Title
Essays on Behavioral Finance and Social Media Technology

Author
Wang, Melissa

Publication Date
2023

Peer reviewed|Thesis/dissertation

Essays on Behavioral Finance and Social Media Technology

by

Melissa Wang

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Anastassia Fedyk, Chair
Professor Dmitry Livdan
Professor David Sraer

Spring 2023

Essays on Behavioral Finance and Social Media Technology

Abstract

Essays on Behavioral Finance and Social Media Technology

by

Melissa Wang

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Anastassia Fedyk, Chair

How does social media change the information landscape and what are the effects on various stakeholders? This dissertation contains two essays that study this question empirically by investigating how social media affects investors, firms, influential individuals, and the stock market.

In the first chapter of the dissertation, I study the implications of social media comment visibility on retail investor consumption and production behavior. I exploit a shock in the comment display sorting algorithm in 2018 on Daily Discussion posts on Reddit's WallStreetBets. By comparing the market reactions to comments before and after the change, I find that going from a more curated but less timely display (Best Regime) to a less filtered but more timely display (New Regime) increases absolute abnormal returns and abnormal retail trading volumes in the five minutes after comment publication. Following the initial five minutes, Best Regime comments see a stronger price drift while New Regime comments see a slight return reversal. These results are driven primarily by firms with small market capitalization and high Robinhood user ownership. In addition, I find that changes in comment display also affect comment production timing and volume.

In the second chapter of the dissertation, I investigate the social media posting behavior of influential individuals, namely CEOs at S&P 1500 firms, in the context of quarterly earnings announcements and provide evidence of CEO strategic behavior. Quarterly earnings announcements followed by an earnings tweet from the CEO correspond with a 1.5-2.6% higher 3-day industry-adjusted announcement return conditional on the same level of earnings surprise and are not followed by return reversals. An intraday event study around earnings tweets shows that CEOs time their tweets to 'take credit' for quarters with positive stock price reactions above and beyond the earnings news. I attempt to tie this behavior pattern to CEO career management concerns and find evidence that suggests that CEOs can reduce their likelihood of being fired conditional on performance measures by leveraging social media.

This dissertation is dedicated to my parents.

Contents

List of Figures

List of Tables

Acknowledgments

Going through a PhD is no joke. It is both a humbling experience as well as an immensely rewarding one. I would not be standing here at the finish line without the unyielding support and guidance from my committee chair, Anastassia Fedyk. I want to thank her for her unwavering support through all the highs and lows over the past half decade.

I also benefited greatly from the support and guidance of my other committee members: Dmitry Livdan and David Sraer. These individuals have all played a great role in my intellectual growth and have graciously given their time to support my studies and career path.

So many other people deserve recognition for their helpful comments and suggestions on my research. I want to thank Matteo Benetton, Terry Hendershott, Ulrike Malmendier, Gustavo Manso, Peter Maxted, Adair Morse, Terry Odean, Christine Parlour, Annette Vissing-Jorgensen, and the rest of the Haas Finance faculty.

But most importantly, I want to thank my friends and family outside of Berkeley. I would not be crossing this finish line without their undying love and support. I am eternally grateful to have these guys in my life - without them, these last 6 years would have been the most isolating years of my life. I want to thank my parents for always pushing me to be the best version of myself. I want to thank my sister for being my biggest cheerleader and supporter through all my self-doubts and concerns. I want to thank all the Sam's (iykyk) in my life for being a steadying anchor through my emotional roller coasters. I want to thank all my girlies (you know who you are) for simply loving me and always wanting the best for me. I want to thank Maisie - though she will never truly understand, she is the best thing that has happened to me in the last 5 years and is my very best furry friend. And lastly, I want to thank Blaine, who has been with me through thick and thin over the last 12 years. You are my very best friend and my #1 fan, and you make me feel so loved. I cannot wait for the next chapter of our lives, and it's beginning in 3, 2, 1...

Chapter 1

Comment Display Matters: Evidence from Reddit's WallStreetBets

1.1 Introduction

Social media is a growing platform for financial news consumption and production, with over half of U.S. adults in 2022 getting news from social media.[1] Social media differs from traditional media in that there is more feedback and participation from the same population that produces and consumes information. As a result, it enables individuals to go on such platforms and coordinate market participation in a meaningful way. Therefore, *how* these platforms display incoming posts and comments plays an important role in information consumption and production, and can have a meaningful impact on investor behavior—particularly retail investors—and price responses. However, it is challenging to find a clean setting in which to identify the effect of information visibility on market behavior in a social media setting.

One particularly influential social media platform for trading is Reddit's WallStreetBets, which is an investment discussion forum that came into mainstream prominence in early 2021 due to retail investors using the platform to coordinate market participation that contributed to the Gamestop short squeeze and has been tied to retail investor trading.[2] I exploit a unique change in comment display on WallStreetBets in 2018 to estimate the effect of comment visibility on retail investor behavior and the dynamics of price discovery. Specifically, I take advantage of a shock to readers in the display of incoming comments within the Daily Discussion posts on WallStreetBets.[3] A new Daily Discussion post is created at the beginning

[1]https://www.pewresearch.org/journalism/fact-sheet/social-media-and-news-fact-sheet/

[2]V. Fedyk, 2022 and Eaton et al., 2022 find evidence that retail investors trading on the Robinhood app are more likely to purchase stocks that were recently discussed on WallStreetBets.

[3]Prior empirical strategies to estimate the causal impact of media employ the exogenous variation of newspaper strikes (Peress, 2008), news arrival through weather-related disruptions (Engelberg and Parsons, 2011), targeted earnings announcement news to a subset of Yahoo Finance users (Lawrence et al., 2018), Robinhood outages (Eaton et al., 2022), and front-page news article placement on the Bloomberg terminal

of every trading day and is pinned to the top of the WallStreetBets homepage. Given that Daily Discussion posts are featured prominently on the WallStreetBets homepage and dedicated to intraday trading discussion, the comment activity in these particular posts is a natural setting in which to examine how comment display might influence retail investors.

Prior to 7/26/2018, comments in Daily Discussion posts were default sorted by Best, which sorts comments by the lower bound of the Wilson score confidence interval, which depends on the fraction of upvotes and the number of total votes a comment receives. After a comment is published, it gets placed at the bottom of the page for a few minutes until the Best sorting algorithm gets enough voting information to position the comment. This means that comments that receive prominent positioning experience a slight delay before getting moved to the top. And once comments make it to the top, those comments often remain at or near the top for a prolonged duration. The idea behind this comment sort is to curate which comments are likely to be of some value and position them so that they have top visibility.

Starting on 7/26/2018, the default comment sort changed to New, so that at any point in time the newest comment is displayed at the top of the page. As time goes on, newer comments displace older comments at the top of the page, and the older comments continue to get pushed lower and lower. The implications behind the New sort is that every comment receives prominent visibility immediately after publication, but only for a short period of time (comments stay in a top 10 slot for only a few minutes). While the first comments users see now are the most timely ones in that given moment, there is no filtering to weed out noise.

This identification strategy yields a couple of key predictions for market reactions following comments in the two sorting regimes motivated by the literature on limited attention, information diffusion, and heterogeneous beliefs.[4] First, comments published while New was the default sort (hereafter New Regime comments) are more likely to appear at the top of the page immediately after publication compared to comments published while Best was the default sort (hereafter Best Regime comments), inducing a higher share of immediate attention. Therefore, I predict that New Regime comments are accompanied by larger abnormal retail trading and absolute abnormal returns compared to Best Regime comments immediately (within minutes) after comment publication.

Second, New Regime comments are displaced from near the top of the page after a few minutes and continue to drop down with time. Best Regime comments, on the other hand, are more likely to move up the page by then and remain there or continue to move up over time, inducing a higher share of attention over a delayed and prolonged period. Thus, I predict that Best Regime comments initially underreact and are accompanied by a stronger return continuation compared to New Regime comments. And since New Regime comments at the top of the page are not filtered, it is likely that investors initially overreact. So I

(A. Fedyk, 2022).

[4]For models of limited attention, gradual information diffusion, and heterogeneous beliefs, see, for example, Hong and Stein, 1999, Peng and Xiong, 2006, Hong and Stein, 2007, DellaVigna and Pollet, 2009, and Hirshleifer et al., 2011.

predict that New Regime comments see a partial return reversal after comments drop from the top of the page.

My empirical findings confirm these predictions. To help with identification, I restrict the time period to the 6 months surrounding the comment sorting display change (4/25/2018 to 10/25/2018) in my main specification. I find that switching from the Best Regime to the New Regime induces 10.1% higher abnormal retail trading and 6.7-basis-points higher absolute abnormal returns (a relative increase of 31.5%) within 5 minutes of comments posting after controlling for firm size, contemporaneous comments, and prior comments, as well as including industry and half hour fixed effects.

Two additional findings further support that my results are attributed, at least in part, to changes in comment visibility. First, consistent with information posted on WallStreetBets being consumed predominantly by retail investors, these results are driven by comments mentioning firms with high retail ownership using Robinhood stock ownership data. Second, I find that these results are stronger for small firms. My results are robust to a difference-in-difference approach where the treatment group is the Daily Discussion comment sample set and the comparison group is comprised of all other WallStreetBets ticker-tagged comments in posts where Best was the only default comment sort during my sample period. Put together, these results suggest that firms where retail trading matters most—smaller firms and firms with high retail ownership—are most responsive to the comment display change.

Following the initial 5-minute return, Best Regime comments are accompanied by a stronger return continuation over the subsequent 10 to 25 minutes compared to New Regime comments. This is consistent with Best Regime comments receiving increased visibility and thus attention after a few minutes compared to New Regime comments that drop from the top of the page by then. On the other hand, New Regime comments see a slight return reversal over the corresponding time period, consistent with prior studies documenting short-term price reversals[5] and prior studies showing that retail investors are prone to initial overreactions when new is prominently displayed.[6] The price paths for comments from the two sorting regimes eventually converge after approximately 30 minutes, suggesting that the average comment informativeness and quality between the two regimes did not change significantly. Once again, this result is driven by firms with high retail ownership and small market capitalization. Altogether, these results are consistent with the overall effect of the comment display change speeding up information incorporation while also swinging the pendulum from an initial underreaction to an initial slight overreaction.

I conduct several additional tests to address alternative explanations. While I interpret my results as the comment display change affecting comment visibility and subsequently changing retail trading and price formation, there are other events that might affect the relationship between comments and trading. I first conduct a placebo test to examine whether the time period after the comment display change was generally more volatile by looking

[5]See Atkins and Dyl, 1990, Ederington and Lee, 1995, Fung et al., 2000, Chordia et al., 2002, Zawadowski et al., 2006, Heston et al., 2010 and A. Fedyk, 2022.

[6]Antweiler and Frank, 2006 find overreaction and subsequent reversal to Wall Street Journal corporate news articles.

at the market response to news articles outside of social media that were unaffected by the Daily Discussion display change. I compare the market response to news articles found on Nasdaq, Dow Jones, Reuters, MarketWatch, Business Insider, and CNBC before and after the Daily Discussion display change and find no significant differences in immediate retail trading, immediate absolute returns, or return continuation. This result suggests that the display change did not coincide with a more general change in the response to news or a trend in trading activity.

I next check for pre-trends. A concern is that comments and subsequent market responses are both a reaction to some other event prior to the comments that has nothing to do with comment visibility. If that were the case, then I would expect economically similar or perhaps larger differences in retail trading and returns between regimes prior to comment publication. I rerun my immediate market response tests but change the dependent variable to capture the 5 minutes prior to comment publication. I find no difference in immediate abnormal retail trading and a small but weakly significant difference in immediate absolute abnormal returns between New Regime comments and Best Regime comments prior to comment publication. This evidence is consistent with comment visibility playing a role in subsequent market reactions.

I next consider similar subreddits with potential overlapping coverage and address the concern that my main results may be driven by comments on other forums. I compare the timing of Daily Discussion ticker coverage to comments in Reddit's Investing and Stocks subreddits (hereafter r/investing and r/stocks, respectively) and document that the tickers discussed on WallStreetBets, specifically in Daily Discussion comments, are uniquely timed. This suggests that my main results cannot be attributed to other subreddit comments.

In addition to the comment display change affecting comment visibility, I also document that the change affected comment production and discuss potential implications and whether they could explain my main results. I find that comment timing and comment volume changed significantly after the comment display change. In terms of comment timing, I find that the majority of Best Regime comments are published earlier in the day while New Regime comments are more evenly distributed across the day. This change in comment timing is consistent with the Best sorting algorithm favoring earlier comments in terms of visibility and the New sorting algorithm providing equal visibility to all comments. I further confirm this by showing that comments at the top of the page at the end of the day are more likely to come from morning comments in the Best Regime and are more likely to come from evening comments in the New Regime. This indicates that comment authors noticed the change in comment display and shifted their production behavior to potentially maximize comment visibility.

I attempt to link the change in comment timing to market behavior and find that immediate retail trading and volatility are lower in the afternoon versus the morning in the Best Regime, consistent with afternoon comments having little chance of getting seen prior to the comment display change. After the regime change, retail trading and volatility increases in the afternoon versus the morning. This suggests that the change in when comments are produced during the day changed when retail trading and volatility occurred throughout the day.

Since Best Regime comments posted later in the day are not as likely to be seen compared to morning comments, there is less of an incentive to post in the afternoon. On the other hand, all New Regime comments receive prominent positioning upon publication regardless of when they post. And since New Regime comments stay at the top of the page for only a brief period of time, authors seeking to capture more attention will want to post more frequently. Therefore, I expect to see more comments produced in the New Regime, which I document. I also confirm that this increase did not coincide with a more general increase in attention to WallStreetBets as non-Daily Discussion comment volume on WallStreetBets remained stable over the sample period.

I now explore whether readers are responding to changes in comment production between the two regimes as opposed to changes in comment visibility. One possible result of the increase in comment volume is a decrease in comment informativeness and quality, which would mean that the price response for Best Regime comments should exceed that of New Regime comments. I examine the contents of the comments using machine learning and compare the distribution of topics between regimes. I employ a topic modeling algorithm that utilizes an embedding approach well-suited for short-text documents to assign comments to a topic cluster, which I then inspect and assign a name to. I find that the majority of comments pertain to buying and selling tickers, and discussions of stock performance and news. After the comment display change, there are slightly fewer comments that discuss buying and selling, and an increase in comments that mention WallStreetBets-specific jargon such as 'mooning,' 'praying,' and 'drilling.' The increase in jargon could be indicative of an increase in speculative, noisy comments that might decrease comment quality after the comment display change, but that is not supported by the market data. I document that the price paths for Best Regime and New Regime comments converge after approximately 30 minutes and do not differ over subsequent hours. This suggests that, though the terminology in comments changed over time, the average comment quality did not change as a result of the display change.

The increase in comment volume could also be associated with an overall increase in attention to Daily Discussion comments, which would lead to more retail trading and volatility, which are the same immediate predictions according to the comment visibility channel. I distinguish between the increased attention channel and the comment visibility channel by repeating my main tests for varying time windows surrounding the display change. My results are robust to a shortened time window surrounding the display change where it is less likely that changes in comment production could affect trading behavior. It is therefore unlikely that my results are primarily driven by higher overall attention unrelated to the comment visibility change.

My paper provides two main contributions. First, I provide evidence that social media comment visibility plays a direct role in retail trading and price formation dynamics. Specifically, I show that the Daily Discussion comment sort moving from a more curated and less timely display to a less filtered and more timely display induces higher immediate retail trading and volatility, and faster market reactions. I also document a novel result highlighting the importance of comment visibility on comment production dynamics. These findings fit into the growing body of literature that underscores the importance of information display. News positioning on the Bloomberg terminal affects the speed of information incorporation for institutional investors (A. Fedyk, 2022). Prominently featured stocks on the Robinhood trading app affect Robinhood user trading behavior (B. Barber et al., 2022). Targeted earnings announcement news to a subset of Yahoo! Finance readers affect the pricing of earnings (Lawrence et al., 2018). Price information display affects investor behavioral biases like the disposition effect (Frydman and Wang, 2020; Loos et al., 2020). And prominently displaying various metrics affects investor preferences for funds (Choi et al., 2010; Kaniel and Parham, 2017; Kronlund et al., 2021). Along with this literature, my results suggest that the way that information is delivered—especially on social media platforms—plays an important role in coordination efforts, of which the benefits and harms remain an ongoing debate.

Second, my findings add to the more general body of literature debating the market consequences of social media. Several studies find evidence that certain types of social media, including WallStreetBets, provide investment value (Chen et al., 2014; Jame et al., 2016; Bartov et al., 2018; Cookson and Niessner, 2020; Eaton et al., 2022; V. Fedyk, 2022; Kogan et al., 2022). Other studies show that social media intensifies behavioral biases or spreads stale news (Heimer, 2016; Chawla et al., 2017; Bali et al., 2021; Pedersen, 2021; Cookson et al., 2022). Farrell et al., 2022 exploit the editorial delay between report submission and publication on the crowdsourced investment research platform Seeking Alpha and find that Seeking Alpha has a distinct influence on the intensity and direction of retail trading. Bradley et al., 2022 study a subset of posts on WallStreetBets labeled "Due Diligence" and find that those posts positively predict one-month ahead returns at the beginning of their sample period but correspond to returns that subsequently reverse in recent years. I add to this body of literature by confirming that Daily Discussion posts in WallStreetBets predict retail trading at the daily-level. And I show that comment display on WallStreetBets can influence the intensity and direction of attention-related behavioral biases (overreaction and underreaction).

The rest of this chapter is organized as follows. Section 1.2 provides further background on WallStreetBets and lays out my empirical strategy. Section 1.3 details the data collection process. Section 1.4 presents my empirical findings. Section 1.5 concludes the paper.

1.2 WallStreetBets and the Comment Display Change

To capture the causal effect of comment visibility, I exploit an exogenous change in comment sorting in Daily Discussion posts on WallStreetBets.

WallStreetBets Background

Reddit is one of the most visited social media platforms.[7] Registered users can submit content in the form of posts or comments, which are then voted up or down by other users. The Reddit community is made up of a collection of user-created and user-moderated forums called subreddits, and each forum is dedicated to a particular topic. WallStreetBets is one such subreddit that focuses on highly speculative stock and option trading strategies. Founded in 2012, WallStreetBets has since grown to 12.4 million members as of August 2022. While WallStreetBets is not the only subreddit to focus on trading strategies, it has been widely recognized as a central player in retail investor coordination during high-attention events such as the 2021 Gamestop short squeeze and its comment activity far exceeds that of similar subreddits.[8]

At the top of a subreddit, its moderators (volunteers who create and uphold guidelines) can pin up to two posts at any given point in time. This means that pinned posts remain at the top of the subreddit's homepage regardless of other incoming posts until the moderator decides to unpin them. By examining historical snapshots of the WallStreetBets homepage using Wayback machine, I find that WallStreetBets starts pinning Daily Discussion posts at the beginning of 2018. A new Daily Discussion posts is created and pinned to the top of the WallStreetBets homepage around 2-3 hours prior to market open every trading day. Soon after markets close, Daily Discussion posts gets unpinned and removed from the top of the homepage.

Comment Display Change

Moderators in a subreddit set the default comment sort for posts.[9] There are 6 different types of comment sorts - Best (sorted by the lower bound of the Wilson score confidence interval), Top (sorted by comment score), New (sorted by newest), Controversial (sorted by comments with a large number of upvotes and downvotes), Old (sorted by oldest), and Q&A (sorted by comments that the author of the post commented on). The default comment sort for all WallStreetBets posts was Best from at least 1/1/2018 until 7/25/2018.

Best sort is a confidence sort ranking algorithm that calculates the lower bound of the Wilson score confidence interval, which is defined as:

$$BestScore_c = \frac{\hat{p_c} + \frac{z^2}{2n_c}}{1 + \frac{z^2}{n_c}} - \frac{z}{1 + \frac{z^2}{n_c}} \sqrt{\frac{\hat{p_c}(1 - \hat{p_c})}{n_c} + \frac{z^2}{4n_c^2}} \tag{1.1}$$

[7]As of March 2022, Reddit is the 9th most visited website globally and the 6th most visited website in the U.S., according to Semrush.

[8]For example, on 10/12/2018, r/WallStreetBets had 5944 comments while similar subreddits like r/investing and r/stocks had 839 and 398, respectively.

[9]Once users click into a post, they can then manually change the comment sort via a drop-down menu.

where $\hat{p_c}$ equals the proportion of upvotes comment c currently has, n_c equals the number of total votes comment c currently has, and z is the $1 - \frac{\alpha}{2}$ quantile of a standard normal distribution. For the 95% confidence level used in Best sort, $\alpha = 0.05$, so $z = 1.96$.

This algorithm treats comment vote count as a statistical sampling of a hypothetical full vote by everyone. In other words, given the proportion of upvotes comment c currently has, there is a 95% chance that the proportion of upvotes if everyone who saw comment c voted on it is at least what? The more votes a comment gets, the closer the 95% confidence score gets to the actual score. For example, if a comment has 10 upvotes and 1 downvote, Reddit might have enough confidence to place it above a comment with 40 upvotes and 20 downvotes. The algorithm constantly updates with incoming comments and votes.

With the Best sort, incoming comments initially start near the bottom of the page. Once some votes come in, the Best sorting algorithm updates comment scores and moves comments around accordingly. That means that there is a delay (minutes) between when comments are published and when comments get moved up or stay down. While not every comment eventually spends time at or near the top of the page, comments that do end up at the top often receive prominent positioning for some period of time (usually between 10 minutes to a few hours). This means that readers that navigate to a post that is sorted by Best will often see that the first few comments are from at least a few minutes ago.

Starting on 7/26/2018, the default comment sort for just Daily Discussion posts switched to New while the default comment sort for all other posts remained Best. New sort displays the newest comment at the top of the page at any point in time. This sort provides users a more timely glimpse into what tickers and trading strategies are being discussed, but without any filtering to weed out potential noise. The average comment during the New Regime remains near the top of the page (one of the first 10 comments on the page) for a few minutes (up to 5) before dropping down. My identification strategy exploits this change in comment sorting display from Best to New in Daily Discussion posts.

Empirical Predictions

This identification strategy yields a couple of key predictions for market reactions following Best Regime comments versus New Regime comments. The theoretical motivation is based on standard models of limited attention and gradual information diffusion that assume that only a fraction of investors are attentive to a comment signal at any point in time, investors do not fully take into account other investors' beliefs and information sets, and investors have heterogeneous beliefs (comment signals influence some investors more than others).

First, comments during the New Regime are more likely to appear at the top of the page compared to comments during the Best Regime immediately after publication, inducing a higher share of immediate attention. This should lead to more retail trading volume. And in a model of heterogeneous beliefs, if naive investors dominate in markets and arbitrageurs face short-sale constraints, prices should move with volume. However, if arbitrageurs dominate in markets, then there should be little price response despite increased trading volume.

Prediction 1 (Immediate Reaction): *New Regime comments are accompanied by larger retail trading volumes and absolute returns compared to Best Regime comments immediately (within minutes) after comment publication.*

After the initial few minutes, comments in the New Regime drop from the top of the page and continue to drop down with time. Comments in the Best Regime, on the other hand, are more likely to move up the page after a few minutes and remain there or continue to move up over time, inducing a higher share of attention over a delayed period.

Prediction 2a (Best Regime Underreaction): *Best Regime comments are accompanied by stronger return continuations compared to New Regime comments.*

Since comments that appear at the top in the New Regime are not filtered, it is likely that investors overreact initially to New Regime comments.

Prediction 2b (New Regime Overreaction): *The price reaction to New Regime comments partially reverses after comments drop off from the top of the page.*

I use an immediate time window of 5 minutes after comment publication and return continuation windows of [5min, 15min] and [5min, 30min] following comment publication.

1.3 Data

To test my predictions, I collect Daily Discussion comment data and merge with market data, Robinhood user stock ownership data, and traditional news data.

WallStreetBets data

I collect Daily Discussion comments between 1/1/2018 - 3/1/2019 using the Pushshift API, which retrieves real-time Reddit data and retains posts and comments deleted by the author. I first collect all posts on WallStreetBets in that time period. For each post, I collect the following information: post id, datetime, post author, post title, link flair (a tag to categorize the content of the post), post score (upvotes minus downvotes), upvote ratio, comment count, and post body text. I then identify Daily Discussion posts by searching for "Daily Discussion" in the post title and collect all comments for those posts.[10]

For each comment, I collect the following information: comment id, datetime, comment author, comment parent id, post id, comment body text, and comment score (upvotes minus downvotes). One thing to note is that comment and post scores are reported as-of the date of collection (early 2022). As a result, one limitation with the data is that I do not observe intraday changes in comments scores or comment positioning.[11]

To test my predictions, I am interested in comments that mention tickers. I take the following steps to identify ticker mentions: (1) remove punctuation from comments, (2)

[10]For completeness, I end up collecting all WallStreetBets posts and comments between 1/1/2018 - 3/1/2019. In addition, I collect all posts and comments for two similar subreddits - r/investing and r/stocks.

[11]I can recreate intraday comment positioning according to the New sort, but I am unable to do that for the Best sort.

identify words with either 2-5 uppercase letters or words with a $ followed by 1-4 letters, (3) remove words with 2-5 uppercase letters if they are in a custom list of stopwords.

Figure 1.1 shows the daily number of unique tickers mentioned in Daily Discussion comments in the 3 months prior to and 3 months after the comment sort change (4/25/2018 - 10/25/2018). There is a large variation in the number of tickers mentioned in a given day, ranging from 4 tickers on 7/5/2018 to 103 tickers on 10/4/2018 with the mean (median) over my sample period of 38 (29). While Daily Discussion posts account for less than 1% of all WallStreetBets posts over that 6-month time period, Daily Discussion comments comprise upwards of 50% of all WallStreetBets comments on a given day. The most mentioned tickers are SNAP, MSFT, AMD, AMZN, and AAPL.

Note that only top-level comments (and not comment replies) affect the sorting algorithm. So I further restrict my comment sample set to top-level comments. For tighter identification, I focus on comments published in the 3 months prior to and 3 months after the comment sort change (4/25/2018 - 10/25/2018, hereafter referred to as my sample period). And to be able to measure retail trading after comments, I keep comments published during market hours. Lastly, I restrict my comment sample set to comments that mention a single ticker.

Out of the $43,406$ top-level Daily Discussion comments between 4/25/2018 and 10/25/2018, $33,570$ are during market hours. Of those top-level market hour comments, $17,861$ have a ticker tag and approximately 85% of those ticker-tagged comments have a single ticker tag ($15,145$). My final comment sample set is at the ticker-minute level (since multiple comments mentioning the same ticker can occur in the same minute) and contains $13,923$ observations mentioning 479 unique tickers with available data in CRSP and TAQ.

Daily Discussion Comment Topics

I now analyze the text of Daily Discussion comments using machine learning. Specifically, I employ topic modeling, which identifies common topics within a collection of unstructured texts. The topic modeling algorithm I use is called BERTopic, originally proposed by Grootendorst, 2022. BERTopic generates topics using sentence embeddings and is particularly well-suited for my dataset because it works well with short-text documents. The more widely-used topic modeling technique is Latent Dirichlet Allocation (LDA), but LDA is better suited for longer text documents such as entire news articles. And since BERTopic uses an embedding approach, it relies on the original structure of the text and so little-to-no pre-processing is necessary for the comments.

BERTopic proceeds in 3 steps. First, a sentence transformer embeds each comment by assigning it a numerical representation of the meaning behind the comment. I use the 'all-MiniLM-L6-v2' sentence transformer that has a dimension of 384 and has been pre-trained on over 1 billion sentence pairs, of which over 70% of them come from Reddit comments.

Second, dimension reduction is performed on the comment embeddings using uniform manifold approximation and projection (UMAP) to reduce the dimension from 384 to 2 or 3, and then the hierarchical density-based spatial clustering of applications with noise (HBDSCAN) algorithm is used to cluster semantically-similar comments.

Lastly, a class-based term frequency inverse document frequency (c-TF-IDF) algorithm is used to identify key words and terms that are relevant to a particular cluster by comparing the importance of terms within each cluster.

I run BERTopic on a sample set containing all comments on WallStreetBets during my sample period. Figure A.1 depicts a intertopic distance map generated by the BERTopic model, and it can be seen that there are 9 clusters created. After careful inspection, I assign a name to each of the 9 clusters and Table 1.1 displays the top 10 words in each cluster along with the relative frequencies of comments that belong to each cluster, separately for all WallStreetBet comments, Daily Discussion comments, and top-level single ticker-tagged market hour Daily Discussion comments.

The majority of WallStreetBets comments belong to two topics: 'Buy and Sell' and 'Stock Performance and News.' The most common words in 'Buy and Sell' comments are related to trading such as buying, selling, calls, and puts. The most common words in 'Stock Performance and News' comments are related to news such as earnings, or stock price movements such as a price dip. These two topics account for around 68% of topics on WallStreetBets and Daily Discussion comments, and over 83% of topics in the top-level single ticker-tagged market hour Daily Discussion comment sample set.

Some of the other identified topics contain WallStreetBets-specific jargon. One topic called 'Mooning' is comprised of comments mentioning a variation of that word or containing a rocket ship emoji. This phrase is used to express confidence in the performance of a chosen stock. Another topic is called 'Autistic,' which is comprised of comments mentioning a variation of that word. On WallStreetBets, that word is supposed to describe someone that does due diligence and knows what they are doing. Oftentimes, though, members use the term more generally as a blanket term to refer to themselves and fellow members. And another topic called 'Pray, Drill, Drop' is comprised of comments either indicating or speculating that the stock price is declining or will decline. Table A.1 provides representative examples of Daily Discussion comments that belong to each topic.

Market data

In order to test my empirical predictions, I merge my comment data with market data. I obtain daily-level market data from CRSP, minute-level ticker trade and quote price data from TAQ, and minute-level retail trading data from TAQ data using the Boehmer et al., 2021 algorithm for identifying retail trades. The Boehmer et al., 2021 algorithm identifies trades as retail if the trade takes place off-exchange (exchange code "D" in TAQ) and at a price just below or above a round penny (since retail traders typically receive a small fraction of a cent price improvement over the National Best Bid or Offer (NBBO) for market orders).

Robintrack data

To further test my empirical predictions, I merge my comment data with retail investor stock ownership data from Robinhood. I download Robinhood investor ownership data from Robintrack, which is publicly available at https://robintrack.net/data-download. This dataset covers the time period from 5/2/2018 to 8/13/2020.[12] Observations contain the number of Robinhood investors holding a security at approximately hourly intervals. For this paper, I am interested in the number of Robinhood investors holding a security at the end of day and the overnight change in Robinhood investors holding a security, so I only keep the first and last available observations that day (after 9 AM ET and prior to 4 PM ET, respectively).

Traditional news data

To account for other news events and to use as a placebo test, I collect news article data. I obtain news articles from the Finnhub API, which aggregates financial news articles from publicly available websites such as Nasdaq and Reuters, and ticker tags them for ease of access. I collect news articles over the same time period as my WallStreetBets dataset. For each news article, I collect the following information: article id assigned by Finnhub, datetime, news source, article headline, article summary, url, and ticker(s) mentioned in the article. I limit the sample to common stocks (CRSP share codes 10 and 11) with available data in CRSP, TAQ, and Robintrack. My article dataset contains 3,256 ticker-tagged articles covering 1,160 tickers from 123 news sources.

1.4 Results

Do WSB Comments Matter?

Before comparing market reactions to comments around the comment display change, I first establish that Daily Discussion comments on WallStreetBets matter and that investors pay attention to them irrespective of the display change. In this section, I consider all ticker-tagged Daily Discussion comments in my sample period.

Determinants of Daily Discussion Comments

I first examine what drives Daily Discussion comment production. Prior research finds that retail investors gravitate towards high-attention stocks.[13] As it is likely that the people who consume WallStreetBets comments are primarily retail investors, I investigate whether high-attention stocks in the market drive Daily Discussion comments. I estimate the following

[12] Note, that the starting coverage date for Robinhood ownership data is 5/2/2018, which is about a week later than the start of the time period in my main specification. To preserve the number of comments prior to the display change, I only drop observations between 4/25/2018 and 5/2/2018 when Robinhood-related variables are used.

[13] Retail investors' preferences for attention-seeking stocks was first documented by Grinblatt and Keloharju, 2001, Seasholes and Wu, 2007, and B. M. Barber and Odean, 2008.

regression:

$$WSBCommentVolume_{i,d} = exp\begin{pmatrix} \beta_0 + \beta_1 WSBCommentVolume_{i,d-1} + \\ \beta_2 NewsArticleVolume_{i,d-1} + \\ \beta_3 |Ret|_{i,overnight} + \beta_4 |Ret|_{i,[d-2,d-1]} + \\ \beta_5 |Ret|_{i,[d-5,d-2]} + \\ \beta_6 AbnRetailVol_{i,d-1} + Controls + \epsilon_{i,d} \end{pmatrix} \tag{1.2}$$

where $WSBCommentVolume_{i,d}$ is the number of Daily Discussion comments that mention ticker i on date d. $NewsArticleVolume_{i,d-1}$ is the number of traditional news articles from Finnhub that mention ticker i on date $d-1$. $|Ret|_{i,overnight}$, $|Ret|_{i,[d-2,d-1]}$, and $|Ret|_{i,[d-5,d-2]}$ are absolute values of returns for ticker i overnight, from date $d-2$ to $d-1$, and from date $d-5$ to $d-2$, respectively. $AbnRetailVol_{i,d-1}$ is the abnormal retail trading volume, which is defined as the log of $RetailVol_{i,d-1}/\frac{1}{40}\sum_{k=1}^{40} RetailVol_{i,d-1-k}$, where $RetailVol_{i,d-1}$ is equal to the number of retail trades for ticker i on date $d-1$ and $\frac{1}{40}\sum_{k=1}^{40} RetailVol_{i,d-1-k}$ is the average number of retail trades for ticker i over the 40 previous days. All regression specifications include controls for firm size as well as ticker and date fixed effects, and the sample set consists of ticker-date observations comprised of firms with at least one comment mention over the sample period.

Table 1.2 reports results from the Poisson regression. I choose to estimate a Poisson regression because the outcome variable is count-based, zero bounded, and highly right skewed, making the traditional approach of adding a constant to the outcome variable and estimating a log-linear OLS regression unsuitable since the coefficients would have no meaningful economic interpretation.[14]

Column (1) of Table 1.2 indicates that previous day Daily Discussion comments, previous day news articles, extreme overnight returns, and previous day abnormal retail volume for ticker i positively and significantly predict Daily Discussion comment volume for ticker i on date d. The estimated coefficient on $WSBCommentVolume_{i,d-1}$ is 0.007, significant at the 1% level. This implies that, holding all else constant, an additional Daily Discussion comment mentioning ticker i on date $d-1$ is expected to increase Daily Discussion comment volume for ticker i on date d by 0.7% ($exp(0.007) - 1 = 0.007$). The estimated coefficient on $|Ret|_{i,overnight}$ is 0.217, significant at the 1% level. This implies that, holding all else constant, a 1% increase in overnight absolute returns is expected to increase Daily Discussion comment volume by 24.2% ($exp(0.217) - 1 = 0.242$).

In addition, recent research finds a relationship between WallStreetBets and Robinhood user trading activity.[15] I add in an explanatory variable for the magnitude of the change in overnight Robinhood users holding ticker i ($|RobinhoodUserHoldingChange|_{i,overnight}$), which is defined as the absolute value difference between the number of Robinhood users

[14]For further discussion of appropriate econometric approaches to working with count-based data in finance, see Cohn et al., 2022.

[15]V. Fedyk, 2022 and Eaton et al., 2022 find evidence that Robinhood investors exhibit a strong propensity for trading stocks mentioned on WallStreetBets.

holding ticker i prior to market open on date d and the number of Robinhood users holding ticker i prior to market close on date $d-1$. Column (2) of Table 1.2 reports the regression results. The coefficient on $|RobinhoodUserHoldingChange|_{i,overnight}$ is positive and significant, suggesting that tickers with extreme Robinhood trading are discussed more heavily in Daily Discussion posts. The coefficients on the previous explanatory variables remain similar in economic magnitude and statistical significance.

Put together, these results are consistent with Daily Discussion comments discussing recent high-attention, newsworthy stocks. There is also a persistence of tickers being discussed if they were discussed previously. This suggests that WallStreetBets users pay attention to market news.

Do Daily Discussion Comments Predict Retail Trading?

I now address whether Daily Discussion comments impact markets at all. Recent research looking at financial social media has found that social media content predicts market outcomes such as trading volume, returns, and earnings surprises.[16] I add to these findings by confirming that Daily Discussion comments on WallStreetBets also contain investment value.

To attempt to capture whether the audience reading WallStreetBets reacts to Daily Discussion comments, I examine abnormal retail trading volume around comment publication. My outcome variable is $AbnRetailVol_{i,d}$, which is defined as the natural logarithm of the number of retail trades for ticker i on date d (measured during market hours) scaled by the average number of retail trades for ticker i over the 40 previous days. My main explanatory variable is an indicator variable for whether a Daily Discussion comment was made about ticker i on date t ($WSBComment_{i,d}$), where the comment could be from prior to market opening or during market hours (9:30 AM - 4 PM ET).

