Is Firms' Social Media Engagement Informative about Firm Performance?

Atul Singh

September 4, 2019

Is Firms' Social Media Engagement Informative about Firm Performance?

Abstract

In this paper, I examine whether the volume of a firm's tweets and its followers' engagement is informative to capital market participants and financial intermediaries, namely investors and analysts. My data comprises of 178,236 firm-quarters (46,449 Tweet firm-quarters) and approximately 17.50 million firm-initiated tweets collected from the Primary Twitter sites of 2,229 public US firms between 2006 and 2017. I find that the volume of a firm's tweets and the followers' engagement during a quarter predicts the firm value during that period. The results also suggest that changes in tweet (engagement) volume are informative to investors and the information gets impounded in the stock prices concurrently. I also find evidence that followers' engagement is more informative than the firm's tweet volume for predicting firm-value. My findings further indicate that analysts may be using this additional information in the firm's tweet (engagement) volume to make more accurate earnings and sales forecast, which reduces the Tweet firm's unexpected earnings and unexpected sales growth. In additional analysis, I find that the level of tweets (engagement) helps predict a firm's earnings and sales whereas changes in tweet (engagement) volume incrementally explain the firm's sales growth and this may be the source of additional information to investors and analysts.

1. Introduction

In recent times, an increasing number of firms have started using social media to disseminate information and communicate directly with their stakeholders. Jung et al. (2018) show that almost 50% of S&P 1500 firms have Twitter or Facebook accounts. Twitter, in particular, has emerged as the most popular social media platform[1] for dissemination of information by firms. There has also been an increasing interest in accounting literature about why and how firms use Twitter. Most papers focus on one category of tweets - earnings announcements. However, earnings-related or financial news tweets constitute only a small proportion of the total volume of a firm's tweets. Indeed, firms may use Twitter to disseminate information as well as to engage directly with all their stakeholders – customers, investors, suppliers, employees, etc. Twitter provides a unique platform for dissemination because it facilitates real-time two-way communication and feedback between a firm and its followers while limiting the size of the message.

In this paper, I study the aggregate information in firms' tweets and followers' engagement on their official Twitter accounts. Specifically, I examine whether the volume of a firm's tweets and its followers' engagement is informative to capital market participants and financial intermediaries, namely investors and analysts. Additionally, I test whether the tweet (engagement) behavior of a firm (followers) predicts the firm's quarterly earnings, sales and sales growth. I use a broad hand-collected sample of approximately 17.50 million firm-initiated[2] tweets, and 108 million retweets, 164 million likes and 14.9

[1] Jung et al. (2018) report that 47% and 42% of the S&P 1500 firms use Twitter and Facebook, respectively as of January, 2013.

[2] I collect tweets from the primary official Twitter account of the firm, the link for which appears on the home webpage of the firm. These tweets are initiated by firms themselves and, hence, I refer to them as firm-initiated tweets.

million replies by followers, from the Primary[3] Twitter sites of 2,229 public US firms from 2006 to 2017. The full sample has 178,236 firm-quarters (46,559 Tweet firm-quarters)[4] and 6,268 public US firms.

Social media in general and Twitter, in particular, facilitates real-time and publicly viewable communication and engagement of the firm with its followers who can share their opinions with the firm through direct messages and feedback. Firms tweet about myriad topics such as sales and promotion, customer fulfillment, financial disclosures, corporate disclosures, new product launches, CSR initiatives, etc. These tweets generate a varying amount of interest from the followers of the firm's Twitter account; a follower may respond by liking, retweeting, or replying to a particular tweet – collectively defined as engagement – or may choose to ignore the tweet. The level of engagement with a firm's tweet is the collective positive response or feedback of the followers to that tweet and represents the overall enthusiasm of the followers for the firm and its products and offerings.

I hypothesize that the information in the volume of a firm's tweet (engagement) volume gets reflected in the firm value or stock prices concurrently. Efficient Markets Hypothesis (EMH) states that all available public information is impounded immediately in the market prices fully (Fama 1970). Therefore, the aggregate information in tweets and followers' engagement should get reflected in the stock prices as these are in the public domain. This means that the tweet (engagement) volume should be value relevant and contribute positively to the firm value. However, researchers in accounting and finance have also shown that there are notable exceptions to EMH such as PEAD (Ball and Brown, 1968; Bernard and Thomas, 1990), accrual anomaly (Sloan, 1996; Xie, 2001), etc. Some of the possible explanations for this could be limited investor attention due to costly processing and information complexity (Bloomfield, 2002; Hirshleifer et al., 2002; Hirshleifer et al., 2003) or underreaction due to slow diffusion of information

---

[3] I define Primary Twitter account as the main official Twitter account which appears on the webpage of a firm. In addition, the firm may have other Twitter accounts as well. In this study, I consider tweets of Primary Twitter accounts only and, henceforth, refer to them as Primary Twitter accounts or just as Twitter accounts.

[4] I define a Tweet firm-quarter as a quarter in which the firm has a primary Twitter account and tweets.

(Hong and Stein, 1999).[5] Therefore, ex-ante it is not clear whether the aggregate information in the firm's tweets (engagement) will be incorporated in the firm value during the period, beyond the concurrent information already contained in other known sources of information such as press, analyst forecasts, and voluntary disclosures.

