Journal of Accounting and Economics 75 (2023) 101517

Contents lists available at ScienceDirect

Journal of Accounting and Economics

journal homepage: www.journals.elsevier.com/journal-of-accounting-and-economics

Earnings Virality☆

Brett Campbell [a], Michael Drake [b], Jacob Thornock [b, *], Brady Twedt [c]

[a] Yale School of Management, Yale University, USA
[b] BYU Marriott School of Business, Brigham Young University, USA
[c] Lundquist College of Business, University of Oregon, USA

ARTICLE INFO

Article history:
Received 22 July 2020
Received in revised form 4 May 2022
Accepted 20 May 2022
Available online 2 June 2022

JEL classification:
G12
G14
M41

Keywords:
Earnings virality
Viral
Social media
Earnings announcements

ABSTRACT

We examine the determinants and market consequences associated with earnings announcements going viral on social media, a phenomenon we label "earnings virality." Using a comprehensive panel of historical Twitter data, we find that the typical earnings announcement receives relatively little social media coverage, but others go viral on social media, quickly reaching the feeds of millions of people. We find that viral earnings announcements generally have Twitter content that is more extreme in tone and contains less unique content. Further, earnings virality is positively associated with revenue surprises, investor recognition, retail investor ownership, and retail investor trading around the announcement. Earnings virality appears to be detrimental to markets, as it coincides with lower market liquidity and slower price formation. Overall, our evidence suggests that user-driven dissemination through social media platforms, when amplified and taken to extreme levels, may be harmful to markets.

1. Introduction

A piece of news is said to *go viral* when it spreads through the social media landscape rapidly from one individual to another until it is ultimately viewed by a large group of people. The concept of virality is relatively new, originating in the field of marketing in the late 1990s. However, the concept of virality has become commonplace with the proliferation of social media. While virality has been studied across a variety of fields, including marketing, network science, communications, and information science, the concept has received relatively little attention in the context of corporate financial news. Our broad objective is to investigate the determinants and capital market effects associated with social media virality for what is arguably the firm's most important recurring information event: the earnings announcement.

☆ We thank Brigham Young University, the University of Oregon, and Yale University for financial support. We are grateful to Joanna Wu (editor) and an anonymous reviewer for outstanding guidance throughout the review process. We are also grateful for helpful comments from Ashleigh Bakke, Ed deHaan, Robbie Moon, Christina Zhu, as well as workshop participants at Boston University, Brigham Young University, Georgia State University, University of Buffalo, University of Kansas, University of New South Wales, University of Texas Arlington, and the Virtual Disclosure Research Working Group. We also thank Christina Zhu for generously sharing code, Daphne Armstrong and Ben Yates for research assistance, and Adam Durfee and Devin Knighton of Y-Digital for guidance and advice.

* Corresponding author.
E-mail addresses: brett.campbell@yale.edu (B. Campbell), mikedrake@byu.edu (M. Drake), thornock@byu.edu (J. Thornock), btwedt@uoregon.edu (B. Twedt).

https://doi.org/10.1016/j.jacceco.2022.101517
0165-4101/

We argue that virality is a construct distinct from others in the literature, such as general investor attention or traditional media coverage, for several reasons. First, virality is a social phenomenon facilitated by social media platforms that are designed to rapidly spread information. Within social networks, users push information to other users' feeds, who can then reshare it with others almost instantaneously and at virtually no cost. Second, virality is fueled by the masses—everyday people who suddenly take an interest in an event or topic and re-share it within their networks. Third, virality does not arise gradually over time or from general attention to a topic, but rather it is typically triggered by a specific event. These points of distinction are important because they suggest that the determinants and capital market effects associated with social media virality may differ from those associated with these related constructs. Our objective is to examine virality in the context of earnings announcements, a phenomenon we call "earnings virality."

Our empirical analyses are based on social media activity for a broad sample of approximately 82,000 quarterly earnings announcements from 2010 to 2019. We employ an expansive panel of historical Twitter activity from Crimson Hexagon (now Brandwatch), a large digital consumer research and marketing analytics company. We augment these data with tweet-level data from the Twitter Application Programming Interface (API), which allows us to compute content-based measures. Overall, the data are well suited to identify viral earnings announcements because they allow us to observe the timing, content, and extent of the social media activity.

We first provide a descriptive portrait of earnings virality by comparing tweet characteristics associated with earnings announcements that go viral to those that do not. On a relative basis, the tweets of viral earnings announcements generally use more extreme tone—that is, they contain words that are more negative or more positive than neutral. Tweets of viral earnings announcements also use fewer financial words and are less likely to contain links to other content than are tweets of non-viral announcements. These findings indicate that, on average, the nature of tweets associated with viral earnings announcements is more extreme and less substantive. Furthermore, social media users that participate in viral earnings announcements are more likely to both retweet and to issue more tweets per user.

Our first set of formal hypotheses relate to three event- and firm-level factors that we predict will be associated with earnings virality. The first factor relates to the content of the earnings announcement, namely the magnitude and sign of the earnings and revenue surprises. The emerging literature on social sharing suggests that content that is more surprising, interesting, controversial, or practically useful is more likely to go viral (see, e.g., Berger and Milkman 2010, 2012; Rudat et al., 2014; Kümpel et al., 2015). It is unclear, however, whether good or bad news will be more likely to trigger virality. While some research suggests that content with a positive valence or "goodness" is more likely to go viral (Bakshy et al., 2011), other research suggests that bad news travels faster than good news within social networks (Naveed et al., 2011; Hornik et al., 2015). We find that the magnitude of the revenue surprise, but not the earnings surprise, is positively associated with virality. This result suggests that large revenue surprises tend to draw the kind of social media attention that leads to viral diffusion, while earnings appear to be less relevant in this context. We also find modest evidence that virality is more strongly associated with positive revenue surprises than negative revenue surprises.

We next predict that earnings virality will be positively associated with investor recognition. This prediction is motivated by the idea that investors are more likely to pay attention to companies on social media that are recognizable either because they are familiar (Merton 1987; Huberman 2001; Lehavy and Sloan 2008) or because they are flashy and highly visible (Odean 1999; Barber and Odean 2008). Consistent with this prediction, we find that the incidence of earnings virality is higher for business-to-consumer firms and those with high revenue growth. These findings – together with our findings that earnings virality is associated with revenue surprises, but not earnings surprises – suggest that virality is associated with recognizable firms for whom revenue growth is a particularly important performance indicator.

The third factor we examine relates to the audience of the earnings announcement. Virality is a construct that is inextricably linked to the people who rapidly share information across their social networks. Accordingly, we predict that earnings announcements of firms with relatively higher retail investor ownership are relatively more likely to go viral than those with lower retail investor ownership. This prediction stems from the idea that in order for something to go viral on social media, it must draw the attention of the masses. Further, prior research argues that social media is a channel through which retail investors become informed and that the link between social media and market activity is strongest for retail investors (Chen et al., 2014; Farrell et al., 2020; Gomez et al., 2020). Consistent with this prediction, we find that earnings virality is positively associated with retail investor stock ownership.

Having examined the determinants of earnings virality, we explore its relation to general investor activity and then test formal predictions about its relation with market outcomes. We begin by investigating the association between earnings virality and general measures of investor activity around the earnings announcement. We find that earnings virality is positively associated with abnormal trading volume and price volatility. We further find that earnings virality is positively associated with abnormal retail trading and Robinhood holdings, suggesting that earnings virality not only draws the attention of retail investors, but also influences their trading activity. Finally, we find that earnings virality is negatively associated with professional investor information search activity on Bloomberg terminals. In sum, the evidence suggests that earnings virality coincides with increased trading activity, price volatility, and an influx of retail investor trading.

Our second formal hypothesis relates to the association between earnings virality and market outcomes, with a focus on liquidity and speed of price formation. One advantage of these particular market outcomes is that they are less susceptible to reverse-causality concerns than are general measures of investor activity (e.g., volume, volatility) because they are either difficult for investors to obtain in real-time or their measurement extends to the period after earnings go viral. Ex ante, it is unclear whether earnings virality will be positively or negatively associated with these market outcomes. On the one hand, prior theoretical research argues that increased information dissemination will reduce the information advantage of informed investors, which leads to positive market benefits such as increased liquidity and faster price formation (Grossman and Stiglitz 1980; Diamond and Verrecchia 1991; Bloomfield 2002). Empirical research supports this theory with evidence that greater dissemination of news is associated with capital market benefits across a variety of market outcomes (e.g., Bushee et al., 2010; Blankespoor et al., 2014; Twedt 2016; Drake et al., 2017). On the other hand, other theoretical models demonstrate that, under some conditions around public news events, increased dissemination can increase information asymmetry, thereby decreasing liquidity (Copeland and Galai 1983; Glosten and Milgrom 1985; Kim and Verrecchia 1994). Further, earnings virality may slow down the price formation process if it triggers adverse social interactions, such as correlated noise trading (e.g., Shiller 1984; Baker and Wurgler 2006), failure to account for the repetition of information (DeMarzo et al., 2003), trading on signals that have already been priced (Barber and Odean 2013), or free riding on friends' information (Han and Yang 2013). These competing arguments suggest that the relation between earnings virality and certain market outcomes is an open empirical question.

With respect to liquidity, we find that earnings virality is positively associated with abnormal spreads and negatively associated with abnormal depths, suggesting that earnings virality coincides with reduced liquidity around earnings announcements. Regarding the speed of price formation, we find that earnings virality is negatively associated with a measure of intra-period efficiency (IPE) and the ratio of short-to long-window returns, suggesting the earnings virality also coincides with slower price formation. Thus, our evidence indicates that earnings virality is associated with negative market outcomes. This result is novel because it reveals a setting when greater dissemination of public information potentially harms investors by reducing liquidity and impeding the market from efficiently processing corporate news.

One challenge with our market-based tests is the potential for endogeneity. In addition to the inclusion of a broad set of control variables in our models, including controls for current investor activity, traditional press coverage, and prior social media coverage, we take several steps to address this issue. First, we employ a quasi-exogeneous setting that leverages two important changes Twitter made to its newsfeed algorithm in 2016. These changes increased the likelihood that earnings news goes viral on Twitter and that some users could consume the viral earnings news on a delay if they missed it initially. Importantly, these changes effect social media *virality*, but are unlikely to affect other related constructs such as general investor attention or traditional press coverage. The results indicate that the negative market consequences that coincide with earnings virality are stronger (more negative) after the changes implemented by Twitter in 2016. Second, we use the impact threshold approach suggested by Larcker and Rusticus (2010) to examine the potential impact of unobserved confounding variables. The results suggest that, in order to render our results insignificant, an omitted variable would need to have an impact greater than that of any of our control variables, which includes variables such as firm size and traditional press coverage. Finally, we employ a variety of research designs, including cross-sectional tests that isolate variation within social media content, tests that account for firms that frequently go viral, and tests that isolate variation within firms. Although this triangulated approach using various tests and methodologies lessen concerns about endogeneity (e.g., Glaeser and Guay 2017; Heckman and Singer 2017; Armstrong et al., 2021), we acknowledge that our results still primarily rely on association-based tests.

Another important challenge with our setting relates to ensuring that our models can empirically separate social media virality from related constructs, such as general investor attention and traditional media coverage. With respect to this issue, our tests reveal a clear contrast between the market effects of social media virality and those of these other constructs. Specifically, we find a negative association between earnings virality and market outcomes, but generally find a positive association between professional press coverage and these same outcomes. Thus, while social media virality appears to be detrimental to markets, heightened traditional press coverage is not. We further address this concern by conducting cross-sectional analyses that isolate variation that exists *within tweet content*, but not within these other constructs. We find that the detrimental pricing effects of social media virality are generally stronger when the tweets contain more extreme tone, and negative tone in particular. We also find that these detrimental effects are stronger when the tweets contain fewer financial words, when they do not include a hyperlink to other online content, or when each user issues a greater number of tweets. Together, these tests suggest that the documented market effects are more likely linked to earnings virality than to other constructs.

We contribute to the emerging literature on the role of social media in markets by examining the concept of social media virality. While there is a large and growing literature on the market effects of *old media*, Miller and Skinner (2015), Blankespoor et al. (2020) and Hirshleifer (2020) call for additional research on the effects of *new media*, with its potential for a large set of amateur users to create or share content about firms. Recent research examines company usage of social media (e.g., Blankespoor et al., 2014; Lee et al., 2015; Blankespoor 2018; Nekrasov et al., 2020), as well as the provision of analysis and research on social media platforms, such as Seeking Alpha (e.g., Chen et al., 2014; Drake et al., 2020; Gomez et al., 2020). Much of the prior work examining social media demonstrates its positive market influence, primarily due to increasing visibility and broader dissemination of important financial information. However, recent empirical evidence is now starting to examine whether social media potentially degrades market quality. For example, a recent study by Jia et al. (2020) documents that social media can act as an informational rumor mill, which impedes price discovery. Our results complement those of Jia et al. (2020) by providing evidence on a separate but important mechanism—namely, viral information dissemination, through which social media can be potentially counter-productive to market liquidity and price efficiency.

The evidence in this study is timely given the recent academic attention on the influence of social interactions on the pricing of stocks. Indeed, Hirshleifer (2020) calls "social finance" a new economic paradigm, which he identifies as the missing chapter in better understanding how social forces influence market behaviors. Our results are consistent with social compounding, in which the views of a single investor can be compounded across social networks and have large market effects. For instance, we find that the detrimental price effects of earnings virality are heightened when average tweet attributes are more extreme and less substantive, consistent with social compounding of these attributes. Further, the dramatic price swings for meme stocks such as GameStop and AMC in early 2021 demonstrate how social media virality can coordinate the trading activities of retail investors and perhaps contribute to a degradation of market quality. These extreme events are a novel market phenomenon and highlight the uniqueness of earnings virality as an economic construct with potentially unique market consequences.

2. Earnings virality: conceptualization and hypotheses development

2.1. Earnings virality

We define virality as the extreme person-to-person spread of a particular piece of information via social media.[1] We argue that virality, as a concept, is a unique and novel addition to the accounting literature in three ways. First, virality is characterized by user-driven information diffusion. Social media platforms are particularly well suited to facilitate rapid diffusion because they tether users together into networks, which enables information to spread almost instantaneously from user to user, both within and across user networks. This spread of information is characterized by an element of speed, which social media facilitates by pushing information to users' feeds. These features are, to a degree, what sets social media apart from traditional media that generally must be sought out (or pulled) proactively by users and is typically not supported by such social networks that facilitate rapid information sharing.

Second, virality results in content being consumed by a large and broad audience. Because of its wide availability and low cost, social media has become the media of the masses—the most prominent social media platforms report participation by billions of people. Social media is costless for users to acquire and use, and social media apps currently make up five of the ten most frequently downloaded apps.[2] As a result, social media garners participation from a much broader set of users than other forms of media.

Third, virality is typically triggered by an event (e.g., statement is made, opinion is expressed, news is announced) that sets off the rapid and extreme sharing of the content. This particular feature of virality is what sets it apart from general attention. For example, certain types of individuals or companies (e.g., Elon Musk, Tesla) garner lots of general attention because of who they are or what they do. However, it generally takes an event (e.g., Elon Musk breaking the Cybertruck's window) to trigger social media virality. To be sure, there are certain entities whose events are more likely to go viral, simply because of their size, visibility, or prominence in social media. In these instances, virality is likely influenced by factors related to the triggering event, as well as the entity experiencing it.

Having developed the general notion of social media virality, we now apply it to our specific setting − the release of corporate earnings news. We define earnings virality as the extreme spread of earnings announcement information across social media networks that is rapidly viewed by a large number of people. While prior research examines determinants and market consequences associated with traditional media coverage, the unique features of social media virality in this setting suggest that earnings virality may be associated with different determinants and market consequences. In the next section, we discuss related literature before formally developing our hypotheses.

2.2. Prior research

While the concept of earnings virality is new to the literature, a number of recent studies have begun to examine the associations among social media activity, financial information, and market activity. One set of papers examines social media activity *by firm managers* (e.g., via Netflix's corporate Twitter account, @netflix) and how it is used to disseminate information to stakeholders. For example, Blankespoor et al. (2014) examine Twitter usage by technology companies and find that additional Twitter dissemination is associated with reductions in information asymmetry. Jung et al. (2018) examine the use of Twitter by S&P 1500 firms and find that social media use is lower when companies have bad news to report. Lee et al. (2015) examine social media use for product recalls and find that this communication attenuates negative price reactions. Finally, Nekrasov et al. (2020) examine the company's inclusion of earnings-related visuals (e.g., graphics, illustrations) in tweets around the earnings announcement and find that the increased use of visuals is related to higher earnings response coefficients and reduced price drift.

Another set of papers examines whether social media activity *by firm outsiders* impacts the firm's information environment and market pricing. For example, Bollen et al. (2011) link the aggregate mood of the public, as reflected in tweets, with returns to the Dow Jones Industrial Average index. They find that this aggregate mood is predictive of index returns. Curtis et al. (2014) show that the level of social media activity on Twitter and StockTwits around earnings announcements is

[1] Our definition of virality is consistent with that found in general dictionaries and how the phrase is used in common parlance. Prior academic literature does not provide a formal definition of virality of which we are aware.

