Tweeting in the Dark: Corporate Tweeting and Information Diffusion

Isabella Wolfskeil*

December 31, 2019

Abstract

Is there a link between corporate information dissemination on social media and valuations? Are social media reshaping the diffusion of corporate information? After constructing a novel and comprehensive dataset of over 7 million tweets posted by S&P 1500 firms, I adopt text analysis methods and find that firms with negative earnings surprises have higher announcement returns if they tweet about earnings news. This result is concentrated among firms with higher retail investor ownership and larger social media networks. I also find evidence that firm-initiated tweets increase investors' fundamental information acquisition and the speed of information diffusion to investors. The findings are consistent with firms managing investors expectations and utilizing social media to expedite the diffusion of corporate information, encouraging more efficient market reactions.

*Bocconi University, Department of Finance. Address: Via Röentgen 1, 20136 Milan, Italy. Email: isabella.wolfskeil@phd.unibocconi.it. I am grateful for comments from Mariano (Max) Massimiliano Croce, Miles Gietzmann, and Stefano Rossi. I thank conference participants at Showcasing Women in Finance – EU (in collaboration with Luxembourg School of Finance and AFFECT), seminar participants at Bocconi University, and discussants at the 35th EAA Doctoral Colloquium. I also thank the Finance Theory Group (FTG) PhD Summer School and CFRA Workshop on Natural Language Processing in Financial Markets for insightful comments and discussions. Any remaining errors are my own. Keywords: information diffusion, investor sentiment, limited attention, price formation.

1. Introduction

Social media has changed the way firms communicate with investors by giving them a direct, instantaneous, and network-enhanced communication channel. Furthermore, social media are gaining an outreach as relevant as that of traditional information intermediaries, like business press and financial analysts. Given these new trends, my study asks the following questions: is there a link between corporate information dissemination on social media and valuations? Can firm-initiated social media increase the fundamental information acquisition and speed of information diffusion? More broadly, are social media helping to expedite the diffusion of corporate information, encouraging more efficient market reactions?

To answer these questions I construct a novel and comprehensive dataset of tweets by S&P 1500 firms. Among social media platforms, Twitter is already connecting 300 million monthly active users. My dataset aggregates over 7 million individual tweets and represents the complete tweeting history of more than $1,000$ firms between January 2014 and December 2017.

When quantifying firms' tweeting behaviors I focus on earnings-news tweets and the relative importance of tweets in the network of Twitter users. To identify earnings-news tweets I use standard textual analysis in the literature.[1] To capture the impact of tweets in the network of Twitter users I account for both the size of the firm's network of followers and the number or retweets each tweet receives. The first main result reveals that firms with negative earnings surprises have higher announcement returns if they tweet about their earnings announcement. This effect

[1] See, for example, Bartov et al. (2017), Jung et al. (2017)

is concentrated among firms with low institutional ownership (i.e., high retail investor ownership) and among firms with a large social media networks. This result suggests that firms are able to successfully manage investors valuations, especially if investors are unsophisticated and if the firm has a large social network.

To study the implications of firm-initiated social media on investors' fundamental information acquisition I use the SEC's EDGAR Log File Data Set. This relatively new data set provides a direct measure of investor demand for financial reports. The results show that tweeting about earnings-news is associated increased fundamental information acquisition in the days surrounding earnings announcements.

To better understand what kinds of information are in firm-initiated tweets I investigate the ability of tweets to predict quarterly earnings surprises. I find that tweets contain little incremental information to predict firms' cash flows beyond the consensus.

Short-run continuation in returns has been explained theoretically and empirically by gradual diffusion of information (Hong and Stein (1999), Hong et al. (2000)). If tweeting about earnings news increases the speed of information diffusion to the market, then momentum in returns should decrease. Consistent with this hypothesis, I find firms that tweet about earnings news more consistently have less short-run continuation in returns. This result suggests that tweeting can increase the speed of information diffusion to investors. These findings are broadly supportive of the view that of the view that social media help facilitate price discovery after news releases.

The results in this paper are related to the literature on investor sentiment (De Long et al. (1990), Barberis et al. (1998), Baker and Wurgler (2006), Kumar and Lee (2006), Tetlock (2007)). Standard asset-pricing theories suggest that the current price of a stock reflects the present value of its future cash flows. According to this view, the correlations in the returns of assets are caused by correlations in changes of the assets' fundamental values. Theories that rely on investor sentiment generally explain the comovements in asset prices with demand shocks; the correlated trading activities of noise traders induce movements in asset prices which arbitrage forces may not fully offset. Investor sentiment is generally attributed to individual, retail investors (e.g., Lee et al. (1991) Barber et al. (2008))). I find that the effect of tweeting on announcement returns is concentrated among firms with low institutional ownership (i.e., high retail ownership). This result is consistent with models of investor sentiment.

The results in this paper are related to the literature on inattention in finance (Cohen and Frazzini (2008), DellaVigna and Pollet (2009), Hirshleifer and Teoh (2003), Hong and Stein (1999), Peng and Xiong (2006)). Standard asset-pricing models typically assume that markets distill new information and incorporate it into their expectations at lightning speed. In reality, such distillation and estimation is limited by investors' cost of acquiring and processing information. If investors information processing capacity is not infinite, then there are a number of reasons Twitter may increase investor attention. First, the 280 character limit on tweets, approximately 45 words, can potentially increase the salience of the information. Salience determines which information will most likely grab one's attention and have the greatest influence on one's perception of the world. Second, unlike many other important information channels such as business press, analysts' reports, and newswire services, Twitter is free, reducing the upfront costs of acquiring corporate information. Finally, Twitter is a push technology, i.e., firms can initiate the information transaction rather than waiting for investors to request the information. Consequently, potential investors who may not otherwise seek out the information can have it at their finger tips. Consistent with Hong and Stein (1999), which predict that momentum in returns should be more pronounced in stocks where information diffuses more slowly, I find firms that tweet about earnings news consistently have less short-run continuation in returns. This result suggests that tweets potentially increase investor attention and the speed of information diffusion to investors.

This study contributes to the emerging literature studying the role social media plays in financial markets. Blankespoor et al. (2013) examine how the use of social media is used to improve firms' information environments. They find evidence that firms can reduce information asymmetry by more broadly disseminating news via Twitter. Jung et al. (2017) find that firms use social media to strategically disseminate financial information. Bhagwat and Burch (2016) find that firms' use of Twitter can increase the magnitude of earnings announcement returns. The current literature has brought attention to the importance of social media for investor communication. This paper complements these prior studies by promoting a novel investigation of the link between corporate information dissemination on social media and firm valuations.

This study also contributes to the broad literature studying the impact of the media on asset prices (Dyck and Zingales (2003), Veldkamp (2006), Tetlock (2007), Fang and Peress (2009), Fedyk (2018)). This line of work studies information intermediaries, by contrast I study firm-initiated communication.

The rest of the paper is organized as follows. Section 2 discusses the regulatory setting of disclosure using social media. Section 3 describes the database. Section 4 details the empirical methodology and results. Section 5 presents robustness analyses. Finally, Section 6 concludes.

2. Institutional Background

The SEC has embraced social media and other information technologies in an effort to promote widespread access to corporate information (SEC, 2013b). Following a controversial Facebook post by the CEO of Netflix, the SEC officially stated that social media can be used as a channel for the disclosure of material, nonpublic information and provided guidance on the application of Regulation Fair Disclosure (Reg. FD) to social media (SEC, 2013b).[2] Disclosures made through social media channels fall under the umbrella of Reg. FD, therefore firms' must pursue steps to alert investors, the market, and the media to their social media platforms and disclosure practices (SEC, 2013a). Disclosures made through social media channels must also be truthful and accurate, as do all other forms of disclosure from listed firms and other regulated

[2]On July 3, 2012, the CEO of Netflix, Reed Hastings, posted the following message to his personal Facebook page: "Congrats to Ted Sarados, and his amazing content licensing team. Netflix monthly viewing exceeded 1 billion hours for the first time ever in June. When House of Cards and Arrested Development debut, we'll blow these records away. Keep going, Ted, we need even more!". The nonpublic information disclosed in the tweet, 1 billion hours, represented a 50% increase in viewing hours from Netflix's January 25, 2012 announcement. Netflix's stock price rose from $70.45 at the time of Reed Hastings's Facebook post to $81.72 at the close of the following trading day. Because material and nonpublic information was exclusively disclosed through Facebook and Netflix had not previously informed shareholders that the CEO's Facebook page would be used to disclose nonpublic information, Reed Hastings's was found in violation of Regulation FD.

firms (Stein, 2018).[3] Moving forward the importance of social media will likely grow as more firms begin to disclose information on social media.

In this paper I investigate the use of social media around earnings announcement events. Due to the importance of information released during earnings announcements, communication of earnings news is carefully regulated. The SEC requires most listed companies to file a Form 10-Q (quarterly financial report) within 40 days of the end of the quarter.[4] Companies typically file said reports and Form 10-Ks (annual financial reports) in the last two days of the required filing period (Amir and Livnat, 2005). In the days leading up to the earnings announcement firms can discuss their preliminary earnings results on social media as long as the firm jointly files a Form 8-K (current report), notifying the SEC and market participants to the information disclosure.[5] Because of the careful regulation around earnings announcements, it is likely that firms will only disclose earnings news on social media if it is accompanied by and official disclosure with the SEC, i.e., Forms 8-K, 10-K, or 10-Q.

3. Data

The general goal of this paper is to examine the role of social media in the dissemination of corporate information. From a practical point of view, however, there are many

[3]In 2018 Elon Musk was charged with securities fraud for a misleading tweet. On August 7, 2018 Musk tweeted "Am considering taking Tesla private at $420. Funding secured." The SEC's complaint alleges that Musk knew that the potential transaction was uncertain and violated antifraud provisions of the federal securities laws (SEC, 2018).

[4]Non-accelerated filers with less than a public float of $75 million are granted 45 days.