I also predict that the volume of tweets by a firm and the followers' engagement or a change in them provides incremental information to analysts enabling them to make more accurate sales and earnings forecasts. Analysts process and interpret news and information about a firm and the economy from all public, as well as private, sources and come up with their estimates of the firm's expected earnings and sales for a period (Lang et al., 1996; Rogers et al., 1997; Bowen, Davis et al., 2002). The consensus estimate of all the analysts following a firm is often considered as the market's expectation of a firm's earnings and sales during a period (O'Brien 1988; Zhang 2008; Lobo et al., 2017). An analyst may also be paying attention to firms' dissemination of information on Twitter and other social media accounts as part of this information-gathering process. Therefore, the volume of a firm's tweets and its followers' engagement, or a change in them, may provide incremental information to the analysts about the likely business performance of the firm during the period. This implies that conditional on a firm using Twitter, the firm's unexpected earnings and unexpected sales growth should have a negative relationship with the level of (change in) tweets and engagement.

The issue of whether the firm's (followers') tweets (engagement) are informative at the aggregate level is of particular interest to investors and analysts. Extant literature has examined specific categories of tweets using event studies (Blankespoor et al., 2014; Lee et al., 2015). However, whether firms' tweets

[5] In addition, Hales (2007) shows that investor preferences can also significantly influence the manner in which information is processed and affect their expectations of future earnings performance. This could lead them to make investment decisions which may not be in their best interest. However, I focus only on firm-initiated tweets and, hence, this is not relevant to my study.

convey any new information to the market participants and financial intermediaries at the aggregate level remains an empirically unexplored question. I use five different measures to proxy for the aggregate information contained in firm-initiated tweets and followers' engagement[6]. I apply Fama-MacBeth (1973) monthly cross-sectional regressions to test whether change in tweet (engagement) volume is priced concurrently by the capital market. I use OLS multiple linear regression with year-quarter and firm fixed effects for the other empirical analysis and control for the other public sources of information about the firm – press releases, newspaper articles, and analyst coverage.

I find that the volume of tweets by a firm and engagement of its followers during a quarter incrementally predicts the firm value during that period and this association is strongly positive. One standard deviation increase in *LOG (TWEETS) (LOG (ENGAGEMENT))* is associated with an increase of approximately 3.25% (6.13%) increase in the firm-value measured as *TOBIN'S Q*. The results also suggest that changes in tweet (engagement) volume are informative to the capital market participants and get impounded in the stock prices concurrently. Therefore, changes in tweets (engagement) volume help explain the observed cross-sectional differences in returns beyond the priced common risk factors. My study also suggests that the followers' engagement volume may be more informative to the stock market than the firm's tweet volume. The engagement level, which is a positive response by the followers to the firm's tweets, represents only a lower bound on the potential reach and viewership of a firm's tweets and can, therefore, be considered as a measure of the efficacy of a firm's dissemination effort or as representing the overall enthusiasm of the followers about the firm and its products and services.

My findings further indicate that analysts may be using this additional information in tweet (engagement) volume to make more accurate earnings and sales forecast, thereby reducing the firm's unexpected earnings and unexpected sales growth. It also seems that the analysts are especially able to

[6] See Section 3.2 for a detailed description of what each of these five measures represents.

make better forecasts using the aggregate information in tweet (engagement) volume when there is a negative earnings surprise. They also appear to be paying more attention to engagement volume in comparison to tweet volume when making their forecasts. In further analysis, I find that the source of this additional information could be that the level of tweets (engagement) helps predict a firm's earnings and sales whereas a change in the volume of tweets (engagement) incrementally explains the firm's sales growth.

The results are robust to using three different samples – full sample of all Tweet and non-Tweet firm-quarters, a subsample of firm-quarters excluding firms which never create a Twitter account, and a subsample of only Tweet firm-quarters[7] – as well as to using alternative measures of dependent and independent variables. There might be a concern that the results may be driven by a few industries. Therefore, I repeat the analysis for earnings and unexpected earnings using median industry-adjusted (SIC 2-digit) tweet/engagement volume and find consistent results. The results are qualitatively similar when I use industry fixed-effects instead of firm fixed-effects.

My paper makes significant contributions to five different strands of literature. Firstly, it contributes to a growing body of accounting literature that studies social media; why and how firms use it and how it affects the capital markets. One strand of this literature examines the determinants and market consequences of firms disseminating information through their official Twitter accounts (Blankespoor et al., 2014; Lee et al., 2015; Jung et al., 2018; Crowley et al., 2018). Another stream of this literature studies the information content of third- party tweets[8] about firms' earnings, products, or stocks and whether it predicts a firm's future sales and stock returns. These studies use the concept of 'Wisdom

[7] The quarter in which a firm first starts tweeting and all subsequent quarters are referred to as Tweet firm-quarters. For example, if a firm creates a Twitter account in Feb 2015 but starts to tweet in May 2015, then all quarters from April 2019 onwards will be classified as Tweet firm-quarters for the firm; quarters before April 2015 will be classified as non-Tweet firm-quarters.