[2] https://www.cyberclick.net/numericalblogen/top-10-most-downloaded-apps-of-2020-so-far.

positively associated with market reactions to the news. Bartov et al. (2018) examine the content of individual tweets just prior to earnings announcements and find that it helps predict the upcoming earnings news and the market reaction to the news. Tang (2018) demonstrates the predictive ability of tweeted product opinions for sales growth. Jiang and Shen (2017) examine social media coverage of several prominent business scandals. Ang et al. (2020) show that crowd wisdom in social media is predictive of merger outcomes, while Jia et al. (2020) provide evidence of a distortive rumor mill caused by social media during the M&A process. While much of prior research documents that social media can improve information dissemination and in turn market pricing, Jia et al. (2020) is among the first to document detrimental social media pricing effects.

Our study contributes to these streams of literature by introducing the concept of social media virality in the context of earnings announcements. While prior research in accounting examines traditional media dissemination and, more recently, social media activity, our focus is on the construct of *extreme* social media dissemination of earnings information. That is, we provide evidence on the concept of virality that is familiar to us all, but has not yet received consideration with respect to financial information events.

2.3. Hypothesis 1 — the determinants of earnings virality

Our first objective is to provide insights into why some earnings announcements go viral and others do not. Broadly speaking, we consider three main areas of determinants: the content of the earnings announcement, firm characteristics, and audience characteristics.

A growing body of social science research examines factors associated with the sharing of information via social networks (see Kümpel et al., 2015 for a review of this literature). This literature provides evidence that virality is positively associated with content that is surprising, interesting, controversial, or practically useful (see, e.g., Berger and Milkman 2010, 2012; Rudat et al., 2014). The evidence also suggests that articles that evoke high-arousal emotions such as awe, anger, or anxiety are also more likely to go viral (Berger and Milkman 2012), as are articles that contain positive valence or goodness (Bakshy et al., 2011). The common theme underlying these findings is the idea that individuals share information via social media when they believe the content will be more surprising or valuable to others. This motivates our first determinants hypothesis, which relates to the content or news being disclosed in the earnings announcement, with a specific focus on the recent financial performance of the company. The content is surprising to the extent that it deviates from expectations. Our prediction is that surprising financial information, principally larger earnings surprises and revenue surprises, will trigger viral social media activity.

In addition, we predict that earnings virality will relate to the release of value-relevant information. To capture the concept of value-relevance, we rely on research showing that the sign of earnings news is differentially relevant to firm value. On the one hand, positive (negative) earnings news is considered less (more) transitory and thus more (less) useful (e.g., Hayn 1995). The social science literature provides evidence that positive news is more frequently shared on social media (e.g., Bakshy et al., 2011). On the other hand, negative earnings news may be more sensational and/or appeal to investors' loss aversion, increasing the likelihood that the earnings announcement will go viral.

The above literature motivates our prediction that earnings virality will be associated with the unexpected magnitude and direction of the news. We state this hypothesis in the null form as follows:

H1a. Earnings virality is not associated with the magnitude or sign of the earnings or revenue surprise.

Our next hypothesis regarding the determinants of earnings virality is motivated by the literature on investor recognition and familiarity. Merton (1987) develops an asset pricing model which shows that, in the presence of incomplete information, rational investors prefer and hold stocks for which they are better informed, holding fundamentals constant. A key insight from his model is that investors construct portfolios based on securities with which they are familiar. Consistent with this theory, Huberman (2001) posits that "familiarity breeds investment" (p. 678) because investors feel more comfortable investing in companies that are more recognizable to them. We conjecture that this familiarity bias will extend into investors' tendencies to consume and share company information online via social media. In other words, familiarity breeds sharing.

Furthermore, certain transitory firm characteristics, such as high revenue growth, could also attract social media attention because the characteristics make firms more recognizable to investors. Because investors have limited time and resources, they must selectively allocate their attention across investment options (Hirshleifer and Teoh 2003). This creates a "formidable search problem" (Barber and Odean 2008, p. 786) for investors, and Odean (1999) argues that one way investors address this problem is by limiting their attention to stocks that are attention-grabbing. Barber and Odean (2008) provide empirical support for this idea, documenting that investors, and more specifically retail investors, are especially active in stocks that grab their attention. Consistent with this view, Boehmer et al. (2021) find that retail investors aggressively purchase the stocks of flashy, high-growth firms. This discussion suggests that earnings virality will be more likely in firms that are more recognizable to investors, either because they are familiar or because they are attention-grabbing. We state this prediction in the alternative form as follows:

H1b. Earnings virality is positively associated with investor recognition.

Our third hypothesis regarding the determinants of earnings virality examines whether the audience of the earnings announcement is associated with earnings virality, with a specific focus on the ownership base of the firm. For content to go viral via social media, it must appeal to a broad set of investors, but not just any investor—the content must appeal to those investors that are most likely to consume and participate in social media. Prior research suggests that retail investors are most likely to use social media platforms and that the link between social media and market activity is strongest for firms with high retail investor ownership (Chen et al., 2014; Farrell et al., 2020; Drake et al., 2020; Gomez et al., 2020). This leads to our prediction that earnings virality will be positively associated with retail investor ownership. We state this prediction formally in the alternative form as follows:

H1c. Earnings virality is positively associated with retail investor ownership.

2.4. Hypothesis 2 − earnings virality, liquidity, and price formation

Our second objective is to understand the relation between earnings virality and market outcomes, as measured by liquidity and speed of price formation.[3] Ex ante, it is unclear whether earnings virality will be positively or negatively associated with these market outcomes. Theoretical research argues that broad dissemination of public information reduces the information advantage of informed traders, which in turn lowers information asymmetry and increases liquidity (Diamond and Verrecchia 1991). Consistent with this argument, Blankespoor et al. (2014) provide empirical evidence that increased dissemination by the company's Twitter account is associated with increased liquidity (i.e., lower spreads and greater depths) in technology firms. Other models suggest that news broadcasted to a wider audience gets the information to a more diverse set of investors, which can improve the efficiency of the price discovery process (Grossman and Stiglitz 1980; Bloomfield 2002). Numerous papers provide empirical support for these arguments. For example, Bushee et al. (2010) find that increased dissemination of news by business press reporters reduces information asymmetry (as measured by spreads and depths) around earnings announcements. Twedt (2016) and Drake et al. (2017) find that broader dissemination of news by the professional business press is associated with faster price formation after earnings guidance and earnings announcements, respectively. These findings suggest that the extreme dissemination reflected in earnings virality could potentially help markets to be more liquid and to incorporate news more efficiently into prices.

However, other theoretical and empirical work demonstrates that increased dissemination can also have negative market effects when it encourages differences of opinions among investors or exacerbates information asymmetry. For example, the model of Kim and Verrecchia (1994) shows that, under some conditions, public disclosures can simultaneously induce increased trading volume at the announcement and *decreased* market liquidity (as reflected in higher bid-ask spreads). The decrease in liquidity occurs because some traders process the news into private, and potentially diverse information, that can exacerbate information asymmetry. This suggests that earning virality could decrease market liquidity if it induces market makers to suspect that a more diverse set of investors are potentially trading. Verrecchia (2001) discusses the possibility of a liquidity premium occurring when agents demand a premium for trading against unknown counterparties, which could occur if virality leads to a more diverse set of investors. In addition, the models of Copeland and Galai (1983) and Glosten and Milgrom (1985) also predict that liquidity will decrease around earnings announcements when market makers anticipate facing a more diverse set of investors that could potentially be informed. Empirical research provides some support for these theories. Specifically, Lee et al. (1993) find an increase in spreads and decrease in depths around earnings announcements when the announcement is more newsworthy, consistent with suppliers of liquidity responding to information asymmetry risk. Similarly, Skinner (1993) and Patel (1993) provide evidence of increased information asymmetry around earnings announcements when the announcements are more informative. Thus, if earnings virality encourages differences of opinions from a broader set of traders, then it could induce decreases in market liquidity.

It is also possible that earnings virality could impede the speed of price formation. The issue here, in part, relates to *how* information is transmitted to others, a phenomenon often described as social information transmission or social learning (e.g., Hirshleifer 2020). In social learning, people convey information one-to-another through conversations, opinion-sharing, and observations of others' actions (Golub and Sadler 2017). Perhaps nowhere is social learning more salient than in social media, where investors learn from the discussion, opinions, and revealed trades of other investors. As a result, investors often receive low-precision signals earlier than high-precision signals (Dugast and Foucault 2018) or investors invest more in learning about other traders than in learning from fundamentals (Banerjee et al., 2018), either of which can slow price discovery.

Virality also raises the potential for distortive network effects within social networks as information is repeated and reinforced across members in the system at the exclusion of other considerations, including whether or not the information has already been priced. As noted in Hirshleifer (2020), these network effects potentially trigger social compounding, where even small biases have large effects as they are rapidly socially transmitted. Moreover, enclosed networks can also lead to

[3] We focus on liquidity and the speed of price discovery for several reasons. First, the empirical measures used to capture these constructs are less susceptible to concerns of reverse causality because they are not easily observable in real-time and because they are often measured after the event period. Second, these constructs are important facets of price efficiency and market quality in their own right. Third, our use of these measures follows prior research (e.g., Blankespoor et al., 2018), which uses these constructs and measures for similar reasons.

naïve learning, in which agents simply update beliefs based on their neighbors' opinions (Golub and Jackson 2010)—as a result, investors overweight public opinion and underweight fundamental information, which can slow price formation. Recent theory work also suggests that social communication can induce investors to free ride on friends' opinions, which can lead to excess trading and reduced market quality (Golub and Jackson 2010; Han and Yang 2013). Any of these distortive network effects associated with social media sharing could potentially slow down the speed of price formation and degrade market quality.

It is important to note, however, that the arguments for a negative relations between earnings virality and market outcomes, are not necessarily attributable to the actions of less sophisticated retail investors. More sophisticated investors also likely consume social media, either for their own information or to anticipate retail trading demand. Indeed, in 2013, a Twitter feed was incorporated into Bloomberg terminals, and in 2018, Bloomberg expanded its Twitter data offering, stating that " [o] ur customers tell us that Twitter data is a vital part of their information-driven trading strategies, helping them uncover early trends and changes in sentiment" (Bloomberg 2013, 2018). Indeed, the Wall Street Journal recently reported that “85% of hedge funds and 42% of asset managers are now tracking retail-trading message boards …” (McCabe 2022). This suggests that sophisticated professional investors may respond to the retail investor activity they observe on social media with more informed trading, thereby improving market outcomes.

These arguments suggest that earnings virality could be either positively or negatively associated with liquidity and the speed of price formation. This leads to our second set of hypotheses, which we state in the null form as follows:

H2a. Earnings virality is not associated with liquidity.

H2b. Earnings virality is not associated with the speed of price formation.

3. Data, measures, and descriptive statistics

3.1. Data and sample

We use a unique dataset that allows us to identify viral earnings announcements and observe variables that capture several of its important elements. These data come from Crimson Hexagon (now part of Brandwatch), a large digital consumer research and marketing analytics company. For our purposes, the primary advantage of Crimson Hexagon is that it provides firehose access to an extensive panel of historical Twitter data.[4] The data provided by Crimson Hexagon includes tweet identifiers, user identifiers, and sentiment and tone scores using their own proprietary algorithm.

We acquire Twitter data from Crimson Hexagon for all members of the Russell 3000 index during the January 2010 to December 2019 period. We download all tweets that include a particular company's stock ticker using the “cashtag,” which is a common feature in Twitter where users identify stock tickers with a leading $ symbol (e.g., $MSFT is the cashtag that refers to Microsoft's stock). Our focus on tweets with cashtags reduces the likelihood that tickers and companies are misclassified and has been used in extant research (e.g., Jia et al., 2020). We further limit our sample to tweets that mention only one cashtag to make it more likely that the tweet is about a specific company's earnings announcement. We then collect several measures that capture the content of the tweets by querying Twitter's public API using the tweet identifier provided by Crimson Hexagon.[5] This allows us to obtain information about the content of each tweet, including the tweet text, date, and time, and whether the tweet was a retweet. We also obtain user-level data that includes the number of followers, username, and other qualitative attributes.[6]

We also obtain all quarterly earnings announcement dates during our sample period from Compustat and IBES, only retaining those with matching announcement dates across the two databases. We then require assets greater than zero, analyst following of at least one, and non-missing values for the observation to be included in the sample. The intersection of these sample requirements with our Twitter data leads to a final sample of 82,232 quarterly earnings announcement observations from 2909 unique firms.

[4] There are several benefits to focusing on Twitter data. For example, specific features of Twitter (e.g., emphasis on dissemination, pushed feeds, constrained character limits) are highly amenable to virality. Twitter allows us to measure extreme information dissemination with reduced likelihood that the effects of the social media posts are being confounded by lengthy additional commentary. Further, Twitter is frequently either the source of virality or plays a role in a piece of information going viral because Twitter often assists information posted on other social media platforms (e.g., a video posted on YouTube) going viral. Zhou et al. (2015) document a faster response to Twitter than to Facebook posts in corporate disclosure, consistent with Twitter being more likely to generate virality. Finally, we observe literally billions of cashtag tweets during our sample period, suggesting that Twitter is an important venue for social investing. Still, we acknowledge that virality can certainly occur on other social media platforms. Thus, our focus on Twitter does constrain the generalizability of our results.

[5] The Twitter API has several endpoints. The public endpoint for performing keyword searches limits the user to the past seven days. A separate public endpoint allows the user to pull historical data if the user already has tweet identifiers, which we obtain from Crimson Hexagon. Thus, Crimson Hexagon provides a list of tweets, while the Twitter API provides the text and details of those tweets.

[6] Unfortunately, the user data is measured as of the time we query the API, not the time when the tweet occurred. As we have no way to recover historical user data, we must rely on the assumption that the user data on the date we queried the API is similar to that on the date they published the tweet. We control for time trends in all models.

3.2. Empirical measures

3.2.1. Measuring earnings virality

Our primary measure of earnings virality is based on a combination of the volume of tweets, including retweets, and the number of users who view the tweets. We include retweets in our measure because the concept of retweeting, or resharing in general, is part of what fuels the speed and depth of dissemination on social media. For each earnings announcement in our sample, we aggregate the total number of original tweets and retweets that include the firm's cashtag on days 0 and + 1 and label this variable *#tweets*.[7] This measure proxies for proactive discussion about the firm on Twitter at the time of its earnings announcement. Next, we obtain the number of followers for each user that tweeted and aggregate this total over the earnings announcement window (days 0 and + 1) as well. We label this variable *#feeds*. This measure captures the total viewership of the tweets issued at the earnings announcement. We then construct our key variable of interest, *Viral Earnings*, using both *#tweets* and *#feeds*. Specifically, we rank both *#tweets* and *#feeds* into deciles each year. We then set *Viral Earnings* equal to one for earnings announcements that are in the top yearly decile of both variables, and to zero otherwise.[8] We discuss results using alternative measures of earnings virality below.

3.2.2. Measuring determinants of earnings virality

To test H1a, we examine the association between earnings virality and the content of the earnings announcement using several variables that capture financial performance surprises. First, we examine the absolute value of the earnings surprise using the median consensus analyst earnings forecast as the benchmark, scaled by price (*Abs Earn Surprise*). Second, we examine the absolute value of the revenue surprise, using the analyst consensus revenue forecast as the benchmark (*Abs Rev Surprise*).[9] To capture the value relevance of the news, we use a proxy based on the sign of the news for both revenue and earnings surprises. *Neg Earn Surp* and *Neg Rev Surp* are indicators set equal to one if the company's announced earnings surprise or revenue surprise, respectively, are negative, and to zero otherwise. We also include interactions of the magnitude and sign of the earnings and revenue surprises.

To test H1b, we examine the association between earnings virality and four proxies for investor recognition. We use two proxies for firm familiarity, including an indicator variable for consumer-facing firms (*BtoC*) and another for technology firms (*Tech*). We also use two proxies for high-growth firms or growth expectations. *High Rev Growth* is an indicator variable that equals one for firms in the top quintile of seasonal, quarterly revenue growth, and equals zero otherwise. *Low BTM*, is an indicator variable that equals one for firms in the bottom quintile of the ratio of book value of equity to market value of equity, and equals zero otherwise.

To test H1c, we examine the association between earnings virality and retail investor ownership. Our proxy is based on retail investor stock ownership using the level of non-institutional ownership in a given firm. Specifically, *Retail Own* is equal to one minus the percentage of outstanding shares held by institutional investors.

Our determinants model also includes a broad set of control variables that may be associated with earnings virality, which we group into five constructs: traditional media coverage, general social media coverage, general market attention, sophisticated investor attention, and general firm attributes. We discuss the specific control variables related to each construct in turn.

To capture traditional media coverage, we include three variables related to the count of the Dow Jones Newswires and Wall Street Journal articles published about the firm, as in Bonsall et al. (2020). The variables are *Press Cov[0,1]*, *Press Cov[-7,-1]* and *Press Cov[-30,-8]*, which capture the logged number of published articles during the announcement period, the week before the announcement, and the month prior to but ending one week before the announcement, respectively. Next, we include two measures of general social media coverage measured before the earnings announcement, *Tweets[-7,-1]* and *Tweets[-30,-8]*, which are the logged number of tweets for the firm in the week before the announcement and in the month prior to but ending one week before the announcement, respectively. In addition, we include measures of social media viewership before the earnings announcement, *Feeds [−7,-1]* and *Feeds[-30,-8]*, measured as the logged number of followers for users that tweeted about the firm during those same windows.

