeling differences contribute to disagreement and to trading volume. Cookson and Niessner (2020) provide evidence that differential interpretation accounts for about half of the dispersion of opinion. 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.

[5]For example, Chang et al. (2014) provides evidence of slow information diffusion in Chinese financial markets, showing that linguistically-diverse areas express more diverse opinions than linguistically-similar areas. 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.

[6]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.

before the 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% (> 9,000) 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:

$$Follow\ Bear_{i,s,t+1} = \xi_t + \gamma_s + \lambda_i + \varepsilon_{i,s,t} \tag{1}$$

in which the dependent variable $Follow\ Bear_{i,s,t+1}$ is an indicator for whether user i followed more bearish than bullish users about symbol s on day $t+1$. The regression includes day (ξ_t), user (λ_t) and symbol (γ_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 provides an interesting contrast. The bullish echo chamber stocks tend to be pure play stocks in very particular markets: six of the top ten bullish echo chambers are 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 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.

3.3 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.

3.3.1 Choosing Information Sources

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

$$Follow\ Bull_{i,s,t\to t+k} = \gamma_i + \eta_{s,t} + \beta_1 Declare\ Bull_{i,s,t=0} + \varepsilon_{i,s,t} \tag{2}$$

where the dependent variable $Follow\ Bull_{i,s,t\to t+k}$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 $Declare\ Bull_{i,s,t=0}$, which is an indicator equal to 1 if user i declared bullish sentiment about stock s on date t. The coefficient of interest β_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 account for individual heterogeneity (e.g., optimism or rate of following), we include a person fixed effect γ_i in some specifications and a person-symbol fixed effect $\gamma_{i,s}$ 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 β_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-person 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 β_1 coefficient in column (1) corresponds to a 41 percent increase in this likelihood, with a similar magnitude (a 39 percent increase, or 1.78 percentage points) over the next 5 days in column (2). Both are strongly statistically significant.

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 (3) 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 (4) through (7) consider the subset of observations for which the focal user follows at least one new user after their sentiment declaration. Column (4) indicates that users are 12 percentage points more likely to follow other same-sentiment users. Columns (5) through (7) 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 Table A.3. 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.

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.

3.3.2 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 $SPOT"). We construct indicator variables for whether a user i buys ($Buyer_{i,s,t}$) or sells ($Seller_{i,s,t}$) the security s on date t.

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:

$$Follow\ Bull_{i,s,t\to t+k} = \gamma_i + \eta_{s,t} + \beta_1 Declare\ Bull_{i,s,t=0} + \beta_2 Declare\ Bull_{i,s,t=0} \times trade_{i,s,t=0} + \varepsilon_{i,s,t} \tag{3}$$

where the dependent variable $Follow\ Bull_{i,s,t\to t+k}$ 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 the security ($Declare\ Bull_{i,s,t=0} \times trade_{i,s,t=0}$). In this specification, the coefficient β_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 purchase*. The coefficient on the interaction β_2 captures the change in the baseline selective exposure rate if the user also declares that a purchase was made.

Table 4 presents the results from estimating equation (3). The odd columns of the table present the baseline estimates without the interaction, whereas the even columns also introduce the interaction with buying. The baseline estimate in column (1) shows a very 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 at 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.

3.3.3 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:

$$Bullish\ user\ impressions_{i,s,t+k} = \gamma_i + \eta_{s,t} + \beta_1 Declare\ Bull_{i,s,t} + \varepsilon_{i,s,t} \tag{4}$$

where $Bullish\ user\ impressions_{i,s,t+k}$ 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. $Declare\ Bull_{i,s,t}$ is the same indicator of declared bullish sentiment we use in our follower specifications. The coefficient of interest β_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.[7]

[7] 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.2 for details.

3.3.4 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:

$$Likes\ of\ bullish\ messages_{i,s,t+k} = \gamma_i + \eta_{s,t} + \beta_1 Declare\ Bull_{i,s,t} + \varepsilon_{i,s,t} \tag{5}$$

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 $Declare\ Bull_{i,s,t}$ is the same indicator of declared bullish sentiment we use in our previous regressions. The coefficient of interest β_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 security-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.

3.4 Heterogeneity and Mechanisms

Next, we turn to evaluating two 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 an investor's experience and activity.

3.4.1 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 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:

$$\begin{aligned} follow\ Bull_{i,s,t+1\rightarrow t+2} = \gamma_i + \eta_{s,t} + \beta_1 Declare\ Bull_{i,s,t} \\ + \beta_2 Declare\ Bull_{i,s,t} \times EA\ day_{s,t+1} + \varepsilon_{i,s,t} \end{aligned} \quad (6)$$

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 the day after the sentiment declaration day (day $t+1$). 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. When we split this main effect out separately for bull follows and bear follows, we observe that the increase in selective exposure is driven by both types of connections – bulls follow bulls (column 4) and bears follow bears (column 7) to a greater extent when earnings news arrives. 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 merely reflect an increase in StockTwits activity on these days, nor an increase in optimism, because the specifications employ day x symbol fixed effects. Instead, we find that both bulls and bears are more likely to put themselves in echo chambers (i.e., follow users of the same declared sentiment) when earnings news arrives.

We then ask whether selective exposure is further driven by the content of the news. To evaluate 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 $t+2$). In a regression that only includes earnings days in column (5), 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 (8), we find no change in selective exposure for bears on earnings days with positive information. This latter finding is consistent with bears more likely to seek other bears on earnings days but not being sensitive to the content of the earnings news.

3.4.2 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.[8] 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 this phenomenon could have real financial market consequences, a question we address in the next section.

In columns (5) through (7), 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 *Declare Bull* with an indicator for whether the user is highly active – proxied by following more than the median number of users (column 5), having more than the median number of followers (column 6), and posting more than the median number of messages (column 7). All three columns deliver a consistent message: more active users exhibit greater selective exposure. In all three specifications, the main effect is also 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.

[8] Though StockTwits users self-report their experience, Cookson and Niessner (2020) validate the StockTwits experience data.

4 Selective Exposure and Market Outcomes

In this section, we present our findings on how selective exposure to information affects returns and trading volume.

4.1 Returns

First, to consider how being in an echo chamber affects subsequent returns, we analyze the relationship between declarations of bullish or bearish sentiment and the *ex post* returns on those stocks. We estimate:

$$Abnormal\ return_{i,s,(t+1 \to t+\tau)} = \beta_0 Bull_{i,s,t} + X'_{s,t}\delta + \gamma_t + \phi_{i,s} + \eta_{s,month} + \varepsilon_{i,s,t} \tag{7}$$

where i indexes users, s indexes stocks and t days. The dependent variable *Abnormal Return* is percentage abnormal return for stock s (stock return minus CRSP value-weighted market return) in the forward-looking window from 1 to τ days after user i makes a bullish (or bearish) declaration about stock s. The main tests employ two time windows: a five-day window $(t+1, t+5)$ and a ten-day window $(t+1, t+10)$.[9] The specifications include date fixed effects, user-stock fixed effects, and stock-month fixed effects. The vector of controls $X_{s,t}$ includes abnormal returns for the preceding thirty trading days, and we cluster standard errors by user, and by permno-day.