[8] Third- party tweets are between individuals and are not on the official Twitter accounts of firms.

of Crowds'[9] to explain the predictive power of third- party-generated tweets (Tang 2018; Bartov et al., 2018). Most of these studies are event studies, with the exception of Tang (2018), focus on one category of tweets or on individual tweets around a specific event and draw their inferences using a small sample of firms. My study extends this literature by showing the information value of firm-initiated tweet (engagement) volume to investors and analysts over the long window using a comprehensive sample of firm-initiated tweets (all publicly listed US firms between 2006 to 2017).

Second, the study contributes to the literature which study efficiency of capital markets and whether it impounds all available information in the stock prices or there is a significant underreaction due to limited attention (Ball and Brown, 1968; Fama 1970; Bernard and Thomas, 1990; Sloan, 1996; Bloomfield, 2002; Hirshleifer et al., 2002; Hirshleifer et al., 2002). I demonstrate that investors seem to paying attention to changes in tweet (engagement) volume and impound this information into stock prices concurrently.

Third, my paper also extends prior studies which have examined the use of the different medium by firms for voluntary disclosure and dissemination. Prior papers have examined management guidance, conference calls, press releases, company website and supplementary financial statement releases (Coller et al., 1997; Rajgopal et al. 2003; Bushee et al., 2010; Matsumoto et al., 2011; Michaely et al., 2016). My paper studies the use of Twitter as another medium of disclosure and dissemination of information and its consequences for firm performance.

My study also adds to the literature on the role of financial and non-financial leading indicators in predicting future earnings and firm value such as market penetration, air pollution index, customer satisfaction scores, order backlog, web traffic and customer ratings (Amir and Lev, 1996; Ittner and

[9] Wisdom of Crowds refers to the aggregation of information provided by many (non-expert) individuals which may often predict outcomes more precisely than experts as the individuals may be coming from diverse backgrounds and are ,therefore, less likely to herd.

Larcker, 1998; Deng et al., 1999; Hughes, 2000; Trueman et al., 2001; Rajgopal et al., 2003; Luo et al., 2013) in firm valuation. My paper highlights another source of nonfinancial information - the volume of firm-initiated tweets and followers' engagement - that could be informative about the firm's future financial performance to investors and analysts.

Finally, my paper also contributes to the literature in marketing and information systems which focus on social media and its consequences for firms (Schniederjans, et al., 2013; Rishika et al.,2013; Luo et al., 2013; Rui et al., 2013; Yu et al., 2013; Gong et al., 2017). I show that the volume of a firm's tweets and followers' engagement is positively associated with the firm value, stock returns earnings, sales, and sales growth.

The rest of the paper is organized as follows: I discuss Literature review and Hypotheses development in Section 2; Sample, Data collection, Variable Construction and Research Design in Section 3; Descriptive Statistics in Section 4; Empirical Results in Section 5; and finally Conclude with my findings in Section 6.

2. Literature Review and Hypothesis Development

2.1 Literature Review

In the last ten years, social media has emerged as one of the most popular platforms of communication between people. Consequently, an ever-increasing number of firms have started using social media for dissemination of firm-related information to investors, customers, employees, and other stakeholders. Twitter[10], arguably, has emerged as one of the most popular social media platforms. Kang, Hosseini, Savickas, and Singh (2019), hereafter referred to as HKSS, show that close to 52% of publicly listed US firms have official Twitter accounts as on Dec 31, 2017. This new medium of information

[10] In its 2017 10-K filing ,Twitter disclosed that it had 330 million average monthly active users (MAUs) in the three months ended December 31, 2017.As of Dec. 31,2018 ,new age high-tech firms such as Google and Facebook had approximately 20.5 million and 13.5 million followers, respectively and can tweet information and engage directly with them.

dissemination has also generated a great deal of interest from accounting researchers. One strand of literature examines the determinants and market consequences of firms disseminating information using Twitter. Blankespoor, Miller, and White (2014) is one of the first studies in this area. They show that firms can reduce information asymmetry by more broadly disseminating their news using Twitter. They find that additional dissemination of firm-initiated news via Twitter is associated with lower bid-ask-spreads and greater abnormal depth around earnings announcement and the results hold mainly for less visible firms. Lee, Hutton, and Shu (2015) examine how corporate social media affects the capital market consequences of firms' disclosure of negative news in the context of product recalls. Their results suggest that corporate social media attenuates the negative price reaction to product recall disclosures. Interestingly, their study also indicates that the level of control a firm has over its social media content has a role to play in the attenuation benefits. The attenuation benefits are lessened with the arrival of Twitter as the firms do not exercise complete control over the content of their Twitter accounts. Both these papers demonstrate the important role of social media, in general, and Twitter, in particular, as a medium of information dissemination by firms, over and above the coverage by the business press[11], which the capital market pays attention to.