Table 1.3 reports OLS regression results measuring the relationship between daily abnormal retail trading and Daily Discussion comment publication. To attempt to establish some predictive interpretation going from Daily Discussion comments to abnormal retail trading, I control for lagged abnormal retail trading and lagged absolute returns to capture reactions to other events prior to date d. It is also possible that retail investors are reacting to some other form of news. To account for that, I control for contemporaneous news article coverage. The regression specifications include controls for firm size as well as ticker and date fixed effects, and the sample set consists of ticker-date observations comprised of firms with at least one comment mention over the sample period.

Column (1) of Table 1.3 confirms that Daily Discussion ticker mentions either prior to market open or during market hours are associated with a 15% increase in abnormal retail trading volume on date d. This result suggests that retail investors are trading in response to either the comments themselves or trading in response to news coming out that day that the comments are also talking about.

[16]See, for example, Chen et al., 2014, Cookson and Niessner, 2020, Eaton et al., 2022, and Kogan et al., 2022.

If comments are responsible for at least some of the increased retail trading volume, then it is likely that the results would be concentrated in firms with less volume and liquidity, and with fewer large institutional investors. In column (2), I interact a high Robinhood user ownership dummy (defined as equal to 1 if ticker i is in the top 20% of firms based on the previous day's Robinhood users holding ticker i scaled by market capitalization) with my Daily Discussion comment dummy ($WSBComment_{i,d}$). This measure of Robinhood user ownership assumes that each Robinhood user holding represents an equal number of dollars held.[17] The effect on abnormal retail trading is approximately 50% larger for firms with higher Robinhood user ownership, significant at the 1% level.

In column (3), I interact a small firm dummy (defined as equal to 1 if ticker i is in the bottom 20% of firms based on the previous day's market capitalization) with my Daily Discussion comment dummy. The effect on abnormal retail trading is 79% larger for small firms compared to larger firms, significant at the 1% level. These results are consistent with Daily Discussion comments playing some role in the increased abnormal retail trading volume.

Robustness For robustness, Table A.2 reruns the regression in column (1) of Table 1.3, replacing the outcome variable with retail trades as a percentage of all trades ($RetailPct_{i,d}$) in column (1), absolute abnormal return ($|AbnRet|_{i,d}$) in column (2), and abnormal option trading volume ($AbnOptionVol_{i,d}$) in column (3).

I conjecture that retail investors are more likely to participate and respond to Daily Discussion comments, and as a result the response of Daily Discussion comments to retail trading should be stronger. Instead of looking at abnormal retail trading, column (1) of Table A.2 measures retail trades as a percentage of all trades. I find that ticker-days with Daily Discussion ticker mention are associated with a 0.219% increase in retail trading relative to all trading on date d, significant at the 5% level. This represents a relative increase of 1% (the average percentage of retail trading during this period was 21.7%).

Column (2) of Table A.2 examines the relationship between comments and absolute abnormal returns, where abnormal returns are daily returns in excess of the contemporaneous value-weighted average return among all tickers in the sample. If the market already incorporated the information contained in the comments, then the expected absolute return should be zero. I find return volatility is positive and significant on days with Daily Discussion ticker mentions. The absolute abnormal return is 31.7-basis-points higher on days where ticker i is mentioned in a Daily Discussion comment, significant at the 1% level. This represents a relative increase of 21.7% (the daily absolute abnormal return during this period was 1.46%).

Another way that readers of Daily Discussion comments might react to the comments is by trading options since there are mentions of buying and selling calls and puts in the

[17]In untabulated results, the results remain unchanged if I assume that each Robinhood user holding represents an equal number of shares held.

comments. While there are novel measures to capture retail option trading,[18] they rely on information introduced after my sample period. Due to that limitation, I use a measure for aggregate option volume, which is defined as the natural logarithm of option share volume for ticker i on date d scaled by the average daily option share volume in the 40 previous days. Daily option share volume is collected from OptionMetrics. I find that Daily Discussion ticker mentions are associated with a 25% increase in abnormal option volume, significant at the 1% level. While this measure includes both option trading from retail investors as well as other investors, this result further supports that retail investors react to Daily Discussion comments.

Altogether, these results indicate that Daily Discussion comments are associated with increased trading activity (especially among retail investors) and increased price volatility, suggesting that Daily Discussion comments matter. I will now investigate how these relationships change based on how comments are displayed.

Market Impact of Comment Display Change

Having established that Daily Discussion comments are important, I now conduct my main and novel test on comment visibility. The change in default comment sorting provided a shock to the positioning mechanics of Daily Discussion comments. I use this change to compare the market reactions to comments made in the two sorting regimes to identify the overall effect that comment visibility has. For my main analysis section, I restrict my comment sample set to top-level single ticker-tagged market hour Daily Discussion comments.

Immediate Response to Comments

I begin by comparing immediate abnormal retail trading volumes and absolute abnormal returns associated with comments in the 3 months after the display change (New Regime) versus in the 3 months prior to the display change (Best Regime). I first estimate the following regression:

$$AbnRetailVol_{c,i,d,[t,t+5min]} = \beta_0 + \beta_1 NewRegime_c + Controls + \epsilon_{i,d} \tag{1.3}$$

where $AbnRetailVol_{c,i,d,[t,t+5min]}$ is defined as the abnormal retail trading volume for ticker i on date d in the 5 minutes after comment c posts. $NewRegime_c$ is a dummy variable equal to 1 if comment c was published during the New Regime (7/26/2018 and after) and equal to 0 if published during the Best Regime (7/25/2018 and prior). Regression specifications include controls for same-ticker comment volume and firm size as well as industry and half hour fixed effects, and the sample set consists of ticker-minute observations with at least one comment mentioning a given ticker in a given minute. And since some comments are sometimes preceded by other Daily Discussion comments mentioning that same ticker, I control for lagged same-ticker comment mentions in the 5 minutes prior to comment c.

[18]See, for instance, Bryzgalova et al., 2022.

Table 1.4a presents regression results for abnormal retail trading. Without any controls in column (1), New Regime comments see 14.9% higher abnormal retail trading compared to Best Regime comments, significant at the 1% level. After controlling for firm size, same-minute ticker comments, and lagged comments, and adding in industry and half hour fixed effects, the difference in abnormal retail trading decreases to 10.1%, significant at the 1% level.

The third column in Table 1.4a adds in the high Robinhood retail ownership dummy and interacts it with the New Regime dummy. The difference in abnormal retail trading between New Regime and Best Regime comments increases by 36.7% for firms with high Robinhood retail ownership, significant at the 1% level. Firms with low Robinhood retail ownership see a differential retail trading response of -1.1%, which is both statistically insignificant and economically small. This is consistent with information posted on WallStreetBets being primarily consumed by retail investors, particularly the type of retail investor that trades on the Robinhood app.

The fourth column in Table 1.4a replaces the Robinhood retail ownership dummy with the small firm dummy and interacts it with the New Regime dummy. I find that the differential retail trading response to the regime change is stronger in smaller firms, with smaller firms seeing a 46.2% higher abnormal retail trading difference, significant at the 5% level. Put together, these results suggest that firms where retail trading matters most - smaller firms and firms with high retail ownership - are most responsive to the comment display change.

Having shown that retail trading is responsive to the comment display change, I now test whether prices respond to the comment display change by estimating the following regression:

$$|AbnRet|_{c,i,d,[t,t+5min]} = \beta_0 + \beta_1 NewRegime_c + Controls + \epsilon_{i,d} \tag{1.4}$$

where $|AbnRet|_{c,i,d,[t,t+5min]}$ is the absolute value of the abnormal return for ticker i on date d in the 5 minutes after comment c posts. The rest of the regression specification is identical to Equation 1.3.

Table 1.4b presents regression results for absolute abnormal returns. Column (1) shows that New Regime comments see 10.4-basis-points higher absolute abnormal returns in the first 5 minutes compared to Best Regime comments without any controls, significant at the 1% level. In column (2), I control for firm size, same-minute comments, and lagged comments, and add in industry and half hour fixed effects. The absolute return difference decreases slightly to 6.7 basis points, significant at the 1% level. To measure the economic significance, the average absolute return for Best Regime comments is 21.3 basis points, so a 6.7-basis-points increase represents a relative increase of 31.5%. This increase seems large, but I next show that the bulk of the effect is concentrated in small firms and firms with high retail ownership.

Column (3) in Table 1.4b adds the high Robinhood retail ownership dummy and interacts it with the New Regime dummy. The immediate absolute price response difference between New Regime comments and Best Regime comments is 20.9-basis-points larger for high Robinhood retail ownership firms than for low Robinhood retail ownership firms, significant at the 1% level. Firms with low Robinhood retail ownership see a differential price response of 1.3 basis points, which is both statistically insignificant and economically small.

Column (4) in Table 1.4b replaces the Robinhood retail ownership dummy with the large firm dummy and interacts it with the New Regime dummy. I find that small firms see 65.3-basis-points higher absolute abnormal returns compared to larger firms in the New Regime versus Best Regime, significant at the 1% level. Larger firms see a differential absolute abnormal return of 4.8 basis points, significant at the 1% level.

These immediate reaction results are robust to a difference-in-difference approach using the Daily Discussion comment sample set as the treatment group and adding in non-Daily Discussion comments on WallStreetBets where there was no change in default comment sort as the control group. Table A.3 reports the results. The control group consists of non-Daily Discussion top-level single ticker-tagged market hour comments that are not published within 5 minutes of a Daily Discussion comment mentioning the same ticker. Column (1) shows that immediate abnormal retail trading and absolute returns are higher for Daily Discussion comments after the comment display change. Columns (2) and (3) add in a triple interaction between the New Regime dummy, Daily Discussion comment dummy, and high Robinhood ownership (small firm) dummy. The coefficient on the triple interaction terms are all positive and statistically significant. This confirms that retail trading and volatility increases even more so for Daily Discussion comments about high Robinhood ownership (small) firms after the comment display change.

Taken together, these results are consistent with increased immediate comment visibility in the New Regime compared to the Best Regime, resulting in more immediate retail trading and price responses.

Return Continuation

Since Best Regime comments are more likely to receive prominent positioning after a delay of, on average, a few minutes while New Regime comments published at the same time drop from the top of the page by that point, I expect that Best Regime comments see a larger return continuation compared to New Regime comments. Conditioning on the first 5-minute abnormal return, I compare price dynamics afterwards by running the following regression:

$$\begin{aligned} AbnRet_{c,i,d,[t+5min,t+t_j]} = \beta_0 + \beta_1 AbnRet_{c,i,d,[t,t+5min]} + \beta_2 NewRegime_c + \\ \beta_3 AbnRet_{c,i,d,[t,t+5min]} x NewRegime_c + Controls + \epsilon_{i,d} \end{aligned} \tag{1.5}$$

where $AbnRet_{c,i,d,[t+5min,t+t_j]}$ is the delayed abnormal return for ticker i in the first $t_j \in \{15, 30\}$ minutes excluding the initial 5 minutes after comment c on date d. $AbnRet_{c,i,d,[t,t+5min]}$ is the initial 5-minute abnormal return and $NewRegime_c$ is a dummy variable equal to 1 if comment c posted in the New Regime and 0 if comment c posted in the Best Regime. $Controls$ include firm size, same-minute comments, lagged comments, industry fixed effects, and half hour fixed effects. Standard errors are clustered by ticker and date.

Table 1.5a presents results for $t_j = 15$ minutes with varying control variables. The coefficient on $AbnRet_{c,i,d,[t,t+5min]}$ not interacted with the New Regime indicator is positive and significant at the 1% level across all specifications, indicating that Best Regime comments see a positive return continuation. With all controls, Best Regime comments see a return continuation of 15.7% from five minutes after publication to 15 minutes after. New Regime comments, on the other hand, do not see a positive return continuation. The interaction term is negative and significant at the 1% level. Conditioning on the initial 5-minute abnormal return, New Regime comments induce 20.8% less continuation in returns compared to Best Regime comments over the corresponding time period. This result is consistent with Best Regime comments receiving more visibility after a slight delay and New Regime comments losing visibility after few minutes.

The return continuation results increase when looking at $t_j = 30$ minutes (Table 1.5b). Best Regime comments see a 20.9%-21.1% return continuation from 5 minutes to 30 minutes after comment c posts conditioning on the initial 5-minute abnormal return, significant at the 10% level. New Regime comments induce a 26.6%-26.9% lower return continuation compared to Best Regime comments, significant at the 5% level. This suggests that Best Regime comments are more visible over a longer period of time and further supports that New Regime comments disappear from view after a few minutes.

To investigate whether these results are driven by retail investor trading, I slice my comment data by Robinhood retail ownership and rerun the return continuation regression for $t_j = 30$ minutes separately for high and low retail ownership. Table 1.6 columns (1) and (2) report the results. High RH are firms on date d in the top 20% of Robinhood retail ownership as of date $d - 1$ among all firms in my sample and Low RH are all other firms. The return continuation after 30 minutes for Best Regime comments mentioning high Robinhood retail ownership firms is 56%, significant at the 5% level. In comparison, the return continuation for Best Regime comments mentioning low Robinhood retail ownership firms is a statistically insignificant and economically smaller 8.3%. Furthermore, the return continuation for New Regime comments mentioning high RH retail ownership firms is 61.8% lower, significant at the 5% level.

I separately slice my comment data by firm size and rerun the return continuation regression for $t_j = 30$ minutes separately for large and small firms. Table 1.6 columns (3) and (4) report the results. Small firms are in the bottom 20% among all firms in my sample according to previous day's market capitalization and larger firms are all other firms. The return continuation after 30 minutes for Best Regime comments mentioning larger firms is a statistically insignificant and economically small 2.3% while the return continuation for small firms is a statistically significant 90.5%. In addition, the return continuation for New Regime comments mentioning large firms is a statistically insignificant 7.8% lower compared to Best Regime comments, while the return continuation for small firms is a statistically significant 98.2% lower. These return continuation results support the immediate return and volume results—firms where retail investors are more likely to have the most impact exhibit the strongest sensitivity to changes in comment visibility.

Figure 1.2 graphs cumulative abnormal returns following comment publication sliced by

(1) the direction of the initial 5-minute abnormal return and (2) the comment display regime. Consistent with Table 1.4, New Regime comments are accompanied by larger immediate abnormal returns in both the positive and negative directions. Following the first 5 minutes, Best Regime comments see a stronger drift up until 30 minutes afterwards, consistent with Table 1.5. New Regime comments, on the other hand, see little to negative drift over the same time period. Though the standard error bars widen, the graph suggests that the difference in market reaction between New Regime comments and Best Regime comments disappears after approximately 30 minutes. I further confirm the convergence by looking at return continuations from 30 minutes after comment publication to 1-3 hours conditioning on the initial 30-minute abnormal return, and find no significant difference in price paths between Best Regime and New Regime comments after the first 30 minutes. The interaction coefficients in Table 1.7 are all statistically insignificant. This suggests that the average comment information between the two regimes are similar and that the net effect of the comment display change sped up the speed of information incorporation.

Because comments that appear at the top of the page in the New Regime are not filtered, I predict that investors, on average, overreact to New Regime comments. While my regression results in Table 1.5 indicate that New Regime comments see a partial return reversal, I run an additional test with the following regression on New Regime comments:

$$PriceImpact_{c,i,d,[t+5min,t+30min]} = \beta_0 + \beta_1 PriceImpact_{c,i,d,[t,t+5min]} + Controls + \epsilon_{i,d} \quad (1.6)$$

where $PriceImpact_{c,i,d,[t+5min,t+30min]}$ is the price change from the prevailing mid-quote 5 minutes after comment c posts to the mid-quote 30 minutes after the comment and $PriceImpact_{c,i,d,[t,t+5min]}$ is the price change from the prevailing mid-quote at the time comment c posts to the mid-quote 5 minutes after. I use mid-quote prices as opposed to trade prices here to rule out concerns that any return reversal captured in Table 1.5 can be explained by the spread. Controls are the same as in Equation 1.5.

Table A.4a confirms my prediction. Conditional on the first 5-minute mid-quote price response, the mid-quote price response from 5 minutes after to 30 minutes after a comment is 7.7%-7.9% lower, which is significant at the 1% level. This indicates a partial return reversal after comments in the New Regime drop from the top of the page. I also repeat the price impact regression for Best Regime comments to confirm an initial underreaction to Best Regime comments, and Table A.4b presents results that are consistent with an initial underreaction. Put together, my results suggest that the comment display change affected price efficiency, swinging the pendulum from an initial underreaction in the Best Regime to a slight overreaction in the New Regime.

Alternative Explanations

I now conduct a series of tests to address alternative explanations. While I interpret my results as the comment display change affecting comment visibility and subsequently changing retail trading behavior and price formation, there are other events that might affect the relationship between comments and trading.

First, I want to reiterate that the comment display change was seemingly exogenous to readers and thus to my outcome variables as there was no chatter related to the display change on WallStreetBets in the days leading up to the change. Second, I already showed in the previous sections that my results are strongest in firms where I expect the influence of retail investors reading and reacting to WallStreetBets to be more pronounced.

To more directly address concerns, I first conduct a placebo test looking at the market response to news sources unaffected by the comment display change to examine whether the time period after the display change was generally more volatile. After that, I look at pre-trends leading up to comments to address concerns of reverse causality. I then compare ticker coverage on similar subreddits to address concerns that other news sources might be driving my results.

News Placebo Test The comment display change should have only affected Daily Discussion comment visibility and not news article visibility on sites such as Nasdaq, Dow Jones, Reuters, MarketWatch, Business Insider, and CNBC. I compare the market response to news articles before and after the Daily Discussion display change, where I expect to find no differential effect between the two regimes in either the immediate or short-term. If there still is a differential effect, then that would mean that the display change coincided with more general changes in the market's response to news.

Table 1.8 reports my results. I restrict my news sample set to news articles published on Nasdaq, Dow Jones, Reuters, MarketWatch, Business Insider, and CNBC mentioning tickers in my comment sample set. Table 1.8a shows that the differential immediate response to news articles is not significant for both the 5-minute absolute abnormal return and for the 5-minute abnormal retail trading. Table 1.8b shows that conditional on the initial 5 minute price response, the return continuation from 5 minutes after article publication to 15-30 minutes after is not significantly different for articles published before versus after the Daily Discussion display change. These results suggest that the display change on WallStreetBets' Daily Discussion comments did not coincide with a more general change in the response to news or a trend in trading activity. The display change seems to have only affected the response to Daily Discussion comments on WallStreetBets.

Pre-trends Another concern is that comments and subsequent market responses are both reacting to some other event prior to comments, which has nothing to do with comment visibility. If that were the case, then I would expect to see a similar or perhaps larger difference in immediate market reactions between Best Regime comments and New Regime comments prior to comment publication.

I rerun my main immediate market response tests from Table 1.4 but change the dependent variables to capture 5-minute abnormal retail trading and absolute abnormal returns from 6 minutes to 1 minute prior to comment publication. I restrict my sample set to comments that are posted at least 6 minutes after market open so that I only measure same-day retail trading and absolute price changes.

Table A.5 reports the results. I find that the difference in abnormal retail trading in the 5 minutes leading up to comments between regimes is close to zero and insignificant. In addition, I find that the difference in absolute abnormal returns prior to comments between New Regime comments and Best Regime comments is a small 1.5 basis points, significant at the 5% level. This evidence is consistent with the immediate price and retail trading reactions following comments to be, at least in part, attributed to the comments themselves and not to both reacting to some other event prior to the comment.

Ticker Coverage on Other News Sources I now consider neighboring subreddits with potential overlapping coverage and address the concern that my main results may be driven by contemporaneous comments elsewhere on Reddit. The two most similar subreddits to WallStreetBets is r/investing and r/stocks. I document that tickers discussed on WallStreetBets, specifically in Daily Discussion comments, are uniquely timed compared to other subreddits.

I download comments from r/investing and r/stocks and tag tickers in the same method employed for WallStreetBets comments. Figure 1.3 presents my results, which tabulates the average comment volume of ticker mentions in r/investing and r/stocks in the 60 minutes before to 60 minutes after a Daily Discussion comment is published on WallStreetBets. Each observation corresponds to a minute.

There is no apparent relationship between ticker coverage on Daily Discussion posts and similar subreddits, neither for Best Regime or New Regime comments. This suggests that my main results cannot be attributed to other subreddit comments and that there does not appear to be any cross-subreddit spillover.

I replicate this graph, but for news articles aggregated from Finnhub and find similar results. This suggests that the information contained in Daily Discussion comments is different from the information contained in other news sources.

Comment Production Effects

I now examine how comment production changed around the comment display change. I document the evolution of comment production, discuss potential implications, and address whether changes in comment production could explain my main results.

Comment Production Changes

I first show how comment voting production shifted after the display change. Figure 1.4a plots the percentage of Daily Discussion comments published by hour separately by regime. There is a clear change in the timing of comment production. In the Best Regime, more comments are published earlier in the day. After the display change, New Regime comments are much more evenly distributed across the day, though still slightly more concentrated in the morning. This change in distribution of comment timing is consistent with the Best sorting algorithm favoring earlier comments in terms of visibility and the New sort providing equal visibility to all comments.

I confirm this by looking at the comment publication times of the top 20 comments on the page at the end of the day, separately for the New sort and the Best sort. Figure 1.4b shows that earlier Best Regime comments are more likely to get sorted to the top of the page and remain there until the end of the day. Earlier comments in the New Regime, on the other hand, do not remain at the top of the page at the end of the day, which is consistent with the New sort. This novel result highlights the importance of comment display not just for comment consumption but also for comment production.

The increase in comment visibility in the afternoon in the New Regime could be associated with a change in when trading and volatility occurred for a stock within the day. I find that immediate retail trading and volatility are lower in the afternoon (2-4 PM ET) versus earlier in the day (prior to 2 PM ET) in the Best Regime, consistent with afternoon comments having little chance of getting seen prior to the comment display change (see Table 1.9). After the regime change, retail trading (significant at the 10% level) and volatility (significant at the 5% level) increases in the afternoon versus the morning. This suggests that the change in comment display affected when retail trading and volatility occurred throughout the day.

I now look at how comment volume evolved over the corresponding time period. The change in comment timing suggests that comment producers reacted to the comment display change. Best Regime comments posted later in the day are not as likely to gain prominent visibility compared to morning comments, so there is less of an incentive to post later in the day. On the other hand, all New Regime comments receive prominent positioning regardless of when they post. As a result, I expect that comment volume increases after the comment display change. Figure 1.5 confirms that daily comment volume in Daily Discussion posts increases substantially after the comment display change.

The increased comment volume could also be attributed to an increase in prominence of WallStreetBets or social media in general. I show in that same graph daily comment volume for all non-Daily Discussion posts on WallStreetBets that were not impacted by the comment display change and document visibly no change in comment volume for non-Daily Discussion comments. This indicates that the Daily Discussion-specific comment display change impacted Daily Discussion comment volume and not a more general increase in comment volume over time. For the purposes of this paper, I do not attempt to distinguish between the types of producers and subsequent differential strategy shifts due to the comment display change.

Could Comment Production Changes Drive My Main Results?

Another explanation for my main results is that readers are responding to changes in comment production as opposed to changes in comment visibility. One possible outcome is that increased comment volume decreases comment informativeness or quality. That would mean that the price response for Best Regime comments should exceed that of New Regime comments. I do not find that to be the case. I reiterate that the price paths for Best Regime and

New Regime comments eventually converge after approximately 30 minutes (Figure 1.2), and there is no differential return continuation from 30 minutes after publication to 1-3 hours after, and find no significant difference in price paths after the first half hour (Table 1.7).

I further examine the distribution of topics between Best Regime comments and New Regime comments. Figure A.2 shows that there was a decrease in 'Buy and Sell' comments in the New Regime and an increase in 'Stock Performance and New' comments as well as comments mentioning WallStreetBets jargon such as 'mooning,' 'praying,' and 'drilling.' A Pearson chi-square test to formally compare the distribution of topics between the two regimes confirms that the distributions are significantly different at the 1% level (untabulated). This suggests that while the topics mentioned in comments changed with the comment display change to mention more speculative comments, it did not affect the average comment quality.

The increase in comment volume could also be associated with an overall increase in attention to Daily Discussion comments. This would mean that retail trading and volatility should be higher in the New Regime, which are the same predictions according to the comment visibility channel. To distinguish between the increased attention channel and the comment visibility channel, I repeat my main tests for varying time windows surrounding the display change. Figure 1.6 graphs the difference in immediate retail trading and absolute return responses to New versus Best Regime comments, from 14 days before and after up to 7 months before and after the display change. Though my confidence bands widens with narrower time periods, my results are robust to varying time bands. In particular, since it it less likely that changes in comment production could affect trading behavior when looking at shorter periods, it is unlikely that my results are driven by higher overall attention.

1.5 Conclusion

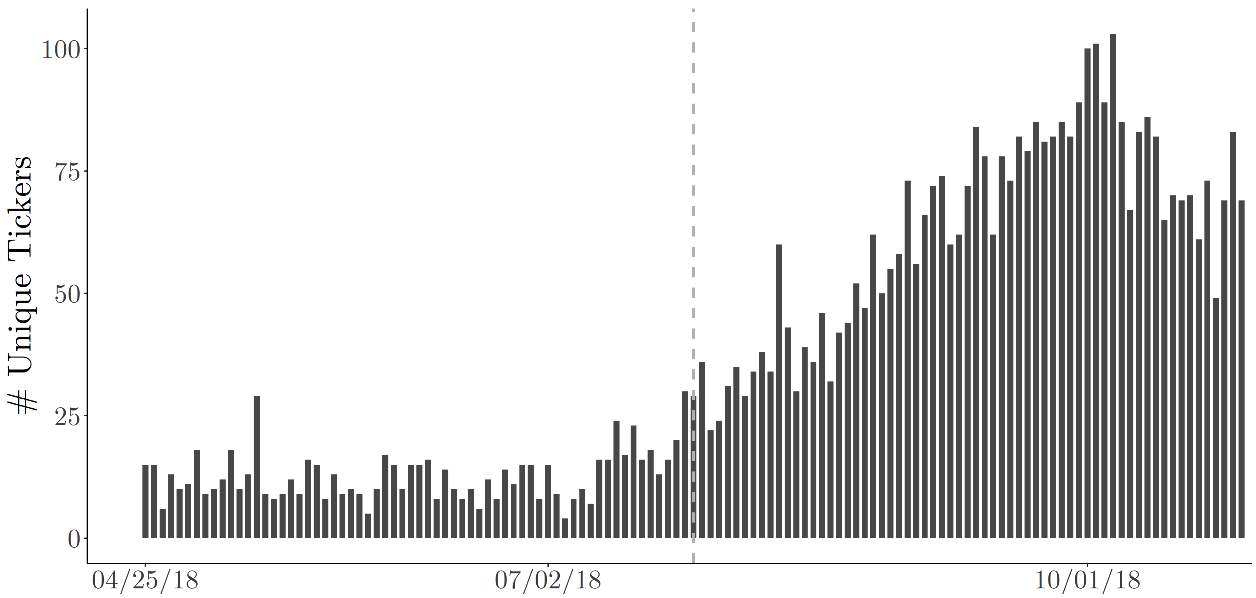
In this paper, I investigate a novel setting of financial information display on a social media platform. I identify a change in default comment display on Daily Discussion posts in WallStreetBets. The change resulted in a more timely and noisy display, which had two effects. First, it increased immediate attention on incoming comments and lead to initial overreactions, but ultimately made prices more efficient. Second, it shifted comment production timing from being heavily morning-skewed to more evenly distributed, and increased comment volume.

My results underscore the importance in the way that information is delivered on social media platforms as it has direct implications on investor and market behavior. In particular, since social media platforms hold some autonomy over how regulated or curated information presentation is, my results highlight the importance of social media regulations to curtail the propagation of potentially harmful and unverified information.

The way that information is displayed also has direct implications on subsequent information production, which can also have market implications. My results on comment production speak to potentially how WallStreetBets was able to increase participation and retail investor coordination that contributed to events such as the Gamestop short squeeze.

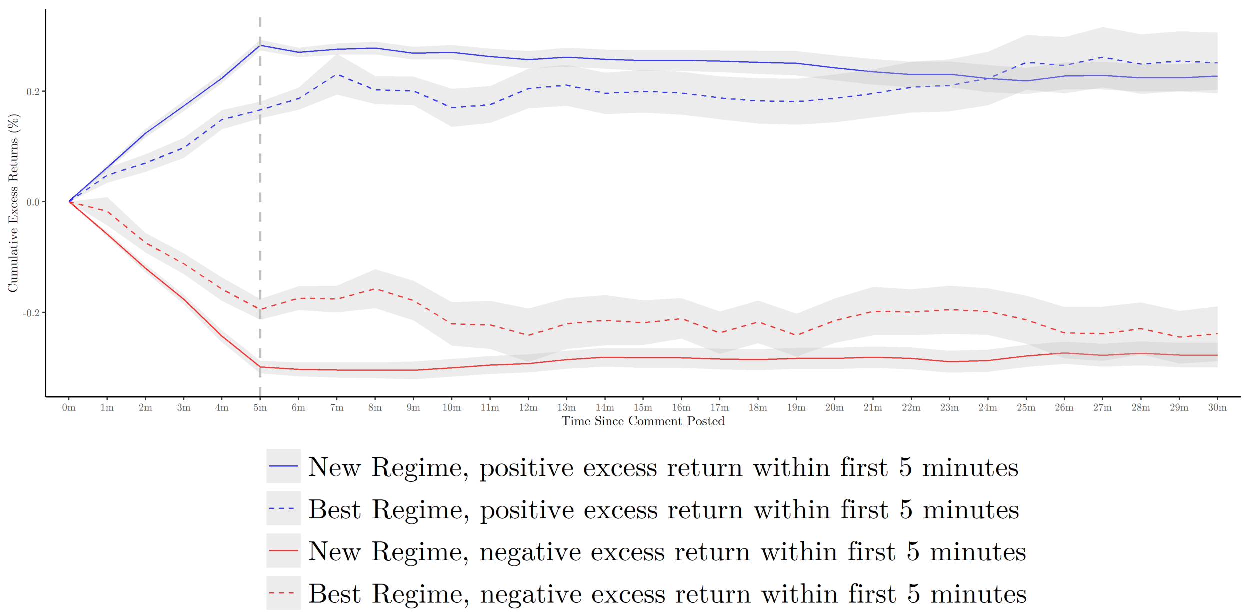
Figures

Figure 1.1: Daily Discussion unique ticker mentions over time



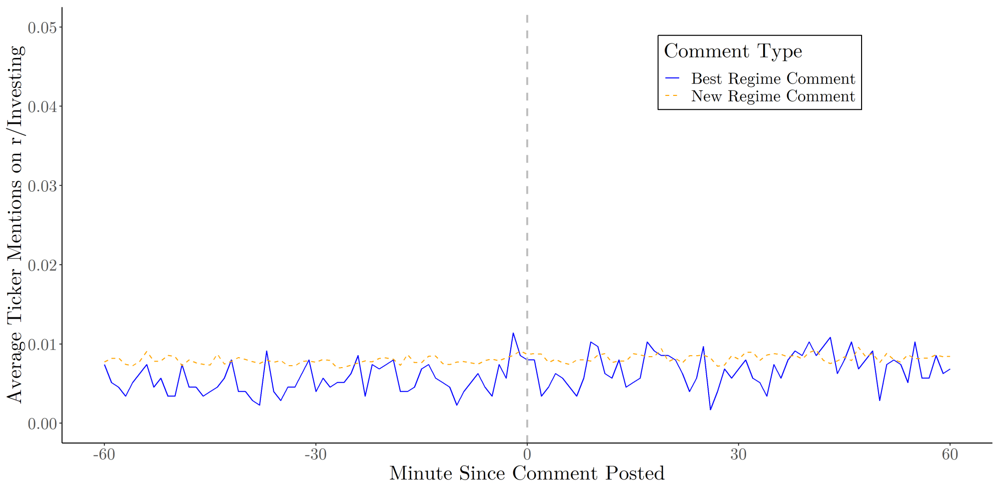
Note: The sample set consists of all Daily Discussion comments with a ticker tag during the 6 months surrounding the comment sort change (4/25/2018 - 10/25/2018). The vertical line indicates the day that the comment display changed from Best to New sort.

Figure 1.2: CAR after Daily Discussion comment publication

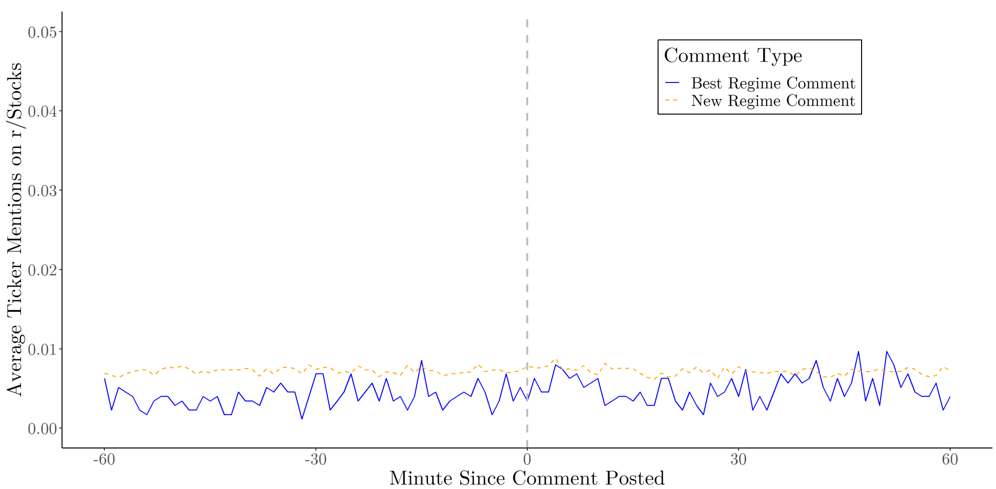


Note: This figure displays cumulative abnormal returns after comment publication sliced by the direction of the initial 5-minute abnormal return and sliced by comment regime for comments in the 6 months (3 months prior and 3 months after) the display change. Blue lines indicate positive ticker-comment events and red lines indicate negative ticker-comment events. Solid lines indicate New Regime comments and dashed lines indicate Best Regime comments. The sample set consists of top-level single ticker-tagged market hour Daily Discussion comments.