The results from estimating equation (7) are reported in column (1) of Table 9. We find an inverse relationship between beliefs on StockTwits and future returns: bullish (bearish) declarations on StockTwits are associated with 1.41% 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 underperformance of 1.85%. The marginal impact of adding additional days does not increase the magnitude, nor does the return reaction revert. This negative return predictability following sentiment declarations suggests opinions on StockTwits are misinformed on average.

If selective exposure worsens decision-making then sentiment declarations made in echo chambers will be associated with weaker *ex post* return performance. To examine this, we calculate the standard deviation of signals received by user i about security s over the preceding thirty days, as-

[9] We skip one day 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 $t+1$ begins with Thursday's close-to-close return, (measured from Wednesday 4pm to Thursday 4pm).

signing a value of 1 to bullish signals and -1 to bearish signals as in Cookson and Niessner (2020). We call this variable *sd received signals*$(30\ days)$. 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 received signals*=0, while a user that saw 2 bullish signals and 2 bearish signals about Tesla would have a *sd received signals*=1. Users in an echo chamber will see a concentration of similar signals and have a low *sd received signals,* while those outside an echo chamber will see a diversity of signals and have a high *sd received signals*.

We add this measure of signal diversity and its interaction with $Bull_{i,s,t}$ to the abnormal return specification (7). The coefficient of interest is on $Bull_{i,s,t} \times sd\ rec.\ signals(30\ days)_{i,s,t}$, which estimates how the return underperformance depends on whether the user has seen greater diversity in signals over the thirty days preceding their declaration. In Table 9 we find a positive interaction coefficient, indicating that declarations made in echo chambers (i.e., less diversity of signals) are associated with greater underperformance. For example, the estimated main effect on $Bull_{i,s,t}$ in column (2) of Table 9 implies that a declaration by a user in a pure echo chamber (i.e., *sd received signals*=0) is associated with 1.67% underperformance 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 prior 30-days (i.e. maximum diversity of signals, *sd received signals*=1) is associated with 0.58% less underperformance over the 5-day return window, reducing the underperformance by more than a third. In column (6), the analogous test for a 10-day window yields an estimated underperformance of 2.31% for sentiment declarations made in an echo chamber, and underperformance is reduced by 1.02% 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 dummies – novice, intermediate and professional – interacted with $Bull_{j,s,t}$. The baseline (omitted) category is users who do not specify their experience. We find a monotonic relationship between experience and underperformance, with professionals outperforming intermediates, who outperform novices.[10] All three categories outperform the baseline – the missing experience category. Importantly, the inclusion of these controls does not affect our conclusion regarding underperformance and echo

[10] When we test the equality of coefficients between $Bull_{i,s,t} \times Novice_i$ and $Bull_{i,s,t} \times Intermediate_i$ in column (3) we reject the null with a p-value of 0.015 (and a p-value of 0.001 for $Bull_{i,s,t} \times Professional_i$).

chambers: the coefficient on $Bull_{i,s,t} \times sd\ rec.\ signals(30\ days)_{i,s,t}$ changes from 0.58% (column 2) to 0.52% (column 4). The inclusion of these experience interactions also has little effect on the coefficient of interest in the 10-day regressions (column 8).

Figure 6 illustrates the dynamics of underperformance in echo chambers. Rather than estimating abnormal returns accumulated over a window, the figure presents the daily abnormal return coefficients for event days 1 through 30. We use the same controls and fixed effects as in equation (7). The underperformance coming from echo-chambers is large in the days after the sentiment declaration, and it decays to approximately zero by event day 10. For those in a pure echo chamber, underperformance begins at -0.46% on day $t+1$ and declines to -0.19% on day $t+10$. By contrast, for users with maximum diversity of signals, the underperformance begins at -0.31% on day $t+1$ and declines to -0.13% on day $t+10$, and is statistically indistinguishable from zero.

The results indicate that the average sentiment declaration on StockTwits is a mis-reaction to information, 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.

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 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 10 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 Operationalizing Selective Exposure

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 $j \in \{2, \ldots, N_{i,s,t}\}$. 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 will be 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; Cookson and Niessner, 2020).

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 $Received\ Disagreement_{s,t}$ 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. We can then estimate the effect of selective exposure on market outcomes as follows:

$$\begin{aligned} Y_{s,t} = {} & \eta_t + \xi_{s,m} + \beta_1 Sender\ Disagreement_{s,t} + \beta_2 Received\ Disagreement_{s,t} \\ & + \beta_3 Received\ Uncertainty_{s,t} + X'_{s,t}\gamma + \varepsilon_{s,t}, \end{aligned} \tag{8}$$

where $Y_{s,t}$ is abnormal log turnover of stock s on date t, and $X_{s,t}$ are time-varying controls.[11] We also include day and stock-month fixed effects (η_t and $\xi_{s,m}$), as well as fixed effects for eight bins capturing the number of messages about a given stock on that day. In this specification, the coefficients of interest are β_2 and β_3. β_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 $\beta_2 < 0$. β_3 measures how increasing received disagreement (i.e. more dispersion of received signals across users – more selective exposure) relates to turnover; we expect β_3 to be positive. These two *Received* measures capture different aspects of selective exposure behavior. Thus, our preferred

[11] Our specifications for *Abnormal Log Turnover* include the same set of control variables as employed in Cookson and Niessner (2020): 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 they are consistent across time.

specification includes both $Received\ Disagreement_{st}$ and $Received\ Uncertainty_{st}$.

4.2.3 Information Silos and Trading Volume

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 ($Received\ Uncertainty_{s,t}$). Table 11 reports the results from estimating the specifications in equation (8). These specifications follow closely the measurement and controls employed in Cookson and Niessner (2020), which helps provide a benchmark for our results.[12] 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 $Received\ Uncertainty_{s,t}$. Column (1), which includes the $Sender\ Disagreement_{s,t}$ measure by itself, provides a somewhat smaller estimate (0.017) to the equivalent specification in Cookson and Niessner (2020): a one standard deviation increase in disagreement increases abnormal turnover by 4% of its mean.

Column (2) adds the $Received\ Disagreement_{s,t}$ 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 $Received\ Disagreement_{s,t}$ is smaller than the coefficient on $Sender\ Disagreement_{s,t}$ (0.004 versus 0.015).

Column (3) includes the average within-newsfeed dispersion ($Received\ Uncertainty_{s,t}$), 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 trading volume. The magnitude of the coefficient on $Received\ Uncertainty_{s,t}$ 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 $Sender\ Disagreement_{s,t}$ increases abnormal turnover by approximately 5.8% of the mean, which is similar to combined effect of the information

[12] We employ a different sample from Cookson and Niessner (2020). Our data range is 2013 to June 2020 (versus 2013 to late 2014), and we consider 1078 stocks (versus the top 100 stocks by StockTwits message volume).

siloing variables (i.e, reducing *Received Uncertainty*$_{s,t}$ and increasing *Received Disagreement*$_{s,t}$ each by one standard deviation).

5 Conclusion

Selective exposure to confirmatory information has been documented in a variety of settings, from politics to religion to vehicle ownership. It appears that once people form a belief about immigration or Christianity or Chevy trucks, they selectively choose information which 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 behave in other settings: by following users who share their beliefs, they build a personalized newsfeed which 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, etc. 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. *Available at SSRN 3335257*.