Next, we include several variables related to general market attention and sophisticated investor attention. We measure general market attention using three variables: abnormal return volatility (*Abn Volatility [0,1]*), abnormal trading volume around the earnings announcement (*Abn Volume[0,1]*), and the cumulative abnormal return for the stock in the month prior to the earnings announcement (*Abn Ret[-30,-1]*). These variables are particularly important controls, as they likely capture the tendency of earnings news to go viral because investors observed large stock price movements, excess trading activity, or both. We next include an indicator for whether earnings are announced after 4:00 p.m. Eastern time (*After Hours*), following the intuition that individuals have more time to engage in social media activities after normal business hours. To capture

[7] We provide formal definitions of all variables in Appendix A.

[8] We use annual decile rankings in measuring earnings virality to account for the general increase in Twitter usage over our sample period; the number of users and tweets are vastly larger in the latter half of the sample relative to the first half. For instance, Twitter recently reported that its worldwide monetizable daily active Twitter users increased 24% year-over-year in 2019 alone. An implication of this trend is that virality potentially means something different in 2019 than it does in 2012. For this reason, we also control for time trends and year fixed effects in all our models.

[9] In an untabulated test, we find that the Pearson correlation between the earnings surprise and revenue surprise to be 0.28, which suggests that they capture different constructs.

institutional attention, we include the number of sell-side analysts following the stock (*#Analysts*) and the total number of unique institutions that own shares in the stock (*Inst Count*). We also include the total number of earnings announcements that occur during the day (*# Firms Announce*), because prior literature provides evidence of differential market activity on busier earnings announcement days compared to less busy earnings announcement days (Hirshleifer et al., 2009; Blankespoor et al., 2020).

We include a number of control variables for different firm attributes that relate to general social media attention. We include the natural log of market capitalization (*Firm Size*) to capture firm prominence. We include a control that captures firm life cycle stage, measured as the number of years the firm appears in Compustat (*Firm Age*). Finally, even though our measure of earnings virality is based on annual-decile ranks, it could still have residual time trend effects if the likelihood of being in *both* the top decile of *#tweets* and *#feeds* changes over time. Accordingly, in addition to including year fixed effects in our models, we include a control for the linear time trend over the sample period (*Time Trend)*.

3.2.3. *Measuring trading activity and market outcomes*

We use a broad set of variables to investigate the relations among earnings virality, trading activity, and market outcomes. Our first set of measures relate to general trading activity. We measure *Abn Volume[0,1]* as the firm's average trading volume during the announcement window (days 0 and + 1), less the firm's average trading volume during the benchmark window (days −30 to −1). We also examine *Abn Volatility[0,1]*, which is the firm's quote-based intraday volatility obtained from TAQ and averaged over the announcement window (days 0 and + 1), less the firm's average quote-based intraday volatility averaged over the benchmark window (days −30 to −1), multiplied by 1 million for interpretability.

We consider two measures of retail investor activity − abnormal retail trading volume and abnormal retail investor holdings. We identify retail trades using the method developed in Boehmer et al. (2021) and used in several recent studies (e.g., Blankespoor et al., 2018; Bushee et al., 2020; Israeli et al., 2020). This method uses trade-level TAQ data to isolate retail trades based on price improvements allotted to retail (but not institutional) traders.[10] *Abn Retail Volume[0,1]* is abnormal retail trading volume, defined as the sum of the firm's daily event-period retail trading volume minus the firm's average retail volume over the prior month. To measure abnormal retail holdings, we employ novel data on retail investor stock holdings from Robinhood Markets Inc. (Robinhood, hereafter), a popular mobile trading platform that offers free and simple trading to millions of retail investors. Following prior research, we obtain daily retail investor holding data from Robintrack, which is an independent website that retrieves data directly from Robinhood. We compute *Abn Robinhood Holdings[0,1]* as the average number of Robinhood users holding a given company's stock over the two-day earnings announcement window scaled by the average number of users holding the stock in the prior month for the same company. These data are only available for the May 2018 to December 2019 period, which covers 15,587 of the earnings announcements in our sample. Finally, we examine one measure of professional investor activity based on attention to firm-specific news on Bloomberg. Following Ben-Rephael et al. (2017), we measure *Abn Bloomberg Activity[0,1]* as the abnormal number of Bloomberg terminal searches for a given firm during the earnings announcement window.[11]

We use the next set of market outcomes to test H2. Following Blankespoor et al. (2018), we examine the relation between earnings virality and two important market outcomes: liquidity and speed of price formation. As recommended in Lee et al. (1993), we measure liquidity by examining both spreads and depths. *Abn Spread[0,1]* is the intraday average percent effective spread obtained from TAQ and averaged across the announcement window (days 0 and + 1), less the average intraday percent effective spread over the benchmark window (days −30 to −1), multiplied by 100.[12] *Abn Depth[0,1]* is the log of the average intraday bid and offer depth obtained from TAQ over the announcement window (days 0 and + 1) divided by the average bid and offer depth over the benchmark window (days −30 to −1).

We measure the speed of price formation using two proxies. First, we use the intra-period efficiency (*IPE*) measure used in prior literature (see e.g., Twedt 2016; Blankespoor et al., 2018; Blankespoor et al., 2020). *IPE* is designed to measure the speed with which information, in this case earnings announcement information, is reflected in price over a specified period of time, [t, T]. In our primary tests, we measure IPE over a one-week period; thus, *IPE* is an area-under-the curve measure that captures the daily proportion of abnormal returns realized up to and including a given trading day, T = 5, starting on the earnings announcement date, t = 0. The faster the return reaches its terminal value on trading day T, the faster the speed of price formation.[13] Following recent research, we adjust IPE to account for overreaction and price reversals (Blankespoor et al. 2018, 2020).

Second, we construct a second measure of the speed of price formation that does not rely on an area under the curve approach. The *Jump Ratio* is the ratio of the immediate announcement period return (days [0,1]) to the post-announcement

[10] We refer the reader to Boehmer et al. (2021) for a detailed discussion of the methods used to identify retail trades. We thank Christina Zhu for generously sharing related code.

[11] We standardize the discrete Bloomberg measure to have continuous values based on the sample distribution characteristics, as in Drake et al. (2020).

[12] We use TAQ-based, intraday measures of spreads and depth, as they likely better capture short-window effects than do other measures that simply rely on open or close prices, as noted in Blankespoor et al. (2020).

[13] More specifically, for each day in the measurement window, we calculate the abnormal buy-and-hold return from day 0 (the earnings announcement date) to that day. Abnormal returns are calculated by adjusting the raw buy-and-hold stock return with the portfolio return from one of 125 benchmark portfolios formed based on quintiles of size, book-to-market, and momentum (5 × 5 × 5) over the same period (e.g., Daniel et al., 1997). We then scale this return by the cumulative buy-and-hold abnormal return for the entire measurement window.

period return of one week.[14] A larger ratio suggests relatively faster price formation. As noted in prior literature, *IPE* and the *Jump Ratio* are susceptible to significant outliers by virtue of their construction (Twedt 2016; Drake et al., 2017). We reduce the influence of outliers by dropping observations where the total period-T return is less than 2% (to account for small denominators, Blankespoor et al., 2018) and by using robust regression to estimate all models where either *IPE* or *Jump Ratio* is the dependent variable.[15]

3.3. Descriptive portrait of viral and non-viral earnings announcements

Before proceeding to our formal hypothesis tests, we provide a descriptive portrait of Twitter activity and tweet content around viral and non-viral earnings announcements to further our understanding of this novel phenomenon. In Fig. 1 (Fig. 2), we present average *#tweets* (*#feeds*) in event time in the 11-day window centered around the earnings announcement, day $t = 0$. The blue solid line depicts daily average social media shares for the top decile of observations, while the red dotted line depicts it for the other 90% of the sample. Both figures reveal a stark difference between the two groups, which are generally similar before and after the earnings announcement, but dramatically different at the earnings announcement. Specifically, in Fig. 1, on the earnings announcement day (day $t = 0$), we observe that the difference between viral and non-viral announcements is nearly 250 tweets, compared to roughly 30 tweets three days before or after. The contrast in Fig. 2 is even more striking − the difference in *Feeds[0,+1]* between viral and non-viral earnings announcements is huge (over 30 times more) at the earnings announcement.

We next examine the content of the tweets associated with viral and non-viral earnings announcements by examining several different characteristics of the average tweet around earnings announcements, including the length, absolute tone, negative words, positive words, use of financial words, inclusion of hyperlinks to other content, retweet status, and participation by unique users. We provide details on how we measure these attributes in Appendix A. In Fig. 3, we present the means of these characteristics for the viral and non-viral earnings announcement groups and note that all differences between the viral and non-viral groups are statistically significant at the $p < 0.01$ level (untabulated). We find that the tweets of viral earnings announcements are generally longer, but only by about two words, which is unsurprising given Twitter's character limits. The viral group also exhibits more extreme tone, as well as more usage of negative and positive words, suggesting that extreme sentiment runs higher on social media for viral earnings announcements.[16] We further find that the content of the viral group generally includes fewer financial words and is less likely to contain a hyperlink to other content online, potentially suggesting that tweets associated with viral earnings announcements contain less substantive content. In

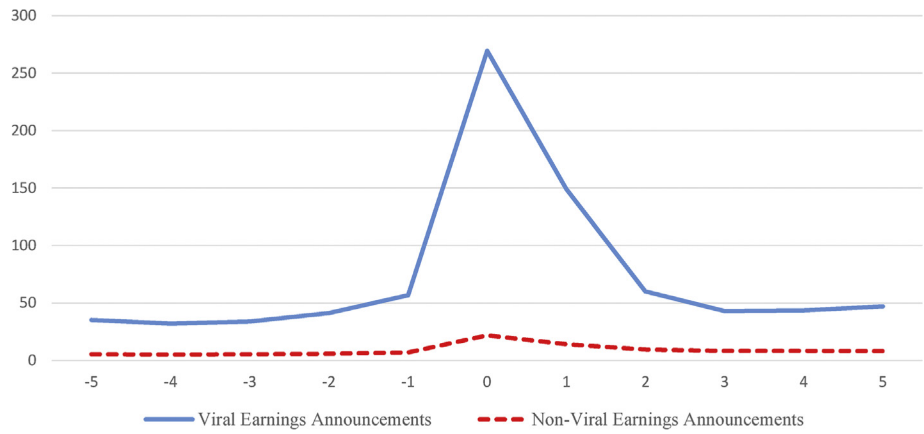


Fig. 1. Average Daily Tweets Around Earnings. This figure plots the average daily number of total tweets (*#tweets*) in event time during the earnings announcement period for viral earnings announcements compared to non-viral announcements.

[14] In robustness tests (untabulated), we use *IPE* and *Jump Ratios* computed over longer horizons of two weeks and one month. The results hold. Throughout the manuscript when we say that a robustness test holds, we mean that the sign of the coefficient is the same as that in our primary tests and that the significance classification (significant with at least a 10% cutoff level versus insignificant) is also the same.

[15] In untabulated analysis, we find similar results when we retain all observations, decile rank *IPE* and *Jump Ratio* and use ordinal logit instead of robust regression, or do not correct *IPE* for overreactions.

[16] We find that the correlation between the signed tone of the tweet and earnings (revenue) surprises is 0.10 (0.10), which provides some validation for our measure of tweet tone.

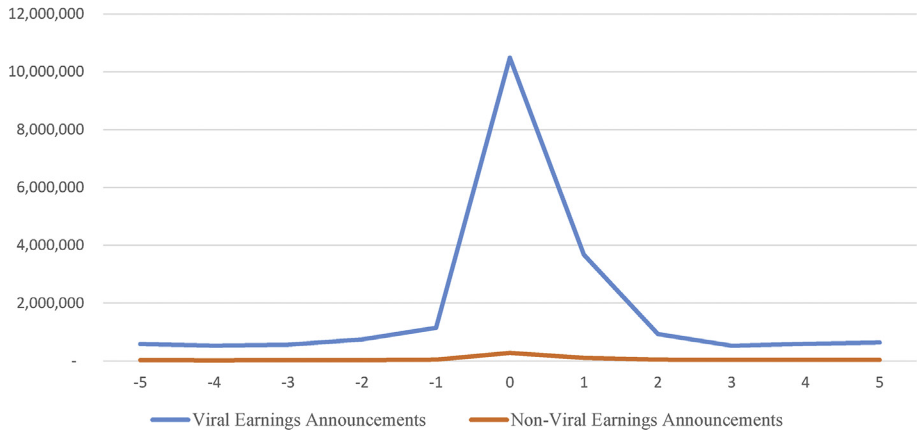


Fig. 2. Average Daily Followers of Tweets Around Earnings. This figure plots the average daily number of total followers (*#feeds*) in event time during the earnings announcement period for viral earnings announcements compared to non-viral announcements.

addition, we find that viral earnings announcements are associated with a greater proportion of retweets, as well as greater tweet intensity (more tweets per user).

To provide additional information about tweet content, we present representative word lists and tweet examples in Appendices B and C, respectively. In the first two columns of Appendix B, we provide a list of the top 20 most common words used in tweets for viral and non-viral earnings announcements. The majority of the most common words in both sets relate to earnings (e.g., earnings, EPS, revenue, beats, etc.); in both groups, the top two words are "earnings" and "EPS". This provides some validation that our data and measures are capturing tweets related to the earnings announcements. In the next three columns of Appendix B, we provide the top 20 most common words for negative viral tweets, positive viral tweets, and neutral viral tweets to provide a sense for the words included in the tweets that invoke these different sentiments. In Appendix C, we provide several examples of tweets categorized as having negative tone, positive tone, or neutral tone.

4. Determinants of earnings virality

In this section, we test our hypotheses regarding the determinants of earnings virality (H1a, H1b, and H1c). We begin by providing descriptive statistics in Table 1, Panel A.[17] We find that 8.2 percent of the earnings announcements in our sample are classified as having gone viral. The median *#tweets* is 26, or 13 tweets per day during the earnings announcement window; by comparison, the median number of tweets in the non-announcement period (*Tweets[-7,-1]*) is 18, or about 2.6 tweets per day during the week prior to the earnings announcement of the median sample firm. We also find that median traditional business press coverage is 0.0 articles in the week leading up to the earnings announcement (*Press Cov[-7,-1]*), but increases to 5.0 articles during the announcement period (*Press Cov[0,1]*), or 2.5 articles released per day.

In Table 1, Panel B, we provide descriptive statistics for viral versus non-viral earnings announcements. We see that the difference in *#tweets* is quite large, with a mean for viral earnings announcements of 279 tweets versus 33 for non-viral earnings announcements over the two-day announcement period. Similarly, the difference in *#feeds* is stark, with viral earnings announcements being viewed by approximately 7.5 million more feeds than are non-viral announcements. We also find that, compared to non-viral announcements, viral announcements are more likely to have larger revenue surprises, higher revenue growth, and lower book-to-market ratios. Viral earnings announcements are more likely to be in consumer-facing or tech-related industries, to have more media articles during and before the earnings announcement, and to have more general social media activity before the earnings announcement. In addition, viral announcements also tend to be observed in larger firms, those with more analyst following, those with greater institutional coverage, and in older firms. In sum, the types of firms that tend to have viral earnings are different on many dimensions than firms with non-viral earnings.

Next, we formally test our hypotheses on factors associated with earnings virality (H1a, H1b, and H1c) using the following model:

[17] We winsorize the top and bottom one percent of all continuous variables to reduce the influence of outliers.

Fig. 3. Viral and Non-Viral Tweet Characteristics. This figure presents means of various tweet and Twitter user attributes across viral and non-viral earnings announcements. All differences in the means of the two groups are statistically significant at the p < 0.01 level. Most measures are scaled by the total number of tweets (#Tweets) during the event-period, as indicated by %. Variable definitions are provided in Appendix A.

$$\text{VIRAL EARNINGS}_{it} = \beta_{FE} + \beta_K\, \text{CONTENT}_{it} + \beta_L\, \text{RECOGNITION} + \beta_M\, \text{RETAIL}_{it} + \beta_N\, \mathit{CONTROLS}_{it} + e_{it}, \tag{1}$$

where:

VIRAL EARNINGS = an indicator capturing earnings virality, defined as earnings announcements in the top annual decile of *#tweets* and *#feeds*;

CONTENT = a vector of *K* content-related variables, including *Abs Earn Surp*, *Abs Rev Surp*, *Neg Earn Surp*, *Neg Rev Surp*, and interactions;

RECOGNITION = a vector of *L* variables related to investor recognition, including *BtoC*, *Tech*, *High Rev Growth*, and *Low BTM*;

RETAIL = variable *Retail Own* that captures retail investor ownership;

Table 1
Descriptive statistics.