[5]It is common practice for firms to disclose preliminary earnings results; Amir and Livnat (2005) find that 80% of firms in their sample consistently issue preliminary earnings announcements—on average, 26 days after quarter-end.

reasons to focus on the Twitter platform. Twitter is a micro-blogging network intended for sharing news, content, and information. Twitter is connecting more than 300 million monthly active users who post, read, and interact with short messages known as "tweets". Unlike many other social media platforms, Twitter has a strong emphasis on real-time information. Twitter enables firms to broadcast financial news directly and instantaneously to a large social network. Twitter is the most widely adapted social media platform by S&P 1500 firms (Jung et al., 2017). Increasingly, investor relations departments are using Twitter to reach investors with messages about earnings announcements, management changes, and public relations crises. A growing number of companies are creating Twitter accounts specifically for investors, for example, Ford Motor Co. (@FordIR), T-Mobile (@TMobileIR) and CVS Health Corp (@CVShealthIR).

A. Data collection and sample selection

To study how social media is reshaping the diffusion of corporate information, I construct a dataset of 7,132,461 individual tweets posted by S&P 1500 firms between January 2014 and December 2017. This firm-tweet data is merged with financial data and market data to relate tweeting activity with firm valuations and short-run continuation in returns.

To gather the data I begin with an initial sample of 2,454 firms, which includes all S&P 1500 firms as of March 2006, the month Twitter was founded. From the starting sample of 2,454 firms, I identified 1,215 firms with active Twitter accounts by manually searching for each account.[6] Of the 1,215 accounts, 489 are verified. The verified feature on Twitter is a signal to the public that an account of public interest is authentic. After gathering the sample of Twitter usernames I assembled a complete history of tweets generated by the 1,215 accounts between January 1, 2014 and December 30, 2017, resulting in a sample of 7,132,461 individual tweets. To isolate firm initiated content that is visible to the firms' followers, I exclude tweets that are reply tweets and retweets.[7] This process reduces the sample to 3,305,257 tweets.

The SEC's EDGAR log file data set is a collection of web server log files that allow researchers to study firm specific web traffic of individuals downloading SEC filings. EDGAR is the central repository for all mandatory SEC filings and the daily level EDGAR search volume for each firm is a direct measure of investors fundamental information acquisition. EDGAR log file data are obtained from James Ryans's webpage. [8]

Quarterly earning announcement dates and analyst consensus forecasts are obtained from Compustat and I/B/E/S, respectively. Daily stock prices are obtained from CRSP, and institutional ownership data is obtained from Thomson Reuters

[6] I started the search on each firms' corporate website, if there was no Twitter handles mentioned on the corporate website I proceeded to search directly on Twitter. The search was conducted in October 2017, therefore the sample is composed of firms which had active Twitter accounts in October 2017.

[7] A reply tweet is a public Tweet directed at a specific person. Reply tweets appear in the feeds of the specific user the firm is replying to, and anyone who follows both the firm that posts the tweet and the user being replied to. It does not appear in the feed of everyone who is following the firm which posts the tweet. A retweet is a repost of another Twitter user's tweet on the firm's own profile. Unlike reply tweets, retweets appear in the feed of everyone who is following the firm that reposts the tweet. However the retweet itself is not original content created by the firm.

[8] The summarized EDGAR log files used in this paper are available for academic use at http://www.jamesryans.com.

13F database. After excluding observations without necessary from data Compustat, CRSP, I/B/E/S, and Thomson Reuters, the final sample includes 1,058 firms and 13,330 firm-quarter observations. A larger sample is used when exploring the effects of tweeting on the speed of information diffusion. This research question is not limited by data availability from I/B/E/S and Thomson Reuters and therefore only excludes observations without necessary data on Compustat and CRSP, resulting in a sample of 1,064 firms and 50,456 firm-month observations.

In my sample, the frequency of tweets over time is relatively flat. The number of firm-quarter observations increases over the sample. This pattern is to be expected because some Twitter users in the sample were not active at the start of the sample period. Appendix Table A3 presents the frequency distribution of tweets and firm-quarter observations by calendar quarter.

The average date firms joined Twitter was in November 2010. The 1st percentile (99th percentile) joined Twitter in June 2007 (May 2017). Twitter users have a mean (median) of 162,642 (6,352) followers and 2,438 (557) friends. Firm-quarters have a mean (median) of 178 (81) tweets. There is considerable heterogeneity across firms' Twitter accounts, this suggests the effect of tweeting may vary by firm. To address this concern I use firm fixed effects and standard errors clustered by firm when estimating Equations (1) and (4). I also include the number of firm followers and the number of retweets when measuring the impact of firm tweets. Appendix Table A4 presents descriptive statistics related to tweet characteristics.

B. Identifying earnings-news tweets

Two of the primary challenges underlying the research design are detecting earnings-news tweets and estimating the importance of individual tweets in the network. Along the lines of prior research I use textual analysis to detect earnings-news tweets.[9] I use a classification scheme based on a dictionary of keywords and phrases; each tweet is considered earnings news if it contains two or more of the terms found in the dictionary.[10]

Using this textual classification approach, I identify 19,148 tweets (5,549 firm-quarters, 783 unique firms) that contain information directly related to earnings announcements. Examples of earnings-news tweets in my sample are provided in Appendix B. As one would expect, earnings-news tweets are concentrated around announcement periods. The number of earnings-news tweets in a 10 days window around the announcement represents, on average, around one-fourth ($0.76/(2.37 + 0.76)$) of all tweets in that period.[11] Figure 1 depicts the relationship between the number of earnings-news tweets and the number of days away from firms' earnings announcement. Most earnings-news tweets take place on the day of the announcement. Figure 2 depicts the relationship between the number of earnings-news tweets on the day of the announcement and the time of day the tweet was posted. Most earnings-news tweets are posted just after lunch or in the late evening.

[9] See, for example, Bartov et al. (2017), Jung et al. (2017)

[10] The dictionary of keywords and phrases can be found in Appendix A.

[11] Summary statistics are provided in Appendix Table A2.

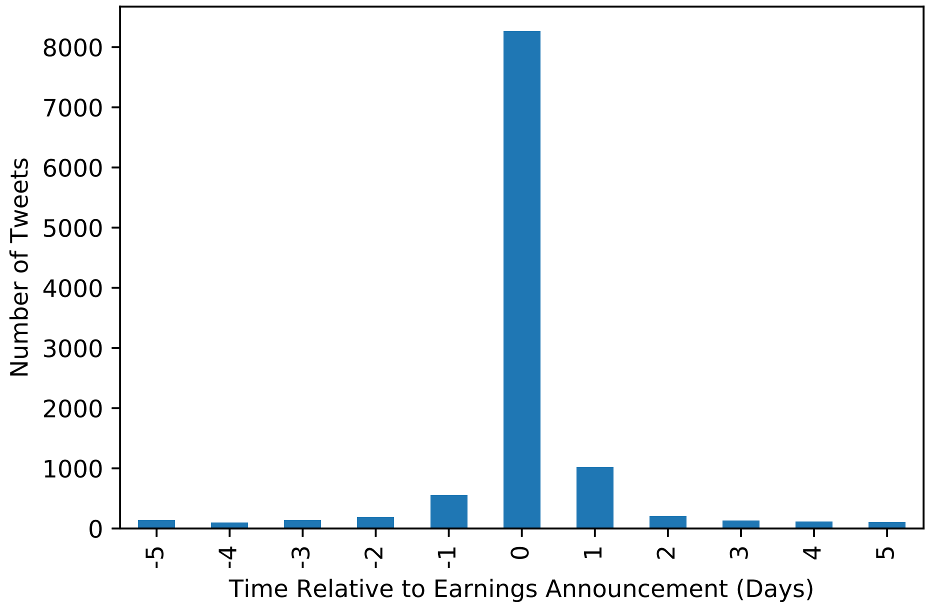


Figure 1: Daily tweeting around earnings announcements. This figure depicts the relationship between the number of earnings-news tweets and the number of days away from firms' earnings announcements. All tweets included in the figure are written during the sample period (Jan. 2014 through Dec. 2017) by S&P 1500 and meet basic minimum word requirements to be considered earnings-news tweets.

C. Measuring network impact

Twitter is an interactive network, therefore it is important to consider the network effects at play when measuring the relative impact of individual tweets. When a tweet is posted by a firm this message is immediately accessible to the firm's followers on their Twitter account. These followers then have the option to retweet and like the tweets; if a tweet posted by a firm is retweeted or liked by one of the firm's followers, user j, then the tweet can be seen by both the firm's followers and user $j's$ followers. As the process of retweeting and liking continues, a tweet can potentially spread through the entire network.

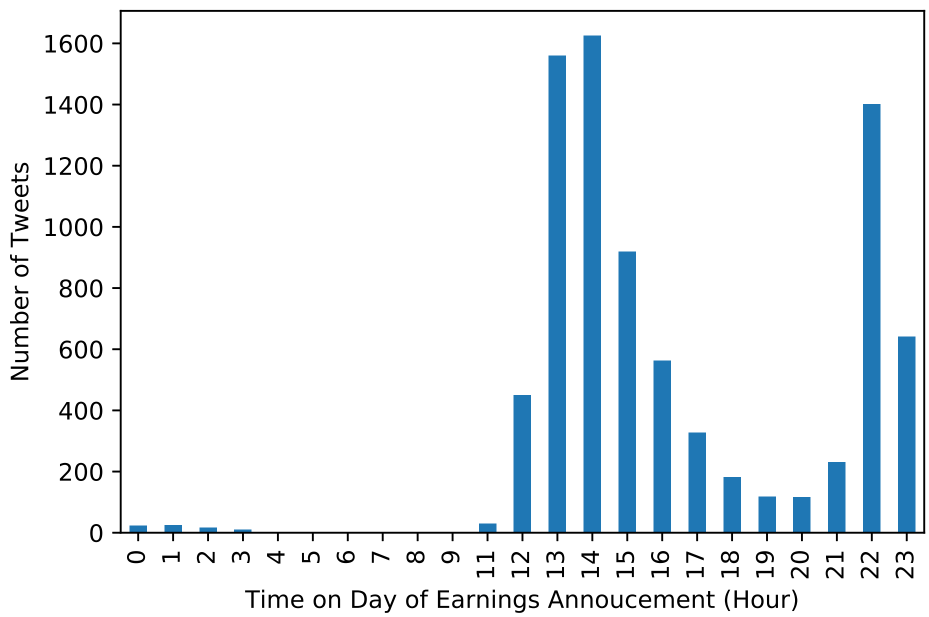


Figure 2: Hourly tweeting around earnings announcements. This figure depicts the relationship between the number of earnings-news tweets on the day of the announcement and the time of day the tweet was posted. All tweets included in the figure are written during the sample period (Jan. 2014 through Dec. 2017) by S&P 1500 and meet basic minimum word requirements to be considered earnings-news tweets (see text for details).