Banerjee, S. and I. Kremer (2010). Disagreement and learning: Dynamic patterns of trade. *The Journal of Finance 65*(4), 1269–1302.

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 (2020). Information Consumption and Asset Pricing. *Journal of Finance Forthcoming*.

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. *Revu d'economie politique 125*(5), 665–685.

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: Reuinifying 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. *Rock Center for Corporate Governance at Stanford University 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. 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. *Available at SSRN 2705545*.

D'Acunto, F., D. Hoang, M. Paloviita, M. Weber, et al. (2020). Effective policy communication: Targets versus instruments.

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 horce 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.

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.

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. *SSRN Working Paper 2439040*.

Olafsson, A. and M. Pagel (2017). The ostrich in us: Selective attention to financial accounts, income, spending and liquidity. *NBER Working Paper No. 23945*.

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.

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.

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.

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

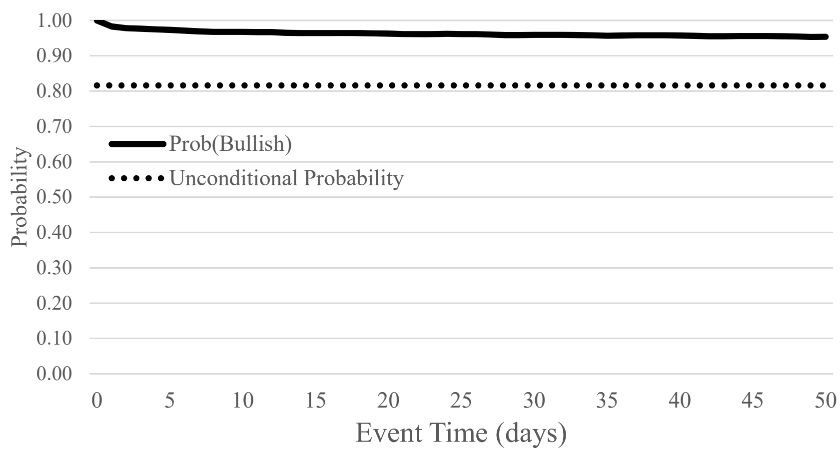
Ziemke, D. A. (1980). Selective exposure in a presidential campaign contingent on certainty and salience. *Annals of the International Communication Association 4*(1), 491–511.

6 Tables and 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.

(a) Sentiment persistence for bullish users



(b) Sentiment persistence for bearish users

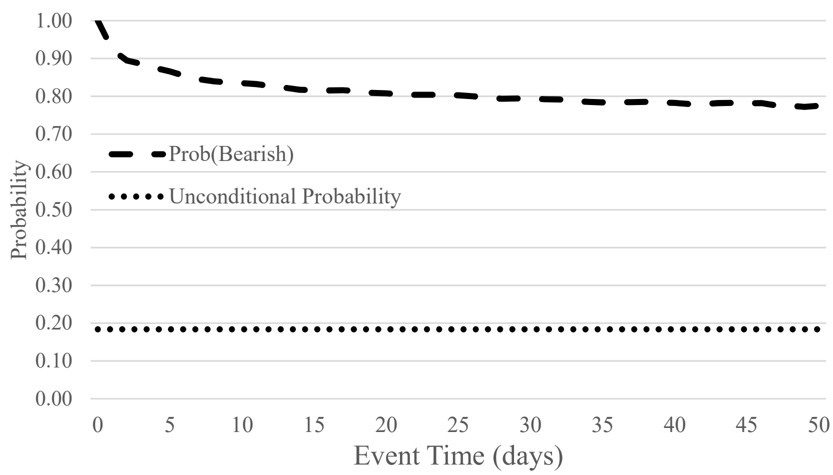
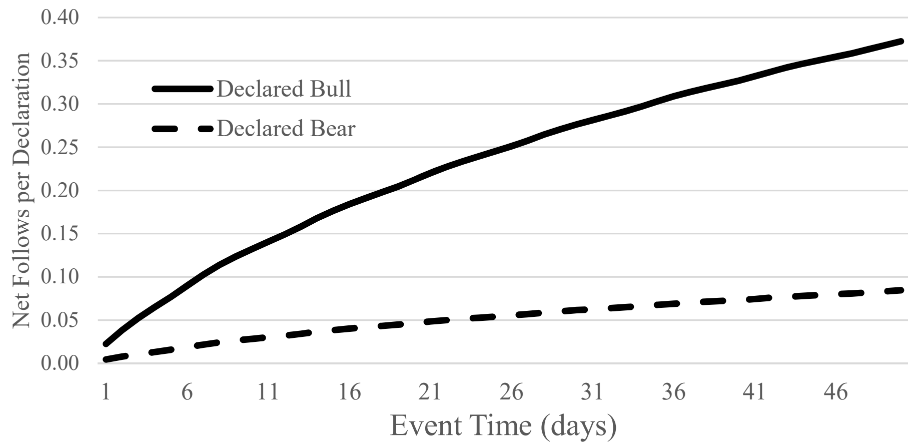


Figure 2: Who do users follow?

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.

(a) Cumulative net new follows of bullish users



(b) Cumulative net new follows of bearish users

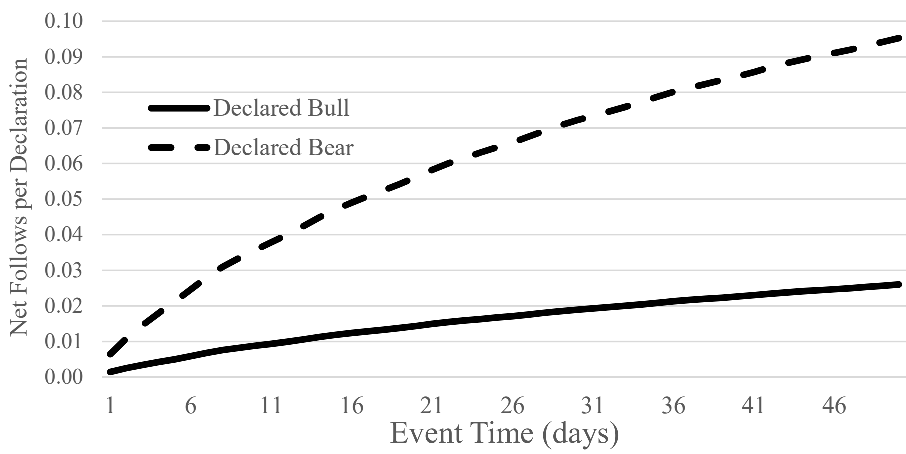
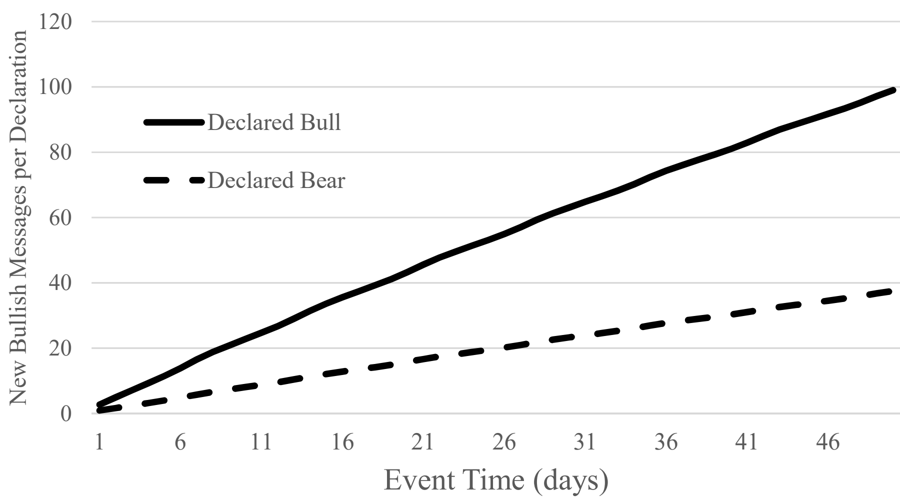


Figure 3: Do users' newsfeeds match their sentiment? Messages

Panel (a) plots the number of bullish messages subsequent to a user declaring as bullish 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.