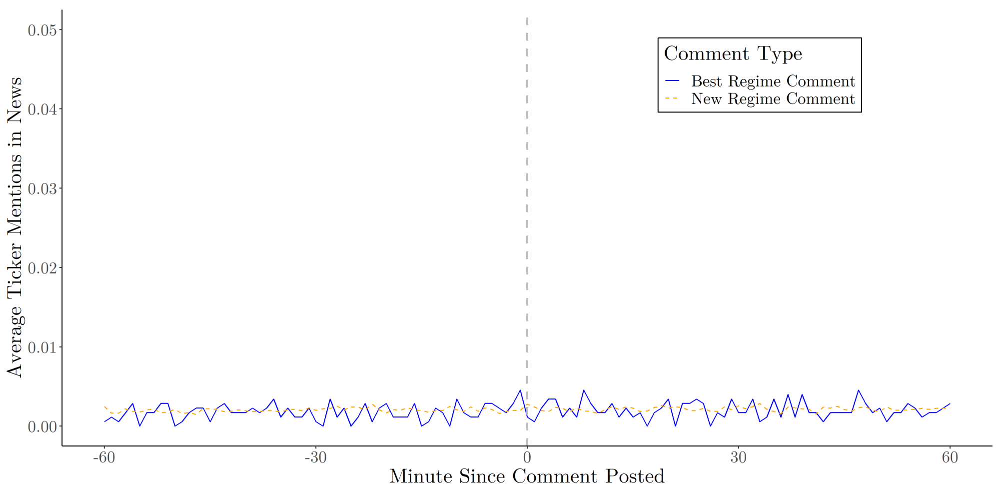
Figure 1.3: Daily Discussion ticker coverage timing



(a) Average coverage in r/Investing comments



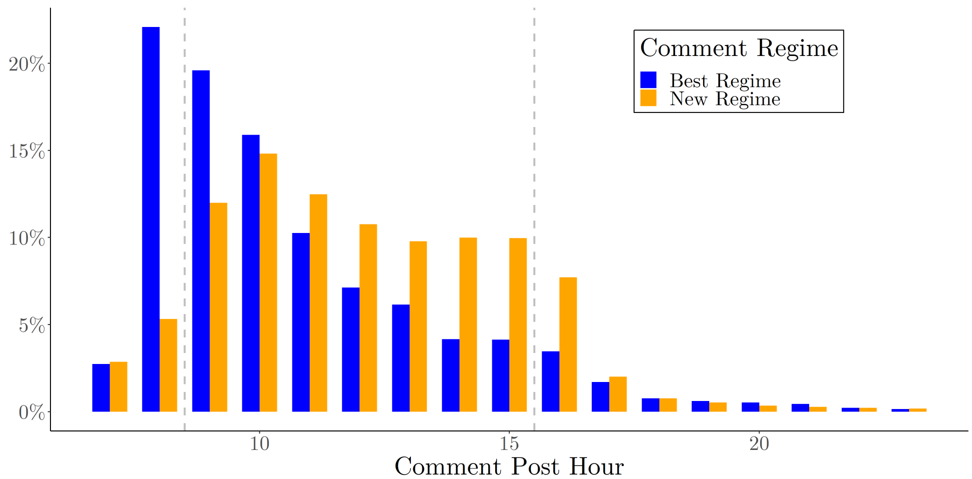
(b) Average coverage in r/Stocks comments



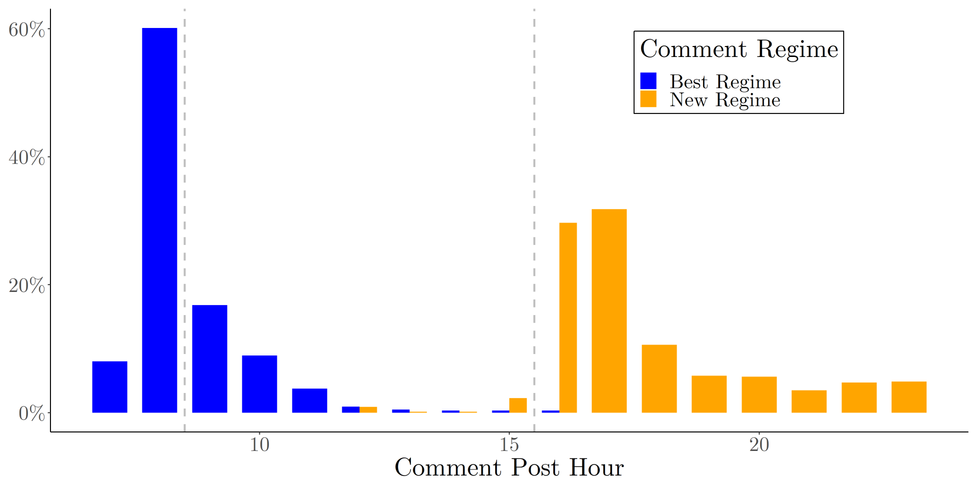
(c) Average coverage in news articles

Note: These figures display the coverage of tickers mentioned in Daily Discussion comments on other subreddits and news sources, by minute relative to Daily Discussion comment publication. Figure (a) tabulates ticker coverage in r/Investing comments. Figure (b) tabulates ticker coverage in r/Stocks comments. Figure (c) tabulates ticker coverage in news articles aggregated from Finnhub. Coverage of Best Regime comments is in blue and coverage of New Regime comments is in orange.

Figure 1.4: Daily Discussion comment hour distribution



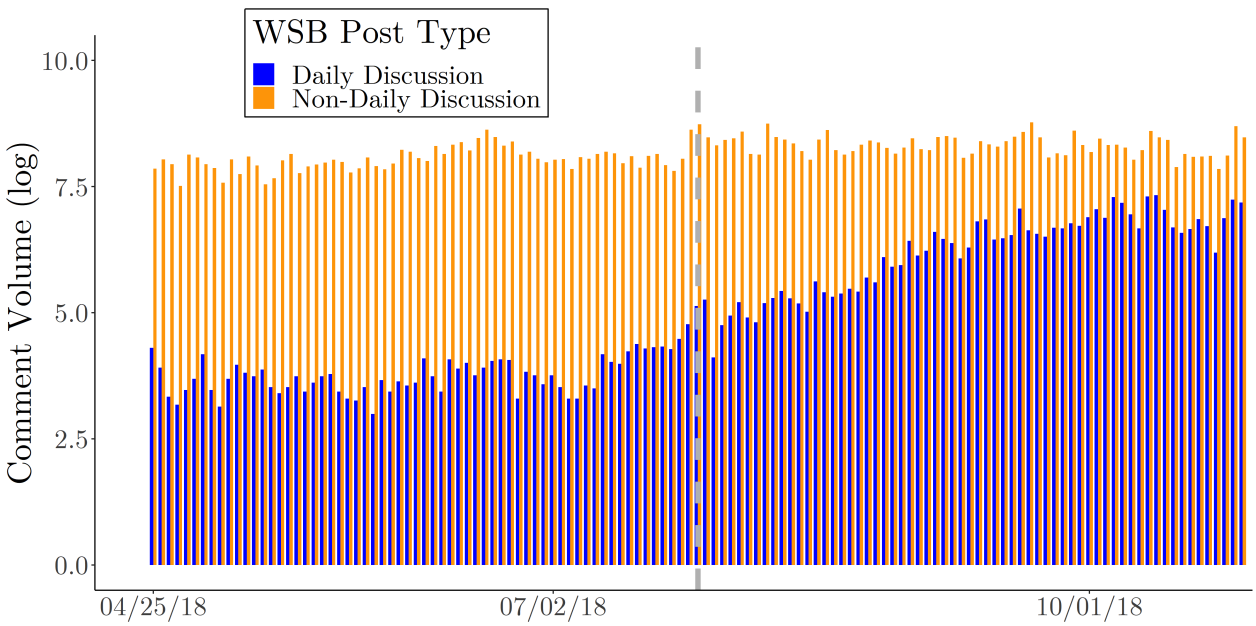
(a) All Daily Discussion comments



(b) Top 20 Daily Discussion comments at the top of the page at the end of each day

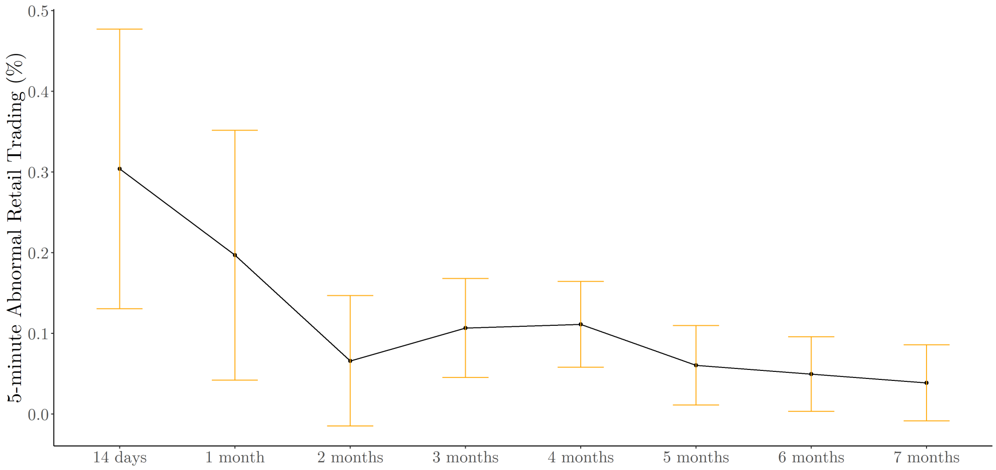
Note: These figures display comment hour distribution for Daily Discussion comments by Best and New Regime. The bar chart plots the percentage of Daily Discussion comments published by hour of day, separately for each regime. The blue bars represent Best Regime comments and the orange bars represent New Regime comments.

Figure 1.5: WallStreetBets daily comment volume

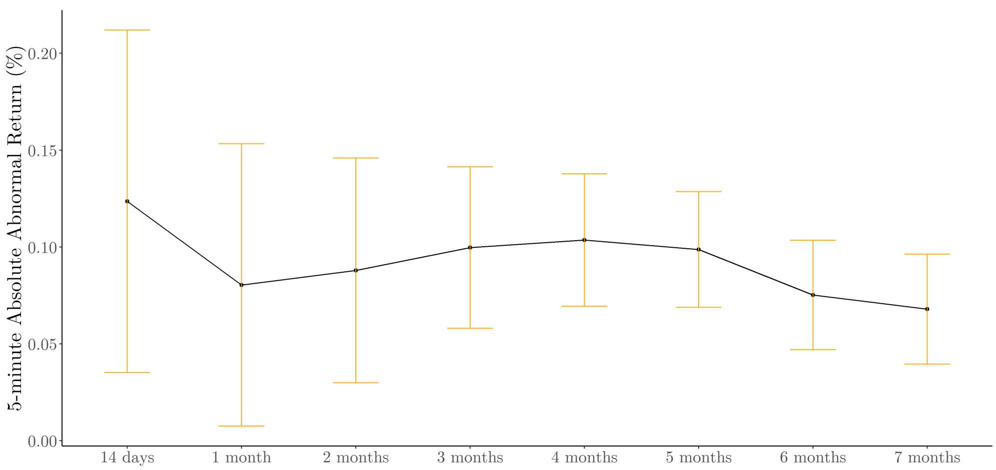


Note: This figure reports daily comment volume for Daily Discussion comments versus non-Daily Discussion comments, reported in log scale. Non-Daily Discussion comments are comprised of all comments from WallStreetBets posts that did not see a change to the default comment sort in my sample period. The vertical line indicates the day that the comment display changed. The blue bars represent Daily Discussion comments and the orange bars represent non-Daily Discussion comments.

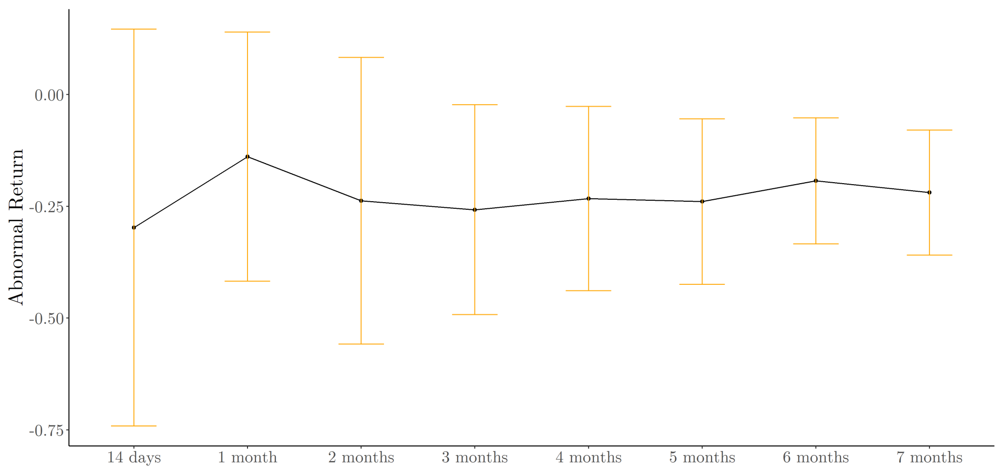
Figure 1.6: Market reactions to New Regime versus Best Regime comments over varying time periods



(a) Immediate Abnormal Retail Trading



(b) Immediate Absolute Abnormal Return



(c) Return Continuation

Note: This figure displays immediate retail trading, immediate absolute return, and return continuation responses to New Regime versus Best Regime comments over different time periods. I plot the differences in immediate and short-term reactions to Daily Discussion comments in the 14 days, 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, and 7 months before and after the comment display change. I estimate regressions from Equations 1.3, 1.4, and 1.5. For immediate retail trading and immediate absolute return, I plot estimates for β_1 in Equation 1.3 and 1.4, respectively, along with 95% confidence intervals. For return continuation, I plot estimates for β_3 (the interaction between New Regime dummy and 5-minute abnormal return) in Equation 1.5 along with 95% confidence intervals.

Tables

Table 1.1: WallStreetBets comment topic frequencies

Topic Name	Top Words	WSB	DD	DD Subset
Buy and Sell	call, bought, buy, at, put, sell, trading, market, money, holding	0.435	0.435	0.523
Stock Performance and News	earnings, up, tendies, fda, approval, shares, dip, profits, tariffs, run up	0.248	0.247	0.314
Mooning	moon, the moon, mooning, to moon, gonna moon, moon tomorrow, rocket emoji, rocket, tomorrow, gonna	0.044	0.044	0.078
Market Conditions	down, up, today, tomorrow, green, red, day, market, money, going	0.177	0.178	0.05
Pray, Drill, Drop	prayer, prayer circle, pray, recession, drilling, drill, halted, drop, dropping, dropped	0.026	0.026	0.019
Portfolio Performance	portfolio, my portfolio, entire portfolio, gains, down, week, loss, lose, your, day	0.023	0.023	0.008
Autist	austistic, austism, autist, autists, me, my autism, mods, thread, wsb, wallstreetbets	0.02	0.02	0.004
Option Expiration	expiry, expiration, expire, worthless, contracts, strike, strike price, iv, high iv	0.019	0.018	0.003
Short	short, shorting, long short, to short, you short, play, plays, next play, your play, what play	0.008	0.008	0.002

Note: WSB includes all WallStreetBets comments in my sample period (4/25/2018 - 10/25/2018). DD includes all Daily Discussion comments in my sample period. DD Subset includes all Daily Discussion comments in my main regression specifications (top-level single ticker-tagged market hour Daily Discussion comments mentioning common stocks with market data in CRSP and TAQ).

Table 1.2: Determinants of Daily Discussion comments

	$WSBCommentVolume_{i,d}$	
	(1)	(2)
$WSBCommentVolume_{i,d-1}$	0.007^{***}	0.006^{***}
	(0.0002)	(0.0002)
$NewsArticleVolume_{i,d-1}$	0.014^{***}	0.012^{***}
	(0.004)	(0.004)
$\|Ret\|_{i,overnight}(\%)$	0.217^{***}	0.178^{***}
	(0.004)	(0.005)
$\|Ret\|_{i,[d-2,d-1]}(\%)$	-0.022^{***}	-0.022^{***}
	(0.003)	(0.003)
$\|Ret\|_{i,[d-5,d-2]}(\%)$	0.005^{***}	0.005^{**}
	(0.002)	(0.002)
$AbnRetailVol_{i,d-1}$	0.558^{***}	0.397^{***}
	(0.015)	(0.016)
$\|RobinhoodUserHoldingChange\|_{i,overnight}$		0.211^{***}
		(0.006)
Firm Size	x	x
Date FE	x	x
Ticker FE	x	x
Observations	78,433	67,014
Note:	*p<0.1; **p<0.05; ***p<0.01	

Note: I estimate a Poisson regression, where $WSBCommentVolume_{i,d}$ is the outcome variable and is defined as the number of Daily Discussion comments that mention ticker i on date d. All regression specifications include controls for firm size as well as ticker and date fixed effects, and the sample set consists of ticker-date observations comprised of firms with at least one comment mention over the sample period (4/25/2018 - 10/25/2018). Standard errors are clustered by date and ticker and reported in parentheses.

Table 1.3: Relationship between Daily Discussion comments and retail trading

	$AbnRetailVol_{i,d}$		
	(1)	(2)	(3)
$WSBComment_{i,d}$	0.150***	0.124***	0.103***
	(0.011)	(0.014)	(0.010)
$WSBComment_{i,d}$ x $HighRHOwnership_{i,d-1}$		0.050***	
		(0.018)	
$WSBComment_{i,d}$ x $SmallFirm_{i,d-1}$			0.079***
			(0.020)
$WSBComment_{i,d-1}$	−0.007	−0.006	−0.004
	(0.009)	(0.010)	(0.009)
$NewsArticle_{i,d}$	0.051***	0.051***	0.051***
	(0.007)	(0.008)	(0.007)
$NewsArticle_{i,d-1}$	0.009	0.007	0.009
	(0.007)	(0.007)	(0.007)
$\|Ret\|_{i,overnight}(\%)$	0.187***	0.196***	0.187***
	(0.006)	(0.007)	(0.006)
$\|Ret\|_{i,[d-2,d-1]}(\%)$	−0.010***	−0.011***	−0.010***
	(0.002)	(0.002)	(0.002)
$\|Ret\|_{i,[d-5,d-2]}(\%)$	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)
$AbnRetailVol_{i,d-1}$	0.509***	0.511***	0.509***
	(0.012)	(0.013)	(0.012)
Firm Size	x	x	x
Date FE	x	x	x
Ticker FE	x	x	x
Observations	78,433	71,609	78,433
R^2	0.548	0.556	0.548
Note:			*p<0.1; **p<0.05; ***p<0.01

Note: I estimate an OLS regression, where abnormal retail trading is the outcome variable and is defined as the log of $RetailVol_{i,d}/\frac{1}{40}\sum_{k=1}^{40} RetailVol_{i,d-k}$. The main predictor variable is $WSBComment_{i,d}$, a dummy variable equal to 1 if there was a Daily Discussion comment mentioning ticker i on date d. All regression specifications include controls for firm size as well as ticker and date fixed effects, and the sample set consists of ticker-date observations comprised of firms with at least one comment mention over the sample period (4/25/2018 - 10/25/2018). Standard errors are clustered by date and ticker and reported in parentheses.

Table 1.4: Differential immediate market reactions following comments posted in the Best Regime versus New Regime

	$AbnRetailVol_{c,i,d,[t,t+5min]}$			
	(1)	(2)	(3)	(4)
$NewRegime_c$	0.149***	0.101***	−0.011	0.088***
	(0.032)	(0.031)	(0.037)	(0.031)
$NewRegime_c$ x $HighRHOwnership_{i,d-1}$			0.367***	
			(0.077)	
$NewRegime_c$ x $SmallFirm_{i,d-1}$				0.462**
				(0.200)
Firm Size		x	x	x
Same-Minute Comment		x	x	x
Lagged Comment		x	x	x
Half Hour FE		x	x	x
Industry FE		x	x	x
Observations	13,923	13,923	13,780	13,923
R^2	0.002	0.091	0.107	0.097
Note:				*p<0.1; **p<0.05; ***p<0.01

(a) Abnormal Retail Trading

	$\lvert AbnRet\rvert_{c,i,d,[t,t+5min]}(\%)$			
	(1)	(2)	(3)	(4)
$NewRegime_c$	0.104***	0.067***	0.013	0.048**
	(0.025)	(0.023)	(0.026)	(0.023)
$NewRegime_c$ x $HighRHOwnership_{i,d-1}$			0.209***	
			(0.055)	
$NewRegime_c$ x $SmallFirm_{i,d-1}$				0.653***
				(0.144)
Firm Size		x	x	x
Same-Minute Comment		x	x	x
Lagged Comment		x	x	x
Half Hour FE		x	x	x
Industry FE		x	x	x
Observations	13,923	13,923	13,780	13,923
R^2	0.001	0.215	0.218	0.224
Note:				*p<0.1; **p<0.05; ***p<0.01

(b) Absolute Abnormal Return

Note: Comparison of abnormal retail trading volume and absolute abnormal returns immediately (within the first 5 minutes) following comments posted in the last 3 months of Best Regime versus first 3 months of New Regime. The sample is restricted to top-level comments posted during market hours. Each observation is a ticker-minute where a comment c mentions ticker i on date d at time t (rounded to the nearest minute). Controls in columns (2)-(4) include firm size, same-minute comment (number of comments mentioning the same ticker in the same minute), lagged comments (dummy equal to 1 if another comment mentions the same ticker in the five minutes prior), industry fixed effects, and half hour fixed effects. Standard errors are clustered by ticker and date and reported in parentheses.

Table 1.5: Short-term return continuation after comments in the Best Regime versus New Regime

	$AbnRet_{c,i,d,[t+5min,t+15min]}$			
	(1)	(2)	(3)	(4)
$AbnRet_{c,i,d,[t,t+5min]}$	0.156***	0.157***	0.157***	0.157***
	(0.056)	(0.056)	(0.056)	(0.056)
$AbnRet_{c,i,d,[t,t+5min]}$ x $NewRegime_c$	−0.207***	−0.208***	−0.208***	−0.208***
	(0.057)	(0.057)	(0.057)	(0.057)
Firm Size				x
Lagged Comment				x
Same-Minute Comment				x
Half Hour FE		x	x	x
Industry FE			x	x
Observations	13,923	13,923	13,923	13,923
R^2	0.005	0.007	0.010	0.011
Note:		*p<0.1; **p<0.05; ***p<0.01		

(a) [5,15 min]

	$AbnRet_{c,i,d,[t+5min,t+30min]}$			
	(1)	(2)	(3)	(4)
$AbnRet_{c,i,d,[t,t+5min]}$	0.210*	0.209*	0.211*	0.211*
	(0.117)	(0.117)	(0.118)	(0.118)
$AbnRet_{c,i,d,[t,t+5min]}$ x $NewRegime_c$	−0.266**	−0.266**	−0.269**	−0.269**
	(0.118)	(0.118)	(0.118)	(0.118)
Firm Size				x
Same-Minute Comment				x
Lagged Comment				x
Half Hour FE		x	x	x
Industry FE			x	x
Observations	13,923	13,923	13,923	13,923
R^2	0.002	0.003	0.007	0.010
Note:		*p<0.1; **p<0.05; ***p<0.01		

(b) [5,30 min]

Note: Columns marked with (1) do not include any controls. Columns marked with (2) include half hour fixed effects. Columns marked with (3) also control for industry fixed effects. Columns marked with (4) also control for firm size, same-minute comments, and lagged comments. Standard errors are clustered by ticker and date and reported in parentheses.

Table 1.6: Short-term return continuation after comments in the Best Regime versus New Regime - sliced

	$AbnRet_{c,i,d,[t+5min,t+30min]}$			
	High RH	Low RH	Large	Small
	(1)	(2)	(3)	(4)
$AbnRet_{c,i,d,[t,t+5min]}$	0.560**	0.083	0.023	0.905**
	(0.270)	(0.096)	(0.130)	(0.384)
$AbnRet_{c,i,d,[t,t+5min]}$ x $NewRegime_c$	$-$0.618**	$-$0.187*	$-$0.078	$-$0.982**
	(0.271)	(0.098)	(0.130)	(0.385)
Firm Size	x	x	x	x
Same-Minute Comment	x	x	x	x
Lagged Comment	x	x	x	x
Half Hour FE	x	x	x	x
Industry FE	x	x	x	x
Observations	5,180	8,600	12,257	1,523
R^2	0.018	0.017	0.009	0.046
Note:				*p<0.1; **p<0.05; ***p<0.01

Note: I rerun the regression from Table 1.5 separately for high retail ownership firms (High RH), low retail ownership firms (Low RH), large firms (Large), and small firms (Small). High RH are firms on date d in the top 20% of Robinhood retail ownership as of date $d-1$ among all firms in my sample and Low RH are all other firms. Small firms are in the bottom 20% among all firms in my sample according to previous day's market capitalization and larger firms are all other firms. Controls include firm size, same-minute comments, lagged comments, industry fixed effects, and half hour fixed effects. Standard errors are clustered by ticker and date and reported in parentheses.

Table 1.7: Longer-term return continuation after comments in the Best Regime versus New Regime

	$AbnRet_{c,i,d,[t+30min,t+t_j]}$		
	$t_j = 1h$	$t_j = 2h$	$t_j = 3h$
	(1)	(2)	(3)
$AbnRet_{c,i,d,[t,t+30min]}$	0.028	−0.110	−0.161
	(0.043)	(0.076)	(0.148)
$AbnRet_{c,i,d,[t,t+30min]}$ x $NewRegime_c$	−0.058	0.033	0.045
	(0.043)	(0.077)	(0.148)
Firm Size	x	x	x
Same-Minute Comment	x	x	x
Lagged Comment	x	x	x
Half Hour FE	x	x	x
Industry FE	x	x	x
Observations	13,923	13,923	13,923
R^2	0.013	0.021	0.017
Note:			*p<0.1; **p<0.05; ***p<0.01

Note: I estimate the following regression:

$$AbnRet_{c,i,d,[t+30min,t+t_j]} = \beta_0 + \beta_1 AbnRet_{c,i,d,[t,t+30min]} + \beta_2 NewRegime_c + \beta_3 AbnRet_{c,i,d,[t,t+30min]} x NewRegime_c + controls + \epsilon_{i,d},$$

where $AbnRet_{c,i,d,[t+30min,t+t_j]}$ denotes the delayed abnormal return of ticker i mentioned in comment c from 30 minutes after publication to $t_j \in \{1, 2, 3\}$ hours after. $AbnRet_{c,i,d,[t,t+30min]}$ denotes the abnormal return within the first 30 minutes of comment c posting. $NewRegime_c$ is a dummy variable equal to 1 if comment c posted in the New Regime. Controls include firm size, same-minute comments, lagged comments, industry fixed effects, and half hour fixed effects. Standard errors are clustered by ticker and date and reported in parentheses.

Table 1.8: Differential market reactions following news articles published in the Best Regime versus New Regime

	$AbnRetailVol_{c,i,d,[t,t+5min]}$	$\lvert AbnRet\rvert_{c,i,d,[t,t+5min]}(\%)$
	(1)	(2)
$NewRegime_c$	0.022	−0.002
	(0.036)	(0.010)
Firm Size	x	x
Lagged Article	x	x
Half Hour FE	x	x
Industry FE	x	x
Observations	832	832
R^2	0.162	0.257
Note:		*p<0.1; **p<0.05; ***p<0.01

(a) Immediate Market Reaction

	$AbnRet_{c,i,d,[t+5min,t+t_j]}$	
	$t_j = 15min$	$t_j = 30min$
	(1)	(2)
$AbnRet_{c,i,d,[t,t+5min]}$	0.007	0.015
	(0.033)	(0.070)
$AbnRet_{c,i,d,[t,t+5min]}$ x $NewRegime_c$	−0.004	−0.023
	(0.050)	(0.107)
Firm Size	x	x
Lagged Article	x	x
Half Hour FE	x	x
Industry FE	x	x
Observations	832	832
R^2	0.070	0.091
Note:		*p<0.1; **p<0.05; ***p<0.01

(b) Return Continuation

Note: Comparison of immediate absolute abnormal returns, immediate abnormal retail trading volume immediately, and return continuations following news articles published in the last 3 months of Best Regime versus first 3 months of New Regime from the following news sources: Nasdaq, Dow Jones, Reuters, MarketWatch, Business Insider, and CNBC. The sample is restricted to news articles during market hours mentioning the same tickers as in my comment sample set. Controls include firm size, lagged articles (dummy equal to 1 if another article mentioned the same ticker in the five minutes prior to the article), industry fixed effects, and half hour fixed effects. Standard errors are clustered by ticker and date and reported in parentheses.

Table 1.9: Differential market reactions following comments posted in the morning versus afternoon

	$AbnRetailVol_{c,i,d,[t,t+5min]}$	$\|AbnRet\|_{c,i,d,[t,t+5min]}(\%)$
	(1)	(2)
$Afternoon_c$	-0.096	-0.138
	(0.103)	(0.097)
$Afternoon_c$ x $NewRegime_c$	0.155*	0.185**
	(0.089)	(0.085)
Firm Size	x	x
Same-Minute Comment	x	x
Lagged Comment	x	x
Half Hour FE	x	x
Industry FE	x	x
Observations	13,923	13,923
R^2	0.137	0.232
Note:		*p<0.1; **p<0.05; ***p<0.01

Note: Comparison of abnormal retail trading volume and absolute abnormal returns immediately (within the first 5 minutes) following comments posted in the afternoon versus morning between comment regimes. Abnormal retail volume and absolute abnormal returns are defined as before (see Table 1.4). $Afternoon_c$ is a dummy variable equal to 1 if comment c posted after 2 PM ET and $NewRegime_c$ is a dummy variable equal to 1 of comment c posted during the New Regime. Controls include firm size, same-minute comments, lagged comments, industry fixed effects, and half hour fixed effects. Standard errors are clustered by ticker and date and reported in parentheses.

Chapter 2

Social CEOs: Informative Signalers or Credit Takers?

2.1 Introduction

Social media is an increasingly popular communication tool through which firms present themselves to stakeholders. Not only has social media expanded the way firms can communicate with the public, social media has also provided CEOs, arguably the most visible face of the firm, a platform to instantaneously communicate with stakeholders without going through traditional intermediaries. While the majority of S&P 1500 firms have a corporate presence on social media channels, relatively few CEOs have established such a presence. This is at odds with the transparency and communication practices that stakeholders desire from CEOs in this day and age.[1,2] Among CEOs with social media, Tesla CEO Elon Musk is arguably the most influential CEO of a publicly-traded firm on Twitter with over 35 million followers.[3] He regularly tweets about a variety of topics including upcoming product launches and off-the-cuff thoughts. On the other hand, former Wells Fargo CEO Timothy Sloan leaves no social media footprint, not even an official LinkedIn profile. These examples illustrate the large heterogeneity among CEOs as to who chooses to use social media and how they choose to use it.

The natural question that follows is how does CEO social media presence and usage impact how CEOs and their respective firms are viewed with stakeholders, and in particular for this paper, investors. Put simply, do CEO social media posts matter for stock prices? This paper compiles a novel dataset of CEO tweets and presents summary statistics about the types of CEOs on Twitter and the content of their posts. I then focus on their tweets in the context of quarterly earnings announcements to assess the impact of tweets on stock prices and find that earnings-related tweets are not an informative signal incremental to the

[1]https://sproutsocial.com/insights/data/social-media-transparency/

[2]https://www.brunswickgroup.com/perspectives/connected-leadership/

[3]Follower count as of July 1, 2020.

earnings and thus do not move prices. Rather, their tweeting behavior is in line with wanting to 'take credit' for superior market reactions incremental to the earnings surprise. I tie this strategic behavior to CEO career management concerns and find preliminary evidence that suggests that strategic CEO Twitter presence is associated with up to a 4.1% reduction in turnover probability, and a reduction in CEO-performance sensitivity. Taken together, these results suggest that CEOs who are social media savvy and play an active role in cultivating their personal brand and communicating their value to stakeholders can extract career benefits.

This paper examines tweets by 243 CEOs comprising the universe of CEOs of S&P 1500 firms on Twitter between 2008 and 2019. I find that CEOs on Twitter tend to be younger, better compensated, and more overconfident in their stock option exercise behavior. They also tend to manage firms that are larger, have more investment opportunities, spend more on research and development and advertising, are headquartered in California or Washington, and belong to the tech industry. Of the over 210,000 tweets collected, the vast majority of them garner little engagement (receiving likes, retweets, and replies). The small fraction of tweets that get the most attention are mostly business-related tweets about corporate image, products, and strategy and performance.

To investigate empirically whether and how tweets matter, I focus on tweets pertaining to quarterly earnings announcements targeted primarily at potential and current investors. Daily event-study methodology of 200 CEOs that tweet around quarterly earnings announcements establishes that within a CEO's tenure at a firm, quarters in which CEOs tweet about earnings following the earnings announcement correspond with a $1.5-2.6\%$ higher industry-adjusted announcement return conditional on the level of earnings surprise. This finding is statistically significant and economically meaningful, and increases with the following CEO tweeting characteristics: earnings tweet intensity, follower network size, and potential audience size. CEO-firm fixed effects allow me to isolate the consequences of CEO tweeting decisions while holding firm idiosyncrasies constant. Moreover, quarters in which CEOs tweet about earnings following the earnings announcement are not followed by a stronger post-earnings announcement drift in either direction suggesting that the higher announcement return in earnings tweet quarters is not associated with an overreaction or underreaction to earnings news or CEO tweets.

One concern with this daily-level analysis is that selection into joining Twitter is endogenous and that the firm-CEO and time fixed effects along with the rich set of controls used in the baseline results do not fully eliminate omitted variable concerns. I address this by using propensity-score nearest neighbor matching and find similar, albeit dampened results.

Because earnings tweets and short-term stock price reactions occur contemporaneously when measured at the daily level, I am not able to establish causality between earnings tweets and announcement returns without looking at more granular data. Either earnings tweets are an informative signal incremental to the earnings ('informative signal' story) or CEOs tweet following superior quarters incremental to the earnings surprise ('taking credit' story). Since tweets are timestamped, I perform an intraday event study around individual tweets to disentangle the relationship between earnings tweets and stock price reactions. Intraday trading data allows me to more cleanly identify in which direction the relationship between earnings tweets and stock price reactions runs. Using 30-minute stock price intervals, I find that CEOs tweet about earnings following a positive market run up. The stock returns after earnings tweets are insignificantly negative, which suggests that earnings tweets do not contain new information. One interpretation is that CEOs tweet about earnings when it is a particularly good quarter incremental to the earnings surprise in order to 'take credit' for the superior news.

I address some challenges to my identification strategy. One concern is that earnings tweets simply coincide with generally positive price movements following the release of earnings news. I mitigate this concern by showing that the average return in the hours after an earnings announcement unconditional on the earnings surprise is significantly lower in quarters in which CEOs do not tweet about earnings.

Another concern is that CEO tweets are a response to the earnings surprise. While it is true that the likelihood of a CEO tweeting about earnings increases with the earnings surprise, I show that the average return in the hours after an earnings announcement is significantly lower in quarters in which CEOs do not tweet about earnings after conditioning on the sign of the earnings surprise. This suggests that the decision to tweet about earnings is in response to the stock price reaction above and beyond the earnings surprise.

I perform a number of additional robustness checks to mitigate concerns of my pooled results. For one, it could be the case that tweets that occur after normal trading hours have different price dynamics compared to tweets posted during normal trading hours since the makeup of market participants and general levels of trading are starkly different (Gregoire and Martineau, 2020). To address this concern, I restrict my sample of earnings tweets to those that occur during normal trading hours and find similar results. Second, it could be the case that earnings tweets from CEOs with more followers actually do cause price movement or increased trading volume. I redo my analysis with only earnings tweets from highly-followed CEOs and the results still hold. It could also be the case that CEOs that tweet about earnings more than once in a single day are tweeting for reasons other than stock price reactions. When I restrict my sample to only single tweet days, I find even stronger results, further supporting my baseline findings.

I address some alternative explanations that might better explain CEO decisions to tweet about earnings. One alternate explanation is that CEOs tweet about earnings in response to firms tweeting about earnings. I do not find that firm earnings tweets preceding CEO earnings tweets lead to any incremental price impact, suggesting that CEOs are not reacting to a price reaction caused by firm earnings tweets. But it is entirely plausible that firms and CEOs co-time their tweets to both capitalize on good news. A second alternate explanation is that CEOs tweet about earnings in response to media articles published about the firm or CEO. I find that there are elevated levels of news articles published on earnings announcement dates with a CEO earnings tweet. Like with firm tweets, I do not find that news articles published prior to CEO earnings tweets cause any incremental price impact, suggesting that CEOs are not reacting to a price reaction caused by news articles. The higher number of news articles that coincide with CEO earnings tweets are likely complementary actions: superior performance induces more media coverage and CEOs to tweet.

Since it appears that CEOs strategically tweet around earnings announcements, I attempt to tie this behavior back to their career by investigating CEO turnover. I find that CEOs active on Twitter in a given year are less likely to be fired the following year. I instrument CEOs being active on Twitter in a given year using the proportion of CEOs that tweet in the same industry in the same state, whether there were any product recalls in the same industry, and whether there was a Superbowl in the same state that year. This suggests that while tweets may not have short-term effects, they might have beneficial effects for the CEO in the long-run.

This paper is most similar to Bhagwat and Burch, 2016, Jung et al., 2018, and Wolfskeil, 2020, who look at tweets from firm corporate Twitter accounts and find that quarters in which firms tweet about earnings are associated with higher announcement returns for certain earnings surprises. In contrast to those studies, this paper focuses on tweets from a CEO's personal Twitter account, covers a longer time period, and finds a similar pattern for all earnings surprises. While those studies rely on daily event studies to claim a causal relationship between firm earnings and announcement returns, I supplement daily-level analysis with intraday trading data to provide a cleaner setting in which to identify the direction in which that relationship runs. While those papers conclude that firm earnings tweets have a non-negative effect on announcement returns, I argue that CEO earnings tweets do not affect announcement returns but simply reflect firm fundamentals incremental to the earnings surprise. I also replicate the same intraday analysis on firm earnings tweets and find that they might have an effect on stock prices, but it is a largely negative one.

This paper also contributes to the broader literature on the relationship between communication and financial markets. One main focus has been on information intermediaries with the overall consensus that they can influence market participant behavior (e.g., Dyck and Zingales, 2003, Antweiler and Frank, 2006, Veldkamp, 2006, Tetlock, 2007, B. M. Barber and Odean, 2008, Fang and Peress, 2009, Lawrence et al., 2018, A. Fedyk, 2019). Another focus has been on corporate communication (see Loughran and McDonald, 2016 for a survey). For instance, oral information content in earnings conference calls (e.g., Mayew and Venkatachalam, 2012, Price et al., 2012, Druz et al., 2020) and tone and readability of written communication in 10-K filings (e.g., F. Li, 2008, Loughran and McDonald, 2014) and earnings press releases (e.g., Demers and Vega, 2011, Davis et al., 2012) have been shown to predict future firm performance and influence analyst estimates. As the communication landscape has evolved with the inclusion of social media as an accepted form of dissemination by the SEC in 2014,[4] to my knowledge, this paper is one of the first to shed light on social media communication and its relationship to financial markets. Practitioners and economists (e.g., Bertrand and Schoar, 2003) agree that CEOs matter for firm performance and value, with surveys of non-CEO executives attributing almost 50% of firm value to the CEO[5] and empirical evidence showing that only CEOs amongst executives are key drivers

[4]https://www.sec.gov/investment/im-guidance-2014-04.pdf

[5]https://www.webershandwick.com/wp-content/uploads/2018/04/ceo-reputation-premium-executive-summary-3.

of firm performance (Bennedsen et al., 2020). Yet there has been little documented evidence of strategic communication behavior directly attributed to the CEO.