To capture the network effects at play I measure the impact of firms' earnings-news tweets ($EarningsTweetImpact$) in three alternative ways. First, I use the number of tweets about earnings as a naive proxy for the impact of the earnings-news tweets in the network. To proxy for the number users that read and process firm's earnings-news tweets, I measure the impact of the tweets as the number of tweets about earnings multiplied by the log number of followers the firm has on Twitter. Finally I consider the popularity of the tweets, to capture how far tweets spread in the network, I measure the impact of the tweets as the number of tweets about earnings multiplied by the log number of retweets.

4. Empirical Design and Results

A. Tweeting and Announcement Returns

Is there a link between corporate information dissemination on social media and valuations? I begin the empirical analysis by examining ability of Twitter to affect investor sentiment. To test this question I examine the association between abnormal announcements returns and firms' tweeting behavior by estimating the following regression:

$$\begin{aligned} CAR_{i,t} =& \alpha + \beta_1 NegativeSurprise_{i,t} + \beta_2 EarningsTweetImpact_{i,t} \\ &+ \beta_3 NegativeSurprise_{i,t} * EarningsTweetImpact_{i,t} \\ &+ \beta_4 X_{i,t} + \theta_i + \psi_t + \varepsilon_{i,t}. \end{aligned} \tag{1}$$

In equation (1), the dependent variable, $CAR_{i,t}$, is the Carhart (1997) cumulative abnormal return for firm i over the three-day window [-1, +1] around the quarterly earnings announcement. $NegativeSurprise_{i,t}$ is an indicator variable equal to one if firm i meets or beats their analysts consensus forecast in quarter t, and zero otherwise. $EarningsTweetImpact_{i,t}$ captures the extent that firm i tweets about its earnings announcement over the three-day window [-1, +1] around the earnings announcement. $NegativeSurprise_{i,t}$ * $EarningsTweetImpact_{i,t}$ is an interaction term, this variable helps capture the impact of a firm's earnings-news tweets, given that the firm beats or misses its consensus forecast.

The control variables, $X_{i,t}$, include $Size_{i,t}$, $B/M_{i,t}$, $AnalystsFollowing_{i,t}$, $\mid SUE \mid_{i,t}$, $Q4_{i,t}$, $Loss_{i,t}$, and $Non\text{-}earningsTweetImpact_{i,t}$. To mitigate the influence of outliers, all continuous variables are winsorized at the 1st and 99th percentiles. Appendix table A2 reports summary statistics on the variables used to estimate equation (1). All variables are defined in detail in Appendix Table A1.

A major concern is that the choice to tweet about earnings news is endogenous. A rich set of fixed effects help to control for determinants of tweeting; year-quarter fixed effects help control for omitted variables related to time varying macro factors, and firm-year (firm) fixed effects control for time-varying (time-invariant) firm characteristics.

Earnings announcement specific characteristics can also bias the estimates. Firms may be more likely to disclose bad news on Fridays than on Mondays through Thursday (DellaVigna and Pollet, 2009). To control for the variation of announcements on different days, I use day-of-week fixed effects. To control for observable announcement specific characteristics I include the variables $\mid SUE_{i,t} \mid$, $Q4_{i,t}$, and $Loss_{i,t}$.

In Table 1, equation (1) is estimated using firm and year-quarter fixed effects. For firms that miss their analysts consensus forecast, the effect of tweeting about earnings in column (2) is $0.139 + 1.345 = 1.484$, which the F-test shows is significant at 1%. The coefficient estimate for *EarningsTweetImpact* in column (2) is 0.139 and is statistically insignificant, meaning for firms that meet their analysts consensus forecast, tweeting about earnings is not associated with a change in the announcement return. The results are robust to different measures of *EarningsTweetImpact* in columns (2) and (3).

The within-group estimates suggest that when the same firm tweets about earnings over different quarters, tweeting has a different effect on announcement returns depending on whether the firm has a positive or negative earnings surprise. Firms with negative surprises, i.e., firms that miss their analyst consensus forecasts, have higher announcement returns if they tweet about earnings news. One way to interpret this finding is that firms are able to successfully manage investors' valuations.

The results are robust to a pooled OLS estimation and rich set of fixed effects. In Table 2, column (1) shows pooled OLS estimates, the coefficient estimate for *NegSurp*EarningsTweetImpact* is 1.266. Including quarter-year and firm-year fixed effects the coefficient estimate increases to 1.450 in column (5). The results are generally consistent across specifications, however the within-group estimates tend to be higher and more statistically significant than the pooled OLS estimates.

A.1. Institutional ownership and network size

In this section I investigate firm specific characteristics that may affect a firm's ability to manage investor expectations using Twitter. Investors have heterogeneous costs of information acquisition and processing, unsophisticated investors tend to have higher costs. Therefore, if tweets reduce the cost of information processing, the evidence of predictability should be stronger or even exclusively concentrated among stocks with high retail ownership. To test this hypothesis, I examine the effects of institutional ownership on announcement return predictability.

I identify the percentage of a firms owned by institutions using 13f data from Thomson Reuters. Based on the average value of total institutional ownership as a percentage of shares outstanding, I sort firms into high or low institutional ownership categories and re-estimate equation (1) on the subsamples of firms.

Table 1: Tweeting and Announcement Returns

This table shows the relationship between cumulative abnormal returns (*CAR*) and firms' tweeting behaviors. In columns (1) and (2) *EarningsTweetImpact* is measured as *EarningsTweetCount*, in column (3) as *EarningsTweetCount*Followers*, and in column (4) as *EarningsTweetCount*Retweets*. To mitigate the influence of outliers, all continuous variables are winsorized at the 1st and 99th percentiles. See Appendix A for variable definitions. Standard errors are clustered at the firm level and provided in parentheses beneath the coefficient estimates. *, **, and *** indicates significance at the 10%, 5%, and 1%, respectively.

	CAR			
	(1)	(2)	(3)	(4)
Negative Surprise	-4.973***	-5.377***	-5.183***	-5.378***
	(0.198)	(0.231)	(0.212)	(0.227)
Earnings Tweet Impact	0.441**	0.139	-0.010	0.009
	(0.205)	(0.215)	(0.177)	(0.021)
*Neg Surp*Earnings Tweet Impact*		1.345***	1.162***	0.140***
		(0.308)	(0.266)	(0.029)
Non-earnings Tweet Count	0.021	0.024	0.067	0.004
	(0.105)	(0.106)	(0.051)	(0.010)
Residual ESV (Ryans)	-0.365	-0.364	-0.370	-0.363
	(0.302)	(0.303)	(0.302)	(0.303)
Absolute SUE	0.268***	0.271***	0.271***	0.272***
	(0.031)	(0.031)	(0.032)	(0.031)
Size	-2.651***	-2.622***	-2.639***	-2.626***
	(0.474)	(0.475)	(0.475)	(0.475)
Loss	-1.905***	-1.889***	-1.897***	-1.887***
	(0.366)	(0.365)	(0.365)	(0.365)
BM	8.237***	8.263***	8.247***	8.259***
	(0.883)	(0.881)	(0.876)	(0.880)
Q4	0.114	0.105	0.097	0.104
	(0.232)	(0.232)	(0.231)	(0.231)
Firm FE	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes
F-test statistic: $\beta2 + \beta3 = 0$		22.869	16.374	26.571
F-test p-value		0.000	0.000	0.000
No. of firms	1009	1009	1009	1009
Adjusted R^2	0.141	0.143	0.142	0.143
Observations	11,483	11,483	11,483	11,483

Table 2: Various Fixed Effects: Tweeting and Announcement Returns

This table shows the relationship between cumulative abnormal returns (*CAR*) and firms' tweeting behaviors. In columns (1) and (2) *EarningsTweetImpact* is measured as *EarningsTweetCount*, in column (3) as *EarningsTweetCount*Followers*, and in column (4) as *EarningsTweetCount*Retweets*. To mitigate the influence of outliers, all continuous variables are winsorized at the 1st and 99th percentiles. See Appendix A for variable definitions. Standard errors are clustered at the firm level and provided in parentheses beneath the coefficient estimates. *, **, and *** indicates significance at the 10%, 5%, and 1%, respectively.

	CAR				
	(1)	(2)	(3)	(4)	(5)
Negative Surprise	-5.139***	-5.142***	-5.383***	-5.621***	-5.620***
	(0.217)	(0.218)	(0.229)	(0.269)	(0.269)
Earnings Tweet Impact	-0.226*	-0.222*	0.121	0.252	0.252
	(0.126)	(0.127)	(0.213)	(0.298)	(0.297)
*Neg Surp*Earnings Tweet Impact*	1.266***	1.255***	1.341***	1.466***	1.450***
	(0.299)	(0.298)	(0.308)	(0.355)	(0.356)
Non-earnings Tweet Count	-0.040	-0.044	0.020	0.090	0.100
	(0.067)	(0.067)	(0.105)	(0.154)	(0.154)
Residual ESV (Ryans)	-0.328	-0.308	-0.328	-0.473	-0.445
	(0.299)	(0.305)	(0.299)	(0.323)	(0.323)
Absolute SUE	0.231***	0.234***	0.269***	0.260***	0.259***
	(0.028)	(0.028)	(0.031)	(0.036)	(0.036)
Size	-0.224***	-0.219***	-2.616***	-4.232***	-4.254***
	(0.043)	(0.043)	(0.423)	(1.171)	(1.168)
Loss	-1.223***	-1.206***	-1.863***	-2.180***	-2.189***
	(0.273)	(0.274)	(0.363)	(0.436)	(0.436)
BM	1.573***	1.576***	7.920***	21.056***	21.068***
	(0.243)	(0.244)	(0.835)	(2.301)	(2.293)
Q4	0.302*	0.171	0.368**	0.233	0.221
	(0.156)	(0.227)	(0.157)	(0.243)	(0.244)
Quarter-year FE	No	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No
Firm-year FE	No	No	No	Yes	Yes
Weekday FE	No	No	No	No	Yes
F-test statistic: $\beta 2 + \beta 3 = 0$	17.072	17.065	22.260	20.260	19.885
F-test p-value	0.000	0.000	0.000	0.000	0.000
No. of firms	1009	1009	1009	976	976
Adjusted R^2	0.113	0.114	0.142	0.158	0.158
Observations	11,484	11,483	11,484	10,863	10,863

Table 3 presents regression estimates for both low and high institutional ownership subsamples. The estimation results reveal that only firms with low institutional ownership and negative surprises having higher announcement returns if they tweet about earnings news. In particular, the coefficient estimate for *NegSurp*EarningsTweetImpact* for the low retail ownership firms in column (1), is 0.944 and statistically significant at 10%. In contrast, the coefficient estimate for the high retail ownership firms in column (4) is 1.358 and statistically significant at 1%.