(a) Number of bullish messages



(b) Number of bearish messages

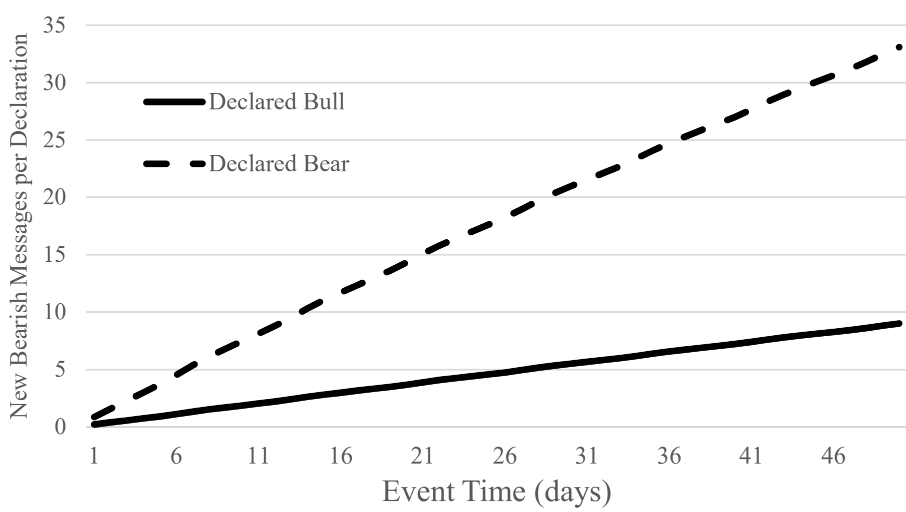
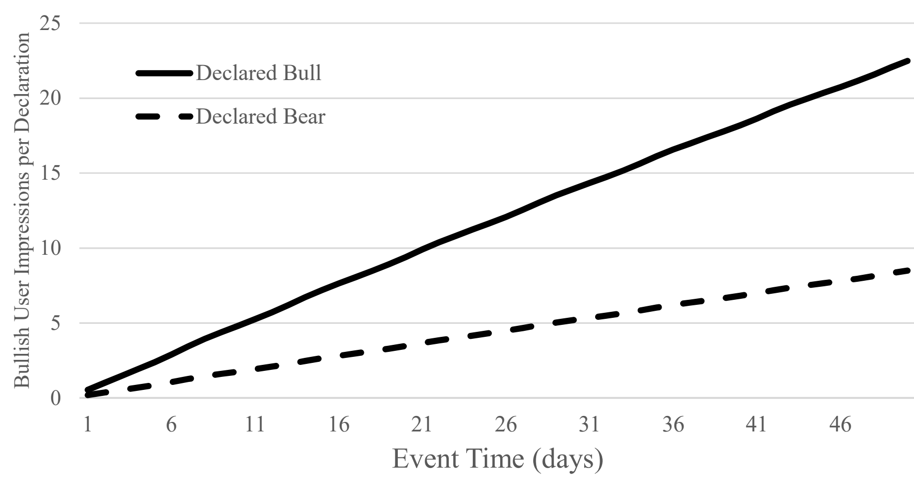


Figure 4: Do users' newsfeeds match their sentiment? User impressions

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.

(a) Number of bullish user impressions



(b) Number of bearish user impressions

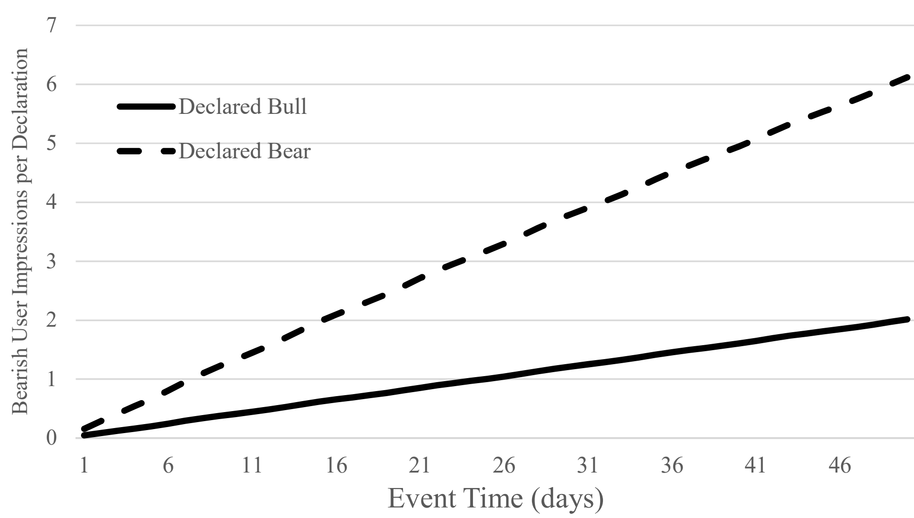
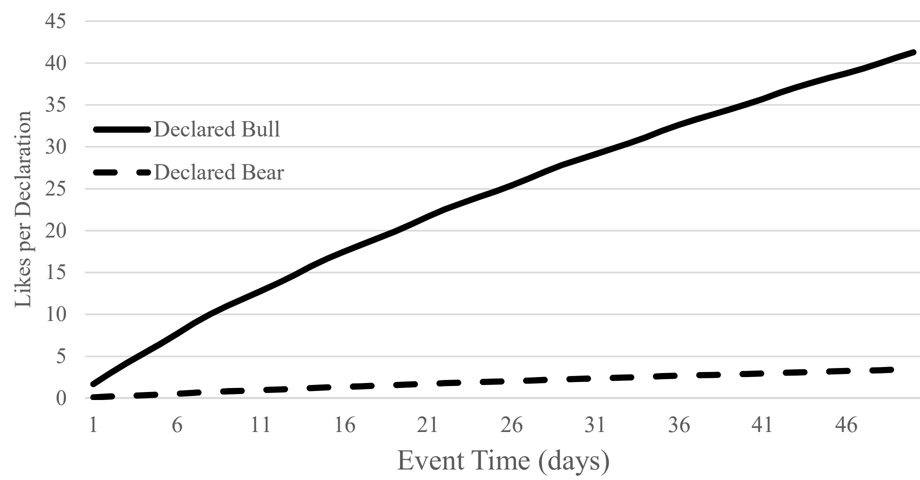


Figure 5: Do users' *likes* match their sentiment?

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.