Variable	Mean	Std. Dev	10th	25th	Median	75th	90th
Panel A: Full Sample							
Viral Earnings	0.082	0.274	0.000	0.000	0.000	0.000	0.000
#tweets	52.6	92.8	2.0	10.0	26.0	53.0	109.0
#feeds	967,796	3,022,758	2906	95,619	242,265	553,511	1,431,262
Abs Earn Surp	0.005	0.012	0.000	0.001	0.002	0.004	0.011
Abs Rev Surp	0.010	0.023	0.000	0.001	0.002	0.008	0.025
Neg Earn Surp	0.309	0.462	0.000	0.000	0.000	1.000	1.000
Neg Rev Surp	0.410	0.492	0.000	0.000	0.000	1.000	1.000
BtoC	0.610	0.488	0.000	0.000	1.000	1.000	1.000
Tech	0.084	0.277	0.000	0.000	0.000	0.000	0.000
High Rev Growth	0.192	0.394	0.000	0.000	0.000	0.000	1.000
Low BTM	0.200	0.400	0.000	0.000	0.000	0.000	1.000
Retail Own	0.289	0.264	0.014	0.091	0.210	0.410	0.708
Press Cov[0,1]	6.525	5.414	0.000	3.000	5.000	8.000	13.000
Press Cov[-7,-1]	0.773	1.572	0.000	0.000	0.000	1.000	2.000
Press Cov[-30,-8]	2.519	3.770	0.000	0.000	1.000	3.000	7.000
Tweets[-7,-1]	36.445	63.368	1.000	6.000	18.000	39.000	78.000
Tweets[-30,-8]	106.130	187.207	3.000	18.000	52.000	114.000	224.000
Feeds[-7,-1]	281,930	854,297	53	11,470	59,803	184,161	489,796
Feeds[-30,-8]	810,271	2,373,575	1104	58,392	215,185	524,801	1,417,971
Abn Volume[0,1]	0.022	0.034	−0.001	0.003	0.010	0.026	0.054
Abn Volatility[0,1]	0.124	0.698	−0.035	0.002	0.023	0.105	0.378
Abn Ret[-30,-1]	0.000	0.091	−0.104	−0.049	−0.001	0.047	0.103
After Hours	0.533	0.499	0.000	0.000	1.000	1.000	1.000
# Firms Announce	224.0	135.9	46.0	111.0	207.0	341.0	414.0
#Analysts	9.4	7.2	2.0	4.0	7.0	13.0	20.0
Inst Count	261.3	277.7	52.0	101.0	171.0	311.0	568.0
Firm Size	8322	21,203	249	590	1762	5643	18,594
Firm Age	24.9	17.4	6.0	12.0	20.0	33.0	55.0
Time Trend	21.9	11.6	5.1	12.4	23.2	32.3	37.0

Panel B: Viral Earnings and Non-Viral Earnings

Variable	Viral Mean	Non-Viral Mean	Diff	P-value
#tweets	278.9	32.5	246.5	0.001
#feeds	7,881,614	353,814	7,527,800	0.001
Abs Earn Surp	0.004	0.005	−0.002	0.001
Abs Rev Surp	0.026	0.009	0.017	0.001
Neg Earn Surp	0.206	0.318	−0.111	0.001
Neg Rev Surp	0.351	0.416	−0.065	0.001
BtoC	0.723	0.600	0.123	0.001
Tech	0.150	0.078	0.072	0.001
High Rev Growth	0.210	0.191	0.019	0.001
Low BTM	0.370	0.185	0.185	0.001
Retail Own	0.276	0.291	−0.014	0.001
Press Cov[0,1]	15.587	5.720	9.867	0.001
Press Cov[-7,-1]	1.928	0.670	1.258	0.001
Press Cov[-30,-8]	5.954	2.214	3.739	0.001
Tweets[-7,-1]	154.101	25.996	128.1	0.001
Tweets[-30,-8]	425.641	77.756	347.9	0.001
Feeds[-7,-1]	1,875,968	140,371	1,735,597	0.001
Feeds[-30,-8]	5,003,109	437,926	4,565,183	0.001
Abn Volume[0,1]	0.059	0.018	0.041	0.001
Abn Volatility[0,1]	0.013	0.133	−0.120	0.001
Abn Ret[-30,-1]	0.002	0.000	0.002	0.034
After Hours	0.505	0.535	−0.031	0.001
# Firms Announce	154.9	230.1	−75.246	0.001
#Analysts	20.0	8.5	11.489	0.001
Inst Count	702.3	222.1	480.2	0.001
Firm Size	45,415	5028	40,387.0	0.001
Firm Age	32.6	24.2	8.36	0.001
Time Trend	22.1	21.9	0.18	0.222

Panel A presents descriptive statistics for the full sample of 82,232 quarterly earnings announcement observations. Panel B presents means for variables across viral and non-viral earnings observations. An earnings announcement is classified as viral if it is in both the top decile by year of number of tweets (*#tweets*) as well as number of Twitter followers (*#feeds*). Unlogged values of variables are provided here for ease of interpretation. Variable definitions are provided in Appendix A.

Table 2
Determinants of earnings virality.

DV=	(1) *Viral Earnings*	(2) *Viral Earnings (raw)*	(3) *Viral Earnings (<3 h)*
Abs Earn Surp	0.070	−0.037	−0.153
	(0.36)	(-0.28)	(-1.28)
Abs Rev Surp	0.799***	0.650***	0.750***
	(4.94)	(4.43)	(4.52)
Neg Earn Surp	−0.001	0.002	0.004**
	(-0.48)	(1.12)	(2.33)
Neg Rev Surp	−0.006***	−0.004***	−0.000
	(-3.02)	(-2.65)	(-0.20)
Abs Earn Surp x Neg Earn Surp	−0.095	−0.060	−0.050
	(-0.58)	(-0.55)	(-0.55)
Abs Rev Surp x Neg Rev Surp	−0.233*	−0.132	−0.070
	(-1.77)	(-1.09)	(-0.57)
BtoC	0.039***	0.024**	0.014
	(2.69)	(2.29)	(1.32)
Tech	−0.003	−0.012	−0.002
	(-0.17)	(-1.00)	(-0.22)
High Rev Growth	0.007*	0.006*	0.005
	(1.95)	(1.79)	(1.57)
Low BTM	0.001	0.000	−0.006
	(0.08)	(0.06)	(-1.13)
Retail Own	0.203***	0.144***	0.106***
	(14.49)	(12.08)	(8.90)
Press Cov[0,1]	0.022***	0.012***	0.012***
	(4.68)	(3.58)	(3.57)
Press Cov[-7,-1]	0.011***	0.013***	0.015***
	(3.94)	(4.82)	(5.57)
Press Cov[-30,-8]	0.001	0.005***	0.007***
	(0.37)	(2.60)	(3.93)
Tweets[-7,-1]	0.070***	0.064***	0.035***
	(20.32)	(16.91)	(11.08)
Tweets[-30,-8]	0.027***	0.021***	0.017***
	(11.82)	(10.76)	(8.59)
Feeds[-7,-1]	−0.008***	−0.014***	−0.004***
	(-12.73)	(-16.68)	(-6.83)
Feeds[-30,-8]	0.000	−0.005***	−0.000
	(0.23)	(-11.14)	(-0.11)
Abn Volume[0,1]	1.790***	1.136***	0.310***
	(19.64)	(12.43)	(4.15)
Abn Volatility[0,1]	0.006***	0.005***	0.004***
	(7.31)	(7.74)	(7.32)
Abn Ret[-30,-1]	−0.009	−0.007	−0.024***
	(-1.13)	(-1.00)	(-4.65)
After Hours	0.031***	0.026***	0.022***
	(5.88)	(6.19)	(5.18)
# Firms Announce	−0.027***	−0.016***	−0.008***
	(-6.78)	(-4.81)	(-2.76)
#Analysts	−0.021***	−0.016***	−0.019***
	(-5.76)	(-5.75)	(-7.50)
Inst Count	0.039***	0.026***	0.020***
	(10.92)	(8.59)	(7.04)
Firm Size	0.035***	0.022***	0.016***
	(12.66)	(10.35)	(8.49)
Firm Age	−0.020***	−0.013***	−0.014***
	(-5.46)	(-4.48)	(-5.03)
Time Trend	0.005***	0.002***	0.001*
	(5.75)	(2.78)	(1.80)
N	82,232	82,232	82,232
Adj. R^2	0.416	0.316	0.266

This table presents results from estimating Model (1) using three alternative measures of earnings virality. Variable definitions are provided in Appendix A. Standard errors are clustered by firm and t-statistics are reported in parentheses. All regressions include year, day of week, and industry (using SIC 2-digit industry definitions) fixed effects. All tests are two-tailed. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

CONTROLS = a vector of *N* controls, including *Press Cov[t,T],Tweets[t,T], Feeds[t,T], Abn Volume[0,1], Abn Volatility[0,1], Abn Ret [-30,-1], After Hours, # Firm Announce, #Analysts, Inst Count, Firm Size, Firm Age*, and *Time Trend*; and

β_{FE} = a set of fixed effects, including year, day of week, and industry (using SIC 2-digit industry definitions).

In model (1), *i* indexes firms and *t* indexes quarterly earnings announcements. We estimate model (1) using a general linear probability model to allow for a broad set of fixed effects; we also cluster the standard errors by firm.[18] In estimating model (1), we do not include firm fixed effects because we are interested in a wide variety of characteristics that explain earnings virality, including several firm characteristics, along with some event characteristics.[19]

We present the model (1) estimation results in Table 2, column (1). Regarding H1a, we find no evidence of an association between earnings news (sign or magnitude) and virality at the earnings announcement. In contrast, the results provide strong evidence that earnings virality is related to revenue surprises. More specifically, we find a positive and significant coefficient on the main effect *Abs Rev Surp* in all three columns. The coefficient on the interaction term *Abs Rev Surp x Neg Rev Surp* is negative, but only marginally significant in column (1).[20] These results suggest that the magnitude of revenue surprises is positively associated with earnings virality, and that the association appears to only modestly differ between bad news revenue surprises and good news revenue surprises. These findings also suggest that extreme social media dissemination is more related to revenue surprises than it is to earnings surprises.

Regarding H2b, we find some support for this hypothesis using two of the four proxies for investor recognition. The coefficients on *BtoC* and *High Rev Growth* are both positive and significant at the $p < 0.10$ level, suggesting that earnings virality is related to firms that are the most recognizable or that experience high revenue growth. We find no evidence that earnings virality is higher for technology firms or for those with high growth expectations (*Low BTM*). Finally, we find support for H1c as the coefficient on *Retail Own* is positive and highly significant. In sum, our results suggest a higher probability of earnings virality for recognizable firms that experience positive revenue surprises and growth and that have a large retail investor base.

In terms of control variables, we find that earnings virality is positively associated with traditional press coverage both in the week before and at the earnings announcement. We also find that prior-period *Tweets[t,T]* is positively related to earnings virality, while prior-period *Feeds[t,T]* is negatively so.[21] We find the earnings virality is positively associated with measures of general market activity, namely abnormal trading volume and abnormal return volatility, suggesting that earnings virality can arise because of variation in stock market activity (e.g., contemporaneous trading activity and price movements relate to virality). Earnings virality is negatively associated with a measure of general market distraction (as measured by the number of concurrent earnings announcements) and positively associated with earnings announced after market hours, which is a time when retail investors likely have more free time. We find mixed results on attention from sophisticated investors as earnings virality is negatively associated with analyst following and positively associated with institutional investor count. Regarding general firm characteristics, we find that the earnings announcements of large firms are more likely to go viral, but also that younger firms are more likely to go viral. Finally, we find evidence of a positive time trend in earnings virality during our sample period.[22]

Next, we use two alternative measures of earnings virality to further test our determinants hypotheses. In the first test, we examine the robustness of our results to a measure of earnings virality that does not rely on yearly decile assignment. We define this new measure, *Viral Earnings (raw)*, as an indicator variable set equal to one for earnings announcements that garner both more than 200 tweets and that are pushed to more than 2 million Twitter users, and to zero otherwise. While these cutoffs are admittedly arbitrary, they do capture extreme social media dissemination, as approximately 4.4 percent of the earnings announcements in our sample are classified as viral using this alternative definition.

In the second test, we use an alternative measure of earnings virality that uses a much tighter, 3-h event window. Specifically, we define *Viral Earnings (< 3 h)* as an indicator equal to one (and zero otherwise) if all of the following three criteria are met: (1) the earnings announcement is in the top annual decile of *#tweets*, (2) it is also in the top annual decile of *#feeds*, and (3) it satisfies both of the first two criteria within 3 h of the public release of the earnings announcement.[23] This group of viral earnings announcements is even more extreme, with only 2.8 percent of the announcements in our sample classified as viral using this definition.

We present the results for these alternative measures of earnings virality in columns (2) and (3) of Table 2. With respect to H1a, we continue to find that revenue surprises are positively associated with virality, but no longer observe a significant asymmetry between good and bad news. With respect to H1b and H1c, we find that our inferences generally hold using these

[18] Greene (2004) provides evidence that including fixed effects in discrete models, such as logistic regression, is problematic and can result in biased estimators.

[19] As discussed in section 2, while virality is generally triggered by an event, there are certain types of companies whose events are more likely to go viral because of who they are in addition to what they announce. The inclusion of firm fixed effects in the model would significantly absorb firm-level variation, which would limit to our ability to test firm-level determinants (deHaan, 2021).

[20] An F-test indicates that the sum of the coefficients on *Abs Rev Surp* and *Abs Rev Surp x Neg Rev Surp* is positive and significant at the 1% level (p-value = 0.001).

[21] The contrasting signs on these controls is likely due to the high correlation between them (the correlations are in the range of 75% and 88%). In an untabulated test, we estimate variance inflation factors (VIFs) for this model. We find that all determinants have VIFS less than 5.0 with the exception of the four social media control variables (*Tweets[-7,-1]*, *Tweets[-30,-8]*, *Feeds[-7,-1]*, and *Feeds[-30,-8]*). When we re-estimate model (1) after removing either *Tweets[-7,-1]* and *Tweets[-30,-8]* or after removing *Feeds[-7,-1]* and *Feeds[-30,-8],* we find qualitatively similar results.

[22] Because our definition of virality is met only when the earnings announcement is in the top annual decile of both *#tweets* and *#feeds,* the positive coefficient on the time trend variable may indicate that firms are increasingly more likely to meet both criteria over time.

[23] The tenor of the results remains consistent if we use an alternative cutoff of 6 h to capture the speed of virality.

Table 3
Earnings virality and investor activity.

DV=	*Abn Volume[0,1]*	*Abn Volatility[0,1]*	*Abn Retail Volume[0,1]*	*Abn Robinhood Holdings[0,1]*	*Abn Bloomberg Activity[0,1]*
Viral Earnings	**0.032*****	**0.068*****	**0.005*****	**206.947*****	**−0.102*****
	(15.83)	**(7.79)**	**(17.69)**	**(9.61)**	**(-6.67)**
Abs Earn Surp	0.218***	0.375	0.030***	−300.470	4.200***
	(8.20)	(0.65)	(8.37)	(-1.00)	(5.76)
Abs Rev Surp	−0.012	0.335**	−0.000	1045.396***	0.886***
	(-0.93)	(2.52)	(-0.06)	(3.92)	(3.81)
Neg Earn Surp	0.004***	0.010	0.000***	−5.343	0.001
	(12.60)	(1.52)	(7.87)	(-1.27)	(0.12)
Neg Rev Surp	0.001***	0.025***	0.000	−1.253	−0.004
	(3.96)	(4.13)	(0.79)	(-0.33)	(-0.62)
Abs Earn Surp x Neg Earn Surp	−0.143***	2.233***	−0.016***	388.659	−0.902
	(-5.39)	(3.16)	(-4.17)	(1.01)	(-1.31)
Abs Rev Surp x Neg Rev Surp	0.013	−0.162	0.003	428.420	0.395
	(1.01)	(-0.82)	(1.43)	(1.17)	(1.50)
BtoC	0.000	−0.031**	0.000	35.043**	−0.061**
	(0.22)	(-2.55)	(0.64)	(1.98)	(-2.21)
Tech	0.007***	−0.050***	0.001***	−24.941	0.053*
	(3.50)	(-3.11)	(2.98)	(-1.04)	(1.69)
High Rev Growth	0.005***	−0.004	0.001***	2.560	0.018*
	(11.10)	(-0.49)	(10.26)	(0.38)	(1.79)
Low BTM	0.004***	0.003	0.000***	3.088	0.023*
	(4.51)	(0.50)	(4.24)	(0.36)	(1.95)
Retail Own	−0.028***	0.075***	−0.002***	163.113***	−0.193***
	(-17.87)	(3.95)	(-7.87)	(7.61)	(-6.22)
Press Cov[0,1]	0.004***	0.003	0.000***	−4.313	0.024***
	(8.73)	(0.77)	(7.23)	(-1.06)	(3.41)
Press Cov[-7,-1]	−0.001***	−0.004	−0.000***	14.452***	−0.029***
	(-3.08)	(-0.98)	(-3.32)	(2.72)	(-4.86)
Press Cov[-30,-8]	−0.002***	−0.001	−0.000***	14.344***	−0.037***
	(-8.25)	(-0.23)	(-6.92)	(3.76)	(-7.62)
Tweets[-7,-1]	0.003***	0.008	0.000***	21.867***	−0.004
	(9.25)	(1.27)	(10.80)	(4.34)	(-0.54)
Tweets[-30,-8]	−0.000	−0.005	0.000	0.799	−0.012*
	(-1.15)	(-0.86)	(0.76)	(0.14)	(-1.67)
Feeds[-7,-1]	0.000**	0.000	−0.000	6.551***	0.002
	(2.54)	(0.06)	(-0.90)	(2.94)	(1.03)
Feeds[-30,-8]	0.001***	−0.005**	0.000***	27.831***	0.011***
	(8.67)	(-2.48)	(6.62)	(7.55)	(4.52)
Abn Volume[0,1]		−0.719***		671.770***	3.889***
		(-9.71)		(4.55)	(28.17)
Abn Volatility[0,1]	−0.001***		−0.000***	0.375	−0.005
	(-8.98)		(-8.40)	(0.24)	(-0.51)
Abn Ret[-30,-1]	−0.009***	−0.352***	−0.001***	−67.160***	0.152***
	(-6.15)	(-9.68)	(-3.71)	(-2.83)	(4.49)
After Hours	−0.000	−0.007	0.000	4.979	0.448***
	(-0.29)	(-1.08)	(0.31)	(0.78)	(40.03)
# Firms Announce	−0.002***	0.017***	−0.000***	−4.756	0.021***
	(-4.78)	(5.05)	(-6.09)	(-1.13)	(3.58)
#Analysts	0.009***	−0.022***	0.001***	−12.579***	0.188***
	(17.62)	(-3.62)	(12.77)	(-2.61)	(16.41)
Inst Count	−0.005***	0.016***	−0.000***	44.358***	−0.021***
	(-10.82)	(3.89)	(-5.06)	(6.73)	(-2.83)
Firm Size	−0.005***	−0.059***	−0.001***	−9.663**	0.167***
	(-13.42)	(-14.00)	(-17.22)	(-2.21)	(25.62)
Firm Age	−0.001***	−0.014***	−0.000***	−23.297***	0.001
	(-3.47)	(-2.79)	(-3.06)	(-5.37)	(0.15)
Time Trend	−0.000***	0.019***	−0.000	−11.731***	−0.032***
	(-3.24)	(8.49)	(-0.04)	(-6.16)	(-12.87)
N	82,232	82,232	82,232	15,547	53,029
Adj. R^2	0.311	0.041	0.293	0.275	0.337