Next, I investigate the effects of the size of a firm's network of Twitter followers on announcement return predictability. If tweets are in fact influencing abnormal stock returns, the evidence of predictability should be stronger among firms with larger networks on Twitter. Based on the size of each firm's network of followers, I sort firms into big or small network categories and re-estimate equation (1).

Table 4 presents regression estimates for both network size subsamples. The estimation results reveal that only firms with large networks on Twitter and negative surprises having higher announcement returns if they tweet about earnings news. In particular, the coefficient estimate for *NegSurp*EarningsTweetImpact* for the large network firms in column (4), is 1.408 and statistically significant at 1%. In contrast, the coefficient estimate for the small network firms in column (1) is insignificant (estimate = 0.630).

B. Tweeting and fundamental information acquisition

Does tweeting encourage fundamental information acquisition? To test this question I examine the association between abnormal announcements returns and firms' tweeting

Table 3: Institutional Ownership: Tweeting and Announcement Returns

This table shows the relationship between the dependent variable, cumulative abnormal returns (*CAR*), and firms' tweeting behavior. The sample is split into high or low insitutional ownership firms (above or below the sample median of total institutional ownership as a percentage of shares outstanding). In columns (1) and (4) *EarningsTweetImpact* is measured as *EarningsTweetCount*, columns (2) and (5) as *EarningsTweetCount*Followers*, and columns (3) and (6) as *EarningsTweetCount*Retweets*. To mitigate the influence of outliers, all continuous variables are winsorized at the 1st and 99th percentiles. See Appendix A for variable definitions. Standard errors are clustered at the firm level and provided in parentheses beneath the coefficient estimates. *, **, and *** indicates significance at the 10%, 5%, and 1%, respectively.

	Low Retail Firms			High Retail Firms		
	(1)	(2)	(3)	(4)	(5)	(6)
Negative Surprise	-6.046***	-5.959***	-6.033***	-4.627***	-4.351***	-4.627***
	(0.330)	(0.305)	(0.326)	(0.310)	(0.280)	(0.306)
Earnings Tweet Impact	0.231	-0.245	0.019	0.160	0.118	0.013
	(0.390)	(0.360)	(0.043)	(0.230)	(0.206)	(0.022)
*Neg Surp*Earnings Tweet Impact*	0.944*	1.380***	0.098	1.358***	0.805***	0.135***
	(0.564)	(0.526)	(0.062)	(0.347)	(0.309)	(0.031)
Non-earnings Tweet Count	0.165	0.060	0.020	-0.078	0.065	-0.007
	(0.165)	(0.082)	(0.017)	(0.133)	(0.063)	(0.013)
Residual ESV (Ryans)	0.053	0.061	0.052	-0.848**	-0.859**	-0.845**
	(0.428)	(0.426)	(0.428)	(0.425)	(0.425)	(0.425)
Absolute SUE	0.247***	0.247***	0.247***	0.304***	0.302***	0.305***
	(0.040)	(0.040)	(0.040)	(0.050)	(0.050)	(0.049)
Size	-3.104***	-3.109***	-3.111***	-1.762**	-1.761**	-1.763**
	(0.667)	(0.660)	(0.667)	(0.701)	(0.701)	(0.701)
Loss	-1.768***	-1.798***	-1.773***	-2.009***	-2.000***	-2.000***
	(0.479)	(0.478)	(0.480)	(0.566)	(0.566)	(0.566)
BM	7.224***	7.260***	7.226***	9.748***	9.685***	9.737***
	(1.145)	(1.140)	(1.144)	(1.287)	(1.278)	(1.285)
Q4	0.237	0.218	0.240	-0.031	-0.043	-0.035
	(0.368)	(0.368)	(0.367)	(0.268)	(0.267)	(0.267)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes	Yes	Yes
F-test statistic: $\beta2 + \beta3 = 0$	5.403	4.465	5.215	15.025	7.792	17.471
F-test p-value	0.020	0.035	0.023	0.000	0.005	0.000
No. of firms	509	509	509	500	500	500
Adjusted R^2	0.148	0.148	0.148	0.139	0.138	0.139
Observations	5,658	5,658	5,658	5,825	5,825	5,825

Table 4: Institutional Ownership: Tweeting and Announcement Returns

This table shows the relationship between the dependent variable, cumulative abnormal returns (*CAR*), and firms' tweeting behavior. The sample is split into large or small network firms (above or below the sample median of Twitter followers). In columns (1) and (4) *EarningsTweetImpact* is measured as *EarningsTweetCount*, columns (2) and (5) as *EarningsTweetCount*Followers*, and columns (3) and (6) as *EarningsTweetCount*Retweets*. To mitigate the influence of outliers, all continuous variables are winsorized at the 1st and 99th percentiles. See Appendix A for variable definitions. Standard errors are clustered at the firm level and provided in parentheses beneath the coefficient estimates. *, **, and *** indicates significance at the 10%, 5%, and 1%, respectively.

	Small Network Firms			Large Network Firms		
	(1)	(2)	(3)	(4)	(5)	(6)
Negative Surprise	-5.436***	-5.473***	-5.339***	-5.172***	-5.172***	-5.063***
	(0.318)	(0.314)	(0.282)	(0.301)	(0.299)	(0.285)
Earnings Tweet Impact	0.180	0.014	0.088	0.103	0.005	-0.132
	(0.310)	(0.038)	(0.349)	(0.243)	(0.022)	(0.166)
*Neg Surp*Earnings Tweet Impact*	0.630	0.095	0.936	1.408***	0.126***	1.212***
	(0.532)	(0.064)	(0.820)	(0.379)	(0.033)	(0.274)
Non-earnings Tweet Impact	0.070	0.011	-0.107	0.088	0.011	0.126**
	(0.154)	(0.019)	(0.125)	(0.133)	(0.012)	(0.054)
Absolute SUE	0.225***	0.225***	0.224***	0.252***	0.252***	0.252***
	(0.039)	(0.039)	(0.039)	(0.041)	(0.041)	(0.041)
Size	-2.288***	-2.290***	-2.247***	-2.629***	-2.632***	-2.653***
	(0.624)	(0.624)	(0.620)	(0.526)	(0.526)	(0.527)
Loss	-1.663***	-1.664***	-1.669***	-1.539***	-1.539***	-1.553***
	(0.462)	(0.462)	(0.464)	(0.410)	(0.409)	(0.407)
BM	8.316***	8.314***	8.297***	6.945***	6.948***	6.973***
	(0.999)	(0.998)	(0.995)	(1.214)	(1.214)	(1.194)
Q4	-0.242	-0.245	-0.246	0.382	0.385	0.388
	(0.343)	(0.344)	(0.344)	(0.262)	(0.262)	(0.262)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes	Yes	Yes
F-test statistic: $\beta 2 + \beta 3 = 0$	2.133	2.595	1.699	19.718	19.385	12.997
F-test p-value	0.145	0.108	0.193	0.000	0.000	0.000
No. of firms	598	598	598	460	460	460
Adjusted R^2	0.147	0.147	0.147	0.118	0.118	0.119
Observations	6,299	6,299	6,299	7,454	7,454	7,454

behavior by estimating the following regression:

$$\begin{aligned} ESV_{i,t} =& \alpha + \beta_1 EarningsTweetCount_{i,t} + \beta_2 NonEarningsTweetCount_{i,t} \\ & + \beta_3 X_{i,t} + \theta_i + \psi_t + \varepsilon_{i,t}. \end{aligned} \tag{2}$$

In equation (2), the dependent variable, $ESV_{i,t}$, is the daily Edgar Search Volume from the SEC's EDGAR web server log file data for firm i over the three-day window [-1, +1] around the quarterly earnings announcement. $EarningsTweetImpact_{i,t}$ captures the extent that firm i tweets about its earnings announcement over the three-day window [-1, +1] around the earnings announcement.

The control variables, $X_{i,t}$, include $Size_{i,t}$, $B/M_{i,t}$, $AnalystsFollowing_{i,t}$, $\mid SUE \mid_{i,t}$, $Q4_{i,t}$, $Loss_{i,t}$, and $\textit{Non-earningsTweetImpact}_{i,t}$. To mitigate the influence of outliers, all continuous variables are winsorized at the 1st and 99th percentiles. Appendix table A2 reports summary statistics on the variables used to estimate equation (1). All variables are defined in detail in Appendix Table A1.

In Table 5, equation (2) is estimated using firm and year-quarter fixed effects. The coefficient estimates for *EarningsTweetImpact* are statistically significant at 5% in columns (2) and (3), indicating that tweeting about earnings is associated with more fundamental information acquisition by investors.

C. Tweeting and the speed information diffusion

Can firm-initiated tweets increase the speed of information diffusion? Momentum in returns has been explained theoretically and empirically by gradual diffusion of information (Hong and Stein (1999), Hong et al. (2000)). Momentum in stock returns

Table 5: Tweeting and Fundamental Information Acquisition

		ESV	
	(1)	(2)	(3)
Earnings Tweet Impact	0.018	0.033**	0.029**
	(0.014)	(0.014)	(0.014)
Non-earnigns Tweet Impact	-0.013*	-0.014*	-0.020***
	(0.007)	(0.007)	(0.007)
Negative Surprise	0.026***	0.033***	0.027***
	(0.009)	(0.010)	(0.009)
Absolute SUE	0.002	0.002	0.001
	(0.001)	(0.001)	(0.001)
Size	0.304***	0.253***	0.264***
	(0.034)	(0.036)	(0.032)
Loss	0.030*	0.045**	0.031**
	(0.017)	(0.018)	(0.015)
BM	-0.033	-0.005	-0.012
	(0.039)	(0.039)	(0.038)
Q4	0.024*	0.038**	0.016
	(0.014)	(0.015)	(0.014)
Analysts	0.001	0.002*	0.002**
	(0.001)	(0.001)	(0.001)
Firm FE	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes
No. of firms	1009	1009	1009
Adjusted R^2	0.797	0.816	0.811
Observations	11,483	11,483	11,483

is a longstanding empirical fact; that is, securities which have performed well over the prior 6-12 months continue to outperform relative to those that did poorly, for the next 3-12 months Jegadeesh and Titman (1993).