(a) User Likes of bullish messages



(b) User Likes of bearish messages

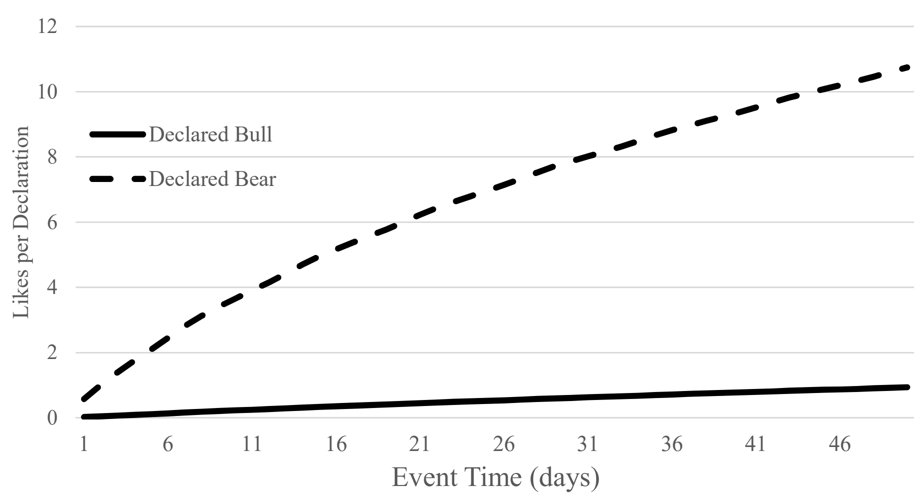


Figure 6: Return predictability of sentiment declarations made in echo chambers versus not

This figure presents the coefficients from equation (7) associated with a user's bullish (bearish) declaration on day $t = 0$. The dependent variable is abnormal returns on days $t + \tau$ (where τ ranges from 1 through 30 days). The black diamonds illustrate the return underperformance that results from trading on users' sentiment declarations. The blue triangles (the estimated coefficients on $bull \times sd\ received\ signals(30\ days)_{i,s,t}$) illustrate the degree to which maximum signal diversity ($sd\ received\ signals(30\ days)_{i,s,t} = 1$) mitigates the return underperformance on each day. The vertical bars represent 95% confidence intervals, clustering standard errors by user and permno-day.

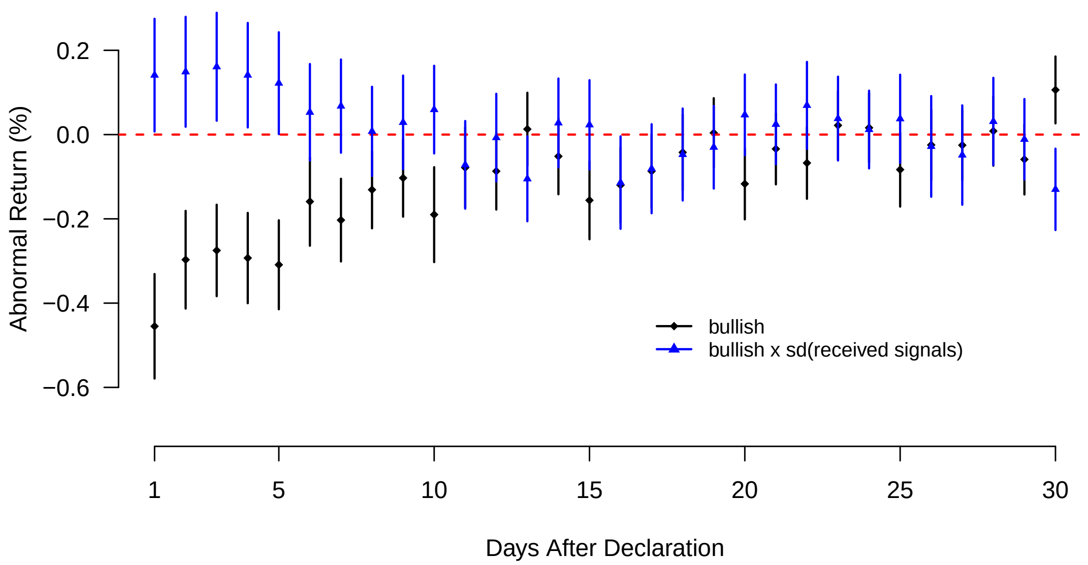


Figure 7: Do users receive *only messages with the same sentiment* more often than would be expected by chance?

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.

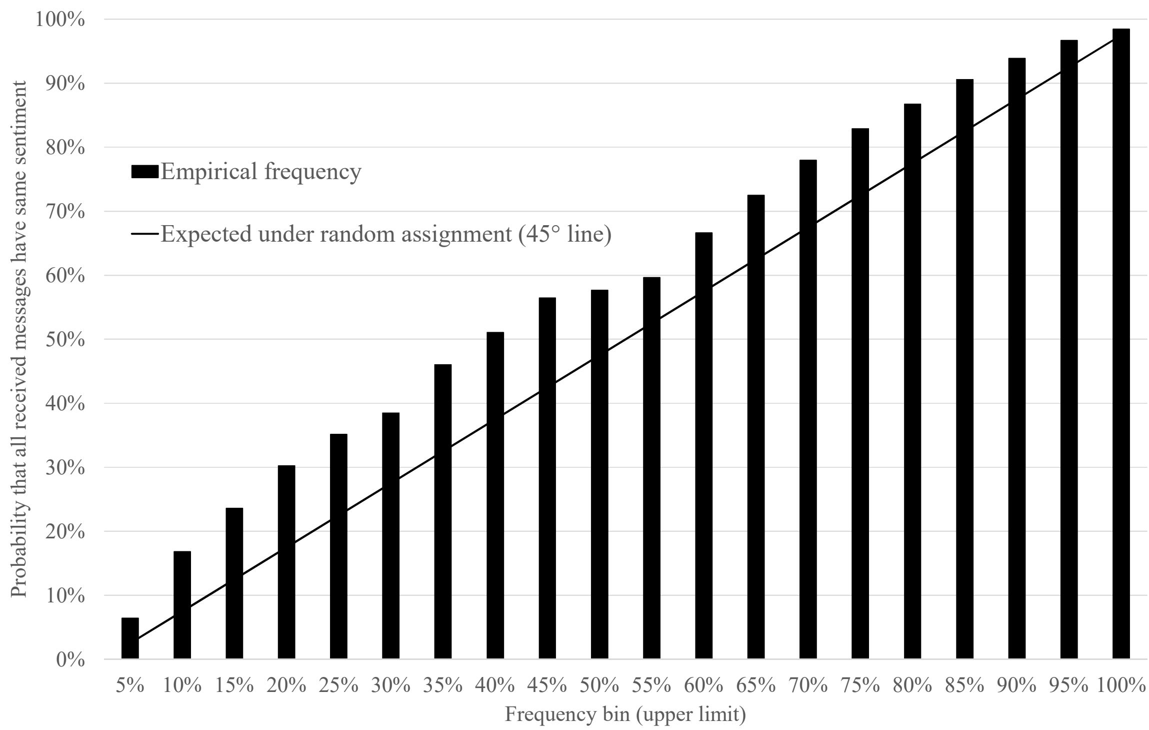


Table 1: Summary statistics

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 11).