A recent paper by Jung, Naughton, Tahoun, and Wang (2018) examines whether firms use social media (Twitter) to strategically disseminate financial information. Using a sample of S&P 1500 firms from 2010 to 2013, the paper shows that firms are less likely to use Twitter to propagate quarterly earnings news when the news is bad and when the magnitude of the earnings forecast errors is greater, consistent with strategic use of Twitter. The paper also studies the determinants of a firm's decision to have a

[11] Bushee et al. (2009) find that business press acts as an information intermediary and plays an important role in disseminating information as well as by creating new information. Their study also suggests that business press reduces information asymmetry around earnings announcements, with broader dissemination of information having a bigger impact.

presence on Twitter. HKSS (2019)[12] use a comprehensive data set of all publicly listed US firms between 2006 and 2017 to explore the determinants of firms having Twitter accounts. They find that the prime determinants of a firm's decision to create a Twitter account are the availability of resources, business complexity and financial information uncertainty, customer engagement and information dissemination, peer pressure and CEO influence, degree of market concentration in the firm's industry and litigation risk. Interestingly, they find that institutional ownership does not influence a firm's decision to use Twitter. Crowley, Huang, and Lu (2018), another recent working paper, studies the discretionary dissemination of financial tweets on Twitter around earnings announcements, accounting filings and other important corporate events by S&P 1500 firms. Their results indicate that firms make discretionary choices in timing and presentation format when disseminating information on Twitter and also incorporate instantaneous feedback from their Twitter account followers into their dissemination strategies.

There is another stream of literature which studies the information content of third- party tweets about firms' earnings, products or stocks and whether it predicts a firm's future sales and stock returns(e.g., Bollen et al., 2011; Mao et al., 2012; Curtis et al., 2016; Tang, 2018; Bartov et al. ,2018). These studies use the concept of 'Wisdom of Crowds' to explain the predictive power of third- party-generated tweets. Tang (2018) examines the predictive ability of third-party-generated product information tweets, aggregated at the firm- level, about firm-level sales. The paper finds that the incremental information content of the aggregate information increases with the extent to which the Twitter comments are representative of the broad customer response to products and brands. In addition, Twitter comments also explain a part of the unexpected component of sales growth. Bartov, Faurel, and Mohanram (2018) also focus on individual tweets around a firm's earnings announcement and study whether aggregate opinion from individual tweets predicts its earnings and announcement returns. Their sample period is

[12] HKSS (2019) uses Duration modeling, also known as Survival Analysis, to examine the determinants of firms' presence on Twitter.

2009-2012 and covers 3,604 firms of the Russell Index. They find results consistent with their conjecture after controlling for concurrent information or opinion from traditional media sources. Their results hold for tweets that convey original information, as well as tweets that disseminate existing information which also underscores the informational role of dissemination.

Prior literature suggests that firms use Twitter to disseminate information and engage with their stakeholders (Blankespoor et al., 2014; Lee et al., 2015; Jung et al., 2018). Most prior accounting studies have focused on the use of Twitter by firms for the dissemination of financial tweets or other corporate disclosures. However, there are other important stakeholders such as customers, suppliers, and distributors who also consume the information disseminated on a firm's Twitter account. Twitter facilitates two-way communication and feedback in real-time and, therefore, its role as a predictor and influencer of future business actions and performance cannot be underemphasized. As suggested by Tang (2018) and Bartov et al. (2018), tweets aggregated at the firm- level may have some incremental information content beyond the other known traditional sources of information such as traditional media and financial intermediaries. Most of these studies, with the exception of Tang (2018), are event studies with a short window focus. Prior studies also suffer from the malaise using a small sample of firms and/or limited time period, due to data collection constraints, thereby raising concerns about the generalizability of the results. My current study uses the aggregate-level information contained in firm-initiated tweets and followers' engagement[13] from a firm's official Twitter account to predict its earnings, unexpected earnings, sales growth, and unexpected sales growth. It is an important issue which has not been explored, to the best of my knowledge. It is, therefore, related to both streams of accounting literature which study tweets – the one which focuses on firm-initiated tweets as well as the one focused on third-party

[13] A Twitter account has followers who follow and respond to the information disseminated on the account. Followers can interact and show their interest to a particular tweet by liking, retweeting, or replying to it, which I collectively refer to as 'engagement'. While anyone can see and respond to tweets, without following the Twitter account, it is reasonable to assume that followers are people or entities who are interested in knowing more about and interacting with the owner of the account.

generated tweets. My sample is also more comprehensive and includes all public US firms and the period of study is from January 2006 to December 2017.

2.2 Hypothesis Development

When a firm creates a Twitter account, it establishes a fast and reliable method of disseminating news and other information to its stakeholders (customers, investors, distributors, etc.).Twitter facilitates a firm's real-time engagement with its followers who can share their views and opinions with the firm through publicly viewable feedback. This provides the firm with a powerful platform to constantly communicate and engage with its stakeholders. For example, the firm may use Twitter to market its products and services or to fulfill a service request from a customer or to make corporate announcements. Twitter, therefore, allows the firm to engage with its stakeholders in a way which traditional modes of communication such as press releases, television, conference calls, etc. do not. When a firm tweets, a follower may respond by liking, retweeting or replying to the tweet – collectively referred to as engagement. The level of engagement to a firm's tweet is the collective positive response or feedback of the followers to that tweet. This also allows a firm to compare the engagement level of followers for two different marketing campaigns or product launches. A high level of engagement represents strong positive feedback whereas lower engagement level may indicate, at best, lukewarm enthusiasm. All firms may not be equally adept at engaging successfully with their customers on social media (Lee et al., 2018). Therefore, the volume of a firm's tweets and the engagement level of the followers aggregated at for a given period may convey incremental information about the firm's likely business performance during that period, where a high level of tweeting and engagement is associated better performance.