If tweeting about earnings news increases the speed of information diffusion to the market, then momentum in returns should decrease. To test this prediction I estimate the following regression:

$$Momentum_i = \alpha + \beta_1 EarningsTweetQuarters_i + \beta_2 X_i + \varepsilon_i. \tag{3}$$

In equation (3), the dependent variable, $Momentum_i$, is a proxy for momentum as it is defined in the empirical asset pricing literature à la Jegadeesh and Titman (1993). $Momentum_i$ is measured as the correlation between the series $ExRet_{i,t}$ and the lagged series $ExRet_{i[t-12,t-2]}$, where $ExRet_{i,t}$ is the monthly excess return of firm i. $EarningsTweetQuarters_i$ is the proportion of quarters a firm tweets about earnings news over the sample period, January 2014 through December 2017. I construct $Momentum_i$ using $t \in$ {January 2014,..., December 2017} to match the sample period.

The controls, X_i, include $Size_i$, B/M_i, and $Analysts_i$ and are measured using the average value over the sample period. All variables are defined in detail in Appendix A. Appendix table A2 presents the descriptive statistics for the variables used to estimate equation (3).

Table 6 presents the results. In columns (1) and (3) *Momentum* is calculated using excess returns relative to 90 day T-bills and in columns (2) and (4) using Fama-French three factor excess returns. In columns (1) and (2) the coefficients are estimated using the full sample of firms and the coefficient estimates for $EarningsTweetQuarters$ are not statistically significant. However, once the sample is restricted to verified accounts only, in columns (3) and (4), coefficient estimates for $EarningsTweetQuarters$ are -0.041 and -0.038 and are statistically significant at 1% and 5%, respectively. This negative relationship suggests firms with verified accounts that tweet about earnings news more consistently have less momentum in returns. This result suggests firms may be able to increase the speed of information diffusion to investors by tweeting about earnings news. This result is consistent with media's role to disseminate information quickly.

5. Additional Analysis

A. Tweeting and cash-flows

Jung et al. (2017) show that firms are more likely to tweet about good news than bad news. If firms are tweeting more about earnings news in the month prior to the announcement than usual, is that a possible signal that the firm may outperform expectations? Can the amount a firm tweets about earnings news help predict the outcome of the announcement beyond the consensus?

To better understand what kinds of information are in firm-initiated tweets, I investigate the ability of tweets in the month prior to earnings announcements to predict firms' earnings. To test the predictive ability of information in tweets I

Table 6: Tweeting and The Speed of Information Diffusion

This table shows the cross-sectional relationship between momentum in monthly stock returns (*Momentum*) and the consistency of tweeting about earnings news. The dependent variable is *Momentum*; in columns (1) and (3) *Momentum* is calculated using excess returns relative to 90 day T-bills and in columns (2) and (4) using Fama-French three factor excess returns. Columns (3) and (4) are estimated using the subsample of firms with verified Twitter accounts. To mitigate the influence of outliers, all continuous variables are winsorized at the 1st and 99th percentiles. See Appendix A for variable definitions. The regression is estimated using OLS with robust standard errors. Standard errors are provided in parentheses beneath the coefficient estimates. *, **, and *** indicates significance at the 10%, 5%, and 1%, respectively.

	All Firms		Verified Twitter Firms	
	(1)	(2)	(3)	(4)
Earnings Tweet Quarters	-0.018	-0.002	-0.041***	-0.040**
	(0.011)	(0.013)	(0.016)	(0.020)
BM	0.022	-0.007	0.038*	0.032
	(0.015)	(0.034)	(0.020)	(0.041)
Analysts	0.001	-0.001	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Size	0.000	0.004	0.004	0.007
	(0.004)	(0.005)	(0.005)	(0.008)
Verified	-0.014	-0.026**		
	(0.010)	(0.011)		
Institutional Ownership	0.010	0.014	0.013	0.003
	(0.032)	(0.036)	(0.049)	(0.066)
Twitter Followers	0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.090*	-0.120**	-0.176**	-0.190**
	(0.053)	(0.050)	(0.073)	(0.094)
R^2	0.054	0.022	0.064	0.047
Observations	1,064	848	443	356

estimate the following regression:

$$SUE_{i,t} = \alpha + \beta_1 EarningsTweetImpact_{i,t} + \beta_2 X_{i,t} + \theta_i + \psi_t + \varepsilon_{i,t}. \qquad (4)$$

In equation (4), the dependent variable, $SUE_{i,t}$, is standardized unexpected earnings. $SUE_{i,t}$ is defined as the I/B/E/S reported quarterly earnings per share (EPS) less the latest I/B/E/S consensus analyst forecast, scaled by either the standard deviation of analysts' forecasts, price per share, or book equity per share. $EarningsTweetImpact_{i,t}$ captures the extent that firm i tweets about its earnings announcement in the month prior to the the quarterly earnings announcement, and is measured as number of earnings-news tweets posted in the window [-30, -1] around the announcement date.

The controls variables, $X_{i,t}$, include $Size_{i,t}$, $B/M_{i,t}$, $AnalystsFollowing_{i,t}$, *Non-earningsTweetImpact*$_{i,t}$, and $InstitutionalOwnership$. To mitigate the influence of outliers, all continuous variables are winsorized at the 1st and 99th percentiles. Table A2 reports summary statistics on the variables used to estimate equation (4). All variables are defined in detail in Appendix Table A1.

Equation (4) is estimated using standard errors clustered by firm, firm fixed effects, and year-quarter fixed effects.

Table 7 presents the results. The coefficients for *EarningsTweetImpact* and *Non-EarningsTweetImpact* are insignificant or weakly significant at 10%. This result suggest that the information contained in tweets before the earnings announcement has no incremental information about firms' future cash flows. In an untabulated result I also control for the sentiment of tweets and the results remain unaffected.

Table 7: Tweeting and Unexpected Earnings

This table shows the relationship between standardized unexpected earnings (*SUE*) and firms' tweeting behaviors. In column (1) *SUE* is measured as unexpected earnings standardized by the standard deviation of analysts' consensus forecasts, in column (2) as as unexpected earnings standardized by price per share of stock at the end of the quarter, and in column (3) as unexpected earnings standardized by the book value of equity per share at the end of the previous quarter. Where unexpected earnings is measured as is the I/B/E/S reported quarterly earnings per share (EPS) less the latest I/B/E/S consensus analyst quarterly EPS forecast just prior to the earnings announcement. *EarningsTweetImpact* captures firms' tweeting behavior in the month prior to the the quarterly earnings announcement, and is measured as number of earnings-news tweets posted in the window [-30, -1] around the announcement date. To mitigate the influence of outliers, all continuous variables are winsorized at the 1st and 99th percentiles. See Appendix A for variable definitions. Standard errors are clustered at the firm level and provided in parentheses beneath the coefficient estimates. *, **, and *** indicates significance at the 10%, 5%, and 1%, respectively.

	SUE		
	(1)	(2)	(3)
Earnings Tweet Impact	0.075*	-0.000	-0.000
	(0.042)	(0.000)	(0.000)
Non-earnings Tweet Impact	-0.011	-0.000	-0.000
	(0.038)	(0.000)	(0.000)
Size	0.129	-0.001	-0.001*
	(0.227)	(0.001)	(0.001)
BM	-0.428	0.002	-0.002***
	(0.282)	(0.002)	(0.001)
Institutional Own.	-1.797**	-0.004	-0.006*
	(0.710)	(0.003)	(0.003)
Analysts	-0.491	-0.001	-0.001
	(0.302)	(0.001)	(0.001)
Firm-year FE	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes
No. of firms	1058	1058	1018
Adjusted R^2	0.150	0.197	0.170
Observations	13330	13330	11149

6. Robustness

The relationship between tweeting on announcement returns is robust to sample selection, different tweeting measures, pooled OLS estimation, and additional fixed effects.

One concern is that some of the Twitter accounts I manually collected are erroneous or fake accounts. To help control for this potential problem I limit the sample to those firms with verified Twitter accounts. Of the 1,215 accounts in my sample, 489 are verified. The verified feature on Twitter is a signal to the public that an account of public interest is authentic. Table 8 shows similar results to those in section 4.A.

In table 1, I show that the results are robust to the measurement of *EarningsTweetImpact*. I measure *EarningsTweetImpact* in three ways: in column (1) I use the number of tweets about earnings as a naive proxy for the impact of the earnings-news tweets in the network of Twitter users, in column (2) I use the number of tweets about earnings multiplied by the log number of Twitter followers the firm has to capture the connectivity of the firm, in column (3) I use the number of tweets about earnings multiplied by the log number of retweets to capture the popularity of the tweets.

In table 2, I show that the results are robust to a pooled OLS estimation and rich set of fixed effects. One concern is that the choice to tweet about earnings news is endogenous; fixed effects help to control for unobservable determinants of tweeting: year-quarter fixed effects for macro factors, firm-year (firm) fixed effects for time-varying (time-invariant) firm characteristics, day-of-week fixed effect for the variation of announcements on different days. The results are generally consistent across specifications, however the within-group estimates tend to be higher and more significant than the pooled OLS estimates.

The relationship between tweeting on unexpected earnings is robust to sample selection, different *SUE* measures, pooled OLS estimation, additional fixed effects, and additional control variables. In Table 9, I limit the sample to those firms with verified Twitter accounts and re-estimate equation 4. The results are similar to those in section 4.A..