The setting of my empirical analysis is quarterly earnings announcements because they are highly-followed, value-relevant, and frequently occurring events in which managers have already been shown to to be strategic prior to the release of earnings. Niessner, 2015 and deHaan et al., 2015 document managers strategically timing bad news around times of lower investor attention. Kirschenheiter and Melumad, 2002 and Harbaugh et al., 2016 document managers strategically smooothing earnings across time and across segments within a report, respectively. Unlike those studies, I study the behavior of CEOs after the release of earnings news as opposed to before and document CEO behavior consistent with strategically 'taking credit' for superior quarters.

The main methodological approach of this paper relates to Bianchi et al., 2020, who use high-frequency identification around Donald Trump tweets to show that Trump influences expectations about monetary policy. This type of discontinuity-based estimation has long been used in asset-pricing papers to identify effects of monetary shocks (Kuttner, 2001, Cochrane and Piazzesi, 2002, Faust et al., 2004, Gürkaynak et al., 2007, Nakamura and Steinsson, 2018). More recently, studies on price formation have employed a high-frequency approach using intraday data to study the effects of after-hours earnings announcements (e.g., Gregoire and Martineau, 2020), analyst conference calls (e.g., Matsumoto et al., 2011), SEC insider trading filing releases (e.g., Rogers et al., 2016, Rogers et al., 2017), and analyst recommendations (e.g., Altınkılıç et al., 2013, E. Li et al., 2015). I use intraday trading data in this paper, but instead focus on establishing causality between CEO earnings tweets and stock price reactions after earnings are realized.

After showing that CEOs do not appear to have a significant effect on stock prices incremental to the earnings through tweeting, my findings that CEOs instead tweet in response to positive stock price movements seems to suggest that CEOs could still possibly extract value from tweeting by taking a more long-term perspective. For instance, can strategic social media presence, such as taking credit for good firm news, help a CEO survive poor firm performance? The general consensus is that CEO turnover is sensitive to firm performance, and also to industry and market performance (e.g., Kaplan and Minton, 2012, Jenter and Kanaan, 2015, Gao et al., 2017, Graham et al., 2020). Dikolli et al., 2014 also shows that the likelihood of being fired decreases with CEO tenure. I add to this literature by introducing a new variable as a predictor of CEO turnover and find a significant and negative relationship between CEOs who tweet and CEO turnover conditional on performance, board, and governance measures.

The rest of this chapter is organized as follows. Section 2.2 provides details on the data collection process and methodology. Section 2.3 presents my empirical findings. Section 2.4 concludes the paper.

---

pdf

2.2 Data and Methodology

The overall goal in this paper is to understand the role of social media in CEO behavior and market reactions. For practical reasons, I focus on the Twitter platform. Twitter was founded in 2006 as a social networking platform meant to connect people and allow them to share their thoughts with a large audience. Users post messages of up to 280 characters (increased from 140 characters in 2017) and other users can then like, reply to, or retweet that post. According to website analytics tool Alexa, Twitter is the 11th most visited website in the US, with 233 million monthly site visits as of June 2019. The average user spends approximately 6.5 minutes on Twitter daily, and Twitter shares a majority overlap in visitors and search keywords with Facebook, Instagram, and LinkedIn. What differentiates Twitter from those other social media platforms and makes it particularly interesting for this research setting is Twitter's strong emphasis on real-time information. In fact, Twitter is the preferred platform for individuals to get their real-time news[6] and has become a preferred medium of choice for potential investors to discuss financial news. For instance, TD Ameritrade stated that it sees more than 80% of its social media chatter volume on Twitter.[7]

The CEO Twitter dataset consists of all tweets from the personal Twitter account of CEOs of S&P 1500 companies between January 1, 2008 until December 31, 2019. To create the CEO Twitter dataset, I start with a list of all CEOs of S&P 1500 companies in office at some point in the aforementioned time period, which comes from the Compustat Execucomp database. I append to the list Tesla CEO Elon Musk because while Tesla was not on the S&P 1500 list as-of December 2019, it joined the S&P 500 soon after and Musk's tweets are among the most covered by traditional media. I then manually identify CEO Twitter accounts using Twitter's name search engine and narrow down potential candidates by examining the following criteria: verification badge, bio description, profile picture, size and identities of followers, identities of accounts they follow, content of tweets. If needed, I also cross reference LinkedIn and company websites to see whether those pages provide a link to the CEO's Twitter account. For example, when I search for 'Marc Benioff,' the CEO of Salesforce.com, I get 10 hits with that exact or similar name. I am able to distinguish his real account from other people or impersonators because the real account is Twitter-verified (a blue badge appears next to account name in bio description), has around one million followers, is followed by salesforce.com's official corporate Twitter account, and says 'ceo@salesforce.com' in his account bio.[8]

Out of the $4,427$ unique CEOs in my search set, I find 330 potential CEO Twitter accounts. I proceed to download all original tweets from those accounts. An original tweet is defined as a post created by the user that contains text, photo, and/or video. Retweets (a post where a user shares another user's tweet without providing additional commentary or context) are excluded from this analysis to focus on original content from a CEO's personal account. I then subset the downloaded tweet dataset to only include tweets posted between

[6]www.americanpressinstitute.org/publications/reports/survey-research/how-people-use-twitter-in-general/
[7]www.thestreet.com/investing/stocks/using-twitter-to-trade-stocks-14491774
[8]www.twitter.com/Benioff

January 1, 2008 and December 31, 2019. Furthermore, I only keep CEO tweets during a CEO's tenure at the firm and accounts with a matching CRSP PERMNO. I end up with $210,393$ tweets from 243 CEO-firms (239 CEOs and 224 firms). Each observation in the data contains the content of a tweet, the date and time it was posted, and the number of likes, replies, and retweets it received as of the download date. As a result, the data only contain tweets that have not been deleted as of the download date and the engagement measures (likes, replies, and retweets) are taken as of the download date.[9] For each CEO account, I also download the historical number of followers using archive.org.[10]

Figure 2.1a depicts the percentage of S&P 1500 CEOs that posted at least one tweet in a calendar year. CEOs of S&P 1500 firms are increasingly on Twitter with over 9% of CEOs on Twitter in 2018 (see Appendix Table B.2). To further get a sense of Twitter usage over time, Figure 2.1b shows the evolution of tweet volume by year. The annual number of tweets are scaled by the number of CEOs that posted at least once in that year. The volume of CEO tweets shot up around 2011, and since then has fluctuated slightly. Interviews from CEOs with Twitter and public relations experts seem to indicate that while some CEOs have a lot of autonomy over their social media accounts, the majority of CEO Twitter posts, especially tweets related to value-relevant events, are drafted with guidance from the firm's communications team with the final draft of the post approved by the CEO.[11] Figures 2.1c and 2.1d suggest that CEO tweets are not solely restricted to business hours and days, though the volume of tweets on weekends and outside of business hours is lower.

Table 2.1 reports summary statistics on CEO tweets. There is large heterogeneity in how active and popular an account is and how much engagement an account receives. CEOs have $279,297$ followers on average, but the median is just $1,818$ followers. The most followed CEO is Elon Musk (Tesla) with over 30 million followers at the end of 2019 while the least followed CEO is Gary Newsome (Health Management) with only 12 followers. The average CEO joined Twitter in 2014, which coincides with the SEC guidance issuance on social media practices. The earliest adopters of Twitter joined in 2008 and include 2 tech firm CEOs (Robert LoCascio of Liveperson and Jonathan Schwartz of Sun Microsystems) and 2 market research firm CEOs (George Colony of Forrester Research and Gail Goodman of Constant Contact). CEOs tweet less than once per day on average, but John Legere (T-Mobile) usually tweets over 20 times a day. As expected by how right skewed the number of followers some CEOs have, some CEOs' tweets receive a lot more attention in terms of likes, retweets, and replies. The CEO who gets the most likes on average is Elon Musk (Tesla) and the CEO who gets the most retweets and replies on average is Warren Buffett (Berkshire Hathaway) even though Buffett has only tweeted 6 times and has not tweeted since 2016.

I next try to make sense of what CEOs tweet about by using textual analysis to categorize each tweet. After manual inspection of a random sample of tweets, I develop the following classification scheme using a dictionary of keywords and phrases consisting of 11

[9] Tweets downloaded date between 1/1/2020 - 1/31/2020.

[10] archive.org is a non-profit library of millions websites and houses older versions of webpages.

[11] https://www.washingtonpost.com/business/2019/02/26/most-ceos-are-rightfully-afraid-twitter-not-elon-musk/

categories: 'products', 'strategy and performance', 'corporate image', 'external validation', 'self promotion', 'customer interaction', 'politics', 'refer to peers', 'information and links', 'general non-business', and 'personal'. Appendix Table B.1 reports the breakdown of tweets by content category, of which 15% were not able to be classified. In terms of volume, tweets belonging to 'customer interaction' and 'information and links' comprise over half of the tweets I was able to classify. However, taking a look at engagement measures yields that the tweets bringing in the most engagement belong to the 'corporate image', 'products', and 'strategy and performance' categories. These three categories are all intrinsically business-related, yet their primary target stakeholder groups are arguably different. 'Corporate image' tweets are aimed at current and prospective employees, 'product' tweets are aimed at current and prospective consumers, and 'strategy and performance' tweets are aimed at current and prospective investors. All these stakeholders groups are critical to firm wealth and performance. For this paper, I investigate tweets targeted at investors.

Earnings-Related Tweets

For the majority of the paper, I will focus on tweets pertaining to quarterly earnings announcements. Going by the tweet categories from above, such tweets fall under the 'strategy and performance' category, which comprise around 2% of all CEO tweets (see Appendix Table B.1). Though this is a small portion of all tweets collected, this set of tweets provides the most ideal setting in trying to measure the effect of tweeting on investor-relevant events. Because quarterly earnings are frequently-occurring and value-relevant events that the firms in my dataset are required to participate in, I am able to construct an unbalanced panel dataset consisting of different CEOs and can observe their tweeting decisions in each quarter within their time at the firm.[12] While other types of tweets fall into the 'strategy and performance' category such as mergers and acquisitions, and board of director hiring announcements, tweets about the information content in earnings announcements are the most frequent and consistent. For instance, in the first 24 hours after an earnings announcement, nearly one-third of all tweet in that period are related to earnings.

To identify earnings-related tweets, I restrict my search period to the 14 days before and after the earnings announcement. This helps to ensure that the keywords used to identify earnings tweets do not misclassify tweets as earnings-related in times when individuals are less likely to be paying attention to firm earnings and when those keywords could be used to talk about something entirely else.[13] I end up with $1,977$ CEO tweets about quarterly earnings of which $1,313$ occur at some point immediately following the earnings announcement up until the end of trading (4 PM ET) the following business day ('on announcement

[12]Very few CEOs that tweet move between firms in my sample, so I am not able to distinguish between CEO and firm fixed effects, but I can use CEO-firm fixed effects to observe CEO decisions holding firm fixed.

[13]The keywords (and its variations) used to identify an earnings-related tweet are: 'earnings', 'eps', 'quarter 1', 'quarter 2', 'quarter 3', 'quarter 4', 'revenue', 'sales', 'income', 'cash flow', 'profit', 'ebit'.

date').[14] Figures 2.2c and 2.2d show that tweets specifically related to earnings news occur much less frequently on weekends and more frequently in the immediate hours prior to and after normal trading hours. This makes sense because almost all quarterly announcements occur after hours (98%). Appendix Table B.5 shows that 477 earnings tweets occur within the first hour after the earnings announcement (or 36% of earnings tweets 'on announcement date') and over half of all earnings tweets 'on announcement date' occur within the first 4 hours after the earnings announcement.

The main focus will be on the earnings tweets that occur 'on announcement date' as opposed to the earnings tweets preceding the announcement ('before announcement date') and tweets that occur further out after the announcement ('after announcement date'). This is because tweets 'before announcement date' are not predictive of and tweets 'after announcement date' are not predicted by earnings news or earnings-related market reactions (see Appendix Table B.12). Appendix Table B.6 breaks down the frequency in which CEOs tweeting about earnings 'on announcement date'. Over half (57%) of the CEOs that have ever tweeted have never tweeted about earnings while the majority of CEOs who have tweeted about earnings do so only sporadically. This variation in tweeting behavior within CEO at a firm is a prime setting in which to observe what prompts CEOs to tweet and how market participants react.

Other Datasets

Quarterly earnings announcement dates and analyst consensus forecasts are obtained from Compustat and I/B/E/S, respectively. The timestamp of quarterly earnings announcements reported in Compustat are typically taken from a newswire source. DellaVigna and Pollet, 2009 have shown that there can be inaccuracies in the reported dates of earnings announcements. To mitigate any reporting errors, I only keep observations in which the announcement date reported in Compustat and I/B/E/S are within one day of each other. Annual accounting variables come from Compustat, market capitalization and daily stock prices come from CRSP, intraday stock prices and volume come from TAQ, CEO demographic and compensation variables come from Execucomp, institutional ownership data come from Thomson Reuters, and governance index[15] and board independence measures come from Insitutional Shareholder Services. For any manually added CEOs in my search set, I fill in missing Execucomp data using Edgar SEC Def 14A filings.

For general summary statistics and CEO turnover analysis, I construct firm-year observations and define a CEO as active on Twitter for a particular year if he tweeted at least once in that year. For daily event study analysis around quarterly earnings announcements, I construct firm-quarter observations and define a CEO as active for a given firm-quarter if the announcement date fell in between the CEO's first and last observed tweet dates. I also

[14]For earnings announcements that occur after 4 PM ET, I define the earnings announcement date as the following business day.

[15]We follow Gompers et al., 2003 and define governance index as the number of governance provisions that restrict shareholder rights.

construct measures to indicate whether or how much the CEO tweets about earnings in a given quarter. I also restrict the firm-quarter dataset to quarterly earnings announcements with at least one analyst forecast in the 90 days before the announcement, so the firm-quarter dataset contains fewer CEO-firms than the firm-year dataset that does not have any analyst coverage requirements. For intraday event study analysis around earnings tweets, I further restrict the number of earnings tweets used because intraday data from TAQ is only available starting in 2010, so I drop tweets before that. I also drop earnings tweets on weekends and on days in which the share price falls below \$5.

Measuring Earnings Surprise

In order to assess how well a firm was able to meet investor expectations, I use the standardized unexpected earnings (SUE) measure. I compute SUE as follows:

$$SUE_{i,q} = \frac{e_{i,q} - E(\hat{e}_{i,q})}{P_{i,q}}$$

where $SUE_{i,q}$ is defined as the actual EPS for firm i in fiscal quarter q ($e_{i,q}$) less the average EPS analyst forecast of the most recent analyst estimates within 90 days prior to the earnings announcement date ($\hat{e}_{i,q}$) all scaled by the share price of the firm from three trading days prior to the announcement (DellaVigna and Pollet, 2009). I focus on the earnings surprise rather than the level because whether a given level of earnings is good news or bad news depends on the level relative to investor expectations. In addition, stock prices days prior to the announcement should reflect all information prior to the earnings announcement so scaling by price accounts for the fact that a given amount of earnings surprise implies different magnitudes depending on the price per share. For example, a positive 10 cent EPS surprise represents a larger surprise if the stock price is valued at \$10 per share than if it is valued at \$100 per share.

When CEOs tweet about earnings, the most common hard number communicated in the 280 character limit is the EPS (see Figure 2.3 for examples). Moreover, the traditional financial press typically reports earnings announcement news as the difference between actual earnings and estimates. So earnings is likely to be the most salient performance measure presented to investors. Oftentimes, these media outlets report unscaled performance measures, so investors may also pay attention to the raw unscaled earnings surprise. My main measure of SUE is price-scaled, but my results are robust to alternative measures of earnings surprise (untabulated).

From the values of SUE calculated, I define a 'good news' quarter when SUE > 0, and a 'bad news' quarter otherwise. Traditionally in earnings announcement effect studies, firms are split into deciles or quintiles based on their SUE measure in each quarter (Hirshleifer et al., 2009). With tweets, exactly meeting expectations (and being above or below that marker) is very salient due to the brevity of tweets. But it is unclear with the traditional decile methodology in which decile(s) the zero SUE firms lie in each quarter. So I modify that methodology slightly by sorting firms in each quarter into one of 7 bins ('news surprise magnitude') based on their SUE. A news surprise magnitude equal to zero means that a firm exactly meets expectations. For firms that do not meet expectations in a quarter, they are sorted into one of two equal-sized bins where a news surprise magnitude equal to -2 represents the very worst news. For firms that exceed expectations in a quarter, they are sorted into one of four equal-sized bins with a news surprise magnitude equal to +4 representing the very best news. This asymmetric binning is done to keep the size of each of the two negative and four positive bins roughly equal due to the number of 'good news' quarters exceeding the number of 'bad news' quarters by about two-to-one.

Identification Strategy

In order to identify the impact of tweets, I use a high-frequency event study approach with the identifying assumption that no other systematic shocks to the market occur within a particular window. The pre-event window is between $T_0 = T_1 - h$ and T_1 where $h \in (2, 6, 12)$ hours and T_1 is the nearest 30 minute time interval prior to tweet time t. The post-event window is between T_2 and $T_3 = T_2 + h$ where $h \in (2, 6, 12)$ and T_2 is the nearest 30 minute time interval after tweet time t. The pre-event price change is measured by computing the raw return between the two price observations that correspond to the trades closest to but not exceeding T_0 and T_1. The post-event price change is measured analogously using the two price observations that correspond with T_2 and T_3. I also compute a measure of price volatility equal to the absolute value of the price change for both the pre-event and post-event time periods. I also compute a trading measure, defined as share turnover, by summing up all trades in the pre-event or post-event periods and scaling by the total number of shares outstanding at the end of the previous trading day. All return and trading measures are purged of time-of-day fixed effects by regressing them on a set of dummy variables for hour and month. The event study is carried out on the residuals from these regressions, which ensures that the results are neither driven nor confounded by time-of-day or seasonal patterns.

2.3 Results

Who Is On Twitter?

Among CEOs Who Tweet

I start by documenting the types of CEOs who tweet. Appendix Figure B.1 shows the distribution of CEOs by their firm headquarters zip code. Appendix Figure B.1a shows that in terms of CEO counts, there are large concentrations of tweeting CEOs in larger metropolitan cities. However, when weighted by follower size, CEOs of firms located in Northern California and Washington are disproportionately over-represented compared to other cities (see Appendix Figure B.1b). When compared to the total universe of S&P 1500

CEO-firms, a larger fraction of CEOs of firms located in California and Washington as well as Massachusetts, Pennsylvania, and Virginia are on Twitter compared to the nationwide average (see Appendix Table B.3a). I further look at the distribution of tweeting CEOs by industry and use the Fama French 12 industry classification to document that twice as many CEOs of tech firms (Business Equipment, Telephone and Television Transmission) are on Twitter while half as many CEOs of firms in highly regulated industries (Finance, Utilities, Oil, Gas, and Coal) are on Twitter compared to the nationwide average (see Appendix Table B.3b).

For CEOs who tweet, Table 2.2 presents correlates between tweeting measures, CEO demographics, and firm demographics. The first set of correlates compares tweeting features and find that the number of followers a CEO has is positively and significantly correlated with the number of replies, retweets, and likes the CEO receives, the volume of tweets sends out, the fraction of days in which the CEO tweets, and the longer the CEO has been on Twitter (Table 2.2a). In other words, the more time the CEO puts into tweeting, the more followers the CEO has. Since all of those tweeting characteristics are positively and significantly correlated with one another, I use the number of Twitter followers to present further correlations between Twitter and other demographics (Tables 2.2b and 2.2c). The number of followers a CEO has is positively and significantly correlated with the CEO's stock ownership and total compensation. The number of followers is also positively and significantly correlated with firm size, leverage, and retail ownership. These results suggest that popular or celebrity Twitter CEOs have a larger stake in the firm, manage larger firms, and manage firms with more retail ownership.

Of the 243 CEO-firms in my sample set, 157 (64.6%) of them started their tenure as CEO after 1/1/2008 (see Appendix Table B.4). Of those 157 CEOs, almost half (46%) of them had tweeted prior to becoming CEO, 27% of their first tweet as CEO came within the first 10 days of starting, and 13% of their first tweet as CEO occurred near an earnings announcement (not mutually exclusive). The remaining 86 CEOs started before 2008 of whom 28% were early Twitter adopters (joined Twitter prior to 2010) and 19% of whose first tweet was near an earnings announcement (5-day window).

Between CEOs Who Tweet and CEOs Who Do Not Tweet

Comparing CEOs who tweet and CEOs who do not tweet yield a clear selection into tweeting (Table 2.3). A CEO is considered active on Twitter in a particular year if the CEO tweets at least once in that year. Some of the comparison variables chosen have been shown to be associated with the utilization of non-mandatory communication channels such as conference calls and presentations (e.g., Bushee et al., 2017): firm size, book-to-market ratio, return on assets, leverage, and analyst coverage. Additional variables included have been used in other studies studying Twitter communication (e.g. Blankespoor et al., 2014): advertising expenditure, research and development intensity, CEO age, and tech firm indicator. Other variables included have been used in traditional corporate finance studies and found to be determinants of differing management behavior and firm practices (e.g., Malmendier and

Tate, 2005, Malmendier and Tate, 2009): CEO gender, CEO tenure, CEO overconfidence, other leadership roles held at the firm, blockholder indicator, institutional ownership, governance, and percentage of shares of the firm owned by the CEO. I also compare which firms have a corporate Twitter account.

CEOs on Twitter tend to manage larger firms, firms with more investment opportunities, firms with better performance and momentum, firms who spend more on advertising and research and development, and firms who also have a corporate Twitter account. CEOs on Twitter also tend to be younger but more tenured, the founder but not chairman or president, highly compensated outside of cash compensation, and overconfident.[16] The baseline summary statistics are constructed by pooling firm-years (Table 2.3a). To give more specific snapshots of given years, I redo the summary tables for 2014 (the year that the SEC issued guidance on corporate social media practices) and 2018 (year with the most CEOs on Twitter) (see Table 2.3b). Overall, the differences between CEOs who tweet and CEO who do not tweet are similar but not as many characteristics are significantly different suggesting that selection into tweeting has not changed drastically over time.

Quarterly Earnings Announcement Event Study

Determinants of Tweeting About Earnings

To get an initial picture of tweeting patterns around earnings announcements, I compare quarters in which CEOs tweet about earnings 'on announcement date' and quarters in which CEOs are considered active on Twitter but do not tweet about earnings 'on announcement date' (Table 2.4). Most of the firm and CEO characteristics used overlap with Table 2.3, but I include additional variables of particular interest in being able to explain the variation in deciding to tweet or not conditional on having created a Twitter account. Quarters in which CEOs tweet about earnings following the announcement have higher analyst coverage, greater media coverage leading up to the announcement date ([-3,-1]),[17] more positive earnings surprises (SUE $>$ 0), and higher 3-day industry-adjusted announcement returns (CAR[-1,+1]). The stock price reaction leading up to the earnings announcement is not significantly different between the two groups and neither is the more granular measure of earnings surprise (SUE). In addition, firms-quarters in which firms also tweet about earnings

[16]The measure of overconfidence used here was developed by Malmendier and Tate, 2005 based on CEO stock option exercise behavior. Overconfidence is measured as the total value per option of in-the-money options scaled by the price at the end of the fiscal year, which gives an indication of the extent to which the CEO retains in-the-money options that are vested. The overconfident dummy variable used here is equal to 1 if a CEO is in the top quartile of the overconfidence measure during that year.

[17]I get the database of news articles from www.gdelt.org. The GDELT project, or Global Database of Events, Language, and Tone, was created by Kalev Leetaru of Yahoo! and Georgetown University and compiles publications from all news sources globally. Daily data is only available for news articles published after 4/1/2013 and intraday data is only available for news articles published after 2/19/2015. For daily-level news data, I look at articles published in the Wall Street Journal, New York Times, CNN, Reuters, and Bloomberg. For intraday-level news data, I look at articles published in the Wall Street Journal and New York Times.

in quarters are overrepresented in quarters in which CEOs tweet about earnings. These summary results provide preliminary evidence that tweeting about earnings could possibly be related to stock price reactions, earnings surprise, media coverage, and firm earnings tweet patterns.

Announcement Return and Tweets

I start by formally establishing a non-causal relationship between earnings tweets and announcement returns. Figure 2.4 provides graphical evidence that a firm's announcement return after taking out month, year, and industry fixed effects reacts differently in quarters in which a CEO tweets about earnings. There is a linear relationship between earnings news surprise magnitude and the announcement return with the announcement return in earnings tweet quarters (defined as having at least one earnings tweet 'on announcement date') higher than quarters with no earnings tweet at almost all points along the x-axis. I measure the relationship between abnormal return and earnings tweets by estimating the following regression:

$$\begin{aligned} CAR_{i,[t-1,t+1]} = \beta_0 + \beta_1 EarningsTweet_{i,(t,t+1]} + \beta_2 NewsSurpriseMagnitude_{i,t} + \\ \beta_3 Controls_{i,t} + \nu_q + \mu_y + \delta_i + \epsilon_{i,t} \end{aligned} \tag{2.1}$$

The dependent variable ($CAR_{i,[t-1,t+1]}$) is the 3-day industry-adjusted cumulative abnormal return measured from the day before the earnings announcement to the day after. The measure of CEO earnings tweeting 'on announcement date' is denoted $EarningsTweet_{i,(t,t+1]}$. I use 4 different measures: earnings tweet dummy (indicator = 1 if CEO tweets about earnings), earnings tweet intensity (fraction of earnings tweets), follower network size (logarithm of 1 plus earnings tweet dummy multiplied by the number of followers), and potential audience size (logarithm of 1 plus earnings tweet dummy multiplied by the number of retweets of earnings tweets). Since a tweet about earnings is considered 'on announcement date' up until the end of normal trading hours (4 PM ET) on the day following the earnings announcement date, the announcement return measure fully incorporates any immediate stock price reactions to the earnings tweets included in my measures of $EarningsTweet_{i,(t,t+1]}$. I do not include tweets in the period [t-1,t] because I am interested in CEO behavior after the announcement. And in order to disentangle whether CEOs tweet because of the earnings surprise or stock price reactions, looking at tweets prior to the realization of the earnings surprise could potentially complicate my findings.

I control for the earnings surprise with $NewsSurpriseMagnitude_{i,t}$. Since there are self-selection concerns on the types of CEOs that select into joining Twitter and select into tweeting about earnings, I control for a number of firm and CEO variables shown to be significantly different in Tables 2.3 and 2.4. These variables include firm size, book-to-market ratio, return on assets, leverage, idiosyncratic risk, momentum, institutional ownership, blockholder presence, analyst following, number of same-day earnings announcements, follower size, and if the firm tweets about earnings. I also include year and quarter fixed effects to control for omitted variables related to time-varying macro factors and CEO-firm fixed effects to control for time-invariant firm characteristics during a CEO's tenure at the firm.

Table 2.5a presents regression results. The estimates on all four measures of $EarningsTweet_{i,(t,t+1]}$ are positive and significant at the 1% level. Column (1) reports an estimate of 0.026 on $EarningsTweet_{i,(t,t+1]}$, which means that when a CEO tweets about earnings 'on announcement date', the 3-day cumulative abnormal return is 2.6% higher than when a CEO does not tweet about earnings (either CEO has a Twitter account but chooses not to tweet or CEO does not have a Twitter account) within the same firm conditional on the earnings surprise and if the firm tweets about earnings. Columns 2-4 use alternate measures of $EarningsTweet_{i,(t,t+1]}$ and the estimates indicate that the more a CEO tweets about earnings, the more popular the CEO is, and the more engagement the earnings tweets receive, the greater the difference the announcement return is between CEO earnings tweet quarters and other quarters.

One possible concern even with CEO-firm fixed effects is that CEOs prior to joining Twitter or CEOs that stopped tweeting are different from CEOs who have a Twitter account but choose not to tweet in a quarter. To address this concern, I re-run the regressions only keeping firm-quarter observations in which the CEO has a Twitter account in a given quarter. Table 2.5b shows that the coefficients are yet again positive and significant at the 1% level, and the magnitude of the estimates is nearly identical.

To further mitigate endogenous selection concerns, I re-run my regressions with a nearest neighbor propensity-score matched sample. The treatment group consists of firm-quarters in which the CEO has a Twitter account and each treatment observation is matched to a firm-quarter in which the CEO does not have a Twitter account based on the following characteristics shown to be significantly different between the two groups: firm size, book-to-market ratio, return on assets, Tobin's Q, leverage, research and development intensity, advertising intensity, tech firm, momentum, idiosyncratic risk, analyst following, firm Twitter, good news quarter, CEO age, CEO tenure, CEO gender, founder, chairman, president, CEO share ownership, CEO compensation, CEO overconfidence, industry, and year. To generate the propensity scores used to match, I run a logit of an indicator variable equal to 1 if the CEO has a Twitter account at the time of the earnings announcement on the match variables and then use the regression results to generate the scores. For each tweeting CEO-firm-quarter observation, I find the most similar non-tweeting CEO-firm-quarter observation (with replacement) by finding the closest propensity score in the same quarter-year, industry,[18] and earnings news type (positive SUE, zero SUE, or negative SUE).

I then separate the treatment group and corresponding matches by whether the CEO tweeted about earnings 'on announcement date' in a firm-quarter. Appendix Table B.8 shows the quality of the matches between firm-quarters in which CEOs tweeted about earn-

[18] I first try to match within the same 4-digit SIC code. If no match is possible, I then try to match within the same 3-digit SIC code. If no match is possible again, I try to match within the same 2-digit SIC code. If there is no match possible, I do not attempt to match further.

ings against their matched counterparts. Of the 23 match variables, of which nearly all were significantly different in the unmatched larger sample, only 6 variables remain significantly different (r&d intensity, analyst coverage, CEO age, founder, shares owned, and total compensation). While the mean analyst coverage measure is significantly smaller for the matched control sample, the median is not significantly different. For the other variables that are not perfectly matched on, I mitigate concerns about the match quality in two ways. First, I adjust for bias in the matching by adjusting the outcome variable (CAR[-1,+1]) in a method proposed by Abadie and Imbens, 2011. Second, I add the significantly different match variables as controls in my regression. I then run regressions on the two subsets (see Appendix Table B.9). The regression in column (1) uses the sample set of firm-quarters in which the CEO tweeted about earnings 'on announcement date' along with their nearest neighbor match. The coefficient on $EarningsTweet_{i,(t,t+1]}$ (dummy variable equal to 1 if CEO tweeted about earnings that quarter and 0 if a nearest neighbor match) in column (1) is positive and weakly significant and the estimate is dampened compared to the baseline regression, but still supports the direction and magnitude of the relationship found between earnings tweets and announcement returns. Column (2) compares quarters in which CEOs are on Twitter but do not tweet about earnings (Twitter Active = 1) against their matched counterparts and find that there is no significant difference in announcement returns further suggesting that there is something indeed going on in quarters in which CEOs tweet about earnings.

I further investigate which types of firms appear to benefit in quarters in which CEOs tweet. Barberis et al., 2016 show that individual investors play a more important role among firms whose price is less subject to arbitrage: smaller stocks, illiquid stocks, stocks with high idiosyncratic volatility, and stocks with low institutional ownership. In addition, firm size and analyst following are often used as proxies for firm visibility (e.g., Blankespoor et al., 2014). One might expect that firms that are less visible and firms that are more impacted by individual investors could benefit the most from reductions in information frictions and might even be more susceptible to the social media strategies of an influential individual. Indeed, I find that the positive difference in announcement returns for CEO tweeting quarters compared to CEO non-tweeting quarters increases for firms that have smaller market capitalization, are more illiquid, have greater idiosyncratic risk, and followed by fewer analysts (see Appendix Table B.10). However, only the interaction effect on idiosyncratic risk is significant.

Combined with the result that the positive difference in announcement returns for CEO tweeting quarters increases with the attention a tweet receives, these results provide weak evidence that firms in which individual investors are believed to play a more important role and firms that receive less attention see the largest relationship between announcement returns and earnings tweets. A possible explanation is that by drawing attention to earnings announcements, prices are simply more efficient. You would then expect to see that if a CEO had not tweeted in a quarter, the initial announcement return would be lower, but the post-earnings announcement drift would be stronger. Another explanation could be that CEOs of firms with larger information asymmetries have a greater ability to manipulate investor expectations. In that case, I would expect initial announcement returns to be followed by a stronger post-earnings announcement drift due to overreaction to good news or underreaction to bad news. Finally, an explanation could be that the less visible a firm is, the more the earnings surprise proxy is not the best measure of how well a firm did in a quarter and that CEO tweeting patterns could be more indicative of a more truthful picture of the health of the firm above and beyond the earnings surprise. This should not result in any return reversals.

I run the following specifications to test for short- and medium-term post-earnings announcement returns:

$$CAR_{i,[t+2,t+5]} = \beta_0 + \beta_1 EarningsTweet_{i,(t,t+1]} + \beta_2 NewsSurpriseMagnitude_{i,t} + \beta_3 Controls_{i,t} + \nu_q + \mu_y + \delta_i + \epsilon_{i,t} \quad (2.2)$$

$$CAR_{i,[t+2,t+30]} = \beta_0 + \beta_1 EarningsTweet_{i,(t,t+1]} + \beta_2 NewsSurpriseMagnitude_{i,t} + \beta_3 Controls_{i,t} + \nu_q + \mu_y + \delta_i + \epsilon_{i,t} \quad (2.3)$$

The results are reported in Table 2.6. The lack of significance on the coefficient for $EarningsTweet_{i,(t,t+1]}$ suggests that there is no underreaction or overreaction incremental to the earnings news in quarters in which CEOs tweet about earnings. This finding is robust to restricting my sample to the nearest neighbor matched sample and to firms with above-median idiosyncratic risk (see Appendix Table B.11).

Combined, these results establish that there is a contemporaneous relationship between earnings tweets and short-term announcement returns, but I cannot cleanly identify the directional effect between the two using daily-level data. Either CEO earnings tweets cause a price effect incremental to the reaction to the earnings announcement or CEOs tweet in response to prices. To further investigate, I will turn to more granular trading data.

Earnings Tweet Event Study

Baseline Results

Based on the evidence from above, two possible explanations fall out. One possible explanation is that earnings tweets are incrementally informative to the earnings news ('informative signal' story). As a result, I would expect to see price movement, increased volatility, and more trading after the tweet. The other possible explanation is that CEOs tweet about earnings in reaction to the stock price movement ('taking credit' story). The 'taking credit' story implies that CEOs only tweet in quarters in which the firm is performing better than the earnings news would suggest in order to signal his value and contribution to the firm's superior performance. As a result, I should see that CEOs time their tweets to follow a positive market run up.

I construct 30-minute price intervals and 30-minute trade intervals including trades before/after market hours (4AM - 8PM) since the majority of earnings announcements (98%) occur after hours and the majority of earnings tweets occur within the first few hours after the earnings announcement. Instead of an event study around earnings announcements, this time I measure and compare price movements and trading volume before and after earnings tweets. I purge time-of-day effects by regressing all performance measures on hour and month fixed effects and using the residuals for the event study.

Figure 2.5 provides graphical evidence that CEOs tweet about earnings after a positive market run up. The cumulative return following the earnings tweet seems to suggest that the earnings tweet has no effect on stock prices. As a point of comparison, tweets about non-earnings topics in the same window after the earnings announcement do not exhibit the same price movements in the pre-event and post-event periods. Figure 2.6 shows that average stock price volatility (absolute price change) and trading volume (share turnover) also increase prior to earnings tweets. Since my main granular measures are in 30-minute intervals, I also construct a minute-by-minute version of the graphs in the 60-minute window around earnings tweets, which further confirm that earnings tweets seem to follow any price impact (see Figure 2.7).

Table 2.7a confirms that the average price change in the 2, 6, and 12 hours prior to an earnings tweet is positive and highly significant (0.362%, 0.847%, 1.090%, respectively) while the average return in the 2, 6, and 12 hours after the earnings tweet is negative but insignificant (-0.048%, -0.113%, -0.129%, respectively). The difference between the average return before and after earnings tweets is significantly different. Volatility and trading volume in the 2 hours after the earnings tweet are not significantly different compared to the pre-event period suggesting that earnings tweets are not doing much for market participant behavior.

One concern is that by not conditioning on the timing of the earnings announcement, the returns measured before the tweet are picking up something more mechanical. I therefore repeat the event study around earnings tweets but this time split up tweets by how far out from the earnings announcement the tweet is (Table 2.7b). I see that for all earnings tweets, conditional on the time of the earnings announcement, the stock return before tweets is positive and highly significant for the majority of tweets. The return for earnings tweets posted 4 to 8 hours after the earnings announcement is not significant but the magnitude (0.601%) is in line with the magnitudes of the other tweets in the pre-event window. All tweets are followed by a negative but insignificant price drop which is in line with the pooled event study. The price volatility is significantly lower in the post-event window except for earnings tweets posted within the first hour of the earnings announcement. This makes sense because price incorporation from the earnings news is still likely to be happening during the first hour. Trading volume is not significantly different between the two event windows for the majority of tweets.