This table presents results from estimating Model (2). Variable definitions are provided in Appendix A. Standard errors are clustered by firm and t-statistics are reported in parentheses. All regressions include year, day of week, and industry (using SIC 2-digit industry definitions) fixed effects. All tests are two-tailed. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

alternative definitions of earnings virality. When taken together, the results of our determinants tests provide consistent evidence that earnings announcements with larger revenue surprises and issued by firms with higher investor recognition and higher retail ownership are more likely to go viral.

5. Earnings virality and investor activity

We next examine the relation between earnings virality and two sets of measures that capture investor activity at the earnings announcement. The first set of measures capture general trading activity and includes abnormal trading (*Abn Volume [0,1]*) and intra-day price volatility (*Abn Volatility[0,1]*). The second set of measures focuses in on the trading activity of specific investor groups. We investigate retail investor activity using two measures, abnormal retail trading volume (*Abn Retail Volume [0,1]*) and abnormal levels of stock holdings in Robinhood (*Abn Robinhood Holdings[0,1]*). We then examine a proxy for institutional investor search activity using professional investor attention to firm-specific news articles on Bloomberg (*Abn Bloomberg Activity[0,1]*) (Ben-Rephael et al., 2017). We estimate these associations using the following model:

$$\text{INVESTOR ACTIVITY}_{it} = \beta_{FE} + \beta_1\,\text{VIRAL EARNINGS}_{it} + \beta_K\,\text{CONTENT}_{it} + \beta_L\,\mathit{RECOGNITION} + \beta_M\,\mathit{RETAIL}_{it} + \beta_N\,\mathit{CONTROLS}_{it} + e_{it}, \quad (2)$$

Where *INVESTOR ACTIVITY* is one of five dependent variables, *Abn Volume[0,1], Abn Volatility[0,1], Abn Retail Volume[0,1], Abn Robinhood Holdings [0,1]*, or *Abn Bloomberg Activity[0,1],* and all other variables are defined previously. As before, *i* indexes firms and *t* indexes quarterly earnings announcements. We again include a set of year, day of the week, and industry fixed effects, and we cluster the standard errors by firm.[24]

We present the model (2) estimation results using abnormal trading volume and volatility as alternative dependent variables in the first two columns of Table 3. In column (1), we find that the coefficient on *Viral Earnings* is positive and significant. The magnitude of the coefficient indicates that abnormal trading volume is 3.2 percent higher around viral earnings announcements than around non-viral announcements. In column (2), we find that earnings virality is also positively associated with abnormal price volatility around the announcement. These results suggest that earnings virality coincides with increased general trading activity.[25]

Regarding investor-specific activity, we find that retail investors are particularly active around viral earnings announcements. More specifically, in column (3), the coefficient on *Viral Earnings* is positive and statistically significant and the magnitude of the coefficient suggests that abnormal retail trading volume is 0.50 percent higher for viral earnings announcements than for non-viral earnings announcements. In column (4), we find additional evidence of increased retail trading activity around viral earnings announcements using abnormal Robinhood holdings. The magnitude of the coefficient indicates that approximately 207 additional Robinhood users choose to hold the stock of a firm whose earnings go viral. In contrast, in column (5) we find that *Viral Earnings* is negatively associated with professional investor information search activity on Bloomberg terminals around earnings announcements. In sum, the results presented in Table 3 provide strong evidence that earnings virality coincides with increased investor activity that includes an influx of trading by retail investors.

5.1. Earnings virality, liquidity, and speed of price formation

Our evidence that earnings virality is associated with more trading activity does not speak to whether earnings virality has a positive or negative association with capital market outcomes. Thus, we now test H2a and H2b, which relate to whether earnings virality helps or hinders liquidity and speed of price formation, respectively.

5.2. Earnings virality and liquidity (H2a)

We investigate the association between earnings virality and liquidity using abnormal spreads and abnormal depths. We test the relation between earnings virality and liquidity using the following model:

$$\text{LIQUIDITY}[0,1]_{it} = \beta_{FE} + \beta_1\,\text{VIRAL EARNINGS}_{it} + \beta_K\,\text{CONTENT}_{it} + \beta_L\,\mathit{RECOGNITION} + \beta_M\,\mathit{RETAIL}_{it} + \beta_N\,\mathit{CONTROLS}_{it} + e, \quad (3)$$

Where *LIQUIDITY[0,1]* is one of two dependent variables, *Abn Spread[0,1]* or *Abn Depth[0,1]* and all variables are as defined previously. As before, we include a set of year, day of the week, and industry fixed effects and we cluster the standard errors by firm. A negative (positive) β_1 in the *Abn Spread[0,1]* (*Abn Depth[0,1]*) regression rejects H2a and suggests that earnings virality is positively associated with liquidity at the earnings announcement. In addition, a positive (negative) β_1 in the *Abn Spread [0,1]* (*Abn Depth[0,1]*) regression also rejects H2a and suggests that earnings virality is negatively associated with market liquidity.

We present the model (3) estimation results in Table 4. We find that *Viral Earnings* is positively associated with abnormal spreads in column (1), suggesting an increase in information asymmetry for viral earnings announcements. We also find that

[24] We do not include firm fixed effects in model (2) because we seek to measure the associated economic effects based on the combined influence of event factors (what the firm announces) and firm factors (who the firm is). If we include firm fixed effects in our main market specifications, then we will have removed a considerable amount of variation related to the firm factor. In Section 7 below, we provide several additional analyses, including tests with firm fixed effects, to investigate whether the economic consequences associated with earnings virality are largely related to firm-level factors.

[25] Our findings that abnormal volume and volatility load significantly as both determinants and outcomes of earnings virality speaks to the potential concern of reverse causality. We discuss this concern in more detail in Section 7.

Table 4
Earnings virality and liquidity.

DV=	(1) *Abn Spread[0,1]*	(2) *Abn Depth[0,1]*
Viral Earnings	**0.019*****	**-0.030*****
	(10.74)	**(-5.02)**
Abs Earn Surp	-0.116*	-0.451**
	(-1.95)	(-2.31)
Abs Rev Surp	-0.066***	-0.973***
	(-3.40)	(-12.27)
Neg Earn Surp	-0.001	0.012***
	(-0.90)	(4.39)
Neg Rev Surp	0.000	-0.008***
	(0.61)	(-3.22)
Abs Earn Surp x Neg Earn Surp	-0.075	0.036
	(-1.00)	(0.15)
Abs Rev Surp x Neg Rev Surp	-0.005	0.174
	(-0.20)	(1.57)
BtoC	-0.004**	-0.032***
	(-2.29)	(-4.78)
Tech	0.004	0.004
	(1.59)	(0.53)
High Rev Growth	0.002**	0.009***
	(2.04)	(3.04)
Low BTM	0.001	0.005
	(0.72)	(1.44)
Retail Own	-0.003	-0.066***
	(-1.28)	(-7.15)
Press Cov[0,1]	0.001*	0.012***
	(1.94)	(5.03)
Press Cov[-7,-1]	-0.000	-0.001
	(-0.34)	(-0.65)
Press Cov[-30,-8]	0.000	-0.006***
	(0.43)	(-3.83)
Tweets[-7,-1]	-0.003***	0.007***
	(-3.83)	(2.87)
Tweets[-30,-8]	0.001	-0.009***
	(1.64)	(-4.15)
Feeds[-7,-1]	0.001***	0.000
	(4.01)	(0.14)
Feeds[-30,-8]	-0.001***	-0.003***
	(-2.98)	(-3.97)
Abn Volume[0,1]	-0.090***	2.070***
	(-6.81)	(40.85)
Abn Volatility[0,1]	0.038***	-0.013***
	(26.78)	(-5.93)
Abn Ret[-30,-1]	-0.046***	0.261***
	(-12.01)	(21.63)
After Hours	0.009***	-0.008**
	(9.82)	(-2.43)
# Firms Announce	0.005***	-0.013***
	(9.10)	(-7.02)
#Analysts	-0.005***	-0.011***
	(-6.52)	(-3.67)
Inst Count	-0.000	-0.012***
	(-0.51)	(-5.20)
Firm Size	-0.003***	0.006***
	(-6.39)	(3.36)
Firm Age	-0.001	-0.002
	(-0.93)	(-0.82)
Time Trend	0.003***	-0.006***
	(11.33)	(-7.11)
N	82,146	82,232
Adj. R^2	0.125	0.130

This table presents results from estimating Model (3). Variable definitions are provided in Appendix A. Standard errors are clustered by firm and t-statistics are reported in parentheses. All regressions include year, day of week, and industry (using SIC 2-digit industry definitions) fixed effects. All tests are two-tailed. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

earnings virality is negatively associated with abnormal depths in column (2), suggesting that when earnings go viral, trading depth is more shallow. Thus, the results reject H2a. The combination of higher abnormal spreads and lower abnormal depth suggests that earnings virality coincides with lower liquidity.

The results in Table 4 are mixed on the associations between liquidity and either traditional media coverage or social media virality. We find that press coverage is positively associated with spreads and thus lower liquidity, which contrasts with our expectations and the findings of prior research (e.g., Bushee et al., 2010).[26] On the other hand, we find that press coverage is positively associated with trading depth, which Blankespoor et al. (2018) argue is perhaps a better measure of liquidity. For depths, our evidence consistently suggests that viral social media and traditional business press coverage have different effects on liquidity – that is, while earnings virality is associated with shallower trading depth, traditional media is associated with greater trading depth.

5.3. Earnings virality and the speed of price formation (H2b)

Next, we examine the relation between earnings virality and the speed of price formation following the earnings announcement. We do this using the following model:

$$\text{SPEED OF PRICE FORMATION}_{it} = \beta_{FE} + \beta_1\ \text{VIRAL EARNINGS}_{it} + \beta_K\ \text{CONTENT}_{it} + \beta_L\ \textit{RECOGNITION} + \beta_M\ \textit{RETAIL}_{it} + \beta_N\ \textit{CONTROLS}_{it} + e, \tag{4}$$

Where *SPEED OF PRICE FORMATION* is one of two dependent variables, *IPE* or *Jump Ratio*. All variables are previously defined. Consistent with our other models, we include a set of year, day of the week, and industry fixed effects and we cluster the standard errors by firm. We also estimate model (4) using robust regression to account for outliers in *IPE* and *Jump Ratio*. A significant β_1 rejects H2b. A positive β_1 suggests that earnings virality increases the speed of price formation, whereas a negative β_1 suggests that earnings virality decreases the speed of price formation.

We present the model (4) estimation results using *IPE* and *Jump Ratio* as the dependent variables in columns (1) and (2) of Table 5, respectively. In both specifications, we find a negative and significant coefficient on *Viral Earnings*. This evidence rejects H2b, indicating that earnings virality slows the speed of price formation immediately after earnings announcements, and suggests that extreme social media dissemination has negative pricing implications.

We also note an important contrast between the effect of social media virality and the traditional media on the speed of price formation. Specifically, we find positive and significant coefficients on *Press Cov[0,1]* in both columns of Table 5, suggesting that press coverage of earnings news increases the speed of price formation, consistent with prior work (e.g., Twedt 2016). In contrast, the negative coefficients on *Earnings Virality* suggest the exact opposite effect for social media virality, i.e., a slowing down of price formation. Together, these findings suggest that increased dissemination of earnings announcement news by the traditional press accelerates price formation, but viral social dissemination of the news is associated with a deceleration of price formation. These contrasting effects support the idea that the construct of earnings virality on social media is distinct from that of traditional media attention examined in extant research.

Taken together, the results presented in Tables 4 and 5 provide consistent evidence that earnings virality can have a negative impact on some market outcomes, as it is associated with increased spreads, decreased depths, and slower price formation.

6. Additional tests and robustness

6.1. Endogeneity

One challenge with our market-based tests is the potential for endogeneity and in particular, concerns related to reverse causality or omitted variables. In dealing with these issues, we follow the recommendations in Glaeser and Guay (2017), Heckman and Singer (2017) and Armstrong et al. (2021) to triangulate results across multiple research designs and tests.

With respect to reverse causality (e.g., higher market activity leads to earnings virality), we acknowledge that this is a clear possibility for the trading activity results tabulated in Table 3, as trading volume and stock price movements could be easily observed by investors, who in turn could spread the news on social media. We address reverse causality concerns in three ways. First, we examine the extent to which tweets reference stock price movements during the earnings announcement window. More specifically, we search all tweets for the presence of one of the following words: “stock”, “share”, “price”, “pop”, “rally”, “hold”, “short”, “squeeze”, or “bounce”. We then calculate the ratio of tweets that include one of these terms to total tweets per day (*Price Tweets%*). Descriptively, we find that 16% of all earnings announcement tweets contain at least one of these words. We then include this ratio in our determinants model (1). We find that the coefficient on *Price Tweets%* is

[26] There are several potential reasons for this finding. First, as trading has become more automated, bid-ask spreads have steadily decreased over time for many firms (e.g., Hendershott et al. (2011), Angel et al. (2015), and Haslag and Ringgenberg (2021)). Second, there are multiple measures of bid-ask spreads, some that capture intraday variation (TAQ or effective spreads) and others that capture end-of-day variation (CRSP or quoted spreads). As noted, prior research does find that press coverage is negatively related to bid-ask spreads, but only using CRSP-based measures (e.g., Bushee et al., 2010). If we use quoted spreads, we find an insignificant coefficient on *Press Cov[0,1]*, suggesting no association between bid-ask spreads and press coverage in our sample. Because of recent trends in bid-ask spreads, as well as difficulty in measuring them, we urge caution in drawing inferences regarding spreads. Instead, we primarily rely on the intuition discussed in Blankespoor et al. (2018) that trading depths more cleanly capture liquidity.

Table 5
Earnings virality and speed of price formation.