However, with an open and interactive social media platform, the firm also relinquishes its full control over the contents being transmitted on its official Twitter account (Lee et al., 2015). Therefore, a firm with a Twitter account also becomes vulnerable as criticism and negative feedback by even a few can be viewed by other followers, investors and analysts. The firm can still influence what gets communicated and discussed on its Twitter account; however, the followers, now, also exercise a great degree of control through the engagement and feedback process. An online platform such as Twitter is also susceptible to manipulation, rumors or negative sentiment by 'interested' parties (Lee et al., 2018; Lee et al., 2015) and most of the communication is qualitative in nature.

Efficient Markets Hypothesis (EMH) states that all available public information is impounded immediately in the market prices fully (Fama 1970). Therefore, the aggregate information in tweets and followers' engagement should get reflected in the stock prices as these are in the public domain. This means that the tweet (engagement) volume should be value relevant and contribute positively to the firm value. However, researchers in accounting and finance have also shown that there are notable exceptions to EMH such as PEAD (Ball and Brown, 1968; Bernard and Thomas, 1990), accrual anomaly (Sloan, 1996; Xie, 2001), etc. Some of the possible explanations for this could be limited investor attention due to costly processing and information complexity (Bloomfield, 2002; Hirshleifer, Lim, and Teoh, 2002; Hirshleifer and Teoh, 2003) or underreaction due to slow diffusion of information (Hong and Stein, 1999).[14] The qualitative nature and sheer volume of the firm-initiated tweets and followers' engagement may make it difficult for the investors to fully process and incorporate it into prices. Therefore, ex-ante it is not clear whether the aggregate information in the firm's tweets (engagement) will be incorporated in the firm value during the period, beyond the concurrent information already contained in other known sources of information such as press, analyst forecasts, and voluntary disclosures. This leads to the following hypotheses stated in the null form:

[14] In addition, Hales (2007) shows that investor preferences can also significantly influence the manner in which information is processed and affect their expectations of future earnings performance. This could lead them to make investment decisions which may not be in their best interest. However, I focus only on firm-initiated tweets and, hence, this is not relevant to my study.

Hypothesis 1A: Ceteris Paribus, a firm's tweet (engagement) volume/ change in tweet (engagement) volume in a given period is not associated with the firm's value/ change in the firm's value during that period.

Hypothesis 1B: Ceteris Paribus, change in a firm's tweet (engagement) volume in a given period is not associated with the firm's stock return during that period.

It has been well documented in accounting and finance literature that analysts perform a valuable role as information intermediaries (Givoly and Lakonishok, 1979; Lys and Sohn, 1990; Francis and Soffer, 1997; Healy and Palepu, 2001; Hilary and Hsu, 2013; Brown, Call, and Sharp, 2015). The analysts process and interpret news and information about a firm and the economy from all public, as well as private, sources and come up with their estimates of the firm's expected earnings and sales for a period (Lang and Lundholm, 1996; Rogers and Grant, 1997; Bowen, Davis and Matsumoto, 2002). The consensus estimate of all the analysts following a firm is often considered as the market's expectation of a firm's earnings and sales during a period (O'Brien 1988; Zhang 2008; Lobo, Song and Stanford, 2017). The market reacts to any unexpected earnings which the firm subsequently reports (Imhoff and Lobo 1992; Chen, Cheng, and Lo 2010; Francis, Schipper, and Vincent 2002; Zhang, 2008). An analyst may also be paying attention to firms' dissemination of information on Twitter and other social media accounts as part of this information-gathering process. Therefore, the volume of a firm's tweets and its followers' engagement, or a change in them, may provide incremental information to the analysts. This additional information may aid the analysts in making better estimates of the expected earnings and sales of the firm. This implies that conditional on a firm using Twitter, the firm's unexpected earnings and unexpected sales growth should have a negative relationship with the level of (change in) tweets and engagement.

On the other hand, analysts may find it hard to fully decipher these tweets because of the large volume and the, predominantly, qualitative nature of information contained in them (Plumee, 2003;

Hoddr, Hopkins, and Wood, 2008; Lehavy, Li, and Merkley, 2011). Additionally, a firm's tweets and followers' engagement may only be adding noise to the information environment of the firm. This may actually make the analyst's forecast worse off than before if the analyst factors in these tweets. This implies that the firm's unexpected earnings and unexpected sales could be higher than when the firm does not have a Twitter account or when the analyst does not pay attention to a firm's Twitter account or social media presence. Therefore, it is an open empirical question whether the volume of (change in) a firm's tweets and followers' engagement is positively or negatively associated with its unexpected earnings or unexpected sales growth during the period. This leads to the following hypotheses stated in the null form:

Hypothesis 2A: Ceteris Paribus, a firm's volume of tweeting (engagement of followers) in a given period is not associated with the firm's unexpected earnings or unexpected sales growth during that period.

Hypothesis 2B: Ceteris Paribus, change in a firm's volume of tweeting (engagement of followers) in a given period is not associated with the firm's unexpected earnings or unexpected sales growth during that period.