Another concern is that even with these shorter windows, I could be picking up a more general market reaction on earnings announcement days and thus tweeting could be entirely unrelated to returns. To see whether my result holds up to this concern, I compare hourly return and volatility measures relative to the earnings announcement between earnings announcements followed by an earnings tweet and earnings announcements not followed by an earnings tweet. Table 2.8a clearly shows that earnings announcements followed by an earnings tweet experience higher returns and larger volatility in the hours after the earnings announcement. Even the return in the hours leading up to the earnings announcement are significantly higher for earnings tweet quarters. When only comparing quarters in which a CEO has a Twitter account (Table 2.8b), I also find that CEOs choose to tweet about earnings in quarters in which the stock price reaction is higher after the earnings announcement though the volatility is insignificantly lower in those quarters. Unlike the comparison against the entire universe of S&P 1500 firm-quarters, there is no significant difference in returns or volatility in the hours leading up to the earnings announcement. Overall, it does not seem likely that tweeting is entirely unrelated to returns.

Furthermore, it is the case that CEOs are more likely to tweet about earnings after the earnings announcements as the earnings surprise increases (see Appendix Figure B.2). Earnings surprise could potentially explain my results, but the daily event study showed that quarters in which CEOs tweet about earnings have higher stock price reactions conditional on the level of earnings surprise. This suggests that CEOs choose to tweet about earnings when that quarter is 'more superior' compared to similar quarters. This is further confirmed by re-running my intraday event study around earnings announcements by splitting up the sample into positive and non-positive SUE quarters (good and bad news, respectively). Table 2.8c shows that in good news quarters, stock price returns in the hours following the earnings announcement are significantly more positive in good news quarters with earnings tweets compared to all good news quarters. The same holds for bad news quarters (Table 2.8d) though the means are not significant after the first 2 hours following the earnings announcement, but the magnitudes are larger than for good news quarters. It appears that CEOs choose to tweet about earnings in response to information incremental to the earnings surprise.

Now that I have established a strong link between stock price reactions and CEO earnings tweets, there are additional possible issues that could be biasing my results. Since my baseline result includes all earnings tweets, I could be mixing in effects from days in which a CEO tweets multiple times. In addition, tweets posted after market hours and tweets posted during normal trading hours might induce different price dynamics even after accounting for time-of-day fixed effects. Earnings tweets by more popular CEOs could also have a different effect than tweets by less followed CEOs. Also, it could be the case that retail investors are reacting to tweets in a different way from more sophisticated investors which is not reflected in the aggregate prices changes. To show that my results are robust to such concerns, I rerun my baseline event study around subsets of earnings tweets (see Appendix Table B.13). None of these concerns seem to alter my findings.

Alternative Explanations

Though it seems pretty clear that CEOs strategically decide to tweet in line with the 'taking credit' story, there could be other reasons that a CEO decides to tweet such as as a result of the firm tweeting from its corporate Twitter account or in response to media articles being published about the firm and/or CEO.

The first alternate explanation is that CEOs tweet about earnings in response to firms tweeting about earnings. As there is anecdotal evidence that for financially-relevant events a firm's communications team or legal counsel aids in drafting up the messages sent out by the firm and CEO, it is possible that firms and CEOs coordinate the content or timing of their messages. I find that firms tweet about earnings 'on announcement date' in approximately 30% of quarters in which CEOs tweet about earnings compared with 13% of all quarters. In about 70% of those quarters in which both firms and CEOs tweet about earnings, the firm tweets before the CEO tweets (median difference in timing of 30 minutes). So either firm earnings tweets move the stock price and the CEO reacts to that by tweeting or firms and CEOs sometimes co-time their tweets to both take credit for the superior news.

To see whether firm earnings tweets causes any incremental stock price reaction, I run an event study around firm earnings tweets that precede CEO earnings tweets. The returns in the 15, 30, and 60 minutes after a firm's earnings tweet are insignificant with mixed signs suggesting that firm earnings tweets do not cause prices to move incremental to the earnings (Table 2.9b). The price volatility and trading volume following a firm's earnings tweet are also not significantly different compared to before the tweet in the first 30 minutes, which further suggests that firm earnings tweets are not causing any incremental price movement. As a result, it does not seem to be the case that CEOs tweet about earnings because of any incremental market effects from firm earnings tweets. It is still possible that CEOs and firms co-time their tweets so that both accounts can take credit for the good news. I repeat the event study on all firm earnings tweets (Table 2.9a). Like with CEO earnings tweets, firm earnings tweets are preceded by positive and significant price changes measured at 6 hours and 12 hours before, but not significant when measured at 2 hours before but still positive. The interesting observation is that, unlike with CEO earnings tweets, the return in the 2, 6, and 12 hours after firm earnings tweets is negative and highly significant. While this seems to suggest that firms might also tweet about earnings to 'take credit' for good price reactions, the tweets seem to also cause a negative price reaction and greater volatility. It is not obvious why this is the case, which further studies can look into this more.

Another reason that a CEO might tweet is in response to media coverage. Due to data limitations, I only observe news article publications starting February 2015 and only from the Wall Street Journal and New York Times. When looking at earnings tweets and quarterly earnings announcement starting in 2015, I find that at least one media article[19] is published about the CEO and/or firm on the same day as an earnings tweet 'on announcement date' in roughly 25% of quarters in which CEOs tweet about earnings (compared to the average

[19] I only keep news articles that mention 3 or less companies or people in order to mitigate concerns about an article mentioning an entity but focusing largely on another entity.

of 7.2% of news articles published on the same day as the earnings announcement). For 55% of those overlapping days, the news article is published prior to the CEO earnings tweet (median difference in timing of 25.7 minutes). Like with firm earnings tweets, either CEOs are reacting to stock price movements caused by media coverage by tweeting or the actions are complementary. I run an event study around news articles that precede CEO earnings tweets 'on announcement date.' Like for firm earnings tweets, news articles are not followed by significant stock price returns in the 15, 30, and 60 minutes following the article publication (Table 2.10). So it does not seem likely that CEOs are tweeting in reaction to a stock price reaction caused by media coverage. Instead, it appears that these are complementary actions: superior performance induces more media coverage and CEOs to tweet.

CEO Turnover-Performance Sensitivity

Since it seems that CEO earnings tweets are not informative and rather a strategic move timed to 'take credit' for good news above and beyond the earnings surprise, I attempt to link this behavior to CEO turnover probability. I ask whether CEOs are able to leverage social media to make them less sensitive to being fired for bad performance. The fraction of S&P 1500 CEOs fired by the end of each fiscal year has remained stable over the past decade with between 8-11% of CEOs turned over each year (Table 2.11). The table displays the number of CEOs present in each fiscal year that are gone by the following fiscal year. The fraction of CEOs turned over of those with Twitter in a given fiscal year appears to be lower (between 0-7% annually) with the number of CEO turnovers where the CEO has Twitter in the single digits each year. Because the number of turnovers of Twitter CEOs is so few, this could lead to small sample bias that underestimates the probability of the event occurring when using maximum likelihood estimation.

In order to determine whether social media has an effect on CEO turnover-performance sensitivity, I instrument CEO tweeting using the fraction of S&P 1500 CEOs that tweet in a given state in a given year, whether there were any product recalls in a given industry that year, and whether there was the Superbowl held in a given state that year.

The first-stage OLS regression partitions variation in CEO tweeting behavior into a predictable component caused by common industry or location factors and a residual CEO-firm specific component. The second-stage regresses an indicator for forced CEO turnover on the predicted and residual components from the first-stage regression. I estimate the second-stage CEO turnover regression using the Cox proportional hazard model. The Cox model flexibly accommodates the fact that the probability that a currently employed CEO is dismissed over the next period is a function of the CEO's tenure as well as other CEO characteristics and control variables. I treat voluntary turnovers as right-censored observations in the estimation.

I first examine how the fraction of S&P 1500 CEOs that tweet in a given state, product recalls in a given industry, and Superbowls held in a given state affect whether a CEO tweets in a given year. My first-stage OLS regressions in Table 2.12 shows that larger proportions of CEOs that tweet in a given year in a given headquarter state significantly increase the probability that a CEO in the same headquarter state tweets at least once in a given year in column (1). A 10 percent increase of the fraction of CEOs that tweet in a given state increases the probability that a CEO in that state tweets by over 9 percentage points. Columns (2) and (3) show similar results for product recalls and Superbowl events. I confirm that the instrumental variables are uncorrelated with the controls by regressing the instrumental variables on the control variables in Table 2.15 and showing no significant relationship between the instrument and control variables.

Table 2.13 reports results for second-stage hazard regression. Column (1) uses the fitted value for CEO tweeting in a given year from column (1) of Table 2.12, column (2) uses the fitted value from column (2) of Table 2.12, and column (3) uses the fitted value from column (3) of Table 2.12. The estimate on the fitted value for CEO tweeting is negative and significant at the 1% level, which suggests that CEO tweeting significantly decreases the odds of the CEO in power in a given year being gone the next year. A CEO tweeting in a given year decreases the odds of CEO turnover the next year by 80-85%.

Since the prior literature primarily uses logit regressions for the likelihood of CEO turnover, I repeat the second-stage turnover analysis using logit regressions and show that my results are not an artifact of using the hazard model in Table 2.14.

2.4 Conclusion

Social media has made transparency between firms and their stakeholders easier than ever before. By collecting a novel dataset of CEO-initiated communication from the social media platform Twitter comprising the universe of S&P 1500 firms from 2008-2019, I examine the communication decisions made by CEOs during their tenure. I find that the firms and CEO characteristics of CEOs who tweet to be significantly different compared to other CEOs along a number of dimensions. While the majority of their tweets do not garner a lot of attention, the ones that do appear to be targeted at three sets of stakeholders: consumers, employees, and investors.

I find that for tweets targeted at investors in the context of quarterly earnings announcements, they coincide with higher announcement returns conditional on the level of earnings surprise compared to quarters in which the same CEO does not tweet about earnings while at the same firm. A high-frequency event study identifies that CEOs time their earnings tweets to follow large price run ups. The insignificant and much smaller negative return following earnings tweets suggest that there is little information contained in the tweets and thus do not move prices or induce more trading.

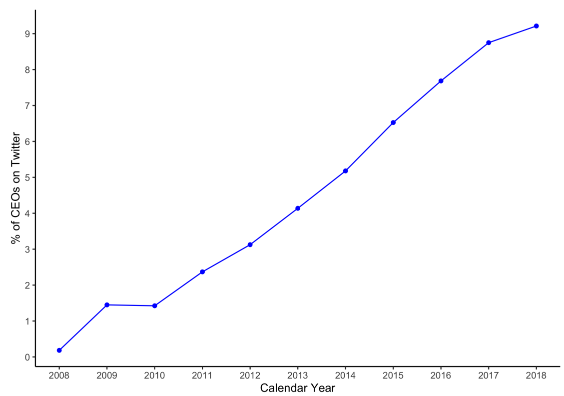
I then attempt to tie this behavior back to the CEO by examining CEO turnover probability. I find evidence that CEOs who tweet are less likely to get fired, suggesting that CEOs can leverage social media to help themselves keep their jobs.

Focusing on earnings tweets left out the vast majority of tweets, of which there are potentially different impacts when considering different stakeholder groups. For instance, CEOs that are transparent and who appear engaged with their employees might aid in employee retention or talent acquisition. Or CEOs that tease product launches or promote current products might see increases in customer sales. Furthermore, CEOs and firms that are vocal about their stances on political and societal issues in times of economic crisis could help mitigate the negative impacts of such shocks.

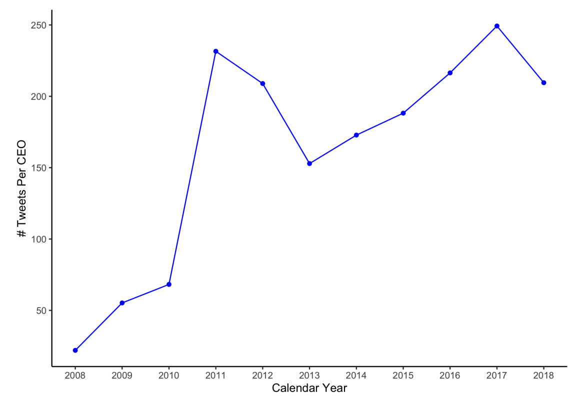
All in all, social media and especially CEO-initiated communication looks to be an interesting channel in which future works can pursue.

Figures

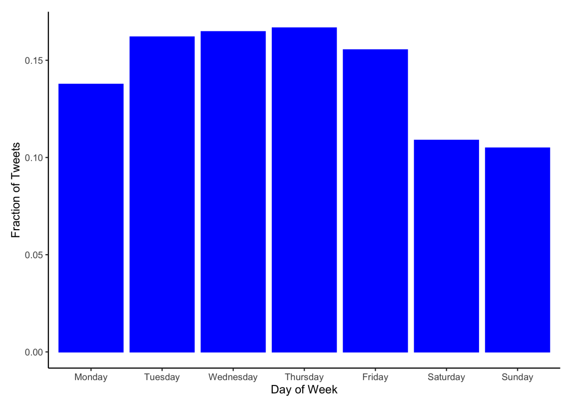
Figure 2.1: CEO tweeting summary statistics - all tweets



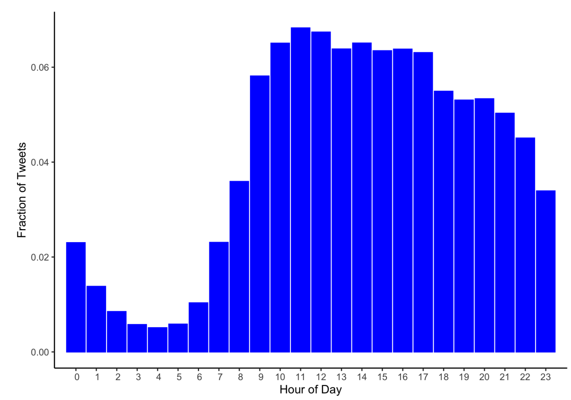
(a) CEOs by Year



(b) Tweets By Year



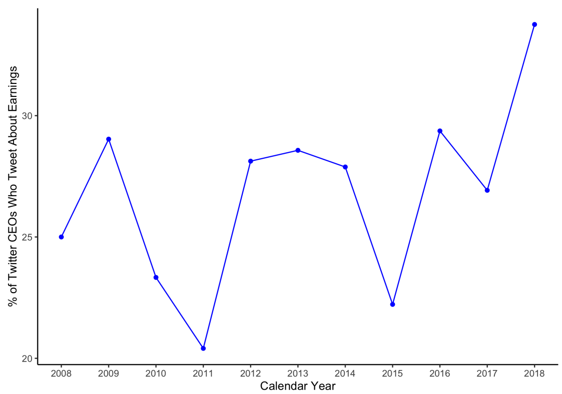
(c) Tweets By Day



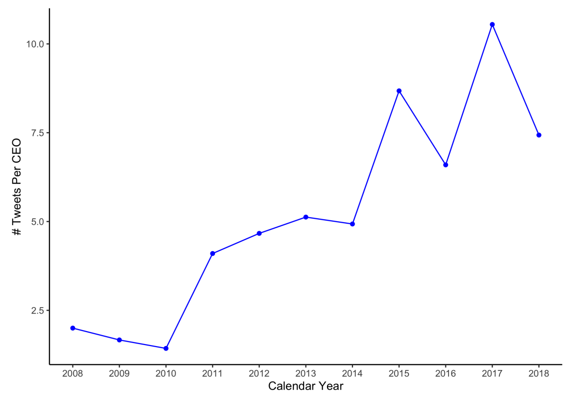
(d) Tweets By Hour

Note: Panel (a) depicts the percentage of S&P 1500 CEOs that posted at least one tweet in a calendar year. Panel (b) shows the annual CEO tweet volume scaled by the number of CEOs that posted at least once in a given year. Panel (c) depicts the fraction of tweets by day-of-week and panel (d) depicts the distribution of tweets by hour.

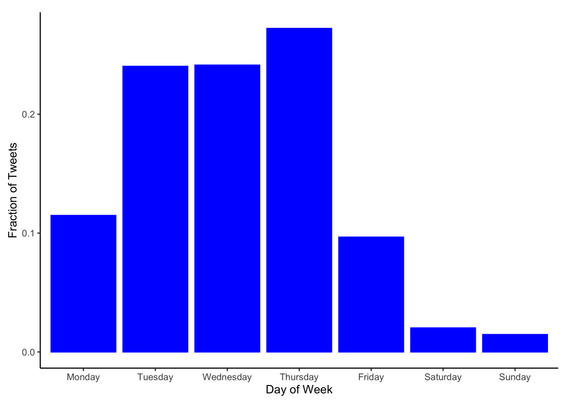
Figure 2.2: CEO tweeting summary statistics - earnings-related tweets



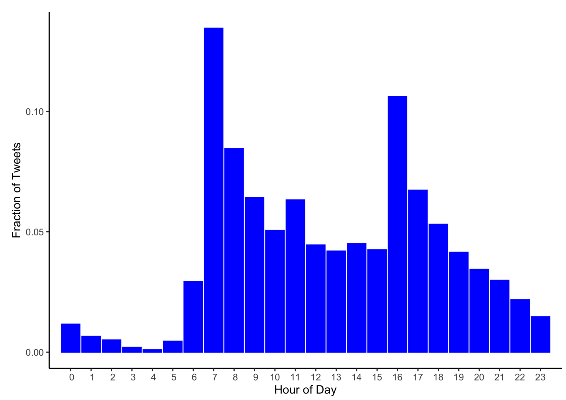
(a) CEOs by Year



(b) Tweets By Year



(c) Tweets By Day



(d) Tweets By Hour

Note: The keywords (and its variations) used to identify an earnings-related tweet are: 'earnings', 'eps', 'quarter 1', 'quarter 2', 'quarter 3', 'quarter 4', 'revenue', 'sales', 'income', 'cash flow', 'profit', 'ebit'. Panel (a) depicts the percentage of S&P 1500 CEOs that posted at least one earnings-related tweet in a calendar year. Panel (b) shows the annual CEO earnings-related tweet volume scaled by the number of CEOs that posted at least once about earnings in a given year. Panel (c) depicts the fraction of earnings-related tweets by day-of-week and panel (d) depicts the distribution of earnings-related tweets by hour.

Figure 2.3: Examples of earnings-related tweets.

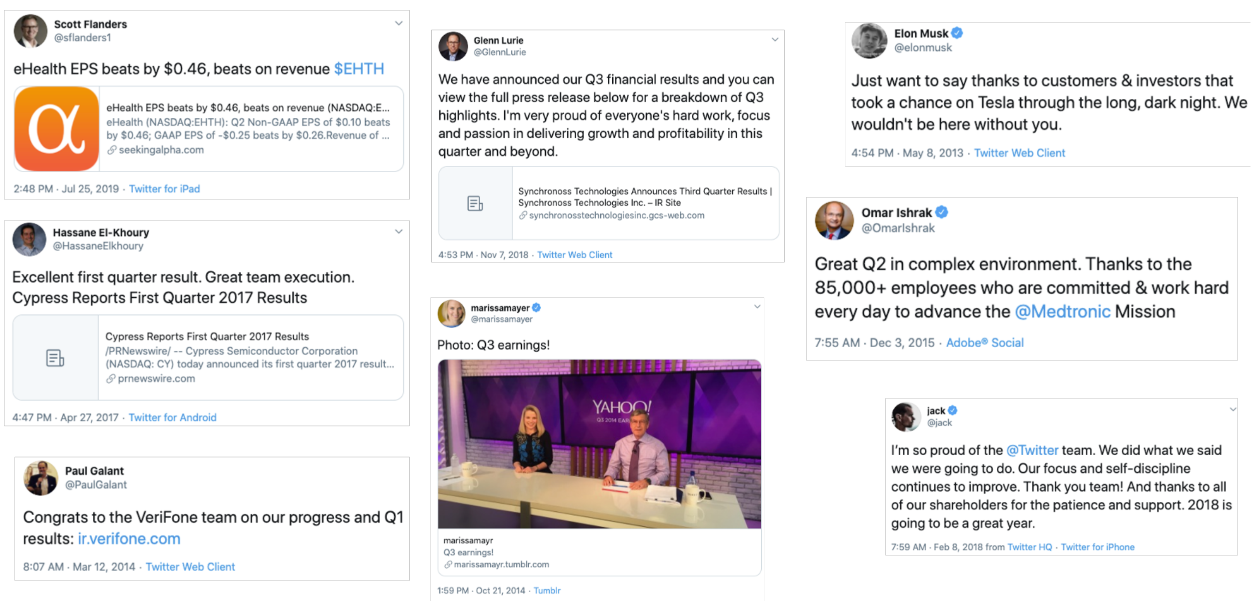
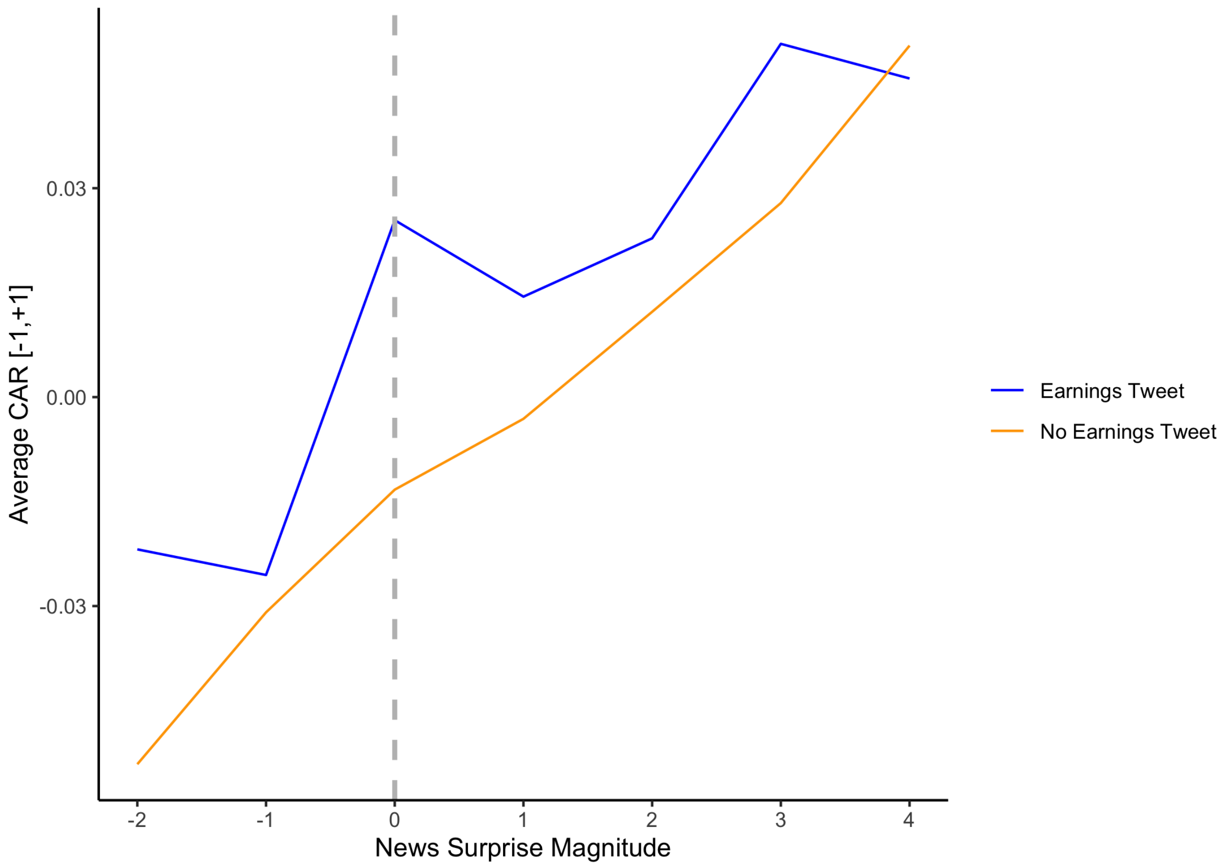
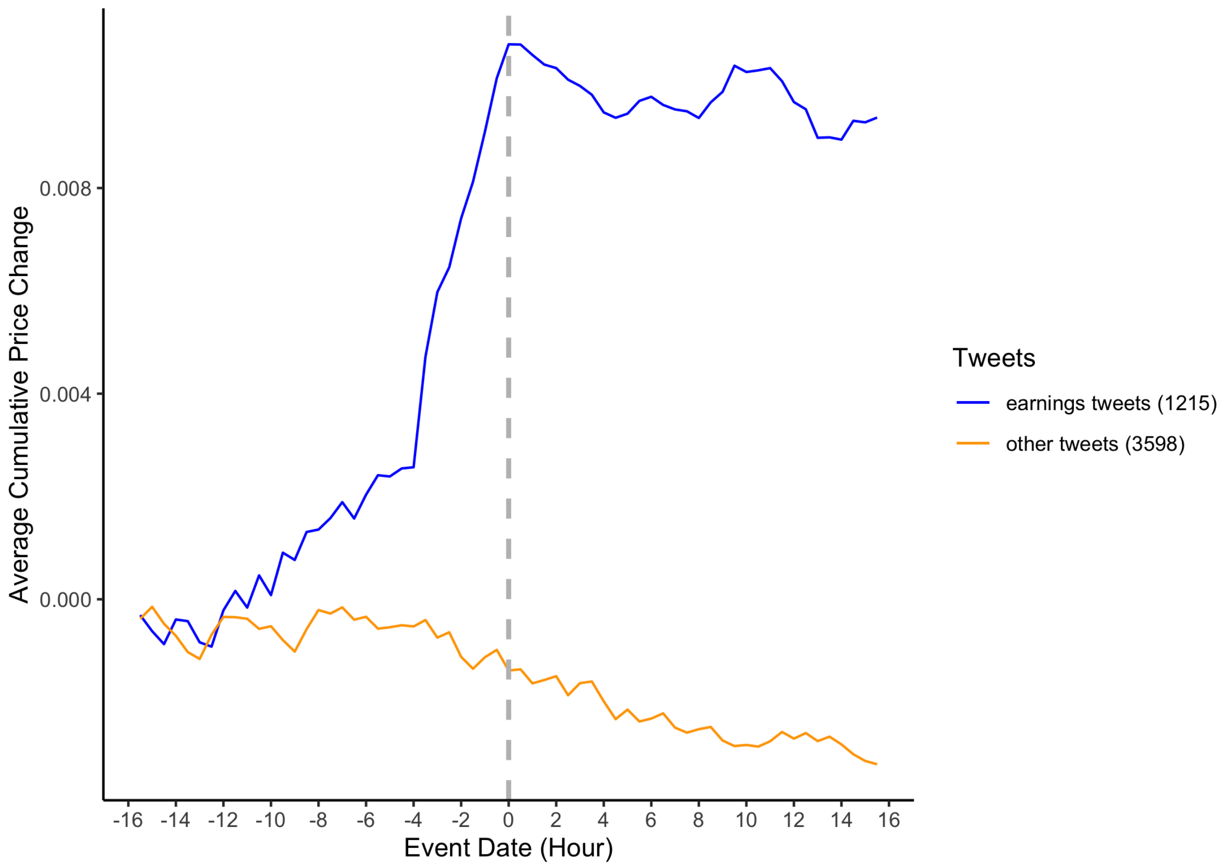


Figure 2.4: Industry-adjusted CAR[-1,+1] by news surprise magnitude



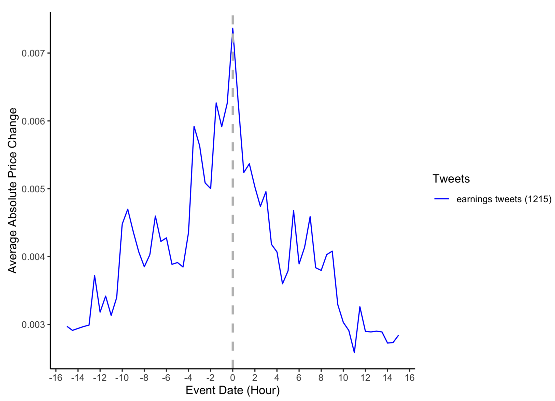
Note: Returns are purged of month, year, and industry fixed effects. News surprise magnitude is equal to zero if a firm exactly meets expectations in a given quarter. Firms that do not meet expectations in a quarter are sorted into one of two equal-sized bins where a news surprise magnitude equal to -2 represents the very worst news. Firms that exceed expectations in a quarter are sorted into one of four equal-sized bins with a news surprise magnitude equal to +4 representing the very best news. The blue line represents quarters in which a CEO tweets about earnings 'on announcement date' and the orange line represents quarters which which a CEO is active on Twitter but does not tweet about earnings.

Figure 2.5: Cumulative returns around CEO tweets around earnings announcements

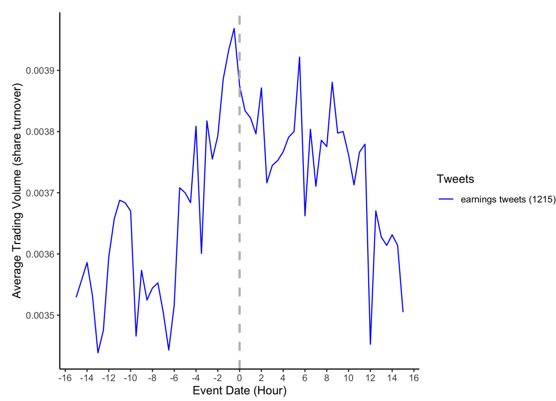


Note: Returns are constructed at 30-minute intervals. The blue line represents earnings-related tweets and the orange line represents non-earnings-related tweets.

Figure 2.6: Volatility and trading around CEO earnings-related tweets - half hour interval



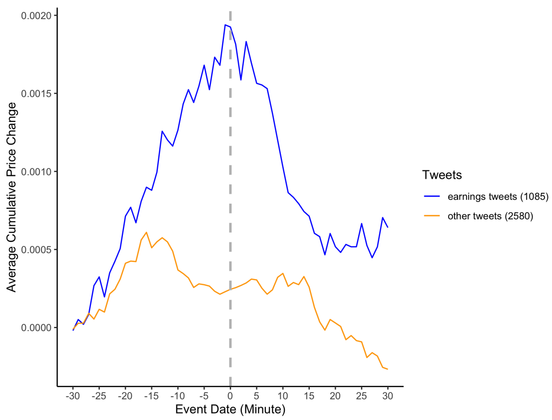
(a) Average Absolute Price Change



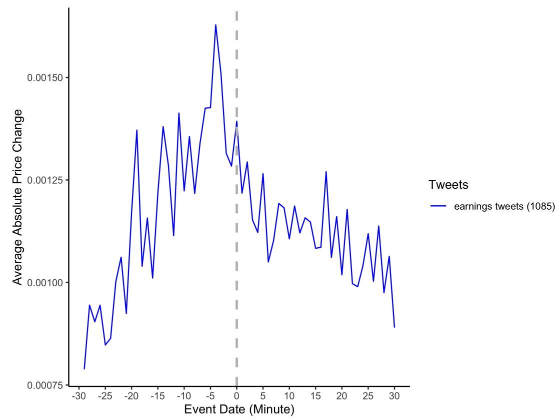
(b) Average Share Turnover

Note: Volatility is the absolute value of the 30-minute price change and trading is share turnover at 30-minute intervals.

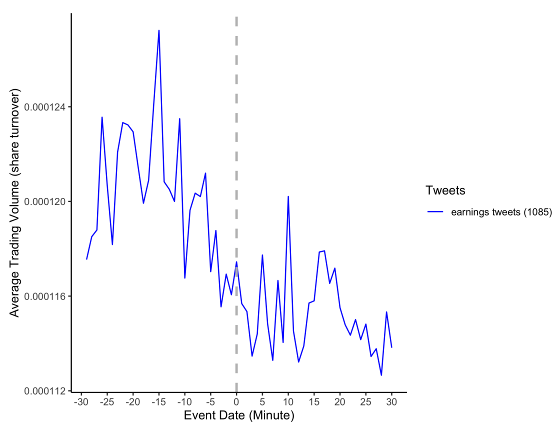
Figure 2.7: Cumulative returns, volatility, and trading volume around CEO earnings-related tweets - minute interval



(a) Average Cumulative Return



(b) Average Absolute Price Change



(c) Average Share Turnover

Note: Volatility is the absolute value of the 1-minute price change and trading is share turnover at 1-minute intervals.

Tables

Table 2.1: CEO tweeting summary statistics

Statistic	N	Mean	Median	St. Dev.	Min	Max
# Followers	243	279,297.10	1,818	2,151,787.00	12	30,369,618
Tweet Tenure (year)	243	3.73	3.34	2.57	0.003	11.15
Start Year	243	2013.67	2014	2.83	2008	2019
Days Tweeted Fraction	243	0.23	0.13	0.25	0.002	1.00
# Tweets per Day	243	0.62	0.19	1.86	0.002	21.52
# Likes Per Tweet	243	234.03	5.91	1,304.78	0.00	14,774.28
# Retweets Per Tweet	243	77.60	2.04	522.28	0.00	7,378.25
# Replies Per Tweet	243	23.49	0.43	150.74	0.00	1,934.62

Table 2.2: Correlates between tweeting measures, CEO demographics, and firm demographics

	# Followers	# Replies Per Tweet	# Retweets Per Tweet	# Likes Per Tweet	# Tweets	Fraction Tweet Days
# Followers						
# Replies Per Tweet	0.231***					
# Retweets Per Tweet	0.276***	0.837***				
# Likes Per Tweet	0.674***	0.665***	0.668***			
# Tweets	0.370***	-0.006	-0.006	0.025		
Fraction Tweet Days	0.278***	-0.036	-0.016	0.048	0.611***	
Twitter Tenure	0.145***	-0.014	-0.015	0.078**	0.093***	0.202***

(a) Correlations with Twitter Characteristics (N=859)

	# Followers	CEO Age	CEO Tenure	Stock Ownership	Total Compensation
# Followers					
CEO Age	-0.017				
CEO Tenure	-0.012	0.354***			
Stock Ownership	0.114***	0.016	0.444***		
Total Compensation	0.640***	-0.015	0.006	0.055	
Salary + Bonus	0.019	0.258***	0.100***	-0.094***	0.015

(b) Correlations with CEO Characteristics (N=859)

	# Followers	Market Cap	Leverage
# Followers			
Market Cap	0.278***		
Leverage	0.118***	0.157***	
Institutional Ownership	-0.188***	-0.178***	-0.079**

(c) Correlations with Firm Characteristics (N=859)

Note: Panel (a) presents correlates between tweeting features, panel (b) presents correlates between the number of Twitter followers a CEO has at the end of a year and other CEO characteristics, and panel (c) presents correlates between the number of Twitter followers a CEO has at the end of a year and firm characteristics. Each observation is a CEO-year in which a CEO is active on Twitter.