DV=	(1) *IPE*	(2) *Jump Ratio*
Viral Earnings	**−0.062***** **(-3.12)**	**−0.052***** **(-5.25)**
Abs Earn Surp	0.263 (0.45)	−0.198 (-0.69)
Abs Rev Surp	−0.376 (-1.40)	−0.187 (-1.40)
Neg Earn Surp	0.032*** (2.97)	0.011** (2.04)
Neg Rev Surp	−0.010 (-1.04)	−0.009* (-1.72)
Abs Earn Surp x Neg Earn Surp	0.169 (0.24)	0.132 (0.38)
Abs Rev Surp x Neg Rev Surp	−0.208 (-0.53)	0.008 (0.04)
BtoC	−0.046** (-2.21)	−0.024** (-2.32)
Tech	0.005 (0.20)	0.003 (0.27)
High Rev Growth	−0.033*** (-2.94)	−0.005 (-0.98)
Low BTM	−0.049*** (-4.30)	−0.018*** (-3.22)
Retail Own	−0.010 (-0.40)	−0.029** (-2.39)
Press Cov[0,1]	0.049*** (6.89)	0.020*** (5.67)
Press Cov[-7,-1]	−0.001 (-0.14)	−0.004 (-0.90)
Press Cov[-30,-8]	−0.034*** (-5.40)	−0.019*** (-6.00)
Tweets[-7,-1]	−0.061*** (-6.33)	−0.017*** (-3.52)
Tweets[-30,-8]	−0.022** (-2.40)	−0.013*** (-3.02)
Feeds[-7,-1]	0.007*** (2.79)	0.002 (1.53)
Feeds[-30,-8]	−0.004 (-1.50)	0.001 (0.76)
Abn Volume[0,1]	6.319*** (45.96)	2.924*** (42.62)
Abn Volatility[0,1]	0.021*** (3.46)	0.009*** (3.00)
Abn Ret[-30,-1]	0.121*** (2.72)	−0.050** (-2.28)
After Hours	−0.478*** (-51.13)	−0.017*** (-3.68)
# Firms Announce	−0.012** (-2.14)	0.004 (1.38)
#Analysts	−0.021*** (-2.62)	0.002 (0.58)
Inst Count	0.018*** (2.79)	0.001 (0.24)
Firm Size	0.085*** (16.48)	0.027*** (10.45)
Firm Age	0.062*** (9.25)	0.020*** (5.98)
Time Trend	0.010** (2.45)	0.005** (2.48)
N	60,288	60,288
Adj. R^2	0.130	0.071

This table presents results from estimating Model (4). Variable definitions are provided in Appendix A. Standard errors are clustered by firm and t-statistics are reported in parentheses. All regressions include year, day of week, and industry (using SIC 2-digit industry definitions) fixed effects. All tests are two-tailed. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

negative and significant at the $p < 0.01$ level (untabulated). Thus, we find no evidence that tweets that reference stock price movements are increasing earnings virality. Second, in all hypothesis tests (namely Tables 4–9), we also include measures of contemporaneous market activity (abnormal trading volume and abnormal return volatility for the event period) in order to

Table 6
Market effects of earnings virality (change in twitter algorithm).

DV=	*Abn Volume[0,1]*	*Abn Spread[0,1]*	*Abn Depth[0,1]*	*IPE*	*Jump Ratio*
Viral Earnings	0.027***	0.034***	−0.021***	−0.008	−0.035***
	(12.73)	(11.48)	(-3.17)	(-0.35)	(-2.83)
Post	0.001*	−0.002	0.056***	0.184***	0.017
	(1.76)	(-1.02)	(8.13)	(6.16)	(1.15)
Viral Earnings x Post	**0.012*****	**−0.032*****	**−0.019****	**−0.113*****	**−0.037****
	(5.36)	**(-10.48)**	**(-2.26)**	**(-3.66)**	**(-2.40)**
Controls	Yes	Yes	Yes	Yes	Yes
N	82,232	82,146	82,232	60,288	60,288
Adj. R^2	0.313	0.127	0.131	0.131	0.071

This table presents results of market tests, examining the effects of earnings virality before and after the changes to the Twitter platform in March 2016. Variable definitions are provided in Appendix A. Standard errors are clustered by firm and t-statistics are reported in parentheses. All regressions include year, day of week, and industry (using SIC 2-digit industry definitions) fixed effects. All tests are two-tailed. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

control for changes in price movements or investor activity that can simultaneously impact both virality and trading activity. Third, for the tests of our formal hypotheses, we employ market outcome variables for which reverse causality is less of a concern, either because they are difficult (if not impossible) for investors to observe or calculate in real-time at the earnings announcement (abnormal effective bid-ask spreads and depths) or their measurement extends to after earnings go viral (e.g., *IPE*, *Jump Ratio*).

Our tests also face concerns about omitted variables, primarily variables that capture unique elements of attention relative to what is included in our models. We conduct two additional tests to account for the potential effects of omitted variables. In the first test, we exploit quasi-exogenous temporal variation in the Twitter platform. We focus on two changes Twitter made to its platform in March of 2016. The first change relates to how tweets appear in the newsfeed. Prior to the change, tweets appeared chronologically in the newsfeed in the order they were posted. Subsequent to the March 2016 change, however, tweets appear according to a proprietary “relevance” algorithm designed to promote certain tweets over other tweets (Koumchatzky and Anryeyev 2017; Oremus 2017).[27] The new algorithm uses machine learning to score tweets based on the probability of user engagement where tweets with higher scores are given greater prominence in user's newsfeed. A senior engineer on the team that developed and tested the new algorithm noted the following: “Overall, we've observed that people who have experienced this new Search results page tend to not only engage more with the Search results but also tweet more and spend more time on Twitter” (Huang 2016).[28] The second change relates to a new feature titled “In Case You Missed It” (ICYMI), which provides a recap of the most popular tweets according to the new algorithm that a user may have missed while not on Twitter. This change increases the possibility that viral earnings announcements could be viewed on an investor's newsfeed on a time delay.

We argue that these two changes increase the likelihood that viral earnings announcements are viewed by an even more diverse set of investors as the viral news will be given priority placement in their feeds and because investors will be given an opportunity to view it at a later point in time in case they missed it initially. This has the potential to increase adverse selection and further reduce liquidity or to slow the incorporation of information into prices, as we observe in our primary tests. In addition, these changes serve as a plausible shock to social media virality, but not to other related constructs such as investor attention or traditional press coverage. Thus, the negative market results associated with earnings virality in our main tests could be even stronger (i.e., more negative) after these two changes made by Twitter to the platform.

We note however, that this test has several limitations. First, the algorithmic change affects all firms in the sample, so no clear control group exists against which to benchmark the results. Second, other changes occurred after 2016 that could also potentially impact earnings virality, including Bloomberg's expanded offering of Twitter data to subscribers in 2018 (discussed previously) and the general increase in Twitter usage observed over time. To help account for other potential temporary changes, we include a trend variable as a control variable, year fixed effects, and we define earnings virality based on an annual decile ranking. Third, while the 2016 changes directly affect our primary construct of interest, earnings virality, it is a generalized pre-versus post-test that cannot isolate a causal effect of earnings virality on market outcomes due to these limitations. The inferences from this test should be interpreted within the bounds of these limitations.

To conduct this test, we create an indicator variable *Post* for all observations subsequent to March 2016, and interacting *Post* with *Viral Earnings*. We examine abnormal trading volume, liquidity, and speed of price formation. We present the results in Table 6. We find that there is increased trading volume after the change and that the detrimental effects of virality on the market measures are stronger after the 2016 changes for three of our four market outcome variables: *AbnDepth[0,1]*, *IPE*, and

[27] https://www.engadget.com/2016-12-21-twitters-search-results-sorted-by-relevance.html; https://tinuiti.com/blog/paid-social/how-the-twitter-algorithm-works-in-2021/.
[28] https://blog.twitter.com/engineering/en_us/topics/insights/2016/moving-top-tweet-search-results-from-reverse-chronological-order-to-relevance-order.

Table 7
Market effects of earnings virality (**cross-sectional tests**).

DV=	*Abn Volume[0,1]*	*Abn Spread[0,1]*	*Abn Depth[0,1]*	*IPE*	*Jump Ratio*
Panel A: Tweet Tone					
Viral Earnings x High Abs Tone	0.049***	0.026***	−0.031***	−0.117***	−0.071***
	(17.37)	(9.07)	(-3.97)	(-4.73)	(-5.80)
Viral Earnings x Low Abs Tone	0.016***	0.013***	−0.028***	−0.005	−0.032***
	(8.65)	(9.15)	(-4.35)	(-0.19)	(-2.60)
Difference	**0.033*** **	**0.013*** **	**−0.003**	**−0.112*** **	**−0.039*** **
Viral Earnings x High Negative	0.048***	0.024***	−0.028***	−0.147***	−0.098***
	(17.88)	(9.64)	(-3.50)	(-5.94)	(-7.94)
Viral Earnings x Low Negative	0.017***	0.015***	−0.031***	0.022	−0.007
	(9.66)	(9.07)	(-4.96)	(0.90)	(-0.58)
Difference	**0.031*** **	**0.009*** **	**0.003**	**−0.169*** **	**−0.091*** **
Viral Earnings x High Positive	0.041***	0.024***	−0.042***	−0.076***	−0.045***
	(15.54)	(8.84)	(-5.91)	(-3.09)	(-3.70)
Viral Earnings x Low Positive	0.024***	0.014***	−0.017***	−0.047*	−0.058***
	(11.69)	(9.27)	(-2.61)	(-1.90)	(-4.70)
Difference	**0.017*** **	**0.003*** **	**−0.017*** **	**−0.029**	**0.013**
Panel B: Tweet Content					
DV=	(1) *Abn Volume[0,1]*	(2) *Abn Spread[0,1]*	(3) *Abn Depth[0,1]*	(4) *IPE*	(5) *Jump Ratio*
Viral Earnings x High Earn%	0.015***	0.016***	−0.025***	−0.030	−0.044***
	(8.50)	(8.49)	(-3.63)	(-1.16)	(-3.49)
Viral Earnings x Low Earn%	0.049***	0.021***	−0.035***	−0.089***	−0.058***
	(19.95)	(9.58)	(-4.98)	(-3.70)	(-4.85)
Difference	**−0.034*** **	**−0.005** **	**0.010**	**0.059** **	**0.014**
Viral Earnings x High Link%	0.016***	0.013***	−0.022***	0.034	−0.024*
	(9.38)	(9.25)	(-3.50)	(1.35)	(-1.93)
Viral Earnings x Low Link%	0.049***	0.025***	−0.038***	−0.157***	−0.081***
	(17.90)	(9.22)	(-4.83)	(-6.34)	(-6.57)
Difference	**−0.033*** **	**−0.012*** **	**0.016***	**0.191*** **	**0.057*** **
Panel C: Tweet Uniqueness					
DV=	(1) *Abn Volume[0,1]*	(2) *Abn Spread[0,1]*	(3) *Abn Depth[0,1]*	(4) *IPE*	(5) *Jump Ratio*
Viral Earnings x High Retweet%	0.023***	0.016***	−0.024***	−0.063**	−0.051***
	(9.76)	(8.95)	(-3.08)	(-2.42)	(-3.89)
Viral Earnings x Low Retweet%	0.040***	0.021***	−0.035***	−0.061***	−0.053***
	(16.98)	(9.36)	(-5.36)	(-2.58)	(-4.48)
Difference	**−0.017*** **	**−0.005** **	**0.011**	**−0.002**	**0.002**
Viral Earnings x High Tweet Intensity	0.034***	0.025***	−0.031***	−0.087***	−0.067***
	(15.38)	(10.23)	(-4.43)	(-3.54)	(-5.47)
Viral Earnings x Low Tweet Intensity	0.031***	0.013***	−0.029***	−0.037	−0.037***
	(14.05)	(7.37)	(-4.45)	(-1.53)	(-3.03)
Difference	**0.003** **	**0.012*** **	**−0.002**	**−0.050***	**−0.030** **

This table presents results of market tests of earnings virality in various cross-sectional analyses based on within-social media factors. Variable definitions are provided in Appendix A. Standard errors are clustered by firm and t-statistics are reported in parentheses. All regressions include year, day of week, and industry (using SIC 2-digit industry definitions) fixed effects. All tests are two-tailed. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 8
Market effects of earnings virality (alternative measures).

DV=	(1) *Abn Volume[0,1]*	(2) *Abn Spread[0,1]*	(3) *Abn Depth[0,1]*	(4) *IPE*	(5) *Jump Ratio*
Panel A: Raw Measure					
Viral Earnings (raw)	**0.032*** **	**0.006*** **	**−0.036*** **	**−0.088*** **	**−0.056*** **
	(11.12)	**(3.70)**	**(-4.83)**	**(-3.58)**	**(-4.62)**
Controls	Yes	Yes	Yes	Yes	Yes
N	82,232	82,146	82,232	60,288	60,288
Adj. R^2	0.296	0.123	0.130	0.130	0.070
Panel B: Speed-Based Measure					
DV=	(1)*Abn Volume[0,1]*	(2)*Abn Spread[0,1]*	(3)*Abn Depth[0,1]*	(4) *IPE*	(5)*Jump Ratio*
Viral Earnings (< 3 h)	**0.013*** **	**0.031*** **	**0.005**	**−0.012**	**−0.026***
	(3.91)	**(8.12)**	**(0.59)**	**(-0.39)**	**(-1.74)**
Controls	Yes	Yes	Yes	Yes	Yes
N	82,232	82,146	82,232	60,288	60,288
Adj. R^2	0.272	0.125	0.130	0.130	0.070

This table presents results of market tests using alternative measures of earnings virality. Variable definitions are provided in.Appendix A Standard errors are clustered by firm and t-statistics are reported in parentheses. All regressions include year, day of week, and industry (using SIC 2-digit industry definitions) fixed effects. All tests are two-tailed. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 9
Market effects of earnings virality (firm effects).

DV=	(1) *Abn Volume[0,1]*	(2) *Abn Spread[0,1]*	(3) *Abn Depth[0,1]*	(4) *IPE*	(5) *Jump Ratio*
Panel A: Often versus Rare Virality					
Viral Earnings (Often)	**0.013*****	**0.026*****	**−0.032*****	**−0.023**	**−0.041*****
	(3.93)	**(8.33)**	**(-3.65)**	**(-0.86)**	**(-2.98)**
Viral Earnings (Rare)	**0.047*****	**0.013*****	**−0.028*****	**−0.087*****	**−0.059*****
	(17.87)	**(6.89)**	**(-4.14)**	**(-3.74)**	**(-5.13)**
Controls	Yes	Yes	Yes	Yes	Yes
N	82,232	82,146	82,232	60,288	60,288
Adj. R^2	0.330	0.125	0.130	0.130	0.071
Within Regression F-Tests	Diff	Diff	Diff	Diff	Diff
(Often) = (Rare)	**−0.034*****	**0.013*****	**−0.004**	**0.064****	**0.018**
Panel B: Dropping Frequently Viral Firms					
Viral Earnings	**0.042*****	**0.017*****	**−0.033*****	**−0.093*****	**−0.067*****
	(16.42)	**(8.82)**	**(-5.16)**	**(-4.16)**	**(-6.00)**
Controls	Yes	Yes	Yes	Yes	Yes
N	79,227	79,141	79,227	58,261	58,261
Adj. R^2	0.317	0.127	0.130	0.129	0.071
Panel C: Firm Fixed Effects					
Viral Earnings	**0.036*****	**0.010*****	**−0.004**	**−0.016**	**−0.038*****
	(20.95)	**(5.20)**	**(-0.59)**	**(-0.63)**	**(-3.01)**
Controls	Yes	Yes	Yes	Yes	Yes
N	82,232	82,146	82,232	60,288	60,288
Adj. R^2	0.521	0.164	0.185	0.208	0.138

This table presents results of market tests, examining the effect of virality separately for those firms that go viral often (above the median within the sample of viral firms) versus rarely (below median) in Panel A, dropping firms that go viral in more than 30 quarters in Panel B, and including firm fixed effects in Panel C. Variable definitions are provided in Appendix A. Standard errors are clustered by firm and t-statistics are reported in parentheses. All regressions include year, day of week, and industry (using SIC 2-digit industry definitions) fixed effects. All tests are two-tailed. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Jump Ratio. This evidence provides additional support for our primary inferences, namely that earnings virality is negatively associated with market outcomes.

In the second omitted variables test, we seek to quantify the potential impact of omitted variables. In particular, we examine the potential impact of unobserved confounding variables using the approach developed in Frank (2000) and recommended in Larcker and Rusticus (2010). This method derives the minimum correlations that would be necessary for an omitted variable to convert a significant result into an insignificant result. To do this, we calculate the Impact Threshold for a Confounding Variable (ITCV) that is the lowest product of the partial correlation between the confounding variable and *Abn Volume[0,1], Abn Spread[0,1], Abn Depth[0,1], IPE,* and *Jump Ratio* and the partial correlation between the confounding variable and *Viral Earnings* that would make the coefficient on *Viral Earnings* insignificant. Following Larcker and Rusticus (2010), we use the impact on *Viral Earnings* of the inclusion of each of the independent variables in the model as a benchmark. Higher values of ITCV indicate greater robustness to omitted variable concerns. In untabulated analyses, we find that for four of the five dependent variables, the ITCV for earnings virality is greater in magnitude (either more positive or more negative) than that of the independent variable with the greatest ITCV in the model. In one case, we find that the ITCV for earnings virality is equal to that of the independent variable with the greatest ITCV in the model. These findings suggest that correlated omitted variables are unlikely to drive our collective results, as the omitted confounding variable would need to have a larger impact than that of any of the 24 other independent variables we include as controls in order to make the observed effect statistically insignificant. However, while this test indicates a lower likelihood that our models omit an important correlated variable, we acknowledge that such an omitted confounding variable may still may exist. Nevertheless, when combined with our other main and additional tests, these findings help narrow the possibility that an omitted variable is significantly influencing the findings.

6.2. Cross-sectional tests

A key assumption in our study is that earnings virality is a distinct construct from traditional media coverage or general investor attention due to its unique features. In this section, we present the results of a set of cross-sectional tests that examine this assumption by isolating variation that exists *within social media*, but not within traditional media coverage or general investor attention. Further, these analyses provide additional insights into the nature and content of the viral social media posts that have the greatest influence on investor activity, liquidity, and speed of price formation.