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

	Totals	Totals	
Users	395,474	Symbols	1,208
Novice	18,771	CRSP (e.g., Tesla)	1078
Intermediate	23,341	Non-CRSP (e.g., Bitcoin)	130
Professional	9,595		
Unclassified	343,767		
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.
Main variables				
Abnormal log volume$_{\mathbf{s,t}}$	0.426	0.209	1.135	421,915
Sender disagreement$_{\mathbf{s,t}}$	0.003	0.175	0.995	421,915
Received disagreement$_{\mathbf{s,t}}$	0.007	-0.691	1.001	421,915
Received uncertainty$_{\mathbf{s,t}}$	-0.002	-0.639	0.995	421,915
Controls				
Std dev. abnormal returns$_{\mathbf{s,(t-5\ to\ t-1)}}$	0.046	0.029	0.077	421,915
Cum. abnormal returns$_{\mathbf{s,(t-5\ to\ t-1)}}$	0.013	-0.001	0.222	421,915
Cum. abnormal returns$_{\mathbf{s,(t-30\ to\ t-6)}}$	0.014	-0.019	0.416	421,915
Log Google ASVI$_{\mathbf{s,t}}$	0.532	0.666	0.448	421,915
$\mathbb{1}$ if Media article$_{\mathbf{s,t}}$	0.266	0.000	0.442	421,915
Num. of messages$_{\mathbf{s,t}}$	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	*Index ETF*
2	Roku	*Technology - Consumer*
3	Beyond Meat	*Technology - Food*
4	Energous Corp	*Technology - Wireless*
5	Tesla	*Automobile*
6	Snap Inc.	*Technology - Mobile app*
7	Bitcoin USD	*Cryptocurrency*
8	AVEO Pharmaceuticals	*Pharmaceutical*
9	Advanced Micro Devices	*Computer processors*
10	SunEdison Inc	*Renewable energy*

Bullish Echo Chambers

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

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

This table examines whether bullish (rather than bearish) users predominantly choose to follow bullish (rather than bearish) posters. Observations are at the user-symbol-day level. We examine a user's 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 $t+1$ (col 1), and days $t+1$ to $t+5$ (inclusive, cols 2 and 3). 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 (4) through (7) are run on a subsample of users that chose to make at least one new net follow in the relevant time window, which extends through $t+50$. We multiply the dependent variable by 100 to aid interpretation of coefficients as percentage points. Standard errors clustered by user are reported in brackets. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

	$\mathbb{1}$ x100 if new follows$_{\mathbf{i,s,t+x}}$ are more Bull than Bear						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Adding User-Symbol FE	Conditional on new follows	Conditional on new follows	Conditional on new follows	Conditional on new follows
	t+1	t+1 $\rightarrow$ t+5	t+1 $\rightarrow$ t+5	t+1 $\rightarrow$ t+5	t+2 $\rightarrow$ t+10	t+11 $\rightarrow$ t+30	t+31 $\rightarrow$ t+50
Declared Bull$_{\mathbf{i,s,t}}$	0.68***	1.78***	1.08***	12.21***	7.70***	3.34***	0.27
	[0.01]	[0.03]	[0.04]	[0.57]	[0.47]	[0.38]	[0.31]
# observations	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	259,476	63,594	77,761	84,419	82,568
R^2	0.12	0.19	0.34	0.72	0.78	0.83	0.87
Unconditional mean (%)	1.64	4.54	4.92	77.03	70.13	60.39	40.85
Effect size (% of mean)	41	39	22	16	11	6	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

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. Observations are at the user-symbol-day level. To focus 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 *Buyer* or *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 to aid interpretation of coefficients as percentage points. Standard errors clustered by user are reported in brackets.**, and *** indicate statistical significance at the 5% and 1% levels.

	$\mathbb{1}$ x100 if net new **Bull** follows$_{i,s,t+1\rightarrow t+5}$ >0				$\mathbb{1}$ x100 if net new **Bear** follows$_{i,s,t+1\rightarrow t+5}$ >0			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline	+ x$\mathbb{1}$ if Buyer	Baseline & UserSymFE	+ x$\mathbb{1}$ if Buyer & UserSymFE	Baseline	+ x$\mathbb{1}$ if Seller	Baseline & UserSymFE	+ x$\mathbb{1}$ if Seller & UserSymFE
Declared **Bull**$_{i,s,t}$	1.72***	1.64***	1.02***	0.95***				
	[0.03]	[0.03]	[0.04]	[0.04]				
Declared Bull$_{i,s,t}$ x $\mathbb{1}$ Buyer$_{i,s,t}$		1.19***		0.93***				
		[0.04]		[0.04]				
Declared **Bear**$_{i,s,t}$					0.50***	0.47***	0.38***	0.36***
					[0.02]	[0.02]	[0.03]	[0.03]
Declared Bear$_{i,s,t}$ x $\mathbb{1}$ Seller$_{i,s,t}$						1.08***		0.78***
						[0.09]		[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
Unconditional mean (%)	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. Observations are at the user-symbol-day 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 followed by the focal user that post with sentiment. The specification follows Equation (4). Standard errors clustered by user are reported in brackets. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

	N impressions$_{\mathbf{i,s,t+1}}$		N impressions$_{\mathbf{i,s,t+2\text{-}t+10}}$		N impressions$_{\mathbf{i,s,t+11\text{-}t+30}}$		N impressions$_{\mathbf{i,s,t+31\text{-}t+50}}$	
	(1) Bullish	(2) Bearish	(3) Bullish	(4) Bearish	(5) Bullish	(6) Bearish	(7) Bullish	(8) Bearish
Declared Bull$_{\mathbf{i,s,t}}$	0.36***		2.02***		3.85***		3.67***	
	[0.02]		[0.09]		[0.17]		[0.18]	
Declared Bear$_{\mathbf{i,s,t}}$		0.12***		0.38***		0.57***		0.46***
		[0.01]		[0.04]		[0.07]		[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^2	0.52	0.49	0.58	0.52	0.60	0.53	0.61	0.52
Unconditional mean (%)	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. Observations are at the user-symbol-day 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 sentiment messages for a symbol-day combination. The specification follows Equation (5). Standard errors clustered by user are reported in brackets. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

	N of Liked $\text{Msgs}_{\mathbf{i,s,t+1}}$		N of Liked $\text{Msgs}_{\mathbf{i,s,t2\text{-}t10}}$		N of Liked $\text{Msgs}_{\mathbf{i,s,t11\text{-}t30}}$		N of Liked $\text{Msgs}_{\mathbf{i,s,t31\text{-}t50}}$	
	(1) Bullish	(2) Bearish	(3) Bullish	(4) Bearish	(5) Bullish	(6) Bearish	(7) Bullish	(8) Bearish
Declared $\text{Bull}_{\mathbf{i,s,t}}$	3.86***		11.10***		15.91***		12.63***	
	[0.10]		[0.39]		[0.77]		[0.81]	
Declared $\text{Bear}_{\mathbf{i,s,t}}$		2.96***		6.12***		6.41***		4.57***
		[0.08]		[0.24]		[0.40]		[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^2	0.30	0.28	0.38	0.30	0.41	0.33	0.45	0.38
Unconditional mean (%)	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. Observations are at the user-symbol-day level. Columns (1) and (2) have as the dependent variable an indicator equal to one if net new follows of bulls strictly 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), (4) and (7) add an interacted indicator for an EA day occurring on day t+1. Columns (3) to (5) (respectively, (6) to (8)) 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 (5) and (8) restrict the sample to EA days and add an interaction with $EA\ day^{+}$. We multiply the dependent variable by 100 to aid interpretation of coefficients as percentage points. Standard errors clustered by user are reported in brackets.**, and *** indicate statistical significance at the 5% and 1% levels.