3. Sample, Data Collection, Variables and Research Design

3.1 Sample and Data Collection

I use a hand-collected sample of tweets, retweets, likes, and replies from the official Twitter accounts of firms for my study[15]. I cover all public US firms listed on NYSE/AMEX/NASDAQ exchanges between 2006 and 2017 and check whether they have an official Twitter account[16] and then use the Twitter Application Program Interface (API) and web-scraping to retrieve the full text of each firm-initiated

[15] I employ the same sample of tweets which has been used for the working paper "Determinants of Firms' Presence on and Use of Twitter: An Empirical Study" by Kang, Hosseini, Savickas and Singh (2019) for my analysis.
[16]A firm shows icons of all the social media platforms on which it has a presence (such as Twitter, Facebook, LinkedIn, Youtube) .This icon is the link to the firm's official account on that social media - in the case of Twitter, I call this the Primary Twitter account of that firm.

tweet[17]. I focus only on the primary Twitter accounts of firms[18] in this paper. The final Tweet data used in my study has approximately 17.50 million tweets by firms, and 108 million retweets, 164 million likes and 14.9 million replies by followers, collected from the primary Twitter sites of 2,229 unique firms for the sample period. This makes it the most comprehensive study, to the best of my knowledge, focusing on firms' Twitter accounts and Tweets. I use quarterly data for my study and the sample period is from the first quarter of 2006 to the last quarter of 2017[19].

I collect quarterly financial data of firms from Compustat, stock and market return data from CRSP, market factors data from Prof. Kenneth French's website, and analyst data from IBES. I also collect newspaper and business press data from LexisNexis. My final data for the full sample comprises[20] of 178,236 firm-quarters (46,559 Tweet firm-quarters) and 6,268 unique publicly listed firms (2,229 unique Tweeting firms).

3.2 Variables Description

In this section, I define the dependent variables and the variables of interest which I use to test the hypotheses in Section 2.2.

Dependent Variables

I use *TOBIN's Q* as a measure of the firm-value to test hypotheses 1A as it represents the ratio of the market value of a firm and the replacement cost of the firm's assets. I use monthly excess stock return (*MON_EXCESS_RETURN*) as the dependent variable to test hypotheses 1B. I compute

[17] Some firms keep their tweets protected and which are visible only to followers. There are 12 such firms in my sample. I remove these firms from my sample.

[18]Each Tweet firm has one Primary Twitter account. Some firms may also have additional Twitter accounts, which I refer to as Secondary Twitter accounts to cater to different regions, investor relations, customer services, recruitment etc. I do not include tweets from these Secondary accounts in my analysis.

[19] The company Twitter was created in March 2006.Starbucks was the first public firm in US to create a Primary Twitter account in November 2006. See https://twitter.com/starbucks for reference.

[20] I eliminate all firm-quarters with missing assets, revenue, EPS, leverage, market value and book value of equity data.

MON_EXCESS_RETURN by subtracting the 1-month Treasury bill rate from the corresponding month's stock return for each firm.

I test Hypotheses 2 A & B using quarterly unexpected earnings (*UE*) and quarterly unexpected sales growth (*U_SALES_GR*) as the dependent variables. I compute quarterly unexpected earnings (*UE*), as the actual EPS for the quarter (reported in I/B/E/S) minus the consensus analyst EPS forecast for that quarter, scaled by previous quarter-end stock price for each firm-quarter and quarterly unexpected sales growth (*U_SALES_GR*), as the actual quarterly sales (reported in I/B/E/S) minus the most recent consensus analyst sales forecast for the quarter, divided by previous quarter's sales for each firm-quarter.

Variables of Interest

The focus of my paper is to examine whether a firm's tweets and followers' engagement aggregated at the firm-quarter level provide any incremental information which can help predict the firm's quarterly earnings, sales growth, unexpected earnings, and unexpected sales growth. I use the following five variables to capture this aggregate level tweet behavior for each firm-quarter:

i. *LOG (TWEETS)*: This measure captures the firm-initiated tweet volume on its primary Twitter account. It represents the volume of information being disseminated by the firm to the public. The tweets could be a combination of financial and non-financial disclosures, marketing campaigns, sales promotion, new product launches, customer service, etc. I aggregate the firm's tweets for each quarter and then take the natural log after adding one to this aggregate.

ii. *LOG (ENGAGEMENT)*: Followers[21] can choose to ignore a firm's tweets or respond to it by liking, retweeting or replying to it – which I collectively refer to as engagement. The volume of engagement to each tweet is representative of the enthusiasm the tweet generates in the followers. It is important

[21] The number of followers are shown on a Twitter account. However, it is a static observation and one cannot observe its time-trend.

to note that engagement is only a lower bound of the extent to which a firm's dissemination has been 'seen' by the intended audience and the excitement or 'buzz' it generates. This measure captures the level of engagement of the stakeholders or followers of the firm's Twitter account for each quarter. It represents how well the firm has been able to communicate with its followers. I add the likes, retweets, and replies for each tweet and then aggregate it for each firm-quarter to compute the engagement. I then take the natural log after adding one to this aggregate engagement.

iii. *RESPONSE*: There may be a concern that larger firms or consumer-facing firms (B2C firms) may have more followers and, therefore, may tweet more and, also, be able to generate more engagement. Also, there might be just a mechanical relationship between the EPS and sales growth and tweet/engagement volume as both might be increasing over time. To allay this concern, I normalize the level of engagement by dividing *LOG (ENGAGEMENT)* by *LOG (TWEETS)*. This measure, *RESPONSE*, then represents the engagement per unit tweet for each firm-quarter.