We focus on three categories that capture variation in social media attributes: tone, content, and uniqueness. For tweet tone, we examine absolute tone, negative words, and positive words. For tweet content, we examine the use of financial words and the inclusion of hyperlinks. Finally, we examine two measures related to uniqueness, including the number of retweets and tweet intensity (tweets per user).[29] We scale all of the measures, except tweet intensity, by total tweets for the announcement, so that they capture the proportion of tweets related to attributes of tone, content, and uniqueness. *Tweet Intensity* is the ratio of total tweets to total users that issued a tweet during the earnings announcement window.[30]

We examine these cross-sectional factors by partitioning each factor into separate high versus low indicator variables based on whether the value is above (*HIGH TWEET ATTRIBUTE*) versus below (*LOW TWEET ATTRIBUTE*) the annual median value. We then interact these indicators with *Viral Earnings*. We specify the model as follows:

$$\text{MARKET VARIABLE}_{it} = \beta_{FE} + \beta_1 (\text{VIRAL EARNINGS}_{it} \text{ x HIGH TWEET ATTRIBUTE}) + \beta_2 (\text{VIRAL EARNINGS}_{it} \text{ x LOW TWEET ATTRIBUTE}) + \beta_K \text{ CONTENT}_{it} + \beta_L \text{ RECOGNITION} + \beta_M \text{ RETAIL}_{it} + \beta_N\, CONTROLS_{it} + e \quad (5)$$

Where *MARKET VARIABLE* is one of five dependent variables: *Abn Volume[0,1]*, *Abn Spread[0,1]*, *Abn Depth[0,1]*, *IPE*, and *Jump Ratio,* with all variables as previously defined. Significant differences between β_1 and β_2 are based on variation within social media, which provides support for our assumption that social media virality is driving our results rather than general investor attention or business press coverage.

We present the model (5) estimation results in Table 7, where we present the results for factors related to tone in Panel A, for content in Panel B, and for uniqueness in Panel C. For parsimony, we only present β_1, β_2, and the test for differences in the table—we suppress the coefficients on control variables and model statistics. In Table 7, Panel A, we find that the negative market effects we observed in our primary tests (i.e., lower liquidity and slower price formation) are stronger when the average absolute tone of the tweets is more extreme and particularly when the tone is negative. We also find that the negative liquidity effects are stronger when the tone of the tweets is positive. These results provide fairly consistent evidence that the average tone of the tweets themselves can impact the market outcomes around viral earnings announcements.

In Table 7, Panel B, we examine the average content of the tweets. We find that the pricing effects associated with viral earnings are less negative when the tweets use more financial words or when they include hyperlinks to other content. This finding suggests that tweets that provide more substance and links to other potentially useful information are less likely to result in detrimental market effects related to earnings virality.

Finally, in Table 7, Panel C, we find mixed evidence of significant differences related to the retweet percentage. However, with respect to tweet intensity, we find that the negative liquidity and speed of price formation effects are more negative when each user issues more tweets during the viral dissemination. This result suggests that viral earnings announcements with a broader user base have a less negative association with price efficiency.

In sum, our cross-sectional results indicate that the negative market effects associated with earnings virality are related to the tone of the tweets, rather than to substantive content being provided. These results are generally consistent across multiple measures of price efficiency. We interpret these results as demonstrating that variation within social media is related to the detrimental effects of earnings virality on price efficiency.

6.3. Alternative measures of earnings virality

In this section, we examine the robustness of our price efficiency results to the two alternative measures of earnings virality discussed previously: *Viral Earnings (raw)* and *Viral Earnings (< 3 h)*. As noted previously, these capture more extreme cases of virality, with only 4.4 and 2.8 percent, respectively, of earnings announcements identified as going viral. We re-perform our market tests (H2a and H2b) using these two alternative measures. For parsimony, we focus on the following market variables, *Abn Volume[0,1]*, *Abn Spread[0,1]*, *Abn Depth[0,1]*, *IPE*, and *Jump Ratio*, with all control variables included as in Tables 4 and 5, but the coefficients suppressed for brevity.

We present the results of these robustness tests in Table 8, with *Viral Earnings (raw)* in Panel A and *Viral Earnings (< 3 h)* in Panel B. In general, we find that nearly all inferences continue to hold using these more restrictive measures of earnings virality. Specifically, we find consistent results indicating that earnings virality is negatively associated with market quality in eight of ten estimations. Thus, results using alternative measures of earnings virality generally corroborate the results of our primary analyses.

[29] We do not examine cross-sectional variation in tweet length because, as presented in Fig. 3, little variation exists in this attribute due to Twitter's character limit.

[30] These tests isolate variation within social media (as opposed to other mediums), which we believe helps narrow the possibility that explanations related to other mediums explain our results. However, it is important to acknowledge that the potential for endogeneity is not entirely eliminated because tweet content can still be influenced by both firm- and event-related characteristics.

6.4. Firm effects

In our final set of tests, we examine the potential influence of firm effects on earnings virality. As discussed previously, we argue that while virality is typically triggered by an event, it can also be influenced by characteristics of the firm. That is, there are certain types of companies whose events are more likely to go viral because of who they are, in addition to what they announce. In other words, the market effects of earnings virality effect are likely a combination of entity- and event-level factors.

In this section, we conduct several tests to investigate whether the economic consequences associated with earnings virality are driven by firm effects only or the combination of firm- and event-level effects. First, we partition our sample of earnings announcements that go viral into two groups—those that often go viral and those that rarely go viral—using a median split on the number of times the firm goes viral in the sample. We then construct two corresponding indicator variables labeled *Earnings Virality (Often)* and *Earnings Virality (Rare)* and include them in our market models.[31] Second, we drop from the sample a set of firms that frequently go viral, which we define as going viral in 30 or more of the 40 quarters in our sample period (or more than 75 percent of the time). If we observe significant market effects for the subsample of firms that rarely go viral, then it provides evidence that the market effects associated with earnings virality are not driven by entity-level factors only. In other words, these tests are designed to address concerns that our results are being driven by the most visible firms of the market, such as Apple or Tesla.

We present the results in Panel A and B of Table 9. In Panel A, we find that the detrimental market effects associated with earnings virality are observed for both firms that often and rarely go viral. We further find that the results are more consistent in the sample of firms that rarely go viral (where we observe all five variables are significant in the expected direction).[32] We find similar evidence in Panel B where we observe significant market effects associated with earnings virality using only the subset of firms that do not frequently go viral. Because the results hold across all subsets of firms, they help ease concerns that market outcomes related to earnings virality are merely reflecting firm effects.

In a final test, we examine whether the results of our market quality tests hold after including firm fixed effects to examine whether the results hold when only within-firm variation is exploited. We present the results in Table 9, Panel C. We find that our volume result is unaffected by the inclusion of firm fixed effects. We also continue to find some evidence of negative liquidity effects and slower price formation effects (in two of the four models).

Taken together, the results presented in Table 9 provide support for the notion that the market effects associated with earnings virality are not merely a function of firm characteristics only, as we also observe the effects for firms that rarely go viral. That is, the link between earnings virality and market outcomes is likely related to both firm- and event-level factors.

7. Conclusion

We examine the determinants and market consequences associated with earnings announcements that go viral on social media. Social media is rapidly becoming an important part of firms' information environments, and it is therefore important to understand its effects on capital markets. Prior research has examined social media content produced by firms (e.g., Blankespoor et al., 2014) or by other intermediaries (Chen et al., 2014; Drake et al., 2017). However, more work needs to be done to better understand the effects of extreme social media dissemination of financial information.

We use a comprehensive panel of quarterly earnings announcements and historical Twitter data to examine the determinants, trading patterns, and market outcomes associated with earnings virality. Viral earnings announcements are different than non-viral earnings announcements along several dimensions. On average, the tweets of viral earnings announcements generally use more extreme tone. They also use fewer financial words and are less likely to contain links to other content. Finally, viral earnings announcements are more likely to be re-tweeted and have more tweets per user, suggesting a lack of original content. These findings indicate that, on average, the nature of tweets associated with viral earnings announcements is more extreme and less substantive.

Regarding the determinants of earnings virality, we find that earnings virality is associated with proxies for revenue surprises (but not earnings surprises), firm recognition, and retail investor ownership. That is, attributes related to the content of the information, the recognizability of the firm itself, and its retail ownership base are positively related to the likelihood a given earnings announcement goes viral.

With respect to investor activity, we find that earnings virality is associated with increased trading volume and volatility, and this increased trading activity appears to be driven in part by retail investors. We further find that earnings virality is associated with negative market outcomes as reflected in lower trading liquidity and slower price formation. Finally, we find

[31] These measures partition the sample of earnings announcements at the median; there are 3410 corresponding earnings announcements are in the often-viral group and 3297 corresponding earnings announcements are in the rarely-viral group. Importantly, there are 99 firms in the frequently-viral group and 575 firms in the rarely-viral group, consistent with the notion that some firms have earnings announcements that nearly always go viral. We further find that the earnings announcements of firms in the often-viral group experience viral earnings announcements the vast majority of the time; 89.7% of the earnings announcements for these firms go viral.

[32] An F-test confirms that the difference in coefficient magnitude is statistically significant using three of the five market outcomes; it is only insignificant using *Abn Depth[0,1]* and the *Jump Ratio*.

that these detrimental market effects of earnings virality are stronger when the social media chatter is more extreme in tone and less substantive.

Altogether, our findings stand in stark contrast to the conclusion in some prior studies that increased dissemination from traditional media positively influences market outcomes (Fang and Peress 2009; Bushee et al., 2010; Drake et al., 2014). In other words, whether extreme media attention has a positive or negative effect on markets depends on the type of media. Increased attention from the legacy traditional media seems to help price efficiency, but viral social media can sometimes hurt it.

Appendix A

Variable Definitions

Variable	Definition
#Analysts	the natural log of the number of analysts following the firm
# Firms Announce	the natural log of the total number of firms announcing earnings that day
#feeds	the number of followers of Twitter users that tweeted about the firm over the two-day window beginning on the earnings announcement date
#tweets	the number of original tweets and retweets that contain the firm's cashtag over the two-day window beginning on the earnings announcement date
Abn Bloomberg Activity[0,1]	the number of news article searches during the earnings announcement window (days 0 and + 1). Bloomberg calculates total activity each 8-h period relative to the trailing 30 days. A score of 3 (4) is assigned if the recent attention exceeds 94% (96%) of the trailing 30 days. Bloomberg aggregates to the daily level based on the maximum value observed during the day. We transform Bloomberg's 0, 1, 2, 3, and 4 scores to continuous values using the conditional means of truncated normal distribution. Under the normal distributional assumption, the corresponding values are −0.350, 1.045, 1.409, 1.647, and 2.154.
Abn Depth[0,1]	the log of the average intraday bid and offer depth obtained from TAQ averaged over the two-day period beginning on the earnings announcement date, minus the same average over the prior month
Abn Retail Volume[0,1]	the sum of daily abnormal retail trading volume over the two-day period beginning on the earnings announcement date, where abnormal retail volume is defined as retail trading volume minus average retail trading volume over the prior month
Abn Ret[0,1]	the size, BTM, and momentum adjusted buy-and-hold return over the two-day period beginning on the earnings announcement date
Abn Ret[-30,-1]	the size, BTM, and momentum adjusted buy-and-hold return over the thirty-day period the day prior to the earnings announcement date
Abn Robinhood Holdings[0,1]	the average number of Robinhood users holding the stock over the two-day period beginning on the earnings announcement date, minus the same average over the prior month
Abn Spreads[0,1]	the intraday average percent effective spread obtained from TAQ and averaged over the two-day period beginning on the earnings announcement date, minus the same average over the prior month, multiplied by 100
Abn Volatility[0,1]	the firm's quote-based intraday price volatility obtained from TAQ and averaged over the two-day period beginning on the earnings announcement date, minus the same average over the prior month, multiplied by 1 million
Abn Volume[0,1]	the sum of daily abnormal trading volume over the two-day period beginning on the earnings announcement date, where abnormal volume is defined as volume minus average trading volume over the prior month
Abs Earn Surprise	the absolute value of *Earn Surp*
Abs Rev Surprise	the absolute value of *Rev Surp*
After Hours	an indicator variable equal to one if the firm announced earnings after trading hours, and zero otherwise.
BtoC	an indicator variable equal to one if the firm is in one of the following Fama-French 48 industry classifications: 1, 2, 3, 7, 9, 11, 13, 18, 23, 27, 28, 29, 30, 32, 33, 34, 35, 41, 42, 43, 44, 45, and 46, and zero otherwise
Earn Surp	the firm's realized earnings minus the most recent median consensus analyst forecast, scaled by price
Feeds[t,T]	the natural log of one plus the number of followers of Twitter users that tweeted about the firm over a given window relative to the earnings announcement date
Firm Age	the natural log of the number of years the firm is in the Compustat database
Firm Size	market value of equity
High Rev Growth	an indicator variable that equals one for firms in the top quintile of quarterly revenue growth, and to zero otherwise
Inst Count	the natural log of one plus the number of institutions holding shares in the firm's stock
IPE	intraperiod efficiency measure of speed with which earnings information is incorporated into price, measured over the five trading days following the announcement of earnings. Calculated as: $\frac{1}{2}\sum_{t=0}^{5}(Abn\_Ret_{t-1} + Abn\_Ret_t)/Abn\_Ret_5 = \sum_{t=0}^{4}(Abn\_Ret_t/Abn\_Ret_5) + 0.5$, adjusted for overreactions following Blankespoor et al. (2018)
Jump Ratio	*Abn Ret [0,1]* divided by the firm's 1-week abnormal return beginning on the earnings announcement date
Low BTM	an indicator variable that equals one for firms in the bottom quintile of the ratio of book equity to market equity and zero otherwise.
Neg Earn Surp	an indicator variable equal to one if *Earn Surp* is less than zero, and zero otherwise
Neg Rev Surp	an indicator variable equal to one if *Rev Surp* is less than zero, and zero otherwise
Press Cov[t,T]	the natural log of one plus the number of traditional media articles written about the firm in the Dow Jones Newswires and Wall Street Journal over a given window relative to the earnings announcement date
Price Tweets%	

(*continued*)

Variable	Definition
	ratio of tweets that include a stock price word (stock, share, price, pop, rally, hold, short, squeeze, or bounce) to total tweets. % indicates that the measure is a percentage of total tweets.
Retail Own	one minus the percentage of the firm's shares held by institutions
Retweet%	the percentage of *#tweets* that are retweets. % indicates that the measure is a percentage of total tweets.
Rev Surp	the firm's realized revenues minus the most recent median analyst forecast, scaled by price and divided by 100
Tech	an indicator variable equal to one if the firm is in one of the following Fama-French 12 industry classifications: Chips, Comps, or Telcm, and zero otherwise
Time Trend	the number of quarters from the first earnings announcement in the sample period until the quarter of the firm's earnings announcement
Tweet Abs Tone	the sum of the positive and negative tone scores provided by Crimson Hexagon, averaged across all tweets that contain the firm's cashtag over the two-day period beginning on the earnings announcement date
Tweet Earn%	the percentage of #tweets that contain at least one accounting word (sales, revenue, earnings, profit, loss, income, and or eps). % indicates that the measure is a percentage of total tweets.
Tweet Intensity	*#tweets* divided by the total number of Twitter users that issued a tweet
Tweet Link%	the percentage of *#tweets* that contain a hyperlink. % indicates that the measure is a percentage of total tweets.
Tweet Negative	the “negative” sentiment score provided by Crimson Hexagon, averaged across all tweets that contain the firm's cashtag over the two-day period beginning on the earnings announcement date
Tweet Positive	the “positive” sentiment score provided by Crimson Hexagon, averaged across all tweets that contain the firm's cashtag over the two-day period beginning on the earnings announcement date
Tweet Words%	the total number of words contained in *#tweets*, divided by *#tweets*. % indicates that the measure is a percentage of total tweets.
Tweets[t,T]	the natural log of one plus the number of tweets that contain the firm's cashtag over a given window relative to the earnings announcement date
Viral Earnings	an indicator variable equal to one if the earnings announcement is in both the highest decile of yearly ranked tweets and the highest decile of yearly ranked Twitter feeds, and zero otherwise
Viral Earnings (< 3 h)	an alternative measure of earnings virality; an indicator variable equal to one if the earnings announcement not only goes viral, but does so within 3 h of the earnings announcement, and zero otherwise
Viral Earnings (Often)	an alternative measure of earnings virality; an indicator variable equal to one if the earnings announcement goes viral, and the firm's earnings go viral more often than the sample median, and zero otherwise
Viral Earnings (Rare)	an alternative measure of earnings virality; an indicator variable equal to one if the earnings announcement goes viral, and the firm's earnings go viral less often than the sample median, and zero otherwise
Viral Earnings (raw)	an alternative measure of earnings virality; an indicator variable equal to one if the earnings announcement has both more than 200 tweets and more than 2 million Twitter feeds, and zero otherwise