Table 2.3: Comparison between tweeting CEOs and non-tweeting CEOs

	Twitter				No Twitter				Difference
	Mean	Median	SD	N	Mean	Median	SD	N	p-value
Firm Characteristics									
Market Cap ($ millions)	27999.90	2796.25	93453.65	864	8916.75	1901.77	26305.59	20405	0.000***
Book-to-Market Ratio	0.39	0.34	0.78	865	0.57	0.51	0.95	20629	0.000***
Return on Assets	0.03	0.04	0.10	865	0.02	0.04	0.33	21225	0.002***
Tobin's Q	2.45	1.83	1.81	865	1.85	1.43	1.98	20628	0.000***
Leverage	0.51	0.48	0.40	821	0.54	0.54	0.39	20299	0.078*
R&D Intensity	0.04	0.01	0.06	865	0.03	0.00	0.07	21243	0.000***
Advertising Intensity	0.03	0.01	0.08	865	0.01	0.00	0.04	21243	0.000***
Tech Firm	0.45	0.00	0.50	865	0.22	0.00	0.42	21244	0.000***
Momentum	0.16	0.15	0.38	819	0.11	0.12	0.46	18407	0.000***
Idiosyncratic Risk	0.05	0.03	0.06	819	0.06	0.03	0.17	18407	0.000***
Institutional Ownership (%)	78.39	81.65	22.13	786	78.80	82.50	21.09	18008	0.615
Blockholder (dummy)	0.95	1.00	0.23	786	0.95	1.00	0.22	18008	0.757
Governance Index	6.00	6.00	1.12	830	6.07	6.00	1.23	20345	0.107
Firm Twitter	0.74	1.00	0.44	865	0.41	0.00	0.49	21244	0.000***
CEO Characteristics									
Age	53.05	53.00	7.52	865	56.38	56.00	7.28	21231	0.000***
Tenure	8.07	6.00	7.62	864	7.49	5.00	7.20	20600	0.029**
Female (dummy)	0.10	0.00	0.29	865	0.04	0.00	0.19	21244	0.000***
Founder (dummy)	0.17	0.00	0.37	865	0.08	0.00	0.27	21244	0.000***
Chairman (dummy)	0.43	0.00	0.50	865	0.56	1.00	0.50	21244	0.000***
President (dummy)	0.61	1.00	0.49	865	0.66	1.00	0.47	21244	0.001***
Shares Owned (%)	3.82	0.63	8.72	861	26.23	0.64	3433.15	20609	0.349
Total Compensation ($ thousands)	11399.76	5070.42	78247.09	865	5678.94	3912.80	6643.08	21236	0.032**
Salary + Bonus ($ thousands)	1063.29	874.18	1329.25	865	1015.60	836.59	1323.50	21244	0.301
Overconfidence (dummy)	0.32	0.00	0.47	865	0.25	0.00	0.43	20937	0.000***

(a) 2008-2019

	2014					2018				
	Twitter		No Twitter		Difference	Twitter		No Twitter		Difference
	Mean	N	Mean	N	p-value	Mean	N	Mean	N	p-value
Firm Characteristics										
Market Cap ($ millions)	23724.87	89	10790.63	1844	0.121	34474.47	144	12468.56	1533	0.020**
Book-to-Market Ratio	0.41	89	0.49	1854	0.095*	0.32	145	0.58	1548	0.017**
Return on Assets	0.03	89	0.03	1920	0.460	0.04	145	0.04	1559	0.949
Tobin's Q	2.58	89	2.00	1854	0.001***	2.51	145	1.95	1548	0.006***
Leverage	0.51	87	0.55	1842	0.395	0.58	133	0.60	1481	0.558
R&D Intensity	0.04	89	0.03	1920	0.006***	0.04	145	0.02	1559	0.002***
Advertising Intensity	0.04	89	0.01	1920	0.003***	0.03	145	0.01	1560	0.001***
Tech Firm	0.43	89	0.22	1920	0.000***	0.42	145	0.19	1560	0.000***
Momentum	0.09	86	0.07	1629	0.390	0.07	136	-0.01	1357	0.025**
Idiosyncratic Risk	0.04	86	0.03	1629	0.182	0.05	136	0.04	1357	0.388
Institutional Ownership (%)	76.71	83	80.85	1648	0.157	69.99	134	70.15	1457	0.920
Blockholder (dummy)	0.89	83	0.96	1648	0.068*	0.97	134	0.97	1457	0.841
Governance Index	5.96	84	6.22	1856	0.053*	6.28	144	6.35	1520	0.382
Firm Twitter	0.63	89	0.52	1920	0.035**	0.93	145	0.66	1560	0.000***
CEO Characteristics										
Age	53.07	89	56.77	1920	0.000***	54.40	145	57.74	1560	0.000***
Tenure	8.31	89	7.65	1853	0.447	7.70	145	7.56	1553	0.826
Female (dummy)	0.11	89	0.04	1920	0.032**	0.10	145	0.04	1560	0.021**
Founder (dummy)	0.20	89	0.07	1920	0.004***	0.12	145	0.05	1560	0.006***
Chairman (dummy)	0.45	89	0.54	1920	0.082*	0.35	145	0.40	1560	0.274
President (dummy)	0.57	89	0.67	1920	0.085*	0.59	145	0.64	1560	0.202
Shares Owned (%)	4.33	88	2.08	1852	0.034**	2.50	145	1.56	1545	0.109
Total Compensation ($ thousands)	7528.29	89	6167.73	1920	0.117	26529.95	145	7477.72	1560	0.228
Salary + Bonus ($ thousands)	1133.83	89	1018.73	1920	0.428	1106.72	145	1063.10	1560	0.525
Overconfidence (dummy)	0.33	89	0.25	1886	0.136	0.31	145	0.24	1556	0.082*

(b) 2014 and 2018

Note: This table compares firm and CEO characteristics between tweeting CEO-firms and non-tweeting CEO-firms. A CEO is defined as being active on Twitter in a given year if he tweets at least once that year. Panel (a) pools CEO-firm-years from 2008-2019. Panel (b) considers 2014 (the year that the SEC issued guidance on corporate social media practices) and 2018 (year with the most CEOs on Twitter) separately.

Table 2.4: Comparison between CEO earnings tweet quarters and non-earnings tweet quarter

	Earnings Tweet In Qtr				No Earnings Tweet In Qtr				Difference
	Mean	Median	SD	N	Mean	Median	SD	N	p-value
CAR[-1,+1]	0.03	0.01	0.09	397	-0.00	-0.00	0.09	2916	0.000***
CAR[-3,-1]	-0.00	-0.00	0.03	397	-0.00	-0.00	0.04	2916	0.308
Price-scaled SUE	0.00	0.00	0.01	418	0.00	0.00	0.01	2668	0.342
Good News Qtr	0.80	1.00	0.40	418	0.67	1.00	0.47	2668	0.000***
Firm Earnings Tweet	0.32	0.00	0.47	439	0.13	0.00	0.33	3006	0.000***
CEO Articles[-3,-1]	0.37	0.00	1.67	383	0.21	0.00	1.34	2450	0.065*
Firm Articles[-3,-1]	2.05	0.00	11.85	383	0.52	0.00	3.64	2450	0.013**
Firm Characteristics									
Market Cap ($ thousands)	23.00	23.29	1.81	438	21.88	21.61	1.92	3002	0.000***
Book-to-Market Ratio	0.47	0.37	0.46	433	0.40	0.33	0.57	2950	0.002***
Return on Assets	0.03	0.03	0.08	435	0.04	0.04	0.11	3003	0.056*
Tobin's Q	2.36	1.65	1.65	433	2.44	1.89	1.75	2950	0.304
Leverage	0.62	0.62	0.38	421	0.48	0.43	0.41	2843	0.000***
R&D Intensity	0.04	0.02	0.06	435	0.04	0.01	0.06	3003	0.394
Advertising Intensity	0.04	0.01	0.06	439	0.04	0.01	0.08	3006	0.966
Tech Firm	0.51	1.00	0.50	439	0.45	0.00	0.50	3006	0.015**
Momentum	0.21	0.20	0.32	395	0.18	0.16	0.36	2850	0.066*
Idiosyncratic Risk	0.06	0.03	0.08	397	0.07	0.03	0.10	2916	0.028**
Illiquidity	0.00	0.00	0.00	397	0.00	0.00	0.00	2916	0.116
Institutional Ownership (%)	71.16	75.02	26.45	349	80.84	83.98	19.43	2712	0.000***
Blockholder (dummy)	0.85	1.00	0.36	349	0.97	1.00	0.18	2712	0.000***
Analyst Following	2.26	2.40	0.87	439	1.81	1.95	0.97	3006	0.000***
CEO Characteristics									
Age	51.83	52.00	7.00	405	52.94	53.00	7.31	2707	0.003***
Tenure	7.40	6.00	6.17	403	8.46	6.00	7.67	2707	0.002***
Female (dummy)	0.11	0.00	0.31	405	0.09	0.00	0.29	2707	0.301
Founder (dummy)	0.18	0.00	0.38	405	0.17	0.00	0.38	2707	0.727
Chairman (dummy)	0.49	0.00	0.50	405	0.42	0.00	0.49	2707	0.004***
President (dummy)	0.55	1.00	0.50	405	0.63	1.00	0.48	2707	0.001***
Shares Owned (%)	3.03	0.55	7.82	401	3.90	0.65	8.59	2696	0.040**
Total Compensation ($ thousands)	24232.06	9960.38	159894.50	405	9666.37	4725.61	62806.55	2707	0.071*
Salary + Bonus ($ thousands)	1304.15	1074.80	1093.76	405	992.90	840.00	1097.12	2707	0.000***
Overconfidence (dummy)	0.40	0.00	0.49	405	0.32	0.00	0.47	2707	0.003***
Follower Size	9.58	9.30	2.38	439	7.69	7.39	2.23	3006	0.000***

Note: This tables compares firm and CEO characteristics between quarters in which CEOs tweet about earnings 'on announcement date' and quarters in which CEOs are considered active on Twitter but do not tweet about earnings 'on announcement date.' 'On announcement date' refers to the period after earnings are announced until the end of trading the following business date.

Table 2.5: Relationship between CEO earnings tweets and abnormal returns around earnings announcements

	CAR [-1,+1]			
	(1)	(2)	(3)	(4)
Earnings Tweet (dummy)	0.026***			
	(0.006)			
Earnings Tweet Intensity		0.023***		
		(0.007)		
Follower Reach			0.003***	
			(0.001)	
Retweet Reach				0.010***
				(0.002)
Other Tweet (dummy)	0.005	0.008	0.005	0.005
	(0.005)	(0.005)	(0.005)	(0.005)
News Surprise Magnitude	0.017***	0.017***	0.017***	0.017***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Firm Earnings Tweet (dummy)	0.004***	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Controls	x	x	x	x
Quarter FE	x	x	x	x
Year FE	x	x	x	x
CEO-Firm FE	x	x	x	x
CEO-Firms	2290	2290	2290	2290
Observations	60,557	60,557	60,557	60,557
R^2	0.187	0.187	0.187	0.187

Note: *p<0.1; **p<0.05; ***p<0.01

(a) Baseline

	CAR [-1,+1]			
	(1)	(2)	(3)	(4)
Earnings Tweet (dummy)	0.026***			
	(0.007)			
Earnings Tweet Intensity		0.024***		
		(0.008)		
Follower Reach			0.003***	
			(0.001)	
Retweet Reach				0.010***
				(0.002)
Other Tweet (dummy)	0.005	0.008*	0.005	0.005
	(0.005)	(0.005)	(0.005)	(0.005)
News Surprise Magnitude	0.017***	0.017***	0.017***	0.017***
	(0.001)	(0.001)	(0.001)	(0.001)
Firm Earnings Tweet (dummy)	0.009	0.009	0.008	0.008
	(0.006)	(0.006)	(0.006)	(0.006)
Controls	x	x	x	x
Quarter FE	x	x	x	x
Year FE	x	x	x	x
CEO-Firm FE	x	x	x	x
CEO-Firms	200	200	200	200
Observations	2,612	2,612	2,612	2,612
R^2	0.199	0.196	0.200	0.201
Note:				*p<0.1; **p<0.05; ***p<0.01

(b) Twitter = 1 Sample

Note: The dependent variable is $CAR_{i,[t-1,t+1]}$, defined as the 3-day industry-adjusted cumulative abnormal return measured from the day before the earnings announcement to the day after. The main predictor variable is $EarningsTweet_{i,(t,t+1]}$, a measure of CEO earnings tweeting 'on announcement date'. $EarningsTweet_{i,(t,t+1]}$ is equal to earnings tweet dummy (indicator = 1 if CEO tweets about earnings) in column (1), earnings tweet intensity (fraction of earnings tweets) in column (2), follower reach (logarithm of 1 plus earnings tweet dummy multiplied by the number of followers) in column (3), and retweet reach (logarithm of 1 plus earnings tweet dummy multiplied by the number of retweets of earnings tweets) in column (4). Control variables include $NewsSurpriseMagnitude_{i,t}$, firm size, book-to-market ratio, return on assets, leverage, idiosyncratic risk, momentum, institutional ownership, blockholder presence, analyst following, number of same-day earnings announcements, follower size, and if the firm tweets about earnings. I also include year and quarter fixed effects as well as CEO-firm fixed effects. Observations are at the firm-quarter level. Panel (a) includes all firm-quarters for S&P 1500 firms and panel (b) includes firm-quarters in which the CEO is active on Twitter.

Table 2.6: Relationship between CEO earnings tweets and abnormal returns after earnings announcements

	CAR [2,5]			CAR [2,30]		
	All Qtrs	Bad News Qtrs	Good News Qtrs	All Qtrs	Bad News Qtrs	Good News Qtrs
	(1)	(2)	(3)	(4)	(5)	(6)
Earnings Tweet (dummy)	0.0002	−0.003	0.002	0.003	0.003	0.005
	(0.003)	(0.007)	(0.004)	(0.008)	(0.018)	(0.009)
Other Tweet (dummy)	0.0004	0.002	0.0001	−0.002	0.017**	−0.008
	(0.002)	(0.004)	(0.002)	(0.004)	(0.008)	(0.005)
News Surprise Magnitude	0.0005***	0.0002	0.001**	0.0003	0.001	0.001
	(0.0001)	(0.001)	(0.0003)	(0.0003)	(0.001)	(0.001)
Firm Earnings Tweet (dummy)	0.0003	0.0001	0.0002	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)
Controls	x	x	x	x	x	x
Quarter FE	x	x	x	x	x	x
Year FE	x	x	x	x	x	x
CEO-Firm FE	x	x	x	x	x	x
CEO-Firms	2290	2147	2236	2290	2147	2236
Observations	60,543	20,874	39,669	60,362	20,805	39,557
R^2	0.069	0.133	0.108	0.101	0.184	0.117

Note: *p<0.1; **p<0.05; ***p<0.01

Note: Regression specifications are similar to Table 2.5a, but the dependent variable is now the industry-adjusted cumulative abnormal return measured from 2 days after the earnings announcement to 5 days after in columns (1)-(3) and to 30 days after in columns (4)-(6).

Table 2.7: Differential intraday market reactions before versus after CEO earnings-related tweets

	2 hours			6 hours			12 hours		
	Before	After	Diff (p-value)	Before	After	Diff (p-value)	Before	After	Diff (p-value)
Return	0.362%***	-0.048%	0.001***	0.847%***	-0.113%	0.000***	1.090%***	-0.129%	0.000***
Standard Error	(0.090%)	(0.078%)		(0.133%)	(0.106%)		(0.166%)	(0.135%)	
# Tweet Bins	804	804		804	804		804	804	
Absolute Return	1.333%***	1.242%***	0.366	2.459%***	1.955%***	0.000***	3.280%***	2.643%***	0.000***
Standard Error	(0.077%)	(0.065%)		(0.105%)	(0.080%)		(0.125%)	(0.098%)	
# Tweet Bins	804	804		804	804		804	804	
Share Turnover	1.558%***	1.532%***	0.480	4.517%***	4.548%***	0.642	8.807%***	9.051%***	0.022**
Standard Error	(0.031%)	(0.019%)		(0.048%)	(0.046%)		(0.070%)	(0.080%)	
# Tweet Bins	804	804		804	804		804	804	

(a) Pooled

				,			Earnings Tweet					
	0-1 Hours After QEA			1-4 Hours After QEA			4-8 Hours After QEA			8+ Hours After QEA		
	1 hour			4 hours			8 hours			12 hours		
	Before	After	Diff (p-value)	Before	After	Diff (p-value)	Before	After	Diff (p-value)	Before	After	Diff (p-value)
Return	0.458%***	0.075%	0.082*	1.402%***	-0.169%	0.000***	0.601%	0.161%	0.333	0.905%***	-0.040%	0.005***
Standard Error	(0.153%)	(0.158%)		(0.366%)	(0.212%)		(0.394%)	(0.226%)		(0.306%)	(0.141%)	
# Tweet Bins	198	198		178	178		129	129		299	299	
Absolute Return	1.020%***	1.383%***	0.051*	3.115%***	1.832%***	0.000***	3.212%***	1.708%***	0.000***	3.642%***	1.734%***	0.000***
Standard Error	(0.138%)	(0.124%)		(0.301%)	(0.162%)		(0.278%)	(0.168%)		(0.227%)	(0.099%)	
# Tweet Bins	198	198		178	178		129	129		299	299	
Share Turnover	0.813%***	0.804%***	0.708	3.049%***	3.005%***	0.646	5.751%***	5.909%***	0.204	9.495%***	8.926%***	0.004***
Standard Error	(0.023%)	(0.010%)		(0.068%)	(0.069%)		(0.090%)	(0.085%)		(0.155%)	(0.121%)	
# Tweet Bins	198	198		178	178		129	129		299	299	

(b) Split by Earnings Announcement Time

Note: This table compares returns, volatility (absolute returns), and trading volume (share turnover) in the hours before versus after CEO earnings-related tweets.

Table 2.8: Differential intraday market reactions between CEO tweeting quarters and other quarters

	Difference(Tweet Quarters - All Quarters)									
	Before Announcement		After Announcement							
	4 hr	2 hr	1 hr	2 hr	3 hr	4 hr	6 hr	8 hr	10 hr	12 hr
Return	0.341%***	0.241%***	0.949%***	1.132%***	1.122%***	1.083%***	1.113%***	1.149%***	1.362%***	1.388%***
Standard Error	(0.122%)	(0.087%)	(0.206%)	(0.223%)	(0.262%)	(0.268%)	(0.275%)	(0.285%)	(0.335%)	(0.346%)
Absolute Return	-0.122%	0.118%	0.823%***	0.679%***	0.672%***	0.246%	0.224%	0.206%	0.300%	0.285%
Standard Error	(0.102%)	(0.078%)	(0.174%)	(0.182%)	(0.200%)	(0.193%)	(0.196%)	(0.200%)	(0.225%)	(0.235%)
# Earnings Announcements (All Qtrs)	51211	51211	51211	51211	51211	51211	51211	51211	51210	51209
# Earnings Announcements (Tweet Qtrs)	337	337	337	337	337	337	337	337	337	337

(a) All S&P 1500 Firm-Quarters

	Difference(Tweet Quarters - All Quarters)									
	Before Announcement		After Announcement							
	4 hr	2 hr	1 hr	2 hr	3 hr	4 hr	6 hr	8 hr	10 hr	12 hr
Return	0.153%	0.115%	0.896%***	1.105%***	1.136%***	1.197%***	1.174%***	1.161%***	1.306%***	1.320%***
Standard Error	(0.129%)	(0.091%)	(0.223%)	(0.242%)	(0.281%)	(0.290%)	(0.299%)	(0.310%)	(0.366%)	(0.377%)
Absolute Return	-0.035%	0.068%	-0.064%	-0.126%	-0.080%	-0.416%**	-0.388%*	-0.419%*	-0.600%**	-0.664%***
Standard Error	(0.108%)	(0.081%)	(0.188%)	(0.198%)	(0.215%)	(0.210%)	(0.213%)	(0.218%)	(0.247%)	(0.257%)
# Earnings Announcements (All Qtrs)	2844	2844	2844	2844	2844	2844	2844	2844	2844	2844
# Earnings Announcements (Tweet Qtrs)	337	337	337	337	337	337	337	337	337	337

(b) CEO Twitter = 1

	Difference(Tweet Quarters - All Quarters)									
	Before Announcement		After Announcement							
	4 hr	2 hr	1 hr	2 hr	3 hr	4 hr	6 hr	8 hr	10 hr	12 hr
Return	0.121%	0.134%	0.783%***	0.895%***	0.789%***	0.634%**	0.645%**	0.692%**	0.683%*	0.729%*
Standard Error	(0.127%)	(0.103%)	(0.246%)	(0.263%)	(0.301%)	(0.305%)	(0.309%)	(0.315%)	(0.369%)	(0.384%)
Absolute Return	-0.127%	0.146%	1.084%***	0.907%***	0.852%***	0.466%**	0.432%*	0.404%*	0.442%*	0.439%*
Standard Error	(0.108%)	(0.093%)	(0.207%)	(0.216%)	(0.232%)	(0.223%)	(0.223%)	(0.223%)	(0.249%)	(0.263%)
# Earnings Announcements (All Qtrs)	33052	33052	33052	33052	33052	33052	33052	33052	33052	33052
# Earnings Announcements (Tweet Qtrs)	265	265	265	265	265	265	265	265	265	265

(c) SUE > 0

	Difference(Tweet Quarters - All Quarters)									
	Before Announcement		After Announcement							
	4 hr	2 hr	1 hr	2 hr	3 hr	4 hr	6 hr	8 hr	10 hr	12 hr
Return	0.395%	0.200%*	0.495%*	0.888%**	0.819%	0.801%	0.852%	0.693%	1.243%	1.277%
Standard Error	(0.360%)	(0.114%)	(0.276%)	(0.392%)	(0.545%)	(0.584%)	(0.634%)	(0.685%)	(0.762%)	(0.767%)
Absolute Return	-0.070%	-0.129%	-0.307%	-0.132%	0.122%	-0.209%	-0.122%	-0.118%	-0.145%	-0.174%
Standard Error	(0.312%)	(0.090%)	(0.240%)	(0.312%)	(0.413%)	(0.421%)	(0.453%)	(0.502%)	(0.501%)	(0.495%)
# Earnings Announcements (All Qtrs)	16559	16559	16559	16559	16559	16559	16559	16559	16559	16558
# Earnings Announcements (Tweet Qtrs)	61	61	61	61	61	61	61	61	61	61

(d) SUE <= 0

Note: This table compares returns and volatility (absolute returns) between firm-quarters in which the CEO tweets about earnings versus other firm-quarters in the hours around CEO earnings-related tweets.

Table 2.9: Differential intraday market reactions before versus after firm earnings-related tweets

	2 hours			6 hours			12 hours		
	Before	After	Diff (p-value)	Before	After	Diff (p-value)	Before	After	Diff (p-value)
Return	0.025%	-0.076%***	0.000***	0.122%***	-0.100%***	0.000***	0.214%***	-0.101%***	0.000***
Standard Error	(0.018%)	(0.021%)		(0.027%)	(0.027%)		(0.032%)	(0.034%)	
# Tweet Bins	16213	16213		16213	16213		16213	16212	
Absolute Return	1.174%***	1.549%***	0.000***	2.150%***	2.198%***	0.107	2.569%***	2.859%***	0.000***
Standard Error	(0.016%)	(0.017%)		(0.021%)	(0.021%)		(0.025%)	(0.026%)	
# Tweet Bins	16213	16213		16213	16213		16213	16212	
Share Turnover	2.174%***	2.193%***	0.199	6.402%***	6.478%***	0.021**	12.423%***	13.056%***	0.000***
Standard Error	(0.009%)	(0.012%)		(0.016%)	(0.029%)		(0.024%)	(0.040%)	
# Tweet Bins	16213	16213		16213	16213		16213	16212	

(a) All Firm Earnings Tweets

	15 min			30 min			1 hour		
	Before	After	Diff (p-value)	Before	After	Diff (p-value)	Before	After	Diff (p-value)
Return	-0.078%	0.002%	0.496	0.029%	0.019%	0.952	0.049%	-0.032%	0.672
Standard Error	(0.077%)	(0.087%)		(0.124%)	(0.098%)		(0.156%)	(0.112%)	
# Tweet Bins	173	173		173	172		173	172	
Absolute Return	0.683%***	0.735%***	0.561	1.058%***	0.985%***	0.523	1.433%***	1.070%***	0.007***
Standard Error	(0.057%)	(0.067%)		(0.094%)	(0.063%)		(0.111%)	(0.076%)	
# Tweet Bins	173	173		173	172		173	172	
Share Turnover	0.125%***	0.122%***	0.264	0.252%***	0.248%***	0.254	0.517%***	0.496%***	0.000***
Standard Error	(0.002%)	(0.001%)		(0.002%)	(0.002%)		(0.005%)	(0.003%)	
# Tweet Bins	173	173		173	172		173	172	

(b) Firm Earnings Tweets That Precede CEO Earnings Tweets

Note: This table compares returns, volatility (absolute returns), and trading volume (share turnover) in the hours before versus after firm earnings-related tweets.

Table 2.10: Differential intraday market reactions before versus after news articles

	15 min			30 min			1 hour		
	Before	After	Diff (p-value)	Before	After	Diff (p-value)	Before	After	Diff (p-value)
Return	-0.118%	0.058%	0.210	-0.126%	0.028%	0.477	-0.296%	0.172%	0.104
Standard Error	(0.082%)	(0.113%)		(0.160%)	(0.148%)		(0.214%)	(0.191%)	
# News Article Bins	90	94		89	91		85	87	
Absolute Return	0.509%***	0.566%***	0.620	0.865%***	0.805%***	0.738	1.170%***	1.171%***	0.999
Standard Error	(0.063%)	(0.097%)		(0.131%)	(0.121%)		(0.175%)	(0.145%)	
# News Article Bins	90	94		89	91		85	87	
Share Turnover	0.454%***	0.440%***	0.635	0.909%***	0.873%***	0.511	1.803%***	1.747%***	0.585
Standard Error	(0.024%)	(0.016%)		(0.046%)	(0.032%)		(0.080%)	(0.065%)	
# News Article Bins	90	94		89	91		85	87	

Note: This table compares returns, volatility (absolute returns), and trading volume (share turnover) in the hours before versus after news articles that precede CEO earnings tweets.

Table 2.11: Annual CEO turnover

	CEO Turnover			CEO Turnover (Twitter = 1)		
Fiscal Year	# CEOs	# CEOs Fired	Fraction	# CEOs	# CEOs Fired	Fraction
2008	134	11	0.08	2	0	0.00
2009	1513	105	0.07	23	2	0.09
2010	1489	120	0.08	25	2	0.08
2011	1264	120	0.09	33	1	0.03
2012	1242	114	0.09	53	2	0.04
2013	1316	116	0.09	65	4	0.06
2014	1432	136	0.09	83	5	0.06
2015	1397	154	0.11	91	7	0.08
2016	1386	151	0.11	107	2	0.02
2017	1429	130	0.09	122	7	0.06
Total	12602	1157	0.09	604	32	0.05

Note: Annual CEO turnover is measured as the number of CEOs present in each fiscal year that are gone by the following fiscal year

Table 2.12: First-stage results for IV regression

	CEO tweet in year t		
	(1)	(2)	(3)
Nearby tweeting CEOs in same industry in year t	0.968***		
	(0.040)		
Product recall in same industry in year t		0.012*	
		(0.007)	
Superbowl in same state in year t			0.013*
			(0.007)
Ind-adj stock return in year t	−0.0004	−0.0005	−0.0005
	(0.003)	(0.003)	(0.004)
Market return in year t	0.008	0.013	0.013
	(0.009)	(0.010)	(0.010)
CEO of retirement age	−0.019***	−0.023***	−0.023***
	(0.006)	(0.008)	(0.008)
CEO overconfident	−0.007	−0.003	−0.003
	(0.006)	(0.008)	(0.008)
CEO with high equity ownership	0.039**	0.039**	0.039**
	(0.016)	(0.020)	(0.020)
Firm size	0.008***	0.010***	0.010***
	(0.003)	(0.003)	(0.003)
Blockholder presence	−0.001	−0.009	−0.009
	(0.013)	(0.020)	(0.020)
High board independence	−0.012**	−0.014*	−0.014*
	(0.006)	(0.007)	(0.007)
Governance index	−0.007**	−0.009**	−0.009**
	(0.003)	(0.004)	(0.004)
CEO tenure	−0.001	−0.001	−0.001
	(0.0005)	(0.001)	(0.001)
Year FE	*x*	*x*	*x*
Industry FE	*x*	*x*	*x*
Observations	14,067	14,067	14,067
Partial F-Stat	10.07	4.23	4.15
R^2	0.369	0.084	0.084
Note:			*p<0.1; **p<0.05; ***p<0.01

Note: First-stage regression results of a dummy variable equal to 1 if a CEO tweets in a given year on the fraction of tweet CEOs in the same industry and state in column (1), whether there was a product recall in the same industry in column (2), and whether the Superbowl was held in a state in a given year in column (3). The reported F-statistic is for the instrument only and is computed using heteroskedasticity-robust standard errors.

Table 2.13: Effect of CEO tweeting on CEO turnover (second-stage hazard regression)

	Forced CEO turnover		
	Nearby tweeting CEOs in industry	Product recall in industry	Superbowl in state
	(1)	(2)	(3)
CEO tweet fitted value	-1.684^{***}	-1.935^{***}	-1.911^{***}
	(0.468)	(0.489)	(0.492)
Ind-adj stock return in year t	-0.960^{***}	-0.970^{***}	-0.970^{***}
	(0.218)	(0.218)	(0.218)
Market return in year t	0.146	0.153	0.152
	(0.738)	(0.742)	(0.742)
CEO of retirement age	-18.413^{***}	-18.433^{***}	-18.432^{***}
	(1, 554.508)	(1, 549.136)	(1, 549.525)
CEO overconfident	-0.657^{***}	-0.652^{***}	-0.651^{***}
	(0.190)	(0.190)	(0.190)
CEO with high equity ownership	-2.572^{***}	-2.560^{***}	-2.562^{***}
	(0.509)	(0.509)	(0.509)
Firm size	-0.120^{**}	-0.118^{**}	-0.118^{**}
	(0.047)	(0.047)	(0.047)
Blockholder presence	−0.185	−0.218	−0.217
	(0.261)	(0.261)	(0.261)
High board independence	0.313^{**}	0.299^{**}	0.299^{**}
	(0.133)	(0.133)	(0.133)
Governance index	−0.032	−0.034	−0.034
	(0.056)	(0.056)	(0.056)
Year FE	x	x	x
Observations	14,067	14,067	14,067
R^2	0.015	0.015	0.015
Max. Possible R^2	0.196	0.196	0.196
Wald Test (df = 10)	$41{,}959.280^{***}$	$41{,}275.440^{***}$	$41{,}320.440^{***}$
LR Test (df = 10)	209.536^{***}	208.524^{***}	208.172^{***}
Score (Logrank) Test (df = 10)	148.739^{***}	149.451^{***}	149.107^{***}
Note:			*p<0.1; **p<0.05; ***p<0.01

Note: This table shows the results from an IV regression of forced CEO turnover in the following year on whether the CEO is active on Twitter in a given year.

Table 2.14: Effect of CEO tweeting on CEO turnover (second-stage logit regression)

	Forced CEO turnover		
	Nearby tweeting CEOs in industry	Product recall in industry	Superbowl in state
	(1)	(2)	(3)
CEO tweet fitted value	-1.668^{***}	-2.003^{***}	-1.984^{***}
	(0.452)	(0.476)	(0.477)
Ind-adj stock return in year t	-0.867^{***}	-0.870^{***}	-0.870^{***}
	(0.211)	(0.212)	(0.212)
Market return in year t	0.072	0.073	0.072
	(0.724)	(0.728)	(0.728)
CEO of retirement age	−15.522	−15.553	−15.552
	(273.266)	(272.821)	(272.853)
CEO overconfident	-0.519^{***}	-0.515^{***}	-0.514^{***}
	(0.192)	(0.192)	(0.192)
CEO with high equity ownership	-1.608^{***}	-1.601^{***}	-1.601^{***}
	(0.522)	(0.522)	(0.522)
Firm size	-0.123^{***}	-0.121^{***}	-0.121^{***}
	(0.045)	(0.045)	(0.045)
Blockholder presence	−0.085	−0.096	−0.096
	(0.265)	(0.265)	(0.265)
High board independence	0.199	0.190	0.190
	(0.135)	(0.135)	(0.135)
Governance index	−0.036	−0.039	−0.039
	(0.058)	(0.058)	(0.058)
CEO tenure	−0.017	−0.017	−0.017
	(0.012)	(0.012)	(0.012)
Constant	−16.144	−16.172	−16.170
	(2,742.499)	(2,733.153)	(2,734.032)
Year FE	x	x	x
Observations	14,067	14,067	14,067
Note:			*p<0.1; **p<0.05; ***p<0.01

Note: This table shows the results from an IV regression of forced CEO turnover in the following year on whether the CEO is active on Twitter in a given year.

Table 2.15: Relationship between instrumental variables and predictor variables

	Dependent variable:			
	Nearby tweeting CEOs in industry	Product recall in industry	Superbowl in state	
	(1)	(2)	(3)	(4)
Ind-adj stock return in year t	−0.0001	0.002*	−0.006	−0.004
	(0.002)	(0.001)	(0.005)	(0.005)
Market return in year t	0.004	0.001	−0.024	−0.012
	(0.005)	(0.002)	(0.029)	(0.020)
CEO of retirement age	−0.004	−0.002	−0.009	−0.013
	(0.005)	(0.002)	(0.007)	(0.009)
CEO overconfident	0.004	−0.001	0.004	−0.0004
	(0.005)	(0.001)	(0.006)	(0.006)
CEO with high equity ownership	0.0003	−0.001	−0.003	0.005
	(0.009)	(0.004)	(0.007)	(0.011)
Firm size	0.002	0.001*	−0.002	0.001
	(0.002)	(0.001)	(0.001)	(0.002)
Blockholder presence	−0.008	0.0004	−0.003	−0.007
	(0.017)	(0.004)	(0.006)	(0.012)
High board independence	−0.002	−0.002	−0.002	−0.0002
	(0.004)	(0.002)	(0.004)	(0.006)
Governance index	−0.002	0.0003	−0.002	−0.001
	(0.002)	(0.001)	(0.001)	(0.003)
CEO tenure	0.00003	0.0003	0.0002	0.001
	(0.0003)	(0.0002)	(0.0003)	(0.0004)
Year FE	*x*	*x*	*x*	*x*
Industry FE	*x*	*x*	*x*	*x*
Observations	14,067	14,067	14,067	14,067
R^2	0.175	0.286	0.622	0.031
Note:				*p<0.1; **p<0.05; ***p<0.01

Bibliography

Abadie, A., & Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics*, *29*(1), 1–11. https://doi.org/10.1198/jbes.2009.07333

Altınkılıç, O., Balashov, V., & Hansen, R. (2013). Are analysts' forecasts informative to the general public? *Management Science*, *59*(11), 2550–2565. https://EconPapers.repec.org/RePEc:inm:ormnsc:v:59:y:2013:i:11:p:2550-2565

Antweiler, W., & Frank, M. Z. (2006). *Do us stock markets typically overreact to corporate news stories?* (Working paper).

Atkins, A. B., & Dyl, E. A. (1990). Price reversals, bid-ask spreads, and market efficiency. *Journal of Financial and Quantitative Analysis*, *25(4)*, 535–547.

Bali, T., Hirshleifer, D., Peng, L., & Tang, Y. (2021). *Attention, social interaction, and investor attraction to lottery stocks* (Working Paper).

Barber, B., Huang, X., Odean, T., & Schwarz, C. (2022). *Attention-induced trading and returns: Evidence from robinhood* (The Journal of Finance (forthcoming)).

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, *21*, 785–818.

Barberis, N., Mukherjee, A., & Wang, B. (2016). Prospect theory and stock returns: An empirical test. *Review of Financial Studies*, *29*, 3068–3107. https://doi.org/10.1093/rfs/hhw049

Bartov, E., Faurel, L., & Mohanram, P. (2018). Can twitter help predict firm-level earnings and stock returns? *The Accounting Review*, *93(3)*, 25–57.

Bennedsen, M., Gonzalez, F., & Wolfenzon, D. (2020). Do ceos matter? evidence from hospitalization events. *The Journal of Finance*. https://doi.org/10.1111/jofi.12897

Bertrand, M., & Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *The Quarterly Journal of Economics*, *118*, 1169–1208. https://doi.org/10.2139/ssrn.376880

Bhagwat, V., & Burch, T. (2016). Pump it up? tweeting to manage investor attention to earnings news. *Economics of Networks eJournal*.

Bianchi, F., Kind, T., & Kung, H. (2020). *Threats to central bank independence: High-frequency identification with twitter* (Working Paper). Duke University.

Blankespoor, E., Miller, G. S., & White, H. D. (2014). The role of dissemination in market liquidity: Evidence from firms use of twitter™. *The Accounting Review*, *89*(1), 79–112. https://doi.org/10.2308/accr-50576

Boehmer, E., Jones, C., Zhang, X., & Zhang, X. (2021). Tracking retail investor activity. *The Journal of Finance*, *76(5)*, 2249–2305.

Bradley, D., Hanousek, J., Jame, R., & Xiao, Z. (2022). *Place your bets? the market consequences of investment research on reddit's wallstreetbets* (Working Paper).

Bryzgalova, S., Pavlova, A., & Sikorskaya, T. (2022). *Retail trading in options and the rise of the big three wholesalers* (Working Paper).

Bushee, B., Jung, M., & Miller, G. (2017). Do investors benefit from selective access to management? *Journal of Financial Reporting*, *2*, 31–61. https://doi.org/10.2308/jfir-51776

Chawla, N., Da, Z., Xu, J., & Ye, M. (2017). *Information diffusion on social media: Does it affect trading, return, and liquidity?* (Working Paper).

Chen, H., De, P., Hu, J., & Hwang, B. (2014). Wisdom of the crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies*, *27*, 1367–1403.

Choi, J. J., D. Laibson, D., & Madrian, B. C. (2010). Why does the law of one price fail? an experiment on index mutual funds. *Review of Financial Studies*, *23(4)*, 1405–1432.

Chordia, T., Roll, R., & Subrahmanyam, A. (2002). Order imbalance, liquidity, and market returns. *Journal of Financial Economics*, *65(1)*, 111–130.

Cochrane, J., & Piazzesi, M. (2002). The fed and interest rates - a high-frequency identification. *The American Economic Review*, *92*, 90–95. https://doi.org/10.1257/000282802320189069

Cohn, J., Liu, Z., & Wardlaw, M. (2022). *Count (and count-like) data in finance* (Journal of Financial Economics (forthcoming)).