iv. *CHANGE_LOG (TWEETS)*: The level of tweets and engagement might be good measures for predicting earnings. However, change of these level variables might be better for studying their association with sales growth, unexpected earnings, and unexpected sales growth which are all change variables. I define *CHANGE_LOG (TWEETS)* as the difference between *LOG (TWEETS)* of current and previous quarters for each firm. This measure has an added advantage that it can take both positive as well as negative values.

v. *CHANGE_LOG (ENGAGE)*: Similar to *CHANGE_LOG (TWEETS)*, I define *CHANGE_LOG (ENGAGE)* as the difference between *LOG (ENGAGEMENT)* of current and previous quarters for each firm. This measure can also take both positive as well as negative values.

3.3 Research Design

[Insert Figure 1 here]

Firms, which have a presence on Twitter, continuously disseminate financial, marketing, customer services and other company-related information through tweets. The followers of the Twitter account also respond to these firm-initiated tweets in real-time, and, therefore, there is an almost uninterrupted flow of publicly viewable information. The firm announces its earnings for the current quarter sometime during the next quarter; after the beginning but before the end of the next quarter. Therefore, the volume of a firm's tweets and followers' engagement during the quarter might be a leading indicator of the firm's performance in that quarter and maybe incrementally informative to capital market participants and financial intermediaries. This additional information is over and above the other known sources of concurrent information such as traditional media, firm's voluntary disclosures, and analysts' forecasts. This is shown schematically in Figure 1.

I examine the informativeness of the volume of firm-initiated tweets (engagement) in two ways. First, I test whether the volume of tweets (engagement) aggregated over a period or a change thereof is value relevant for that period. Next, I verify whether this information is useful for analysts to make more accurate earnings and sales estimates.

I use the following OLS regression model to test Hypothesis 1 A:

$$FIRMVALUE_{i,t} = \beta_0 + \beta_1 TWEET\_VOLUME_{i,t}/CHANGE\_TWEET\_VOLUME_{i,t} + \Sigma\beta_J\ CONTROLS_{i,t} + YEAR\_QTR\ FIXED\ EFFECTS + FIRM\ FIXED\text{-}EFFECTS + \varepsilon_{i,t} \quad (1)$$

where *i* indexes firm and *t* indexes quarter. I use *TOBIN'S Q* to measure the firm value. I predict that $\beta_1 > 0$ for Hypotheses 1A&B.

I use the Fama-French Factor model to test Hypothesis 1B employing Fama-MacBeth (1973)[22] monthly cross-sectional regressions with Newey-West (1987) corrected standard errors for autocorrelation (two lags) used for calculating t-statistics:

$$(R_{i,t} - Rf_t) = \beta_0 + \beta_1 CHANGE\_LOG\ (TWEET)\_MON_{i,t}/CHANGE\_LOG\ (ENGAGE)\_MON_{i,t} + \beta_2 MOM_t + \Sigma\beta_J FAMA\text{-}FRENCH\_FACTORS_t + \varepsilon_{i,t} \quad (2)$$

where *i* indexes firm and *t* indexes month, $R_{i,t}$ is the monthly buy and hold return and Rf_t is 1-month T-bill rate, $(R_{i,t} - Rf_t)$ is the monthly excess stock return, *CHANGE_LOG (TWEET)* and *CHANGE_LOG (ENGAGE)* are as defined in the previous section. I use both Fama-French[23] three factors (Fama, and French, 1983) - and Fama-French five factors (Fama, and French, 2015) and also include Momentum factor (Jegadeesh, and Titman, 1993; Carhart, 1997) to test hypothesis 1B. I predict that $\beta_1 > 0$, which implies that the change in tweet (engagement) volume is informative to the market beyond the Fama-French and Momentum factors.

I use the following OLS regression model to test hypotheses 2 A & B:

$$UE\_EARNINGS_{i,t}/UE\_SALES\_GR_{i,t} = \beta_0 + \beta_1 TWEET\_VOLUME_{i,t}/CHANGE\_TWEET\_VOLUME_{i,t} + \beta_2 TWEET\_VOLUME_{i,t}/CHANGE\_TWEET\_VOLUME_{i,t} * NEG\_UE\_EARNINGS/NEG\_UE\_SALES\_GR_{i,t} + \beta_3 NEG\_UE\_EARNINGS_{i,t} / NEG\_UE\_SALES\_GR_{i,t} + \Sigma\beta_J CONTROLS_{i,t} + YEAR\_QTR\ FIXED\ EFFECTS + FIRM\ FIXED\text{-}EFFECTS + \varepsilon_{i,t} \quad (3)$$

I predict that $\beta_1 < 0$ and $\beta_2 > 0$ for Hypotheses 2 A & B.

[22] I use prior 36 months data to, first, compute the factor betas on a monthly rolling basis for each firm. This ensures that the factor beta used in any month has been computed using previous 36 months data. Next, I perform cross-sectional regressions for each month and, finally, take the time-series average of the slopes in monthly regressions, with Newey-West corrected standard errors used for calculating t-statistics to account for any autocorrelation (two lags).