	𝟙 x100 if follows are more Bull than Bear		𝟙 x100 if net new **Bull** follows >0			𝟙 x100 if net new **Bear** follows >0		
	(1) Baseline	(2) Add x𝟙 if EA day	(3) Baseline	(4) Add x𝟙 if EA day	(5) Only EA days	(6) Baseline	(7) Add x𝟙 if EA day	(8) Only EA days
Declared **Bull**$_{i,s,t}$	1.04***	1.03***	1.03***	1.02***	1.24***			
	[0.02]	[0.02]	[0.02]	[0.02]	[0.21]			
Declared Bull$_{i,s,t}$ x 𝟙 if EA day		0.77***		0.69***				
		[0.11]		[0.11]				
Declared Bull$_{i,s,t}$ x 𝟙 if EA day^{+}					0.57**			
					[0.29]			
Declared **Bear**$_{i,s,t}$						0.21***	0.20***	0.56***
						[0.01]	[0.01]	[0.16]
Declared Bear$_{i,s,t}$ x 𝟙 if EA day							0.36***	
							[0.09]	
Declared Bear$_{i,s,t}$ x 𝟙 if EA day^{+}								-0.01
								[0.21]
# observations	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	29,613	277,040	277,040	29,613
R^2	0.15	0.15	0.21	0.21	0.39	0.45	0.45	0.41
Unconditional mean (%)	2.57	2.57	2.79	2.79	3.90	0.54	0.54	0.79
Main effect size (% of mean)	41	40	37	37	32	38	37	71
User FE	Y	Y	Y	Y	Y	Y	Y	Y
Day x Symbol FE	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 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. The next two columns add an interaction with an indicator for a user with an above median # of people they are following (column (5)), or being followed by (column (6)), as of the latest date each user appears in the sample. Column (7) 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 to aid interpretation of coefficients as percentage points. Standard errors clustered by user are reported in brackets. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

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

Table 9: 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 the following specification:

$$Abnormal\ return_{i,s,t} = \beta_0 Bull_{i,s,t} + X'_{s,t}\delta + \gamma_t + \phi_{i,s} + \eta_{s,month} + \varepsilon_{i,s,t}$$

The dataset and the signals received are at the user-symbol-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 signals received. We keep only StockTwits symbols with CRSP sharecodes of 10, 11 or 12. The dependent variable is cumulative abnormal returns (returns minus the CRSP value-weighted index return) over days $t+1$ to $t+5$ (and $t+10$), where day t is the day of the user's sentiment declaration. In considering future returns we skip day $t+1$ (and relabel $t+2$ as $t+1$, etc.) 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 to aid interpretation of coefficients as percentage points. We cluster standard errors by user and by permno-day; ** and *** indicate statistical significance at 5% and 1%.

	Dependent variable: abnormal return$_{t+1 \to t+5}$				Dependent variable: abnormal return$_{t+1 \to t+10}$			
	(1) Baseline	(2) Add sd(rec. signals)	(3) Baseline for experience	(4) sd(rec. signals) & experience	(5) Baseline	(6) Add sd(rec. signals)	(7) Baseline for experience	(8) sd(rec. signals) & experience
Declared Bull$_{i,s,t}$	-1.41***	-1.67***	-1.75***	-1.97***	-1.85***	-2.31***	-2.36***	-2.75***
	[0.12]	[0.15]	[0.15]	[0.17]	[0.13]	[0.17]	[0.17]	[0.19]
Bull$_{i,s,t}$ × sd rec. signals$_{i,s,t}$		0.58***		0.52***		1.02***		0.94***
		[0.16]		[0.16]		[0.21]		[0.21]
sd received signals(30days)$_{i,s,t}$		-0.30*		-0.25		-0.32		-0.26
		[0.16]		[0.16]		[0.21]		[0.21]
Bull$_{i,s,t}$ × novice investor$_i$			0.01	-0.01			0.32	0.29
			[0.30]	[0.30]			[0.37]	[0.37]
Bull$_{i,s,t}$ × intermediate investor$_i$			0.77***	0.74***			1.08***	1.04***
			[0.17]	[0.17]			[0.23]	[0.23]
Bull$_{i,s,t}$ × professional investor$_i$			1.06***	1.00***			1.46***	1.36***
			[0.19]	[0.18]			[0.23]	[0.22]
Cum. ab. returns$_{s,(t-5 \text{ to } t-1)}$	-5.66***	-5.66***	-5.66***	-5.66***	-8.18***	-8.17***	-8.18***	-8.17***
	[0.62]	[0.62]	[0.62]	[0.62]	[0.86]	[0.86]	[0.86]	[0.86]
Cum. ab. returns$_{s,(t-30 \text{ to } t-6)}$	-2.67***	-2.67***	-2.67***	-2.67***	-3.87***	-3.87***	-3.87***	-3.87***
	[0.39]	[0.39]	[0.39]	[0.39]	[0.54]	[0.54]	[0.54]	[0.54]
# obs.	3,002,998	3,002,998	3,002,998	3,002,998	2,959,065	2,959,065	2,959,065	2,959,065
# clusters (permno × day)	323,271	323,271	323,271	323,271	320,796	320,796	320,796	320,796
# clusters (users)	92,097	92,097	92,097	92,097	90,676	90,676	90,676	90,676
R^2	0.37	0.37	0.37	0.37	0.50	0.50	0.50	0.50
Day FE	Y	Y	Y	Y	Y	Y	Y	Y
User × symbol FE	Y	Y	Y	Y	Y	Y	Y	Y
Symbol × month FE	Y	Y	Y	Y	Y	Y	Y	Y

Table 10: 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 level. We multiply the dependent variable by 100 to aid interpretation of coefficients as percentage points. Standard errors clustered by user are reported in brackets. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

	$\mathbb{1}$ x 100 if all messages received on day t have sentiment that is					
	(1) Bullish	(2) Bullish	(3) Bullish	(4) Bearish	(5) Bearish	(6) Bearish
Declared $\text{Bull}_{\mathbf{i,s,t\text{-}7\ to\ t\text{-}1}}$		8.02^{***}	6.72^{***}			
		[0.20]	[0.18]			
Declared $\text{Bear}_{\mathbf{i,s,t\text{-}7\ to\ t\text{-}1}}$					2.85^{***}	3.04^{***}
					[0.12]	[0.12]
Expected Pr(all Bull) if $\text{random}_{\mathbf{i,s,t}}$	93.01^{***}	78.53^{***}	74.48^{***}			
	[0.35]	[0.32]	[0.64]			
Expected Pr(all Bear) if $\text{random}_{\mathbf{i,s,t}}$				86.58^{***}	79.31^{***}	75.61^{***}
				[0.61]	[0.58]	[0.75]
# 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
Unconditional mean (%)	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 11: Does selective exposure behavior affect trading volume?