In table 7, I show that the results are robust to the measurement of *SUE*. In column (1) *SUE* is measured as unexpected earnings standardized by the standard deviation of analysts' consensus forecasts, in column (2) as as unexpected earnings standardized by price per share of stock at the end of the quarter, and in column (3) as unexpected earnings standardized by the book value of equity per share at the end of the previous quarter. Where unexpected earnings is measured as is the I/B/E/S reported quarterly earnings per share (EPS) less the latest I/B/E/S consensus analyst quarterly EPS forecast just prior to the earnings announcement.

In an untabulated results I re-estimate equation 4 using a pooled OLS model and a rich set of fixed effects and the results remain unaffected. Finally, I control for the sentiment of tweets and the results remain unaffected.

The relationship between the consistency of tweeting and momentum in returns is robust to different momentum proxies. In table 6 I calculate *Momentumn* using both excess returns relative to 90 day T-bills and Fama-French excess returns. In table 10, I measure momentum in three alternative ways. Following Hong et al. (2000) I use the serial correlation coefficient of six-month excess returns (relative

Table 8: Verified Accounts: Tweeting and Announcement Returns

This table shows the relationship between cumulative abnormal returns (*CAR*) and firms' tweeting behaviors for the subsample of firms with verified Twitter accounts. In column (1) *EarningsTweet-Impact* is measured as *EarningsTweetCount*, in column (2) as *EarningsTweetCount*Cardinality*, and in column (3) as *EarningsTweetCount*Retweets*. To mitigate the influence of outliers, all continuous variables are winsorized at the 1st and 99th percentiles. See Appendix A for variable definitions. Standard errors, clustered at the firm level, are provided in parentheses beneath the coefficient estimates. *, **, and *** indicates significance at the 10%, 5%, and 1%, respectively.

	CAR		
	(1)	(2)	(3)
Negative Surprise	-5.442***	-5.465***	-5.303***
	(0.347)	(0.345)	(0.329)
Earnings Tweet Impact	-0.034	-0.005	-0.126
	(0.416)	(0.037)	(0.221)
*Neg Surp*Earnings Tweet Impact*	1.418***	0.136***	1.185***
	(0.420)	(0.037)	(0.328)
Non-earnings Tweet Impact	0.053	0.009	0.064
	(0.196)	(0.018)	(0.070)
Residual ESV (Ryans)	-0.649	-0.646	-0.659
	(0.448)	(0.449)	(0.449)
Absolute SUE	0.350***	0.350***	0.350***
	(0.050)	(0.050)	(0.050)
Size	-4.379**	-4.382**	-4.465***
	(1.703)	(1.703)	(1.702)
Loss	-2.615***	-2.613***	-2.625***
	(0.591)	(0.590)	(0.591)
BM	20.022***	20.021***	20.005***
	(3.510)	(3.513)	(3.525)
Q4	0.454	0.454	0.449
	(0.321)	(0.320)	(0.321)
Firm FE	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes
F-test statistic: $\beta 2 + \beta 3 = 0$	9.089	10.447	7.834
F-test p-value	0.003	0.001	0.005
No. of firms	449	449	449
Adjusted R^2	0.146	0.146	0.146
Observations	6,069	6,069	6,069

Table 9: Verified Accounts: Tweeting and Unexpected Earnings

This table shows the relationship between standardized unexpected earnings (*SUE*) and firms' tweeting behaviors for the subsample of firms with verified Twitter accounts. In column (1) *SUE* is measured as unexpected earnings standardized by the standard deviation of analysts' consensus forecasts, in column (2) as as unexpected earnings standardized by price per share of stock at the end of the quarter, and in column (3) as unexpected earnings standardized by the book value of equity per share at the end of the previous quarter. Where unexpected earnings is measured as is the I/B/E/S reported quarterly earnings per share (EPS) less the latest I/B/E/S consensus analyst quarterly EPS forecast just prior to the earnings announcement. *EarningsTweetImpact* captures firms' tweeting behavior in the month prior to the the quarterly earnings announcement, and is measured as number of earnings-news tweets posted in the window [-30, -1] around the announcement date. To mitigate the influence of outliers, all continuous variables are winsorized at the 1st and 99th percentiles. See Appendix A for variable definitions. Standard errors, clustered at the firm level, are provided in parentheses beneath the coefficient estimates. *, **, and *** indicates significance at the 10%, 5%, and 1%, respectively.

	SUE		
	(1)	(2)	(3)
Earnings Tweet Impact	0.080	-0.000	-0.000
	(0.050)	(0.000)	(0.000)
Non-earnings Tweet Impact	-0.026	0.000	-0.000
	(0.052)	(0.000)	(0.000)
Size	0.077	-0.002	-0.002*
	(0.350)	(0.002)	(0.001)
BM	-0.420	0.004	-0.003**
	(0.466)	(0.004)	(0.001)
Institutional Own.	-1.754	-0.010	-0.009*
	(1.178)	(0.006)	(0.005)
Analysts	-0.842	-0.002	-0.003
	(0.533)	(0.002)	(0.002)
Firm-year FE	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes
No. of firms	439	439	435
Adjusted R^2	0.141	0.309	0.182
Observations	6,121	6,121	5,403

to 90 day T-bills). Also I calculate *Momentumn* using cumulative 3-month excess returns rather than monthly returns. In columns (1) and (2) *MOM* is measured as the correlation between the series $ExRet_{i,[t,t+2]}$ and the lagged series $ExRet_{i,[t-12,t-2]}$, where $ExRet_{i,[t,t+2]}$ is the cumulative 3-month excess return of firm i. In column (1) *MOM* is calculated using excess returns relative to 90 day T-bills and in column (2) using Fama-French three factor excess returns.

7. Conclusion

Social media has changed the way firms communicate with investors by giving them a direct, instantaneous, and network-enhanced communication channel. Furthermore, social media are gaining an outreach as relevant as that of traditional information intermediaries, like business press and financial analysts.

Using a comprehensive dataset of over 7 million tweets posted by S&P 1500 firms between January 2014 and December 2017, I explore the following research questions: Is there a link between corporate information dissemination on social media and valuations? Can firm-initiated tweets increase the speed of information diffusion?

When quantifying firms' tweeting behaviors I use a textual classification based on a dictionary of keywords to detect earnings-news tweets. To measure the relative network impact of firms' tweets I account for how well firms are connected on Twitter and the popularity of their tweets. As Twitter is a social network, it is principal to capture the relative importance of tweets in the network. Through the process of retweeting and liking a tweet can spread through the network of Twitter users and

Table 10: Tweeting and The Speed of Information Diffusion

This table shows the cross-sectional relationship between short-term continuation in returns and the consistency of tweeting about earnings news. The dependent variable is measured in three ways. *MOM* is measured as the correlation between the series $ExRet_{i,[t,t+2]}$ and the lagged series $ExRet_{i,[t-12,t-2]}$, where $ExRet_{i,[t,t+2]}$ is the cumulative 3-month excess return of firm i. In column (1) *MOM* is calculated using excess returns relative to 90 day T-bills and in column (2) using Fama-French three factor excess returns. In column (3) *SCC* is the serial correlation of six-month excess returns (relative to 90 day T-bills). The sample is restricted to firms with verified Twitter accounts. To mitigate the influence of outliers, all continuous variables are winsorized at the 1st and 99th percentiles. See Appendix A for variable definitions. The regression is estimated using OLS with robust standard errors. Standard errors are provided in parentheses beneath the coefficient estimates. *, **, and *** indicates significance at the 10%, 5%, and 1%, respectively.

	MOM		*SCC*
	(1)	(2)	(3)
Earnings Tweet Quarters	-0.048*	-0.057**	-0.017*
	(0.025)	(0.028)	(0.009)
BM	0.024	0.039	0.023***
	(0.028)	(0.044)	(0.005)
Analysts	-0.000	-0.000	0.001*
	(0.001)	(0.002)	(0.000)
Size	0.016**	0.001	-0.003
	(0.007)	(0.010)	(0.002)
Institutional Ownership	0.003	-0.067	-0.018
	(0.084)	(0.100)	(0.024)
Twitter Followers	0.000*	-0.000	0.000
	(0.000)	(0.000)	(0.000)
Constant	-0.284***	-0.116	0.799***
	(0.099)	(0.121)	(0.027)
R^2	0.031	0.022	0.036
Observations	442	352	476

potentially reach a vast number of users.

I start my empirical analysis by examining the association between abnormal stock returns around earnings announcements and firms' tweeting behaviors. Firms with negative earnings surprises, i.e., firms that miss their analyst consensus forecasts, have higher announcement returns if they tweet about their earnings announcement. The effect is concentrated among firms with low institutional ownership and among firms with a large social media network. One way to interpret this finding is that firms are able to successfully manage the expectations of investors, especially if investors are unsophisticated and if the firm has a large social network.

To better understand what kinds of information are in firm-initiated tweets I investigate the ability of tweets to predict quarterly earnings surprises. I find that tweets in the month prior to earnings announcements contain little incremental information to predict firms' cash flows beyond the consensus. The results suggests that, unsurprisingly, firms do not appear to be leaking information to investors before their quarterly announcements; rather, firms are using Twitter to help manage investor sentiment during announcement periods.

In addition to studying how tweeting affects investor sentiment, I investigate whether tweeting can increase the speed of information diffusion to the market. In particular, I examine the association between momentum in returns and the consistency of tweeting about earnings news. The results shows firms that tweet about earnings news consistently have less short-term continuation in returns. This result is consistent with Hong and Stein (1999), which predict that short-run continuation in returns should be more pronounced in stocks where information diffuses more slowly.

This result suggests that tweets can increase the speed of information diffusion to investors.

This study is important to regulators, investors, and firms. The SEC has embraced social media in an attempt to promote full and fair disclosure. The results suggest firms may be able to successfully manage the expectations of investors and increase the speed of information diffusion by utilizing social media, however this is far from a complete explanation of the phenomenon.

References

Amir, E. and Livnat, J. (2005). The economic consequences of (not) issuing preliminary earnings announcement, *Technical report*, Working paper, New York University.

Baker, M. and Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns, *The journal of Finance* **61**(4): 1645–1680.

Barber, B. M., Odean, T. and Zhu, N. (2008). Do retail trades move markets?, *The Review of Financial Studies* **22**(1): 151–186.

Barberis, N., Shleifer, A. and Vishny, R. (1998). A model of investor sentiment, *Journal of financial economics* **49**(3): 307–343.