Cookson, J. A., Engelberg, J., & Mullins, W. (2022). *Echo chambers* (Working Paper).

Cookson, J. A., & Niessner, M. (2020). Why don't we agree? evidence from a social network of investors. *The Journal of Finance*, *75(1)*, 173–228.

Davis, A., Piger, J., & Sedor, L. (2012). Beyond the numbers: Measuring the information content of earnings press release language. *Contemporary Accounting Research*, *29*(3), 845–868. https://EconPapers.repec.org/RePEc:wly:coacre:v:29:y:2012:i:3:p:845-868

deHaan, E., Shevlin, T., & Thornock, J. (2015). Market (in)attention and the strategic scheduling and timing of earnings announcements. *Journal of Accounting and Economics*, *60*. https://doi.org/10.1016/j.jacceco.2015.03.003

DellaVigna, S., & Pollet, J. M. (2009). Investor inattention and friday earnings announcements. *The Journal of Finance*, *64(2)*, 709–749.

Demers, E., & Vega, C. (2011). *Linguistic tone in earnings announcements: News or noise?* (Working Paper). University of Lausanne.

Dikolli, S., Mayew, W., & Nanda, D. (2014). Ceo tenure and the performance-turnover relation. *Review of Accounting Studies*, *19*. https://doi.org/10.1007/s11142-013-9247-6

Druz, M., Petzev, I., Wagner, A., & Zeckhauser, R. (2020). When managers change their tone, analysts and investors change their tune. *Financial Analysts Journal*, 1. https://doi.org/10.1080/0015198X.2019.1707592

Dyck, A., & Zingales, L. (2003). *The bubble and the media* (Working Paper). Harvard Business School.

Eaton, G., Green, T., Roseman, B., & Wu, Y. (2022). *Retail trader sophistication and stock market quality: Evidence from brokerage outages* (Journal of Financial Economics (forthcoming)).

Ederington, L. H., & Lee, J. H. (1995). The short-run dynamics of the price adjustment to new information. *Journal of Financial and Quantitative Analysis*, *30(1)*, 117–134.

Engelberg, J. E., & Parsons, C. A. (2011). The causal impact of media in financial markets. *The Journal of Finance*, *66(1)*, 67–97.

Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *Journal of Finance*, *64*, 2023–2052. https://doi.org/10.2139/ssrn.971202

Farrell, M., Green, T., Jame, R., & Markov, S. (2022). The democratization of investment research and the informativeness of retail investor trading. *Journal of Financial Economics*, *145*, 616–641.

Faust, J., Swanson, E., & Wright, J. (2004). Identifying vars based on high frequency futures data. *Journal of Monetary Economics*, *51*(6), 1107–1131. https://EconPapers.repec.org/RePEc:eee:moneco:v:51:y:2004:i:6:p:1107-1131

Fedyk, A. (2019). *Front page news: The effect of news positioning on financial markets* (Working Paper). UC Berkeley.

Fedyk, A. (2022). *Front page news: The effect of news positioning on financial markets* (The Journal of Finance (forthcoming)).

Fedyk, V. (2022). *This time is different: Investing preferences in the age of robinhood* (Working paper).

Frydman, C., & Wang, B. (2020). The impact of salience on investor behavior: Evidence from a natural experiment. *The Journal of Finance*, *75(1)*, 229–276.

Fung, A. K. W., Mok, D. M. Y., & Lam, K. (2000). Intraday price reversals for index futures in the us and hong kong. *Journal of Banking Finance*, *24(7)*, 1179–1201.

Gao, H., Harford, J., & Li, K. (2017). Ceo turnover–performance sensitivity in private firms. *Journal of Financial and Quantitative Analysis*, *52*, 1–29. https://doi.org/10.1017/S0022109017000126

Gompers, P., Ishii, J., & Metrick, A. (2003). Corporate governance and equity prices. *The Quarterly Journal of Economics*, *118*(1), 107–156. https://EconPapers.repec.org/RePEc:oup:qjecon:v:118:y:2003:i:1:p:107-156.

Graham, J. R., Kim, H., & Leary, M. (2020). Ceo-board dynamics. *Journal of Financial Economics*.

Gregoire, V., & Martineau, C. (2020). *How is earnings news transmitted to stock prices?* (Working Paper). HEC Montreal - Department of Finance.

Grinblatt, M., & Keloharju, M. (2001). What makes investors trade? *The Journal of Finance*, *56*, 589–616.

Grootendorst, M. (2022). *Bertopic: Neural topic modeling with a class-based tf-idf procedure* (arXiv:2203.05794v0571).

Gürkaynak, R., Sack, B., & Wright, J. (2007). The u.s. treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, *54*(8), 2291–2304. https://EconPapers.repec.org/RePEc:eee:moneco:v:54:y:2007:i:8:p:2291-2304

Harbaugh, R., Maxwell, J., & Shue, K. (2016). *Consistent good news and inconsistent bad news* (Working Paper). Indiana University.

Heimer, R. (2016). Peer pressure: Social interaction and the disposition effect. *Review of Financial Studies*, *29*, 3177–3209.

Heston, S. L., Korajczyk, R. A., & Sadka, R. (2010). Intraday patterns in the cross-section of stock returns. *The Journal of Finance*, *65(4)*, 1369–1407.

Hirshleifer, D., Lim, S., & Teoh, S. H. (2009). Driven to distraction: Extraneous events and underreaction to earnings news. *The Journal of Finance*, *64(5)*, 2289–2325.

Hirshleifer, D., Lim, S. S., & Teoh, S. H. (2011). Limited investor attention and stock market misreactions to accounting information. *The Review of Asset Pricing Studies*, *1*, 35–73.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance*, *54(6)*, 2143–2184.

Hong, H., & Stein, J. C. (2007). Disagreement and the stock market. *Journal of Economic Perspectives*, *21(2)*, 109–128.

Jame, R., Johnston, R., Markov, S., & Wolfe, M. (2016). The value of crowdsourced earnings forecasts. *Journal of Accounting Research*, *54*, 1077–1110.

Jenter, D., & Kanaan, F. (2015). Ceo turnover and relative performance evaluation. *Journal of Finance*, *70*(5), 2155–2184. https://EconPapers.repec.org/RePEc:bla:jfinan:v:70:y:2015:i:5:p:2155-2184

Jung, M., Naughton, J., Tahoun, A., & Wang, C. (2018). Do firms strategically disseminate? evidence from corporate use of social media. *The Accounting Review*, *93*(4), 225–252. https://doi.org/10.2308/accr-51906

Kaniel, R., & Parham, R. (2017). Wsj category kings? the impact of media attention on consumer and mutual fund investment decisions. *Journal of Financial Economics*, *123(2)*, 337–356.

Kaplan, S., & Minton, B. (2012). How has ceo turnover changed? *International Review of Finance*, *12*(1), 57–87. https://EconPapers.repec.org/RePEc:bla:irvfin:v:12:y:2012:i:1:p:57-87

Kirschenheiter, M., & Melumad, N. (2002). Can "big bath" and earnings smoothing co-exist as equilibrium financial reporting strategies. *Journal of Accounting Research*, *40*, 761–796. https://doi.org/10.1111/1475-679X.00070

Kogan, S., Moskowitz, T., & Niessner, M. (2022). *Social media and financial news manipulation* (The Journal of Finance (forthcoming)).

Kronlund, M., Pool, V. K., Sialm, C., & Stefanescu, I. (2021). Out of sight no more? the effect of fee disclosures on 401(k) investment allocations. *Journal of Financial Economics*, *141(2)*, 644–668.

Kuttner, K. (2001). Monetary policy surprises and interest rates: Evidence from the fed funds futures market. *Journal of Monetary Economics*, *47*(3), 523–544. https://EconPapers.repec.org/RePEc:eee:moneco:v:47:y:2001:i:3:p:523-544

Lawrence, A., Ryans, J., Sun, E., & Laptev, N. (2018). Earnings announcement promotions: A yahoo finance field experiment. *Journal of Accounting and Economics*, *66*, 399–414.

Li, E., Ramesh, K., Shen, M., & Wu, J. (2015). Do analyst stock recommendations piggyback on recent corporate news? an analysis of regular-hour and after-hours revisions. *Journal of Accounting Research*, *53*. https://doi.org/10.1111/1475-679X.12083

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, *45*(2-3), 221–247. https://EconPapers.repec.org/RePEc:eee:jaecon:v:45:y:2008:i:2-3:p:221-247

Loos, B., S. Meyer, S., & Pagel, M. (2020). *The consumption effects of the disposition to sell winners and hold losers* (Working Paper).

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *Journal of Finance*, *69*(4), 1643–1671. https://EconPapers.repec.org/RePEc:bla:jfinan:v:69:y:2014:i:4:p:1643-1671

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, *54*(4), 1187–1230. https://EconPapers.repec.org/RePEc:bla:joares:v:54:y:2016:i:4:p:1187-1230

Malmendier, U., & Tate, G. (2005). Ceo overconfidence and corporate investment. *Journal of Finance*, *60*(6), 2661–2700.

Malmendier, U., & Tate, G. (2009). Superstar ceos*. *Quarterly Journal of Economics*, *124*(4), 1593–1638. https://doi.org/10.1162/qjec.2009.124.4.1593

Matsumoto, D., Pronk, M., & Roelofsen, E. (2011). What makes conference calls useful? the information content of managers' presentations and analysts' discussion sessions. *The Accounting Review*, *86*, 1383–1414. https://doi.org/10.2308/accr-10034

Mayew, W., & Venkatachalam, M. (2012). The power of voice: Managerial affective states and future firm performance. *The Journal of Finance*, *67*, 1–44. https://doi.org/10.1111/j.1540-6261.2011.01705.x

Nakamura, E., & Steinsson, J. (2018). High-frequency identification of monetary non-neutrality: The information effect*. *Quarterly Journal of Economics*, *133*, 1283–1330. https://doi.org/10.1093/QJE/QJY004

Niessner, M. (2015). *Strategic disclosure timing and insider trading* (Working Paper). Yale School of Management.

Pedersen, L. (2021). *Game on: Social networks and markets* (Working Paper).

Peng, L., & Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics*, *80(3)*, 563–602.

Peress, J. (2008). *Media coverage and investors' attention to earnings announcements* (Working Paper).

Price, S., Doran, J., Peterson, D., & Bliss, B. (2012). Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking Finance*, *36*, 992–1011. https://doi.org/10.1016/j.jbankfin.2011.10.013

Rogers, J., Skinner, D., & Zechman, S. (2016). The role of the media in disseminating insider-trading news. *Review of Accounting Studies*, *21*(3), 711–739. https://EconPapers.repec.org/RePEc:spr:reaccs:v:21:y:2016:i:3:d:10.1007_s11142-016-9354-2

Rogers, J., Skinner, D., & Zechman, S. (2017). Run edgar run: Sec dissemination in a high-frequency world. *Journal of Accounting Research*, *55*(2), 459–505. https://EconPapers.repec.org/RePEc:bla:joares:v:55:y:2017:i:2:p:459-505

Seasholes, M. S., & Wu, G. (2007). Predictable behavior, profits, and attention. *Journal of Empirical Finance*, *14*, 590–610.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, *62*, 1139–1168.

Veldkamp, L. (2006). Media frenzies in markets for financial information. *The American Economic Review*, *96*(3), 577–601. https://doi.org/10.1257/aer.96.3.577

Wolfskeil, I. (2020). *Tweeting in the dark: Corporate tweeting and information diffusion* (Working Paper). Bocconi University.

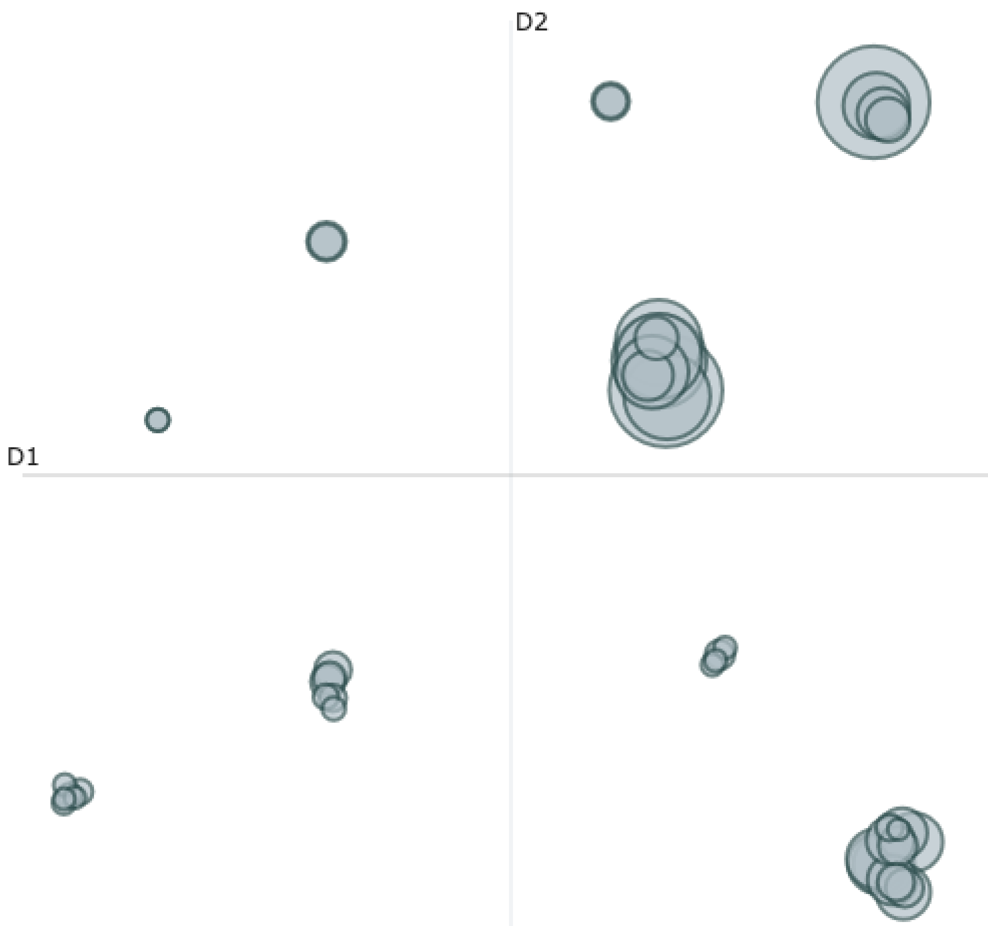
Zawadowski, A. G., Andor, G., & Kertesz, J. (2006). Short-term market reaction after extreme price changes of liquid stocks. *Quantitative Finance*, *6(4)*, 283–295.

Appendix A

WallStreetBets Data

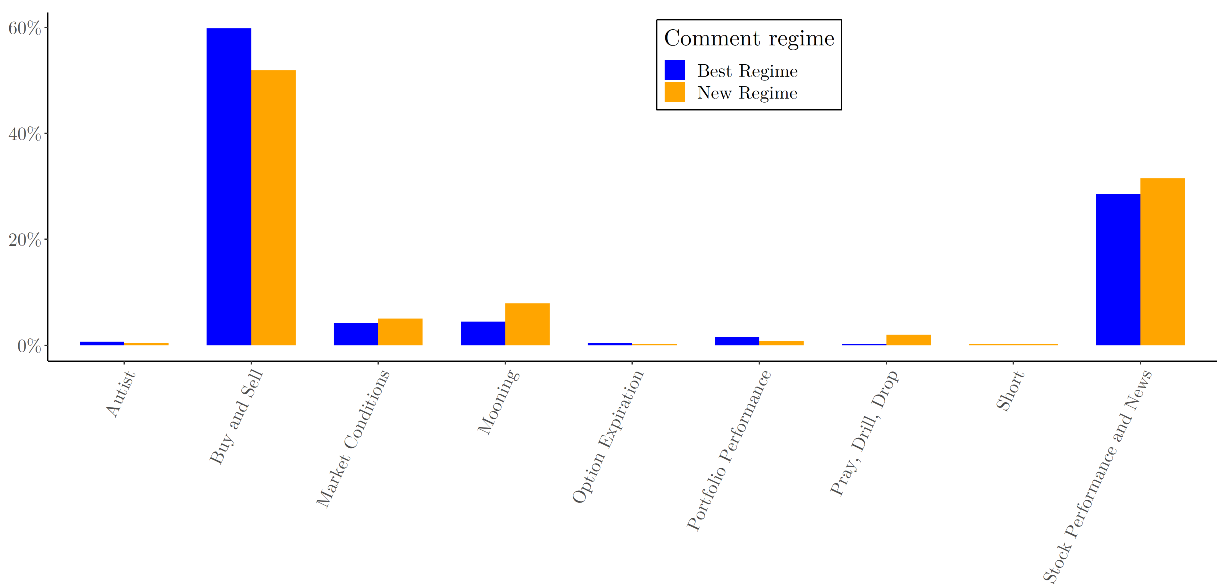
A.1 WallStreetBets Comments Topics

Figure A.1: BERTopic intertopic distance map



Note: BERTopic is run on all WallStreetBets comments in my sample period. BERTopic parameters include using the 'all-MiniLM-L6-v2' sentence transformer, specifying 4 nearest neighbors in the dimensionality reduction step, and specifying a minimum cluster size of 300 when identifying clusters.

Figure A.2: Distribution of topics



Note: Blue bars represent New Regime comments and orange bars represent New Regime comments. The sample set consists of top-level Daily Discussion comments with a single ticker tag during market hours.

Table A.1: Daily Discussion comment examples

Datetime (ET)	Comment	DD Subset?
2018-05-09 12:52:57	should I sell my FB?	Y
2018-05-14 09:33:53	Buy all the $MU you can fucking get, bunch of positive broker reports out today.	Y
2018-08-03 14:08:23	Time to buy the ATVI dip guys, go go go!	Y
2018-08-15 13:19:25	PayPal good call buys at the moment	Y
2018-10-15 15:24:45	I was thinking the 1/18/19 puts are really cheap right now. Looking at the 3 month, seems like a good entry point.	

(a) Buy and Sell

Datetime (ET)	Comment	DD Subset?
2018-09-04 11:48:57	Amazon crashed the market when it hit $1T. What a joke.	Y
2018-09-19 07:19:08	CGC and CRON are both up like 10% too. The fuck is this market smoking	
2018-10-01 16:57:13	Profit is profit	
2018-10-02 14:33:12	if MU hits $48 by the end of the week I'll have tendies for days	Y
2018-10-18 14:49:47	I am absolutely confident AAPL will crush these earnings and do even better relatively for 1/19 earnings. 270+ by march/april. How can you not throw everything at this money churning behemoth with a p/e of 19.	Y

(b) Stock Performance and News

Datetime (ET)	Comment	DD Subset?
2018-05-29 09:38:58	Anyone autistic enough to play HP earnings?	Y
2018-06-14 09:08:47	You deserve your losses for such an autistic play	
2018-06-19 00:41:13	The Autism is strong in this one.	
2018-08-28 10:09:17	AMD action right now is the most intense battle between FOMO autists and putholding autists I've ever seen	Y
2018-10-23 15:17:47	Most autistic yet accurate description, I would give you gold but I'm saving money to buy tendies next week	

(c) Autistic

Datetime (ET)	Comment	DD Subset?
2018-08-15 15:03:29	Sweet HD just turned green!	Y
2018-09-06 11:17:07	Looks like it's a typical red Thursday. Anyone loading up on MSFT or V? IV on them are kinda high though.	Y
2018-10-11 11:45:39	Why are people catching knives on bloody Thursday. Have some respect for tradition	
2018-10-12 09:34:14	sniff So much green, it's beautiful.	
2018-10-12 10:09:35	Only down 30% nbd	

(d) Market Conditions

Datetime (ET)	Comment	DD Subset?
2018-09-12 11:38:35	AMD to the moon at noon.	Y
2018-09-26 15:25:13	AMD is going to the fucking moon tomorrow. Load up now boys and girls. It's fucking time for Su Bae to take everyone's Vcards. New ATH by end of week.	Y
2018-09-27 11:32:09	Damn, FB went to the moon after i got off on the treehouse.	Y
2018-10-12 15:15:07	Nah we're mooning next week	
2018-10-24 08:45:21	No regerts. We holding brother. To the moon or hell	

(e) Mooning

Datetime (ET)	Comment	DD Subset?
2018-05-09 20:21:57	Strike and expiration?	
2018-09-19 13:13:12	I did that a bit back. Guess I'll sell this shit and buy something expiring sooner.	
2018-09-21 10:26:32	When does IV usually hit? thinking about rolling these MSFT 112c 10/26's into longer exp	Y
2018-10-01 13:53:43	My GE puts expire next week and I am up 35 percent. Should i hold them for a few days or exit at EOD	Y
2018-10-22 09:00:36	those should work depending on the expiration.	

(f) Option Expiration

Datetime (ET)	Comment	DD Subset?
2018-08-24 10:03:52	Come on NOC you're basically my portfolio at this point.	Y
2018-10-01 09:30:47	Is your portfolio value +25k? Ive been googling and cant figure it out, this is the only other reason I can think of	
2018-10-01 09:38:25	Really wish ADMP would quit being an anchor on my portfolio. I'd be happy just to break even at this point..	Y
2018-10-05 12:17:49	Go ahead and dump your whole portfolio for the greater good. You will be the mother theresa of WSB	
2018-10-12 10:33:32	Oh god, I can breath again, it's a gorgeous day outside, perfect fall day, my portfolio is up $5.7k. :) Now just need everything up another 20k and I'll be back to where I was two weeks ago.	

(g) Portfolio Performance

Datetime (ET)	Comment	DD Subset?
2018-09-12 14:46:43	alright AAPL, how far we dropping Edit: recession over	Y
2018-09-19 10:52:58	9/21 VKTX puts @15 pray for me	Y
2018-10-04 11:29:11	At least SNAP is sort of drilling to the Earth's core now.	Y
2018-10-18 15:35:02	Power (drill) hour	
2018-10-24 15:05:08	Have mercy	

(h) Pray, Drill, Drop

Datetime (ET)	Comment	DD Subset?
2018-08-22 13:48:07	Is shorting square suicide here?	Y
2018-08-27 15:17:15	lets short BBY?	Y
2018-08-28 00:58:48	Why are you shorting the front side of everything?	
2018-09-18 11:30:13	Should I short TLRY?	Y
2018-10-23 19:12:09	whats ur play?	

(i) Short

A.2 Supplemental Figures and Graphs

Table A.2: Relationship between Daily Discussion comments and retail trading, return volatility, and option trading

	$RetailPct_{i,d}$	$\vert AbnRet\vert_{i,d}(\%)$	$AbnOptionVol_{i,d}$
	(1)	(2)	(3)
$WSBComment_{i,d}$	0.219**	0.317***	0.250***
	(0.102)	(0.033)	(0.020)
$WSBComment_{i,d-1}$	0.214**	0.036	0.040***
	(0.108)	(0.031)	(0.015)
$NewsArticle_{i,d}$	0.065	0.079***	0.082***
	(0.079)	(0.022)	(0.016)
$NewsArticle_{i,d-1}$	0.059	0.028	0.062***
	(0.085)	(0.026)	(0.014)
$\vert Ret\vert_{i,overnight}(\%)$	0.225***	0.345***	0.225***
	(0.040)	(0.017)	(0.007)
$\vert Ret\vert_{i,[d-2,d-1]}(\%)$	0.077***	0.056***	0.011***
	(0.016)	(0.006)	(0.003)
$\vert Ret\vert_{i,[d-5,d-2]}(\%)$	0.057***	0.018***	0.006***
	(0.009)	(0.003)	(0.002)
$RetailPct_{i,d-1}$	0.350***		
	(0.017)		
$AbnOptionVol_{i,d-1}$			0.324***
			(0.010)
Firm Size	x	x	x
Date FE	x	x	x
Ticker FE	x	x	x
Observations	78,433	78,433	67,319
R^2	0.783	0.301	0.310
Note:			*p<0.1; **p<0.05; ***p<0.01

Note: This table presents OLS regression results measuring the relationship between Daily Discussion comments and the percentage of daily trades identified as retail ($RetailPct_{i,d}$) in column (1), absolute value of daily returns ($|AbnRet|_{i,d}$) in column (2), and abnormal option trading volume ($AbnOptionVol_{i,d}$) in column (3). All regression specifications include controls for firm size as well as ticker and date fixed effects, and the sample set consists of ticker-date observations comprised of firms with at least one comment mention over the sample period (4/25/2018 - 10/25/2018). Standard errors are clustered by date and ticker and reported in parentheses.

Table A.3: Differential immediate market reactions following comments in the Best Regime versus New Regime (Diff-in-Diff)

	$AbnRetailVol_{c,i,d,[t,t+5min]}$		
	(1)	(2)	(3)
$NewRegime_c$	0.073	0.160***	0.096*
	(0.049)	(0.057)	(0.051)
$DailyDiscussion_c$	0.334***	0.356***	0.255***
	(0.034)	(0.040)	(0.037)
$NewRegime_c$ x $DailyDiscussion_c$	0.034	−0.150**	0.052
	(0.058)	(0.067)	(0.061)
$HighRHOwnership_{i,d-1}$		0.255***	
		(0.043)	
$NewRegime_c$ x $HighRHOwnership_{i,d-1}$		−0.425***	
		(0.114)	
$DailyDiscussion_c$ x $HighRHOwnership_{i,d-1}$		−0.146*	
		(0.084)	
$NewRegime_c$ x $DailyDiscussion_c$ x $HighRHOwnership_{i,d-1}$		0.697***	
		(0.137)	
Small $\text{Firm}_{i,d-1}$			−0.548***
			(0.049)
$NewRegime_c$ x $SmallFirm_{i,d-1}$			−0.625***
			(0.180)
$DailyDiscussion_c$ x $SmallFirm_{i,d-1}$			0.500***
			(0.090)
$NewRegime_c$ x $DailyDiscussion_c$ x Small $\text{Firm}_{i,d-1}$			0.433**
			(0.197)
Firm Size	x	x	x
Same-Minute Comment	x	x	x
Lagged Comment	x	x	x
Half Hour FE	x	x	x
Industry FE	x	x	x
Observations	17,753	17,321	17,753
R^2	0.134	0.151	0.143
Note:			*p<0.1; **p<0.05; ***p<0.01

(a) Abnormal Retail Trading

	$\lvert AbnRet\rvert_{c,i,d,[t,t+5min]}(\%)$		
	(1)	(2)	(3)
$NewRegime_c$	-0.021	-0.002	-0.019
	(0.034)	(0.039)	(0.034)
$DailyDiscussion_c$	0.021	0.028	0.026
	(0.023)	(0.027)	(0.023)
$NewRegime_c$ x $DailyDiscussion_c$	0.091^{**}	0.012	0.072^{*}
	(0.040)	(0.046)	(0.040)
$HighRHOwnership_{i,d-1}$		-0.146^{***}	
		(0.029)	
$NewRegime_c$ x $HighRHOwnership_{i,d-1}$		-0.035	
		(0.078)	
$DailyDiscussion_c$ x $HighRHOwnership_{i,d-1}$		0.023	
		(0.058)	
$NewRegime_c$ x $DailyDiscussion_c$ x $HighRHOwnership_{i,d-1}$		0.255^{***}	
		(0.094)	
Small $\text{Firm}_{i,d-1}$			-0.192^{**}
			(0.096)
$NewRegime_c$ x $SmallFirm_{i,d-1}$			-0.343
			(0.291)
$DailyDiscussion_c$ x $SmallFirm_{i,d-1}$			0.033
			(0.161)
$NewRegime_c$ x $DailyDiscussion_c$ x Small $\text{Firm}_{i,d-1}$			0.936^{***}
			(0.320)
Firm Size	x	x	x
Same-Minute Comment	x	x	x
Lagged Comment	x	x	x
Half Hour FE	x	x	x
Industry FE	x	x	x
Observations	17,753	17,321	17,753
R^2	0.211	0.217	0.221

Note: *p<0.1; **p<0.05; ***p<0.01

(b) Absolute Abnormal Return

Note: Comparison of abnormal retail trading volume and absolute abnormal returns immediately (within the first 5 minutes) following comments posted in Best Regime versus New Regime using a difference-in-difference approach. Top-level single ticker-tagged market hour Daily Discussion comments comprise the treatment group. Top-level single ticker-tagged market hour non-Daily Discussion WallStreetBets comments where the default comment sort remained Best during my sample period comprise the control group. The control group is further restricted to comments that do not occur within 5 minutes of a Daily Discussion comment mentioning the same ticker.

Table A.4: Overreaction/underreaction to comments in the New Regime and Best Regime

	$PriceImpact_{c,i,d,[t+5min,t+30min]}$			
	(1)	(2)	(3)	(4)
$PriceImpact_{c,i,d,[t,t+5min]}$	-0.077^{***}	-0.078^{***}	-0.079^{***}	-0.079^{***}
	(0.017)	(0.017)	(0.017)	(0.017)
Firm Size				x
Lagged Comment				x
Same-Minute Comment				x
Half Hour FE		x	x	x
Industry FE			x	x
Observations	13,195	13,195	13,195	13,195
R^2	0.002	0.006	0.009	0.010
Note:				*p<0.1; **p<0.05; ***p<0.01

(a) New Regime Comments

	$PriceImpact_{c,i,d,[t+5min,t+30min]}$			
	(1)	(2)	(3)	(4)
$PriceImpact_{c,i,d,[t,t+5min]}$	0.277^{***}	0.292^{***}	0.228^{**}	0.231^{**}
	(0.103)	(0.103)	(0.104)	(0.104)
Firm Size				x
Lagged Comment				x
Same-Minute Comment				x
Half Hour FE		x	x	x
Industry FE			x	x
Observations	728	728	728	728
R^2	0.010	0.041	0.108	0.111
Note:				*p<0.1; **p<0.05; ***p<0.01

(b) Best Regime Comments

Note: I estimate the following regression separately for New Regime comments and Best Regime comments:

$$PriceImpact_{c,i,d,[t+5min,t+30min]} = \beta_0 + \beta_1 PriceImpact_{c,i,d,[t,t+5min]} + controls + \epsilon_{i,d},$$

where $PriceImpact_{c,i,d,[t+5min,t+30min]}$ is the price change from the prevailing mid-quote 5 minutes after comment c posts to the mid-quote 30 minutes after the comment and $PriceImpact_{c,i,d,[t,t+5min]}$ is the price change from the prevailing mid-quote at the time comment c posts to the mid-quote 5 minutes after. Standard errors are clustered by ticker and date and reported in parentheses.

Table A.5: Differential market reactions prior to comments in the Best Regime versus New Regime

	$AbnRetailVol_{c,i,d,[t-6min,t-1min]}$			
	(1)	(2)	(3)	(4)
$NewRegime_c$	0.022	0.005	-0.055	0.017
	(0.027)	(0.027)	(0.041)	(0.028)
$NewRegime_c$ x $HighRHOwnership_{i,d-1}$			0.098^{*}	
			(0.053)	
$NewRegime_c$ x $SmallFirm_{i,d-1}$				-0.094
				(0.080)
Firm Size		x	x	x
Same-Minute Comment		x	x	x
Lagged Comment		x	x	x
Half Hour FE		x	x	x
Industry FE		x	x	x
Observations	13,236	13,236	13,236	13,236
R^2	0.076	0.146	0.147	0.147
Note:		*p<0.1; **p<0.05; ***p<0.01		

(a) Abnormal Retail Trading - 6 minutes prior

	$\lvert AbnRet\rvert_{c,i,d,[t-6min,t-1min]}$			
	(1)	(2)	(3)	(4)
$NewRegime_c$	0.022^{***}	0.015^{**}	-0.002	0.014^{*}
	(0.008)	(0.007)	(0.011)	(0.007)
$NewRegime_c$ x $HighRHOwnership_{i,d-1}$			0.026^{*}	
			(0.014)	
$NewRegime_c$ x $SmallFirm_{i,d-1}$				0.033
				(0.043)
Firm Size		x	x	x
Same-Minute Comment		x	x	x
Lagged Comment		x	x	x
Half Hour FE		x	x	x
Industry FE		x	x	x
Observations	13,236	13,236	13,236	13,236
R^2	0.001	0.325	0.327	0.326
Note:		*p<0.1; **p<0.05; ***p<0.01		

(b) Absolute Abnormal Return - 6 minutes prior

Note: Comparison of abnormal retail trading and absolute abnormal returns in the 5 minutes prior to comments posted in the Best Regime versus New Regime. The sample is restricted to comments posted at least 6 minutes after market open to capture same-day price and retail trading changes.

Appendix B

Twitter Data

B.1 Additional Twitter Summary Statistics

Table B.1: Tweet content distribution

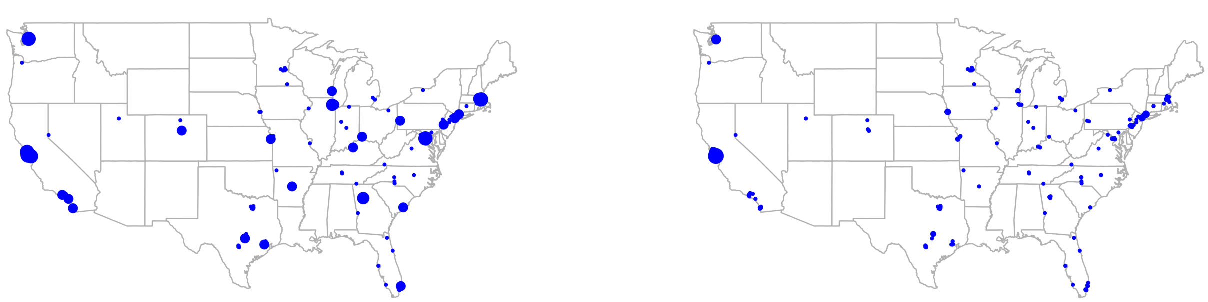
			Median			Engagement: Count x Median		
Tweet Type	Count	Fraction	Replies	Retweets	Likes	Replies	Retweets	Likes
Corporate Image	28,414	0.14	1	3	11	28,414	85,242	312,554
External Validation	3,232	0.02	0	3	11	0	9,696	35,552
Self Promotion	19,908	0.09	0	2	5	0	39,816	99,540
Politics	3,272	0.02	0	1	1	0	3,272	3,272
Customer Interaction	39,128	0.19	1	1	2	39,128	39,128	78,256
Products	3,105	0.01	6	18	47	18,630	55,890	145,935
Strategy and Performance	4,728	0.02	2	11	32.50	9,456	52,008	153,660
General Non-Business	2,784	0.01	1	4	15	2,784	11,136	41,760
Information and Links	51,672	0.25	0	1	2	0	51,672	103,344
Personal	14,930	0.07	1	1	3	14,930	14,930	44,790
Refer to Peers	7,394	0.04	1	1	4	7,394	7,394	29,576
Other	31,826	0.15	1	1	5	31,826	31,826	159,130
Total	210,393	1						

Note: This table displays tweet content distribution by category. Tweets include all S&P 1500 CEO tweets between 2008-2019.

Table B.2: Tweeting CEOs of S&P 1500 firms by year relative to all CEOs on Execucomp

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
# CEOs Active on Twitter	4	31	30	49	64	84	104	126	143	156	157
# CEOs in Execucomp	2181	2138	2107	2069	2047	2031	2009	1931	1862	1784	1705
% of CEOs on Twitter	0.18%	1.45%	1.42%	2.37%	3.13%	4.14%	5.18%	6.53%	7.68%	8.74%	9.21%

Figure B.1: Tweeting CEOs of S&P 1500 firms by location



(a) By CEO Count

(b) By Follower Size

Note: Panel (a) weights by the number of CEOs in each location and panel (b) weights by follower size.

Table B.3: State headquarters and industry breakdown

State HQ	# CEOs (Twitter = 1)	# CEOs	Fraction
California	55	434	0.13
New York	21	230	0.09
Texas	19	296	0.06
Massachusetts	16	120	0.13
Florida	11	108	0.10
Pennsylvania	12	112	0.11
Illinois	8	151	0.05
Virginia	11	88	0.12
Washington	9	46	0.20
Other	81	1185	0.07
Total	243	2770	0.09

(a) State

Fama French 12 Industry Group	# CEOs (Twitter = 1)	# CEOs	Fraction
Business Equipment – Computers, Software, and Electronic Equipment	92	547	0.17
Chemicals and Allied Products	4	78	0.05
Consumer Durables – Cars, TV's, Furniture, Household Appliances	4	66	0.06
Consumer NonDurables – Food, Tobacco, Textiles, Apparel, Leather, Toys	10	142	0.07
Finance	19	511	0.04
Healthcare, Medical Equipment, and Drugs	15	242	0.06
Manufacturing – Machinery, Trucks, Planes, Off Furn, Paper, Com Printing	14	269	0.05
Oil, Gas, and Coal Extraction and Products	4	128	0.03
Other – Mines, Constr, BldMt, Trans, Hotels, Bus Serv, Entertainment	33	311	0.11
Telephone and Television Transmission	15	82	0.18
Utilities	4	96	0.04
Wholesale, Retail, and Some Services (Laundries, Repair Shops)	29	298	0.10
Total	243	2770	0.09

(b) Industry

Note: This table displays state headquarter and industry breakdowns of tweeting S&P 1500 CEOs versus all CEOs. Industries are defined according to the Fama French 12 industry groupings.