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 level; we estimate the following:

$$AbLogTurnover_{s,t} = \beta_1 SenderDisagree_{s,t} + \beta_2 ReceivedDisagree_{s,t} + \beta_3 ReceivedUncertainty_{s,t} + \delta Controls_{s,t} + \varepsilon_{s,t}$$

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, over the entire 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 include abnormal log turnover on day t-1; $MediaAttention_{s,t}$, which is 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) on day t; $LogGoogleASVI_{s,t}$, a measure of abnormal google search volume for the symbol of stock s; Volatility (t-5 to t-1), 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 messages, 1 message, 2, 3, 4, 5-10, 11-30, and over 30 messages. Standard errors separately clustered by stock and day are reported in brackets. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

	Abnormal Log Turnover$_{\mathbf{s,t}}$				
	(1)	(2)	(3)	(4)	(5)
Sender Disagreement$_{\mathbf{s,t}}$	0.017***	0.015***	0.027***	0.024***	0.025***
	[0.001]	[0.001]	[0.002]	[0.002]	[0.002]
Received Disagreement$_{\mathbf{s,t}}$		0.004**		0.008***	0.011***
		[0.002]		[0.002]	[0.002]
Received Uncertainty$_{\mathbf{s,t}}$			-0.014***	-0.016***	-0.016***
			[0.002]	[0.001]	[0.001]
Abnormal Log Turnover$_{\mathbf{s,t-1}}$	0.188***	0.188***	0.189***	0.188***	0.205***
	[0.004]	[0.004]	[0.004]	[0.004]	[0.005]
Media Article$_{\mathbf{s,t}}$	0.128***	0.128***	0.128***	0.128***	
	[0.006]	[0.006]	[0.006]	[0.006]	
Log GoogleASVI$_{\mathbf{s,t}}$	0.347***	0.346***	0.347***	0.346***	
	[0.015]	[0.015]	[0.015]	[0.015]	
Volatility$_{\mathbf{s,(t\text{-}5\ to\ t\text{-}1)}}$	0.117***	0.116***	0.116***	0.115***	
	[0.035]	[0.035]	[0.035]	[0.035]	
Cum. Abnormal Returns$_{\mathbf{s,(t\text{-}5\ to\ t\text{-}1)}}$	0.010	0.010	0.010	0.010	
	[0.012]	[0.012]	[0.012]	[0.012]	
Cum. Abnormal Returns$_{\mathbf{s,(t\text{-}30\ to\ t\text{-}6)}}$	-0.051***	-0.051***	-0.051***	-0.051***	
	[0.009]	[0.009]	[0.009]	[0.009]	
# obs.	421,915	421,915	421,915	421,915	421,928
# 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
Uncond. mean of Abnormal Log Turnover	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

Internet Appendix to:

Echo Chambers

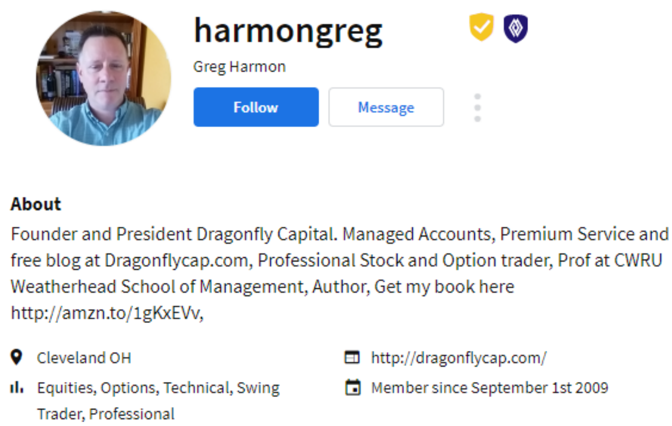
J. Anthony Cookson, Joseph E. Engelberg and William Mullins

[1]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.

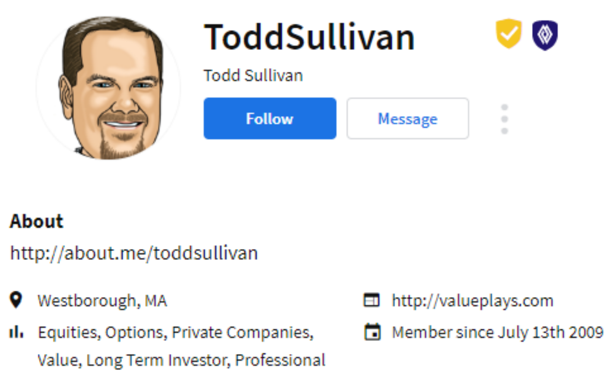
Figure A.1: Examples of StockTwits users

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



(b) Todd Sullivan – Professional Value Investor



(c) Aron Pinson – Professional Fundamental Investor

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: Do Bulls' newsfeeds reflect their bullish sentiment? Messages instead of impressions

This table examines 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 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 clustered by user are reported in brackets. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

	N messages$_{i,s,t+1}$		N messages$_{i,s,t+2\text{-}t+10}$		N messages $_{i,s,t+11\text{-}t+30}$		N messages$_{i,s,t+31\text{-}t+50}$	
	(1) Bullish	(2) Bearish	(3) Bullish	(4) Bearish	(5) Bullish	(6) Bearish	(7) Bullish	(8) Bearish
Declared Bull$_{i,s,t}$	2.57***		12.69***		22.16***		20.60***	
	[0.11]		[0.57]		[1.09]		[1.11]	
Declared Bear$_{i,s,t}$		0.74***		2.71***		3.71***		2.81***
		[0.07]		[0.31]		[0.57]		[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^2	0.43	0.37	0.50	0.43	0.53	0.46	0.54	0.45
Unconditional mean (%)	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.3: Main follows table (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 above median market capitalization as of December 2019 (this was \$3.24bn. Source: Ken French data). Observations are at the user-symbol-day 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 $t+1$ (col 1), and days $t+1$ to $t+5$ (inclusive, cols 2 and 3). 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 4 through 7 are run on a subsample of users that chose to make at least one new net follow in the relevant time window, which extends through $t+50$. We multiply the dependent variable by 100 to aid interpretation of coefficients as percentage points. Standard errors clustered by user are reported in brackets. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

	$\mathbb{1}$ x100 if new follows$_{\mathbf{i,s,t+x}}$ are more Bull than Bear						
	(1)	(2)	(3) Adding User-Symbol FE	(4) Conditional on new follows	(5) Conditional on new follows	(6) Conditional on new follows	(7) Conditional on new follows
	t+1	t+1 $\rightarrow$ t+5	t+1 $\rightarrow$ t+5	t+1 $\rightarrow$ t+5	t+2 $\rightarrow$ t+10	t+11 $\rightarrow$ t+30	t+31 $\rightarrow$ t+50
Declared Bull$_{\mathbf{i,s,t}}$	0.41***	1.04***	0.70***	12.62***	8.64***	4.74***	0.43
	[0.02]	[0.04]	[0.05]	[1.18]	[0.96]	[0.71]	[0.55]
# observations	3,773,560	3,773,560	3,323,876	89,302	145,955	227,802	240,985
# clusters (users)	130,302	130,302	105,810	13,030	17,107	20,441	22,151
R^2	0.10	0.17	0.27	0.73	0.76	0.79	0.84
Unconditional mean (%)	0.84	2.42	2.66	68.44	63.79	56.91	39.77
Effect size (% of mean)	49	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