Bartov, E., Faurel, L. and Mohanram, P. (2017). Can twitter help predict firm-level earnings and stock returns?, *The Accounting Review* .

Bhagwat, V. and Burch, T. (2016). Pump it up? tweeting to manage investor attention to earnings news.

Blankespoor, E., Miller, G. S. and White, H. D. (2013). The role of dissemination in market liquidity: Evidence from firms' use of twitter$^{\text{TM}}$, *The Accounting Review* **89**(1): 79–112.

Carhart, M. M. (1997). On persistence in mutual fund performance, *The Journal of finance* **52**(1): 57–82.

Cohen, L. and Frazzini, A. (2008). Economic links and predictable returns, *The Journal of Finance* **63**(4): 1977–2011.

De Long, J. B., Shleifer, A., Summers, L. H. and Waldmann, R. J. (1990). Noise trader risk in financial markets, *Journal of political Economy* **98**(4): 703–738.

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

Dyck, A. and Zingales, L. (2003). The media and asset prices, *Technical report*, Working Paper, Harvard Business School.

Fang, L. and Peress, J. (2009). Media coverage and the cross-section of stock returns, *The Journal of Finance* **64**(5): 2023–2052.

Fedyk, A. (2018). News-driven trading: who reads the news and when, *Technical report*, Discussion paper, Working paper, Berkeley.

Hirshleifer, D. and Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting, *Journal of accounting and economics* **36**(1-3): 337–386.

Hong, H., Lim, T. and Stein, J. C. (2000). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *The Journal of Finance* **55**(1): 265–295.

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

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency, *The Journal of finance* **48**(1): 65–91.

Jung, M. J., Naughton, J. P., Tahoun, A. and Wang, C. (2017). Do firms strategically disseminate? evidence from corporate use of social media, *The Accounting Review* .

Kumar, A. and Lee, C. M. (2006). Retail investor sentiment and return comovements, *The Journal of Finance* **61**(5): 2451–2486.

Lee, C. M., Shleifer, A. and Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle, *The Journal of Finance* **46**(1): 75–109.

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

SEC (2013a). Report of investigation pursuant to section 21(a) of the securities exchange act of 1934: Netflix, inc., and reed hastings, *Release No. 69279* pp. April 2, 2013.

SEC (2013b). Sec says social media ok for company announcements if investors are alerted, *Press Release* pp. April 2, 2013.

SEC (2018). United States Securities and Exchange Commision vs. Elon Musk. Uninted States District Court Sothern District of New York. Civil Action No. 1:18-cv-8865.
URL: *https://www.sec.gov/litigation/complaints/2018/comp-pr2018-219.pdf*

Stein, K. (2018). Improving information for investors in the digital age. Remarks by Commisioner Karen M. Stein of the SEC at the the Council of Institutional Investors 2018 Fall Conference,New York, NY.
URL: *https://www.sec.gov/news/speech/speech-stein-102318*

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market, *The Journal of finance* **62**(3): 1139–1168.

Veldkamp, L. L. (2006). Media frenzies in markets for financial information, *American Economic Review* **96**(3): 577–601.

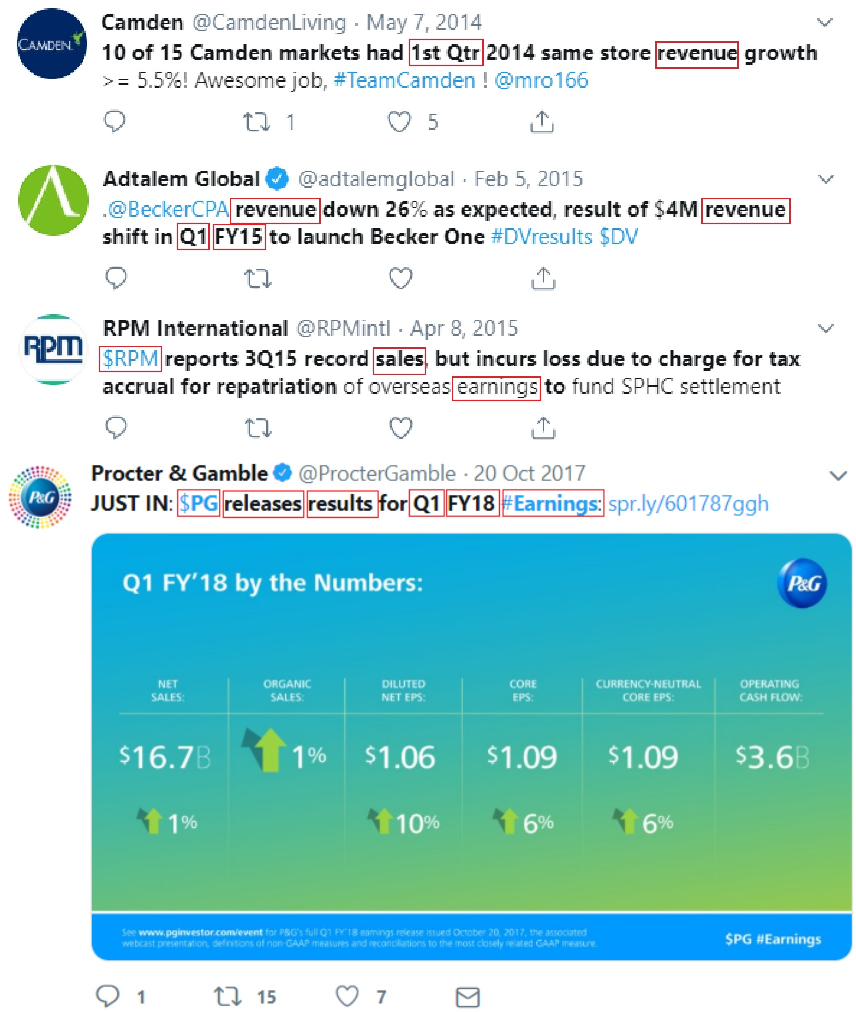
Appendix

A. Earnings announcement keywords

earnings, eps, profit, profits, income, revenue, revenues, sales, net sales, gaap, 1q, 2q, 3q, 4q, q1, q2, q3, q4, qtr1, qrt2, qrt3, qrt4, first quarter, second quarter, third quarter, fourth quarter, 1st quarter, 2nd quarter, 3rd quarter, 4th quarter, quarter, financial results, results, announce, announces, declares, declare, releases, press release, earnings release, fiscal, conference call, earnings call, full year, qtr, year-over-year, year over year, yoy, qoq, cash flow, financial position, dividends, dividend, continuing operations, growth, fy13, fy14, fy15, fy16, fy17, fy18, fy2013, fy2014, fy2015, fy2016, fy2017, fy2018

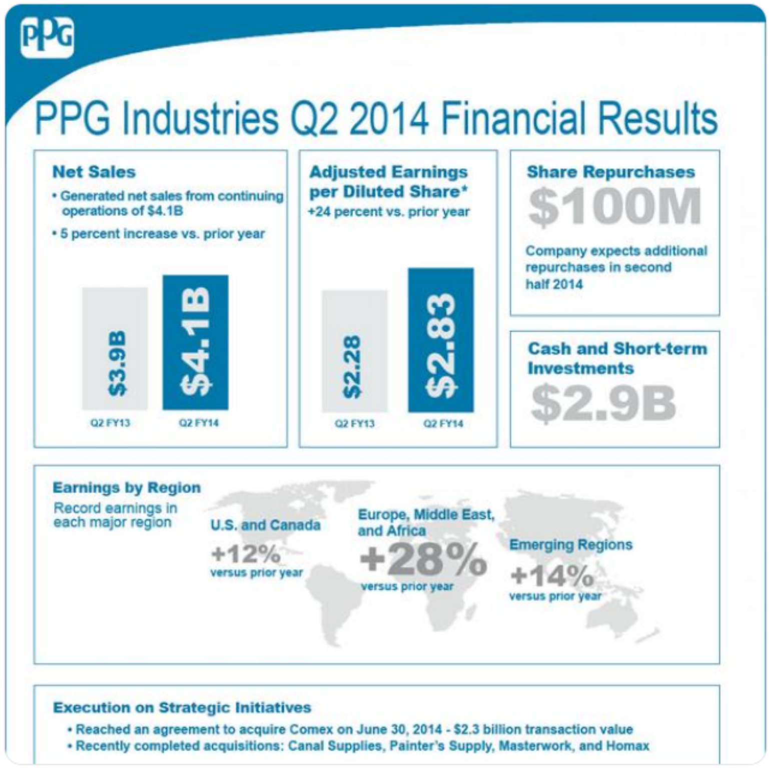
B. Examples of earnings-news tweets identified with text classifier

Earnings announcement keywords are boxed in red.



PPG @PPG · 17 Jul 2014

PPG reports record 2Q net sales of $4.1B, record 2Q #earnings: bit.ly/1jBaC5g $PPG

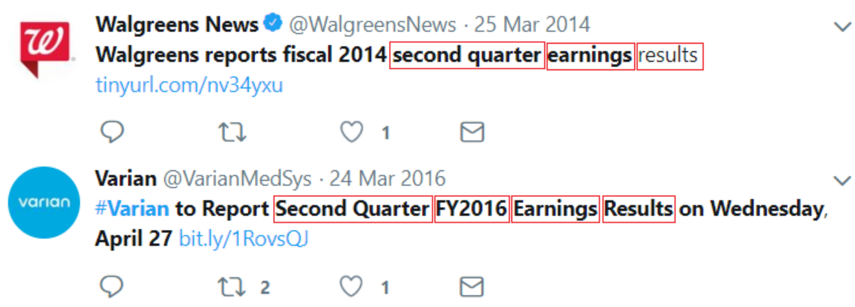


2 3

Oracle @Oracle · 14 Sep 2017

Results from today's Q1 FY18 earnings: Total #cloud revenues were up 51%.