Table B.4: CEO tweeting stylized facts.

	Count	Fraction
# CEOs (Twitter = 1)	243	
Became CEO after 1/1/2008	157	
Tweeted Prior	73	0.46
First Tweet Within 10 Days of Start Date	43	0.27
Tweeted Prior	39	
First Tweet Near Earnings Announcement	21	0.13
First Tweet Within 10 Days of Start Date	9	
Became CEO before 1/1/2008	86	
Early Twitter Adopter (pre-2010)	24	0.28
First Tweet Near Earnings Announcement	16	0.19

B.2 Supplemental Figures and Graphs

Table B.5: Earnings-related tweet timing

Relative to Earnings	# Earnings Tweets	Fraction
[-14,-1) days	324	0.16
[-24,0) hours	149	0.08
[0,+1) hours	477	0.24
[+1,+2) hours	83	0.04
[+2,+3) hours	101	0.05
[+3,+4) hours	65	0.03
[+4,+5) hours	57	0.03
[+5,+6) hours	41	0.02
[+6,+7) hours	33	0.02
[+7,+8) hours	25	0.01
[+8,+24) hours	224	0.11
[+1,+14] days	398	0.20

Note: This table displays earnings-related tweet timing by hour(s) relative to quarterly earnings announcement. The keywords (and its variations) used to identify an earnings-related tweet are: 'earnings', 'eps', 'quarter 1', 'quarter 2', 'quarter 3', 'quarter 4', 'revenue', 'sales', 'income', 'cash flow', 'profit', 'ebit'.

Table B.6: Earnings-related tweet frequency 'on announcement date'.

Frequency	Good News Qtrs	# CEOs	Fraction
Always	0.69	3	0.01
Consistent	0.73	17	0.07
Off-and-On	0.71	80	0.35
Never	0.66	131	0.57
Total	0.69	231	1.00

Note: 'On announcement date' is defined as from earnings announcement until the end of trading (4 PM ET) the following business day. 'Always' is defined as CEOs that tweet about earnings every quarter that they are active on Twitter. 'Consistent' is defined as CEOs that tweet about earnings in at least 60% of quarters in which they are active on Twitter. 'Off-and-on' is defined as CEOs that tweet about earnings at least once, but less frequently than 'consistent' CEOs. 'Never' is defined as CEOs that do not tweet about earnings.

Table B.7: Earning-related tweet industry breakdown

Fama French 12 Industry Group	# CEOs (Earnings Tweet = 1)	# CEOs (Twitter = 1)	Fraction
Business Equipment – Computers, Software, and Electronic Equipment	37	87	0.43
Chemicals and Allied Products	2	4	0.50
Consumer Durables – Cars, TV's, Furniture, Household Appliances	2	4	0.50
Consumer NonDurables – Food, Tobacco, Textiles, Apparel, Leather, Toys	4	9	0.44
Finance	10	19	0.53
Healthcare, Medical Equipment, and Drugs	6	14	0.43
Manufacturing – Machinery, Trucks, Planes, Off Furn, Paper, Com Printing	8	13	0.62
Oil, Gas, and Coal Extraction and Products	0	4	0.00
Other – Mines, Constr, BldMt, Trans, Hotels, Bus Serv, Entertainment	14	32	0.44
Telephone and Television Transmission	6	14	0.43
Utilities	2	4	0.50
Wholesale, Retail, and Some Services (Laundries, Repair Shops)	9	27	0.33
Total	100	231	0.43

Note: Industries are defined according to the Fama French 12 industry groupings.

Table B.8: NN match quality

	Earnings Tweet = 1				No Twitter				Difference
	Mean	Median	SD	N	Mean	Median	SD	N	p-value
Firm Characteristics									
Firm Size	23.07	23.30	1.74	377	22.87	22.85	2.06	377	0.139
Book-to-Market Ratio	0.48	0.37	0.44	377	0.44	0.34	0.36	377	0.211
Return on Assets	0.03	0.03	0.08	377	0.03	0.05	0.17	377	0.728
Tobin's Q	2.36	1.64	1.69	377	2.27	1.75	1.42	377	0.412
Leverage	0.63	0.64	0.37	377	0.59	0.63	0.39	377	0.190
R&D Intensity	0.04	0.02	0.05	377	0.05	0.01	0.09	377	0.043**
Advertising Intensity	0.04	0.02	0.06	377	0.03	0.01	0.06	377	0.387
Tech Firm	0.51	1.00	0.50	377	0.51	1.00	0.50	377	1.000
Momentum	0.22	0.20	0.29	377	0.21	0.17	0.37	377	0.626
Idiosyncratic Risk	0.00	0.00	0.00	377	0.00	0.00	0.00	377	0.907
Analyst Following	2.39	2.40	0.72	377	2.24	2.40	0.77	377	0.007***
Firm Twitter	0.71	1.00	0.45	377	0.72	1.00	0.45	377	0.809
Good News Quarter	0.81	1.00	0.39	377	0.81	1.00	0.39	377	1.000
CEO Characteristics									
Age	51.72	51.00	7.01	377	53.40	54.00	7.49	377	0.002***
Tenure	7.19	6.00	5.59	377	6.88	5.00	6.29	377	0.474
Female (dummy)	0.12	0.00	0.32	377	0.11	0.00	0.32	377	0.819
Founder (dummy)	0.17	0.00	0.38	377	0.11	0.00	0.32	377	0.016**
Chairman (dummy)	0.50	1.00	0.50	377	0.47	0.00	0.50	377	0.467
President (dummy)	0.55	1.00	0.50	377	0.57	1.00	0.50	377	0.558
Shares Owned (%)	2.85	0.42	7.35	377	1.64	0.35	3.53	377	0.004***
Total Compensation ($ thousands)	25506.79	10145.35	165653.66	377	10633.74	7659.71	12010.51	377	0.083*
Salary + Bonus ($ thousands)	1339.47	1125.00	1119.34	377	1262.46	1000.00	1338.86	377	0.392
Overconfidence (dummy)	0.42	0.00	0.49	377	0.38	0.00	0.49	377	0.234

Note: This table displays nearest neighbor match quality. The treatment group consists of firm-quarters in which the CEO has a Twitter account and each treatment observation is matched to a firm-quarter in which the CEO does not have a Twitter account based on: firm size, book-to-market ratio, return on assets, Tobin's Q, leverage, research and development intensity, advertising intensity, tech firm, momentum, idiosyncratic risk, analyst following, firm Twitter, good news quarter, CEO age, CEO tenure, CEO gender, founder, chairman, president, CEO share ownership, CEO compensation, CEO overconfidence, industry, and year. 'Earnings Tweet = 1' refers to treatment group observations and corresponding matches where the CEO tweeted about earnings 'on announcement date' in a firm-quarter. 'No Twitter' are all other observations.

Table B.9: Relationship between CAR and CEOs that tweets relative to their NN match

	adj CAR [-1,+1]	
	(1)	(2)
Earnings Tweet (dummy)	0.015*	
	(0.009)	
Twitter Active (dummy)		0.0004
		(0.003)
Other Tweet (dummy)	0.007	−0.001
	(0.008)	(0.004)
News Surprise Magnitude	0.015***	0.016***
	(0.002)	(0.001)
Firm Twitter	0.001	0.009**
	(0.008)	(0.003)
Firm Size	−0.0004	−0.005***
	(0.003)	(0.001)
Controls	x	x
Quarter FE	x	x
Year FE	x	x
Industry FE	x	x
Matched CEO-Firms	87	208
Observations	713	4,559
R^2	0.196	0.127
Note:	*p<0.1; **p<0.05; ***p<0.01	

Note: Column (1) uses the sample set of firm-quarters in which the CEO tweeted about earnings 'on announcement date' along with their nearest neighbor match. $EarningsTweet_{i,(t,t+1]}$ is a dummy variable equal to 1 if the CEO tweeted about earnings that quarter and 0 if a nearest neighbor match. Column (2) uses the sample set of firm-quarters in which CEOs are on Twitter but do not tweet about earnings (Twitter Active = 1) against their matched counterparts.

Table B.10: Relationship between CAR and CEO earnings tweets

	CAR [-1,+1]			
	(1)	(2)	(3)	(4)
Earnings Tweet (dummy)	0.024***	−0.009	0.072	0.043**
	(0.007)	(0.006)	(0.069)	(0.019)
Other Tweet (dummy)	0.005	0.005	0.005	0.005
	(0.005)	(0.005)	(0.005)	(0.005)
News Surprise Magnitude	0.017***	0.017***	0.017***	0.017***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Firm Earnings Tweet (dummy)	0.004***	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Firm Size	−0.015***	−0.015***	−0.015***	−0.015***
	(0.001)	(0.001)	(0.001)	(0.001)
Idiosyncratic Risk	0.850	0.811	0.850	0.849
	(1.106)	(1.107)	(1.106)	(1.106)
Illiquidity	0.00000	0.00000	0.00000	0.00000
	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Analyst Following	0.0003	0.0003	0.0003	0.0003
	(0.001)	(0.001)	(0.001)	(0.001)
Earnings Tweet (dummy) x Illiquidity	0.083			
	(0.068)			
Earnings Tweet (dummy) x Idiosyncratic Risk		64.182***		
		(8.620)		
Earnings Tweet (dummy) x Firm Size			−0.002	
			(0.003)	
Earnings Tweet (dummy) x Analyst Following				−0.008
				(0.007)
Controls	x	x	x	x
Quarter FE	x	x	x	x
Year FE	x	x	x	x
CEO-Firm FE	x	x	x	x
CEO-Firms	2290	2290	2290	2290
Observations	60,557	60,557	60,557	60,557
R^2	0.187	0.188	0.187	0.187
Note:				*p<0.1; **p<0.05; ***p<0.01

Note: Regression specifications are the same as for Table 2.5a, but with the inclusion of an interaction term between the Earnings Tweet dummy and illiquidity in column (1), idiosyncratic risk in column (2), firm size in column (3), and analyst following in column (4).

Table B.11: Post-announcement return - robustness checks

	CAR [2,5]			CAR [2,30]		
	All Qtrs	Bad News Qtrs	Good News Qtrs	All Qtrs	Bad News Qtrs	Good News Qtrs
	(1)	(2)	(3)	(4)	(5)	(6)
Earnings Tweet (dummy)	−0.001	−0.001	−0.00002	−0.007	−0.013	−0.004
	(0.003)	(0.007)	(0.003)	(0.007)	(0.015)	(0.008)
News Surprise Magnitude	0.0001	0.009	0.001	0.001	0.002	0.005
	(0.001)	(0.006)	(0.002)	(0.002)	(0.012)	(0.004)
Firm Twitter	0.003	0.016*	0.0002	−0.004	0.0002	−0.004
	(0.003)	(0.008)	(0.004)	(0.008)	(0.017)	(0.010)
Firm Size	−0.00002	−0.004*	0.001	−0.005*	−0.008*	−0.004
	(0.001)	(0.002)	(0.001)	(0.002)	(0.005)	(0.003)
Quarter FE	x	x	x	x	x	x
Year FE	x	x	x	x	x	x
Industry FE	x	x	x	x	x	x
Matched CEO-Firms	84	29	78	84	29	78
Observations	713	136	577	713	136	577
R^2	0.070	0.260	0.094	0.065	0.199	0.077

Note: *p<0.1; **p<0.05; ***p<0.01

(a) Nearest Neighbor Match

	CAR [2,5]			CAR [2,30]		
	All Qtrs	Bad News Qtrs	Good News Qtrs	All Qtrs	Bad News Qtrs	Good News Qtrs
	(1)	(2)	(3)	(4)	(5)	(6)
Earnings Tweet (dummy)	0.002	−0.002	0.006	0.006	−0.023	0.014
	(0.005)	(0.008)	(0.006)	(0.012)	(0.025)	(0.014)
Other Tweet (dummy)	0.001	0.008	−0.0002	−0.002	0.024**	−0.009
	(0.003)	(0.007)	(0.003)	(0.007)	(0.011)	(0.008)
News Surprise Magnitude	0.001***	0.0003	0.001**	0.0002	0.001	0.001
	(0.0002)	(0.001)	(0.001)	(0.0004)	(0.002)	(0.001)
Firm Earnings Tweet (dummy)	0.0001	0.001	−0.001	−0.001	0.002	−0.001
	(0.001)	(0.002)	(0.001)	(0.003)	(0.005)	(0.003)
Controls	x	x	x	x	x	x
Quarter FE	x	x	x	x	x	x
Year FE	x	x	x	x	x	x
CEO-Firm FE	x	x	x	x	x	x
CEO-Firms	2178	1920	2051	2178	1920	2051
Observations	30,265	11,367	18,898	30,235	11,360	18,875
R^2	0.097	0.176	0.151	0.141	0.238	0.169

Note: *p<0.1; **p<0.05; ***p<0.01

(b) High Idiosyncratic Risk Firms

Table B.12: Relationship between CEO earnings-related tweets, earnings surprises, and CAR

	Dependent variable:					
	Price-scaled SUE				CAR[-1,+1]	
	(1)	(2)	(3)	(4)	(5)	(6)
Earnings Tweet[-14,-2]	−0.001				0.021	
	(0.001)				(0.014)	
Other Tweet[-14,-2]	0.0002				−0.002	
	(0.0004)				(0.004)	
Earnings Tweet Intensity[-14,-2]		−0.003				0.116*
		(0.003)				(0.061)
Earnings Tweet[-1,0]			0.001			
			(0.001)			
Other Tweet[-1,0]			0.001			
			(0.0003)			
Earnings Tweet Intensity[-1,0]				0.001		
				(0.001)		
Good News Quarter					0.055***	0.055***
					(0.001)	(0.001)
Controls	*x*	*x*	*x*	*x*	*x*	*x*
Quarter FE	*x*	*x*	*x*	*x*	*x*	*x*
Year FE	*x*	*x*	*x*	*x*	*x*	*x*
CEO-Firm FE	*x*	*x*	*x*	*x*	*x*	*x*
Observations	60,557	60,557	60,557	60,557	60,557	60,557
R^2	0.160	0.160	0.160	0.160	0.150	0.150
Note:						*p<0.1; **p<0.05; ***p<0.01

(a) Before Earnings Announcement

	Earnings Tweet[+2,+14]	
	Dummy	Intensity
	(1)	(2)
SUE	0.002	0.002
	(0.007)	(0.005)
CAR[-1,+1]	0.003*	0.0003
	(0.002)	(0.001)
Earnings Tweet[0,+1]	0.042**	−0.0005
	(0.021)	(0.007)
Earnings Tweet[-1,0]	0.076	0.018
	(0.056)	(0.015)
Earnings Tweet[-14,-2]	0.093**	0.012
	(0.039)	(0.009)
Controls	x	x
Quarter FE	x	x
Year FE	x	x
CEO-Firm FE	x	x
Observations	60,557	60,557
R^2	0.334	0.122
Note:	*p<0.1; **p<0.05; ***p<0.01	

(b) After Earnings Announcement

Note: Panel (a) measures the relationship for earnings-related tweets prior to earnings announcements and panel (b) measures the relationship for earnings-related tweets after earnings announcements. *Controls* include firm size, book-to-market ratio, return on assets, leverage, idiosyncratic risk, momentum, institutional ownership, and blockholder presence. All regression specifications include quarter, year, and CEO-firm fixed effects. Standard errors are clustered by firm and reported in parentheses.

Table B.13: Differential intraday market reactions before versus after CEO earnings-related tweets - robustness checks

	2 hours			6 hours			12 hours		
	Before	After	Diff (p-value)	Before	After	Diff (p-value)	Before	After	Diff (p-value)
Return	0.615%***	-0.041%	0.003***	0.996%***	-0.032%	0.000***	1.341%***	0.074%	0.001***
Standard Error	(0.171%)	(0.134%)		(0.223%)	(0.180%)		(0.291%)	(0.235%)	
# Tweet Bins	345	345		345	345		345	345	
Absolute Return	1.576%***	1.384%***	0.307	2.637%***	2.200%***	0.053*	3.619%***	2.984%***	0.026**
Standard Error	(0.152%)	(0.112%)		(0.180%)	(0.136%)		(0.228%)	(0.171%)	
# Tweet Bins	345	345		345	345		345	345	
Share Turnover	1.662%***	1.617%***	0.560	4.688%***	4.744%***	0.656	9.066%***	9.440%***	0.031**
Standard Error	(0.065%)	(0.039%)		(0.092%)	(0.088%)		(0.124%)	(0.121%)	
# Tweet Bins	345	345		345	345		345	345	

(a) Single Tweet Days

	2 hours			6 hours			12 hours		
	Before	After	Diff (p-value)	Before	After	Diff (p-value)	Before	After	Diff (p-value)
Return	0.309%**	-0.009%	0.092*	0.644%***	-0.132%	0.004***	1.432%***	-0.172%	0.000***
Standard Error	(0.156%)	(0.107%)		(0.228%)	(0.134%)		(0.288%)	(0.160%)	
# Tweet Bins	330	330		330	330		330	330	
Absolute Return	1.624%***	1.109%***	0.001***	2.822%***	1.465%***	0.000***	3.641%***	1.838%***	0.000***
Standard Error	(0.129%)	(0.087%)		(0.171%)	(0.107%)		(0.221%)	(0.124%)	
# Tweet Bins	330	330		330	330		330	330	
Share Turnover	1.571%***	1.459%***	0.172	4.627%***	4.440%***	0.161	9.159%***	8.927%***	0.137
Standard Error	(0.072%)	(0.039%)		(0.105%)	(0.083%)		(0.125%)	(0.093%)	
# Tweet Bins	330	330		330	330		330	330	

(b) Normal Trading Hours

	2 hours			6 hours			12 hours		
	Before	After	Diff (p-value)	Before	After	Diff (p-value)	Before	After	Diff (p-value)
Return	0.237%*	-0.203%*	0.010***	1.030%***	-0.124%	0.000***	1.041%***	-0.128%	0.000***
Standard Error	(0.131%)	(0.107%)		(0.214%)	(0.143%)		(0.241%)	(0.180%)	
# Tweet Bins	330	330		330	330		330	330	
Absolute Return	1.335%***	1.081%***	0.073*	2.578%***	1.725%***	0.000***	3.134%***	2.361%***	0.000***
Standard Error	(0.109%)	(0.090%)		(0.170%)	(0.107%)		(0.177%)	(0.124%)	
# Tweet Bins	330	330		330	330		330	330	
Share Turnover	1.563%***	1.535%***	0.589	4.475%***	4.613%***	0.197	8.818%***	9.131%***	0.111
Standard Error	(0.043%)	(0.032%)		(0.073%)	(0.077%)		(0.120%)	(0.155%)	
# Tweet Bins	330	330		330	330		330	330	

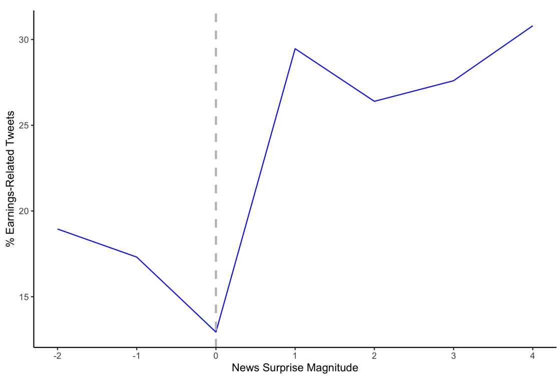
(c) Popular CEOs

	2 hours			6 hours			12 hours		
	Before	After	Diff (p-value)	Before	After	Diff (p-value)	Before	After	Diff (p-value)
Return	0.343%***	-0.040%	0.002***	0.728%***	-0.078%	0.000***	1.005%***	-0.160%	0.000***
Standard Error	(0.094%)	(0.082%)		(0.135%)	(0.111%)		(0.175%)	(0.143%)	
# Tweet Bins	697	697		697	697		697	697	
Absolute Return	1.293%***	1.208%***	0.424	2.310%***	1.880%***	0.002***	3.175%***	2.544%***	0.000***
Standard Error	(0.082%)	(0.068%)		(0.107%)	(0.085%)		(0.133%)	(0.105%)	
# Tweet Bins	697	697		697	697		697	697	
Share Turnover	1.571%***	1.538%***	0.433	4.546%***	4.548%***	0.976	8.864%***	9.020%***	0.159
Standard Error	(0.035%)	(0.022%)		(0.055%)	(0.050%)		(0.079%)	(0.078%)	
# Tweet Bins	697	697		697	697		697	697	

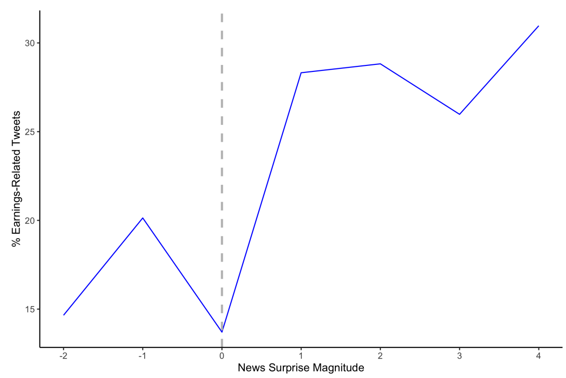
(d) After SEC Guidance

Note: These tables compare returns, volatility (absolute returns), and trading volume (share turnover) in the hours before versus after CEO earnings-related tweets, restricted to single CEO tweet days in panel (a), market trading hours in panel (b), popular CEOs in panel (c), and post-SEC guidance in panel (d).

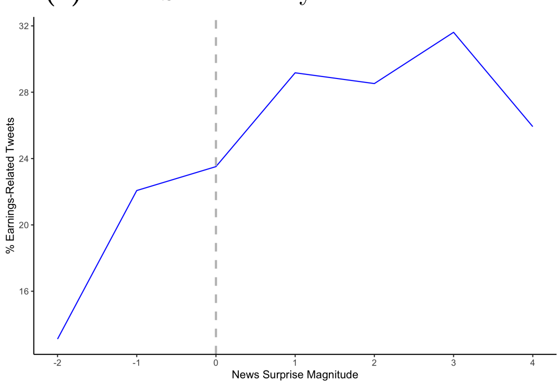
Figure B.2: Earnings tweet fraction By news surprise magnitude



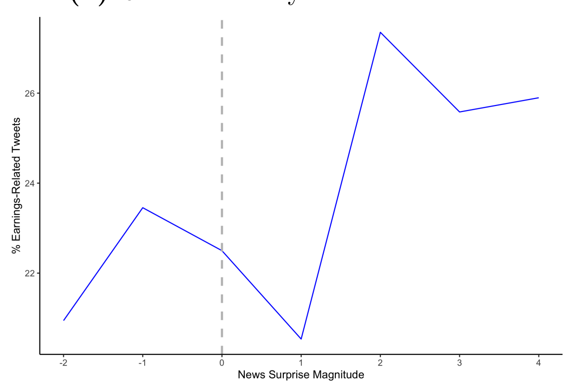
(a) Price-Scaled Analyst-Based Measure



(b) Unscaled Analyst-Based Measure

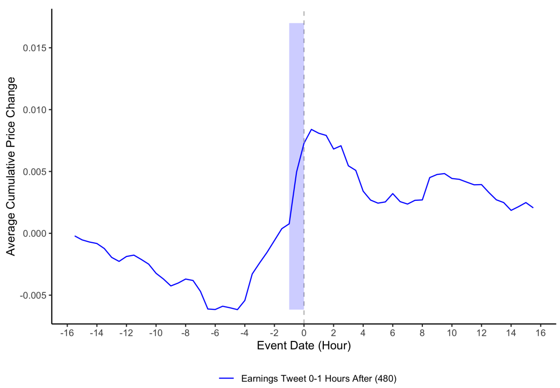


(c) SD-Scaled Analyst-Based Measure

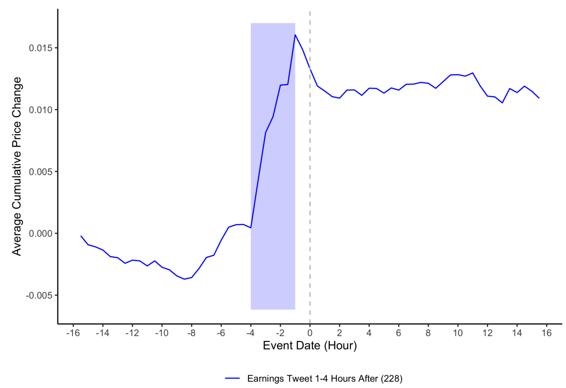


(d) SD-Scaled Random Walk

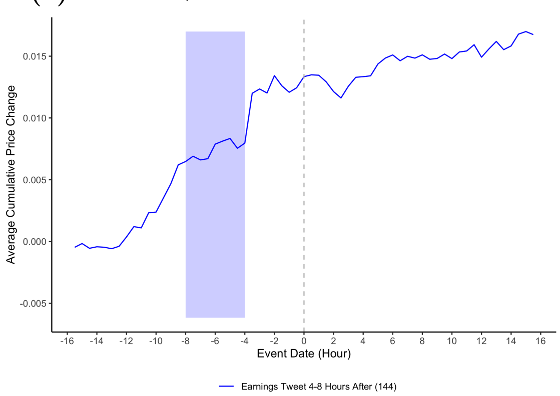
Figure B.3: Average CAR around rweets - by earnings announcement time



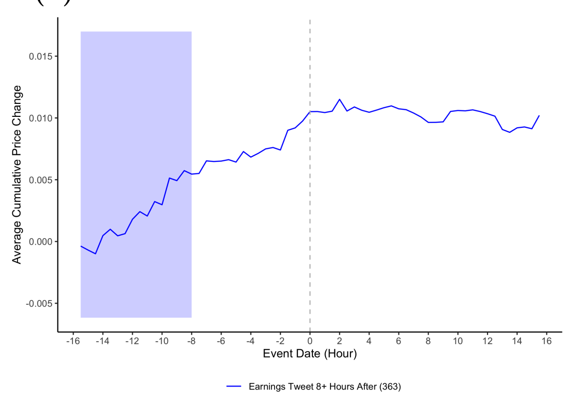
(a) Tweets < 1 Hour After Announcement



(b) Tweets 1-4 Hours After Announcement

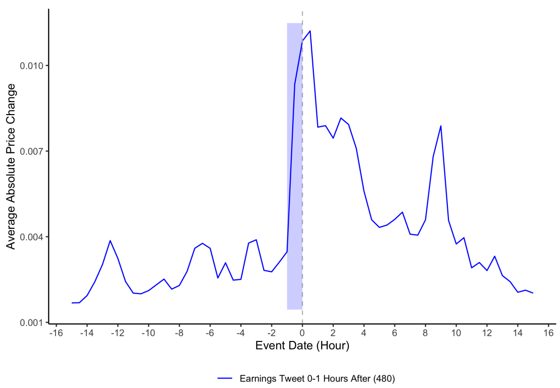


(c) Tweets 4-8 Hours After Announcement

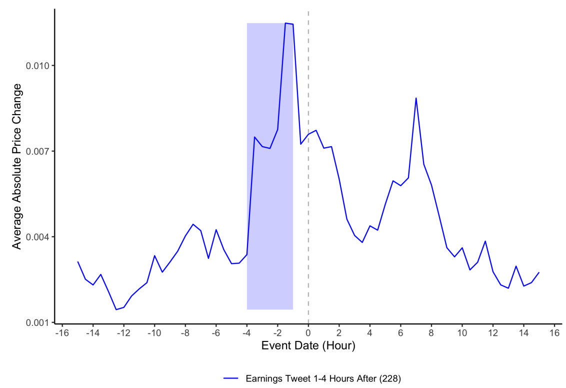


(d) Tweets > 8 Hours After Announcement

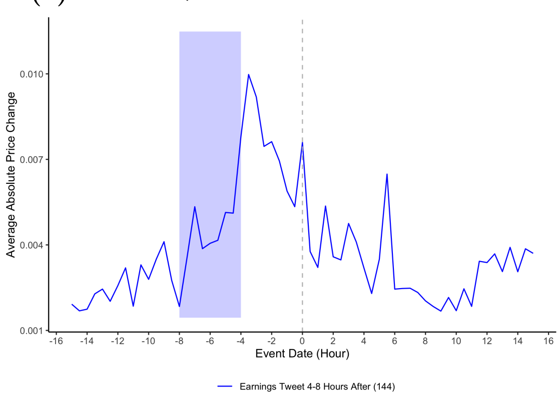
Figure B.4: Average volatility around tweets - by earnings announcement time



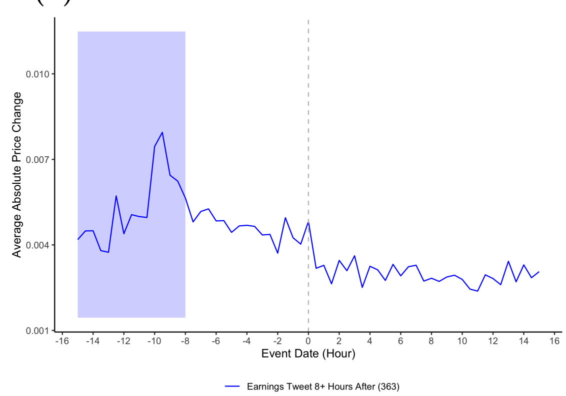
(a) Tweets < 1 Hour After Announcement



(b) Tweets 1-4 Hours After Announcement

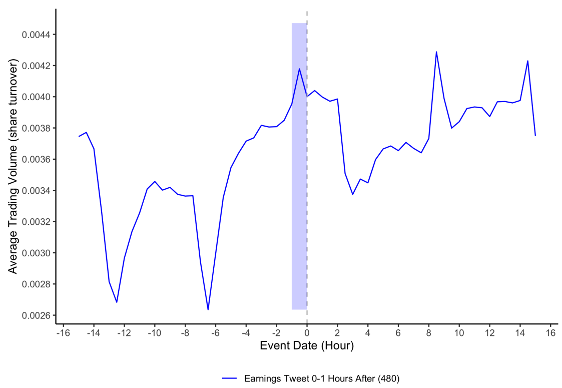


(c) Tweets 4-8 Hours After Announcement

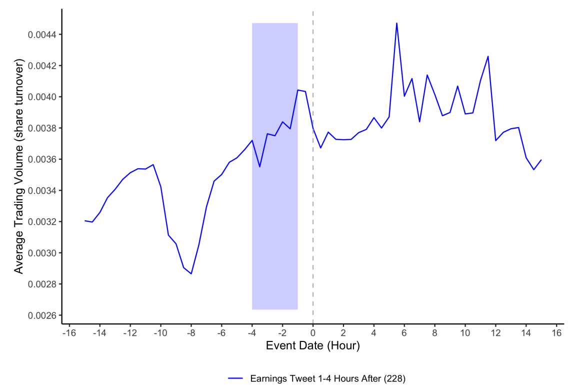


(d) Tweets > 8 Hours After Announcement

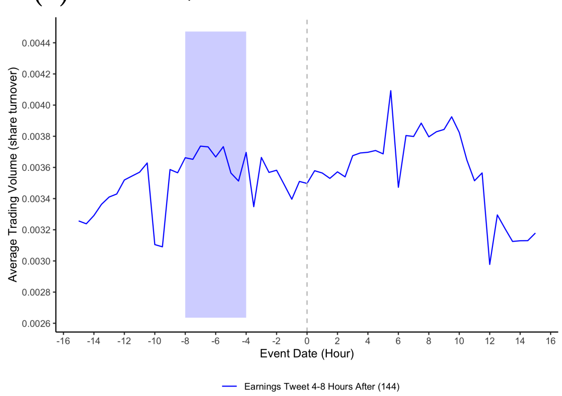
Figure B.5: Share turnover around tweets - by earnings announcement time



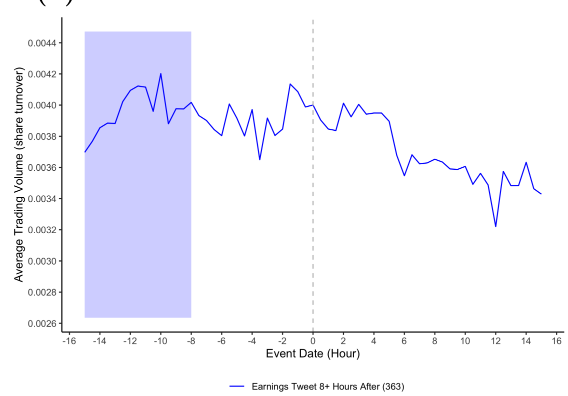
(a) Tweets < 1 Hour After Announcement



(b) Tweets 1-4 Hours After Announcement



(c) Tweets 4-8 Hours After Announcement



(d) Tweets > 8 Hours After Announcement

Table B.14: Firm tweets and CAR

	CAR [-1,+1]	
	(1)	(2)
Earnings Tweet (dummy)	-0.004^{***}	-0.002
	(0.001)	(0.002)
Miss Estimate	-0.073^{***}	-0.075^{***}
	(0.001)	(0.001)
Small Earnings Surprise	-0.031^{***}	-0.032^{***}
	(0.001)	(0.001)
Abs(SUE)	0.005	-0.061
	(0.042)	(0.047)
CEO Earnings Tweet	0.019^{***}	0.023^{***}
	(0.004)	(0.005)
Earnings Tweet (dummy) x Miss Estimate	0.015^{***}	0.016^{***}
	(0.003)	(0.003)
Earnings Tweet (dummy) x Small Earnings Surprise	0.011^{***}	0.011^{***}
”	(0.002)	(0.003)
Miss Estimate x Small Earnings Surprise	0.053^{***}	0.054^{***}
	(0.002)	(0.002)
Earnings Tweet (dummy) x Miss Estimate x Small Earnings Surprise	-0.012^{***}	-0.011^{**}
	(0.005)	(0.005)
Tweet Effect if Miss Estimate = 1 and Small Earnings Surprise = 1	0.011^{***}	0.014^{***}
Tweet Effect if Miss Estimate = 1 and Small Earnings Surprise = 0	0.011^{***}	0.014^{***}
Tweet Effect if Miss Estimate = 0 and Small Earnings Surprise = 1	0.007^{***}	0.009^{***}
Tweet Effect if Miss Estimate = 0 and Small Earnings Surprise = 0	-0.004^{***}	-0.002
Controls	x	x
Quarter FE	x	x
Year FE	x	x
Industry FE	x	
Firm FE		x
Firms	3068	3068
Observations	60,557	60,557
R^2	0.113	0.166
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table B.15: Firm tweets and CAR - sample splits

	CAR [-1,+1]			
	Small Firm Size	Large Firm Size	Low Analyst	High Analyst
	(1)	(2)	(3)	(4)
Earnings Tweet (dummy)	0.0002	−0.001	−0.0005	−0.003***
	(0.003)	(0.001)	(0.003)	(0.002)
Miss Estimate	−0.090***	−0.049***	−0.082***	−0.062***
	(0.001)	(0.001)	(0.001)	(0.001)
Small Earnings Surprise	−0.039***	−0.021***	−0.036***	−0.026***
	(0.002)	(0.001)	(0.001)	(0.001)
Abs(SUE)	−0.095	0.493***	0.044	−0.073
	(0.058)	(0.070)	(0.057)	(0.064)
CEO Earnings Tweet	0.024**	0.018***	0.023***	0.018***
	(0.011)	(0.004)	(0.009)	(0.005)
Earnings Tweet (dummy) x Miss Estimate	0.004	0.007***	0.002	0.018***
	(0.005)	(0.003)	(0.005)	(0.003)
Earnings Tweet (dummy) x Small Earnings Surprise	0.007	0.006***	0.007*	0.010***
	(0.006)	(0.002)	(0.005)	(0.003)
Miss Estimate x Small Earnings Surprise	0.066***	0.036***	0.062***	0.043***
	(0.003)	(0.002)	(0.003)	(0.002)
Earnings Tweet (dummy) x Miss Estimate x Small Earnings Surprise	−0.0003	−0.006	−0.003	−0.012**
	(0.011)	(0.005)	(0.009)	(0.006)
Tweet Effect if Miss Estimate = 1 and Small Earnings Surprise = 1	0.010	0.006	0.005	0.013***
Tweet Effect if Miss Estimate = 1 and Small Earnings Surprise = 0	0.004	0.006***	0.002	0.015***
Tweet Effect if Miss Estimate = 0 and Small Earnings Surprise = 1	0.007	0.005***	0.006	0.007***
Tweet Effect if Miss Estimate = 0 and Small Earnings Surprise = 0	0.000	−0.001	0.006	0.007***
Controls	x	x	x	x
Quarter FE	x	x	x	x
Year FE	x	x	x	x
Industry FE	x	x	x	x
Firms	1980	1805	2597	2385
Observations	30,290	30,267	30,290	30,267
R^2	0.136	0.096	0.133	0.095

Note: *p<0.1; **p<0.05; ***p<0.01