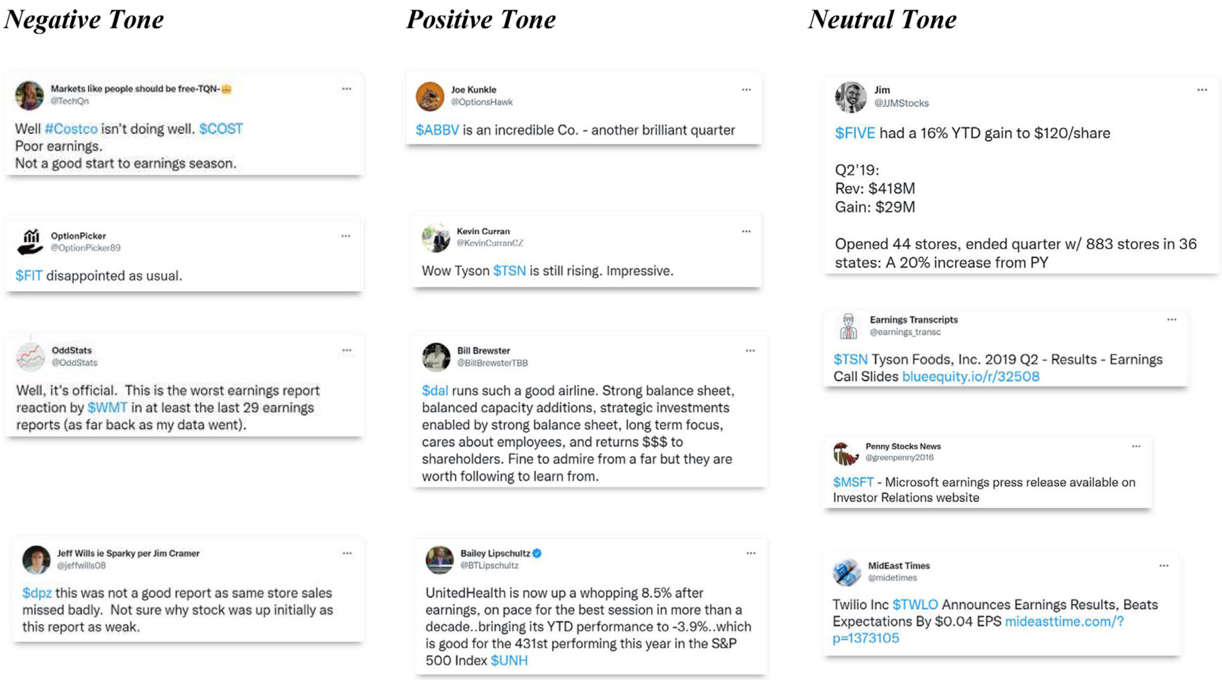
Appendix B

Tweet Word Lists

Viral	Non-Viral	Negative	Positive	Neutral
earnings	earnings	short	good	earnings
eps	eps	loss	beat	results
revenue	results	low	great	eps
stock	beats	miss	strong	beats
sales	reports	weak	growth	consensus
beats	stock	bad	nice	announces
results	revenue	bearish	bullish	revenue
new	consensus	decline	well	quarterly
call	new	drop	wow	estimize
reports	quarter	downgraded	positive	reports
trading	misses	disappointing	love	financial
shares	guidance	negative	better	stock
quarter	quarterly	wrong	pretty	estimates
market	financial	lost	amazing	call
today	announces	downside	happy	misses
read	call	fall/falls	congrats	vs
beat	alert	poor	hope	releases
billion	trading	dip	excited	expectations
guidance	estimates	worse	upside	posts
profit	markets	ugly	bull	via

This appendix provides the top twenty most common words found in viral and non-viral tweets in the first two columns, followed by the top twenty most common words in viral negative, positive, and neutral tweets in the third, fourth, and fifth columns, respectively.

Appendix C. Representative Tweets of Viral Earnings Announcements



This appendix provides representative examples of tweets released during viral earnings announcements. We provide examples for each of three categories: *Negative Tone*, *Positive Tone*, and *Neutral Tone*. Each of these is based on tone scores as categorized by our data provider, Crimson Hexagon.

References

Ang, J.S., Hsu, C., Tang, D., Wu, C., 2020. The role of social media in corporate governance. The Accounting Review, forthcoming.
Angel, J.J., Harris, L.E., Spatt, C.S., 2015. Equity trading in the 21st century: an update. Q. J. For. 5 (1), 1550002.
Armstrong, C., Kepler, J.D., Samuels, D., Taylor, D., 2021. The evolution of empirical methods in accounting research and the growth of quasi-experiments. Working paper.
Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. J. Finance 61 (4), 1645–1680.
Bakshy, E., Hofman, J.M., Mason, W.A., Watts, D.J., 2011, February. Everyone's an influencer: quantifying influence on twitter. In: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, pp. 65–74.
Banerjee, S., Davis, J., Gondhi, N., 2018. When transparency improves, must prices reflect fundamentals better? Rev. Financ. Stud. 31 (6), 2377–2414.
Barber, B.M., Odean, T., 2008. All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Rev. Financ. Stud. 21 (2), 785–818.
Barber, B.M., Odean, T., 2013. The behavior of individual investors. In: In Handb. Econ. Finance, 2. Elsevier, pp. 1533–1570.
Bartov, E., Faurel, L., Mohanram, P.S., 2018. Can Twitter help predict firm-level earnings and stock returns? Account. Rev. 93 (3), 25–57.
Ben-Rephael, A., Da, Z., Israelsen, R.D., 2017. It depends on where you search: institutional investor attention and underreaction to news. Rev. Financ. Stud. 30 (9), 3009–3047.
Berger, J., Milkman, K., 2010. Social Transmission, Emotion, and the Virality of Online Content, vol. 106. Wharton research paper, pp. 1–52.
Berger, J., Milkman, K.L., 2012. What makes online content viral? J. Market. Res. 49 (2), 192–205.
Blankespoor, E., Miller, G.S., White, H.D., 2014. The role of dissemination in market liquidity: evidence from firms' use of Twitter. Account. Rev. 89 (1), 79–112.
Blankespoor, E., deHaan, E., Zhu, C., 2018. Capital market effects of media synthesis and dissemination: evidence from robo-journalism. Rev. Account. Stud. 23 (1), 1–36.
Blankespoor, E., 2018. Firm communication and investor response: a framework and discussion integrating social media. Account. Org. Soc. 68, 80–87.
Blankespoor, E., deHaan, E., Marinovic, I., 2020. Disclosure processing costs, investors' information choice, and equity market outcomes: a review. J. Account. Econ. 70 (2–3), 101344.
Bloomberg, 2013, April 4. Bloomberg integrates live twitter feeds with financial platform. https://www.bloomberg.com/company/press/bloomberg-integrates-live-twitter-feeds-with-financial-platform/.
Bloomberg, 2018, July 18. Bloomberg launches a twitter feed optimized for trading. https://www.bloomberg.com/company/press/bloomberg-launches-twitter-feed-optimized-trading/.
Bloomfield, R.J., 2002. The "incomplete revelation hypothesis" and financial reporting. Account. Horiz. 16 (3), 233–243.
Boehmer, E., Jones, C.M., Zhang, X., Zhang, X., 2021. Tracking retail investor activity. J. Finance 76 (5), 2249–2305.
Bollen, J., Mao, H., Zeng, X., 2011. Twitter mood predicts the stock market. J Comput Sci 2 (1), 1–8.
Bonsall IV, S.B., Green, J., Muller III, K.A., 2020. Market uncertainty and the importance of media coverage at earnings announcements. J. Account. Econ. 69 (1), 101264.

Bushee, B.J., Core, J.E., Guay, W., et al., 2010. The role of the business press as an information intermediary. J. Account. Res. 48, 1–19.
Bushee, B., Cedergren, M., Michels, J., 2020. Does the media help or hurt retail investors during the IPO quiet period? J. Account. Econ. 69 (1), 101261.
Chen, H., De, P., Hu, Y.J., Hwang, B.H., 2014. Wisdom of crowds: the value of stock opinions transmitted through social media. Rev. Financ. Stud. 27 (5), 1367–1403.
Copeland, T.E., Galai, D., 1983. Information effects on the bid-ask spread. J. Finance 38 (5), 1457–1469.
Curtis, A., Richardson, V.J., Schmardebeck, R., 2014. Investor attention and the pricing of earnings news. In: Handbook of Sentiment Analysis in Finance, Forthcoming. http://dx.doi.org/10.2139/ssrn.2467243.
Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. J. Finance 52 (3), 1035–1058.
deHaan, March 30, 2021. Ed, using and interpreting fixed effects models. Available at SSRN: https://ssrn.com/abstract=3699777. https://doi.org/10.2139/ssrn.3699777.
DeMarzo, P.M., Vayanos, D., Zwiebel, J., 2003. Persuasion bias, social influence, and unidimensional opinions. Q. J. Econ. 118 (3), 909–968.
Diamond, D.W., Verrecchia, R.E., 1991. Disclosure, liquidity, and the cost of capital. J. Finance 46 (4), 1325–1359.
Drake, M.S., Guest, N.M., Twedt, B.J., 2014. The media and mispricing: the role of the business press in the pricing of accounting information. Account. Rev. 89 (5), 1673–1701.
Drake, M.S., Thornock, J.R., Twedt, B.J., 2017. The internet as an information intermediary. Rev. Account. Stud. 22 (2), 543–576.
Drake, M.S., Moon, J., Twedt, B.J., Warren, J., 2020. Are social media analysts disrupting the relevance of sell-side analyst research? *Available at*: *SSRN 3456801*.
Dugast, J., Foucault, T., 2018. Data abundance and asset price informativeness. J. Financ. Econ. 130 (2), 367–391.
Fang, L., Peress, J., 2009. Media coverage and the cross-section of stock returns. J. Finance 64 (5), 2023–2052.
Farrell, M., Green, T.C., Jame, R., Markov, S., 2020. The democratization of investment research and the informativeness of retail investor trading. *Available at*: *SSRN 3222841*.
Frank, K.A., 2000. Impact of a confounding variable on a regression coefficient. Socio. Methods Res. 29 (2), 147–194.
Glaeser, S., Guay, W.R., 2017. Identification and generalizability in accounting research: a discussion of christensen, floyd, liu, and maffett (2017). J. Account. Econ. 64 (2–3), 305–312.
Glosten, L.R., Milgrom, P.R., 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. J. Financ. Econ. 14 (1), 71–100.
Golub, B., Jackson, M.O., 2010. Naive learning in social networks and the wisdom of crowds. Am. Econ. J. Microecon. 2 (1), 112–149.
Golub, B., Sadler, E., 2017. Learning in social networks. *Available at*: *SSRN 2919146*.
Gomez, E., Heflin, F., Moon, J., Warren, J., 2020. Can Financial Analysis on Social Media Help Level the Playing Field Among Investors? Evidence from Seeking Alpha. Working paper, Georgia Tech.
Greene, W., 2004. The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. Econom. J. 7 (1), 98–119.
Grossman, S.J., Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. Am. Econ. Rev. 70 (3), 393–408.
Han, B., Yang, L., 2013. Social networks, information acquisition, and asset prices. Manag. Sci. 59 (6), 1444–1457.
Haslag, Peter H., Ringgenberg, Matthew C., September 30, 2021. The demise of the NYSE and NASDAQ: market quality in the Age of market fragmentation. In: Fourth Annual Conference on Financial Market Regulation. Available at: SSRN: https://ssrn.com/abstract=2591715.
Hayn, C., 1995. The information content of losses. J. Account. Econ. 20 (2), 125–153.
Heckman, J.J., Singer, B., 2017. Abducting economics. Am. Econ. Rev. 107 (5), 298–302.
Hendershott, T., Riordan, R., 2011. High Frequency Trading and Price Discovery, vol. 3. Manuscript, University of California, Berkeley.
Hirshleifer, D., Teoh, S.H., 2003. Limited attention, information disclosure, and financial reporting. J. Account. Econ. 36 (1–3), 337–386.
Hirshleifer, D., Lim, S.S., Teoh, S.H., 2009. Driven to distraction: extraneous events and underreaction to earnings news. J. Finance 64 (5), 2289–2325.
Hirshleifer, D., 2020. Presidential address: social transmission bias in economics and finance. J. Finance 4, 1179–1831.
Hornik, J., Satchi, R.S., Cesareo, L., Pastore, A., 2015. Information dissemination via electronic word-of-mouth: good news travels fast, bad news travels faster. Comput. Hum. Behav. 45, 273–280.
Huang, L., 2016. Moving Top Tweet Search Results from Reverse Chronological Order. Engineering. https://blog.twitter.com/engineering/en_us/topics/insights/2016/moving-top-tweet-search-results-from-reverse-chronological-order-to-relevance-order.
Huberman, G., 2001. Familiarity breeds investment. Rev. Financ. Stud. 14 (3), 659–680.
Israeli, D., Kasznik, R., Sridharan, S.A., 2020. Unexpected distractions and investor attention to corporate announcements. *Available at*: *SSRN 3057278*.
Jia, W., Redigolo, G., Shu, S., Zhao, J., 2020. Can social media distort price discovery? Evidence from merger rumors. J. Account. Econ. 70 (1), 101334.
Jiang, May 1, 2017. John (Xuefeng) and Shen, Michael, Traditional Media, Twitter and Business Scandals. *Available at*: *SSRN 2959419*.
Jung, M.J., Naughton, J.P., Tahoun, A., Wang, C., 2018. Do firms strategically disseminate? Evidence from corporate use of social media. Account. Rev. 93 (4), 225–252.
Kim, O., Verrecchia, R.E., 1994. Market liquidity and volume around earnings announcements. J. Account. Econ. 17 (1–2), 41–67.
Koumchatzky, N., Anryeyev, A., 2017, May 9. Using deep learning at scale in Twitter's timelines. https://blog.twitter.com/engineering/en_us/topics/insights/2017/using-deep-learning-at-scale-in-twitters-timelines.
Kümpel, A.S., Karnowski, V., Keyling, T., 2015. News sharing in social media: a review of current research on news sharing users, content, and networks. Social media+ society 1 (2), 2056305115610141.
Larcker, D.F., Rusticus, T.O., 2010. On the use of instrumental variables in accounting research. J. Account. Econ. 49 (3), 186–205.
Lee, C.M., Mucklow, B., Ready, M.J., 1993. Spreads, depths, and the impact of earnings information: an intraday analysis. Rev. Financ. Stud. 6 (2), 345–374.
Lee, L.F., Hutton, A.P., Shu, S., 2015. The role of social media in the capital market: evidence from consumer product recalls. J. Account. Res. 53 (2), 367–404.
Lehavy, R., Sloan, R.G., 2008. Investor recognition and stock returns. Rev. Account. Stud. 13 (2), 327–361.
McCabe, C., 2022. Day traders as 'dumb money?' the pros are now paying attention. Wall St. J. https://www.wsj.com/articles/fund-managers-pay-attention-to-retail-day-traders-11642132135.
Merton, R.C., 1987. A simple model of capital market equilibrium with incomplete information. J. Finance 42 (3), 483–510.
Miller, G.S., Skinner, D.J., 2015. The evolving disclosure landscape: how changes in technology, the media, and capital markets are affecting disclosure. J. Account. Res. 53 (2), 221–239.
Naveed, N., Gottron, T., Kunegis, J., Alhadi, A.C., 2011, June. Bad news travel fast: a content-based analysis of interestingness on twitter. In: Proceedings of the 3rd Inte*rnational Web Science Conference*, pp. 1–7.
Nekrasov, A., Teoh, S.H., Wu, S., 2020. Visuals and Attention to Earnings News on Twitter. *Available at*: *SSRN 3447484*.
Odean, T., 1999. Do investors trade too much? Am. Econ. Rev. 89 (5), 1279–1298.
Oremus, W., 2017. Twitters new order. Inside the changes that could save its business – and reshape civil discourse. Slate. March 5, 2017. Available online at: http://www.slate.com/articles/technology/cover_story/2017/03/twitter_s_timeline_algorithm_and_its_effect_on_us_explained.html.
Patel, S.A., 1993. Intraday Bid-Ask Spreads Surrounding Earnings Announcements and Information Asymmetry. Doctoral dissertation, University of Pennsylvania.
Rudat, A., Buder, J., Hesse, F.W., 2014. Audience design in Twitter: retweeting behavior between informational value and followers' interests. Comput. Hum. Behav. 35, 132–139.
Shiller, R., 1984. Stock prices and social dynamics. Brookings Pap. Econ. Activ. 457–498.

Skinner, Douglas J., March 31, 1993. Bid-ask spreads around earnings announcements: evidence from the NASDAQ national market system. Available at: SSRN: https://ssrn.com/abstract=2143649.
Tang, V.W., 2018. Wisdom of crowds: cross-sectional variation in the informativeness of third-party-generated product information on Twitter. J. Account. Res. 56 (3), 989–1034.
Twedt, B., 2016. Spreading the word: price discovery and newswire dissemination of management earnings guidance. Account. Rev. 91 (1), 317–346.
Verrecchia, R.E., 2001. Essays on disclosure. J. Account. Econom. 32 (1–3), 97–180.
Zhou, M., Lei, L., Wang, J., Fan, W., Wang, A.G., 2015. Social media adoption and corporate disclosure. J. Inf. Syst. 29 (2), 23–50.