5 61 86


Walgreens News @WalgreensNews · 25 Mar 2014
Walgreens reports fiscal 2014 second quarter earnings results
tinyurl.com/nv34yxu
1
Varian @VarianMedSys · 24 Mar 2016
#Varian to Report Second Quarter FY2016 Earnings Results on Wednesday, April 27 bit.ly/1RovsQJ
2
1

C. Variables

Table A1: Variable Descriptions and Data Sources

Variable	Description	Data Source
Twitter Variables		
Earnings Tweet Impact	Measured in one of three ways depending on model: (1) *Earnings Tweet Count*, (2) *Earnings Tweet Count*Followers*, (3) *Earnings Tweet Count*Retweets*	Twitter
Earnings Tweet Count	Number of earnings related tweets during the windows [-10, 1] or [-30, -1] around the quarterly earnings announcement date	Twitter
*Earnings Tweet Count*Followers*	*Earnings Tweet Count*Log(1 + Followers)*. Where *Followers* is the number of followers a firm has as of June 2019.	Twitter
*Earnings Tweet Count*Retweets*	*Earnings Tweet Count*Log(1 + Avg Retweets)*, where Avg Retweets is the average number of retweets for the earnings related tweets during the windows [-10, 1] or [-30, -1] around the quarterly earnings announcement date	Twitter
Non-earnings Tweet Impact	Measured in one of three ways depending on model: (1) *Non-earnings Tweet Count*, (2)*Non-earnings Tweet Count*Followers*, (3) *Non-earnings Tweet Count*Retweets*	Twitter
Non-earnings Tweet Count	Log of 1 + the total number of tweets minus the number of earnings related tweets during the windows [-10, 1] or [-30, -1] around the quarterly earnings announcement date	Twitter
*Non-earnings Tweet Count*Followers*	*Non-earnings Tweet Count*Log(1 + Followers)*. Where *Followers* is the number of followers a firm has as of June 2019.	Twitter

*Non-earnings Tweet Count*Retweets*	*Non-earnings Tweet Count*Log(1 + Avg Retweets)*, where *Avg Retweets* is the average number of retweets for the non-earnings related tweets during the windows [-10, 1] or [-30, -1] around the quarterly earnings announcement date	Twitter
EarningsvTweetvQuarters	Proportion of quarters a firm tweets about earnings related news over the sample period, January 2014 to December 2017	Twitter
*EarningsvTweet Quarters*Followers*	*Earnings Tweet Quarters*Log(1 + Number of Followers)*	Twitter
Twitter Verified	Indicator variable equal to one is a firm's Twitter account is verified by Twitter. When an account is verified by Twitter a blue checkmark appears next to the account name to signal the authenticity of that account.	Twitter
Earnings Announcement Variables		
CAR[-1, 1]	Carhart's cumulative abnormal return in the three day window [-1, 1] around the earnings announcement date	CRSP
Positive Surprise	Indicator variable equal to one if the firm's $SUE >= 0$, and equal to zero otherwise.	IBES
SUE	The firm's actual EPS minus the consensus analyst forecast EPS, standardized by the standard deviation of analysts' consensus forecasts, by price per share of stock at the end of the quarter, or by the book value of equity per share at the end of the previous quarter. Consensus analyst forecast is measured as the median latest analyst forecast in the 90 days prior to the earnings announcement.	IBES
Firm Variables		
Size	Log of total assets (Compustat atq).	Compustat

BM	Book to market value (Compustat ceqq/mkvaltq).	Compustat
Loss	Indicator variable set to 1 if the firm reports a quarterly loss (Compustat $niq < 0$).	Compustat
Analyst	Natural log of one plus the average number of analysts following a given firm during the 90 days prior to the earnings announcement.	IBES
Q4	Indicator variable equal to one if the quarterly earnings announcement is in the fourth fiscal quarter of the year	Compustat
Institutional Ownership %	Total institutional ownership as a percentage of shares outstanding.	Thomson Reuters 13-f
Momentum Variables		
Momentum	Correlation between the series $Ret_{[t,t]}$ and the lagged series $Ret_{[t-12,t-2]}$, where $Ret_{[t,t]}$ is the monthly return of a firm	CRSP

D. Summary Statistics

Table A2: Summary Statistics

This table reports summary statistics for firm-quarter observations used to estimate (1) and (4), and firm-month observations to estimate (3). The sample period is from Q1 2014 to Q4 2017. See Appendix table A1 for variable definitions.

	Mean	SD	P05	Med	P95
Panel A: Variables used to estimate equation (1)					
CAR	0.16	7.51	-11.58	0.19	11.85
Positive Surprise	0.72	0.45	0.00	1.00	1.00
Earnings Tweet Count [-10, 1]	0.76	1.93	0.00	0.00	4.00
Non-earnings Tweet Count [-10, 1]	2.37	1.44	0.00	2.56	4.49
*Earnings Tweet Count*Followers [-10, 1]*	7.61	20.97	0.00	0.00	40.31
*Non-earnings Count*Followers [-10, 1]*	24.05	17.44	0.00	22.89	53.78
*Earnings Tweet Count*Retweets [-10, 1]*	0.74	2.62	0.00	0.00	3.85
*Non-earnings Count*Retweets [-10, 1]*	3.16	3.90	0.00	1.88	11.21
$\mid SUE \mid$	2.58	3.06	0.00	1.60	9.03
Size	8.60	1.76	5.90	8.49	11.74
Loss	0.12	0.33	0.00	0.00	1.00
BM	0.48	0.35	0.08	0.39	1.14
Analysts	2.60	0.62	1.61	2.65	3.50
Q4	0.20	0.40	0.00	0.00	1.00
Panel B: Variables used to estimate equation (4)					
SUE (forecast sd)	1.47	3.61	-3.23	0.97	7.78
SUE (price)	0.00	0.01	-0.00	0.00	0.01
SUE (book equity)	0.00	0.01	-0.01	0.00	0.02
Earnings Tweet Count [-30, -1]	0.31	1.08	0.00	0.00	2.00
Non-Earnings Tweet Count [-30, -1]	3.15	1.60	0.00	3.40	5.35
Institutional Own.	0.85	0.14	0.60	0.86	1.05
Panel C: Variables used to estimate equation (3)					
Momentum	-10.07	12.75	-30.38	-10.84	11.21
Earnings Tweet Quarters	0.34	0.38	0.00	0.18	1.00
*Earnings Tweet Quarters*Followers*	3.14	3.68	0.00	1.48	10.63
Verified	0.42	0.49	0.00	0.00	1.00

Table A3: Distribution of Tweets by Calendar Quarter

This table presents the frequency distributions of tweets and observations by calendar quarter. My sample encompasses 7,132,461 tweets posted by S&P1500 firms with active Twitter accounts as of October 2017. The sample represents represents 16,844 firm-quarters. The number of firm quarter observations increases over the sample. This pattern is to be expected because some Twitter users in the sample were not active at the start of the sample period.

Calendar Quarter	Tweets		Observations	
	N	%	N	%
2014Q1	422,377	5.4%	950	5.0%
2014Q2	428,965	5.5%	968	5.1%
2014Q3	460,574	5.9%	984	5.2%
2014Q4	501,935	6.5%	996	5.2%
2015Q1	463,611	6.0%	1,023	5.4%
2015Q2	481,138	6.2%	1,032	5.4%
2015Q3	489,600	6.3%	1,060	5.5%
2015Q4	570,359	7.3%	1,070	5.6%
2016Q1	433,123	5.6%	1,058	5.5%
2016Q2	435,102	5.6%	1,071	5.6%
2016Q3	420,817	5.4%	1,070	5.6%
2016Q4	454,652	5.9%	1,089	5.7%
2017Q1	428,492	5.5%	1,103	5.8%
2017Q2	378,364	4.9%	1,110	5.8%
2017Q3	384,522	5.0%	1,128	5.9%
2017Q4	378,830	4.9%	1,132	5.9%
All	7,132,461	100.0%	16,844	100.0%

Table A4: Tweet Characteristics

This table presents descriptive statistics related to Twitter users (firms), firm-quarters, and individual tweets. Firms have a mean (median) or 162,642 (6,352) followers. The average date firms joined Twitter was in November 2010. Firm-quarters have a mean (median) of 178 (81) tweets and 1.24 (1) tweets about earnings news. Tweets have a mean (median) of 79 (86) characters, 8 (0) retweets, and 16 (0) likes.

Variable	Mean	Std. Dev	P01	Q1	Median	Q3	P99
Per Twitter User (N = 1,215)							
Number of Followers	162,642	1,383,262	73	1,473	6,352	29,200	2,300,412
Number of Friends	2,438	9,814	0	194	557	1,535	35,626
Date Joined Twitter	Nov2010	-	Jun2007	Apr2009	Jan2007	Jan2012	May2017
Per Firm Quarter (N = 13,350)							
Tweet Count	178.00	430.00	0.00	22.00	81.00	209.00	1438.00
Quarter with Tweets	92%	0.26	0.00	1.00	1.00	1.00	1.00
Earnings Tweet Count	1.24	3.48	0.00	0.00	0.00	1.00	13
Quarter with Earnings Tweets	35%	0.48	0.00	0.00	0.00	1.00	1.00
Per Tweet (N = 7,132,461)							
Number of Characters	79.00	50.00	16.00	18.00	86.00	119.00	217.00
Number of Retweets	8.00	276.00	0.00	0.00	0.00	1.00	149.00
Number of Likes	16.00	753.00	0.00	0.00	0.00	2.00	264.00

Table A5: Distribution of Firms and Observations by Industry

This table presents sample distributions of firm-quarters by SIC code groupings. The sample of firms spans nine SIC code divisions but displays some evidence of industry clustering, e.g. 38% of the sample are manufacturing firms.

		Observations		Firms		Compustat
Sic Codes	Divisions	N	%	N	%	%
1000-1499	Mining	296	0.02	26	0.03	0.11
1500-1799	Construction	218	0.02	16	0.02	0.01
2000-3999	Manufacturing	5,027	0.38	424	0.39	0.24
4000-4999	Trans., Comm., Electric, Gas and Sanitary	1,420	0.11	102	0.10	0.06
5000-5199	Wholesale Trade	439	0.03	41	0.04	0.02
5200-5999	Retail Trade	1,384	0.10	98	0.09	0.03
6000-6799	Finance, Insurance and Real Estate	2,637	0.20	199	0.16	0.40
7000-8999	Services	1,897	0.14	168	0.16	0.11
9900-9999	Nonclassifiable	32	0.00	2	0.00	0.02
	All Industries	13,350	1.00	1,076	1.00	1.00