[23] I obtain the data on monthly factor returns from Professor Kenneth R. French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. I am thankful to Prof. French for making this data available for research.

where *i* indexes firm and *t* indexes quarter , *UE_EARNINGS* is the unexpected earnings for the quarter and *UE_SALES_GR* is the unexpected sales growth for the quarter. *NEG_UE_EARNINGS (NEG_UE_SALES_GR)* is an indicator variables equal to 1 if unexpected earnings (sales growth) is negative.

LOG (TWEETS) /LOG (ENGAGEMENT) /RESPONSE and *CHANGE_LOG (TWEETS) /CHANGE_LOG (ENGAGE)* are used as measures of *TWEET_VOLUME* to test hypotheses 1A, 2A, and 2B. I use *CHANGE_LOG (TWEETS)/CHANGE_LOG (ENGAGE)* as variables of interest to test hypotheses 1B as stock return represents a change in prices. I use size, market to book ratio and leverage as controls for firm characteristics as they might affect the firm's performance. I also control for other known sources of concurrent information about the firm's performance using analyst coverage, press releases by the firm and newspaper articles about the firm during the quarter. I use the lagged dependent variable as an additional control variable in Model 3 as there might be a high correlation between previous and current quarter's unexpected performance. In addition, I control for size, current and previous quarter's ROA[24], loss, and M&A activity in Model 1, and advertising expense[25] and previous quarter's sales in Model 3. I use year-quarter fixed effects to control for any time trends and firm fixed effects to control for any time-invariant firm characteristics. All variables are defined in Appendix A. I cluster the standard errors by SIC 2-digit industry and winsorize all continuous variables at the1% and 99% level.

I apply the models to three different samples. For the first sample, which I refer to as full sample, I use all firm-quarters - Tweet as well as all non-Tweet firm-quarters. This allows me to observe the informational effect of tweets and engagement relative to the control group (non-Tweet) firms. The second sample, a subsample of the full sample, has only Tweet firm-quarters. This allows me to interpret

[24] I use both current and previous quarter's ROA as previous research has shown that the market underreacts to current quarter's earnings.

[25] Advertising expense is reported annually in Compustat whereas I use quarterly data for my empirical analysis. I, therefore, divide the annual advertising expense equally over the four quarters and then scale it by the average total assets for the quarter.

the coefficient of tweet/engagement volume as having an association with firm performance. This is because the firm might be using tweets to disseminate information about its sales promotion, product and service offerings and to engage with customers for providing better customer relationship. This may have an effect on the firm's sales and earnings. The last sample excludes firms which have never created a Twitter account from the full sample. This allows me to observe the informational effect of tweets and engagement for the tweeting firm relative to when that firm did not have a Twitter account. I report most of the results using the first two sample and use the third sample for robustness checks. I have tried to control for other public sources of information such as business press, newspapers and analysts. However, I cannot rule out the possibility of an omitted correlated variable which might be influencing both the tweet volume as well as the firm performance.

4. Descriptive Statistics

[Insert Figure 2 here]

There has been a rapid increase in both the number as well as the proportion of firms which use Twitter to disseminate information and engage with their stakeholders. As shown in Figure 2, the proportion of firms which use Twitter has increased from 0% in 2006 to almost 50% in Dec 2017[26]. This shows that Twitter is a popular social media platform used by firms.

[Insert Tables 1A, B, C, and D here]

Panel A of Table 1 shows the descriptive statistics of the key variable for the full sample of 178,236 firm-quarters - both Tweet as well as all non-Tweet firm-quarters (also includes firms which do not have a Primary Twitter account as of Dec 31, 2017). Panel B of Table 1 shows the descriptive statistics for the smaller subsample of 46,559 Tweet firm-quarters (includes only firms which have a Twitter account and

[26] HKSS (2019) document that 52% of US public firms had a Twitter account as of Dec 31, 2017.The proportion of firms using Twitter for dissemination is slightly lesser than this. One of the reasons could be that 73 new firms joined Twitter in 2017 and there might be a gap between when a firm joins Twitter and the time it starts tweeting.

started tweeting during the sample period. 'Tweet' firms tweet 375 times, on average, every quarter and are able to generate an average of 6150 likes, retweets and replies (engagement) by the followers. There is wide variation in the use of Twitter by firms - firms at 25% (75%) percentile have 16 (231) tweets and 9(534) engagement per quarter. On average, there is an increase in tweets (engagement) of 8 (138) per firm-quarter. Again, there is a wide variation in the distribution of these variables as firms at 25% percentile experience a decrease in tweets (-18) and engagement (-13) whereas firms at 75% have a net increase in tweets (23) and engagement (57) every quarter. In comparison, these 'Tweet' firms have, on average, 14 press releases, appear 58 times in newspapers and have 9 analysts (untabulated) following them every quarter. This suggests that firms might be using Twitter to disseminate information not only about voluntary disclosures and corporate announcements but for other purposes as well. Panel C shows the time-trend of Tweet firm-quarters, total tweets, average tweets, total engagement and average engagement over the sample period of 2006[27] to 2017. This shows that there has been an explosive growth in the usage of Twitter by firms as well as followers' engagement. However, there has been a dip in the number of tweets in 2017.[28]

[Insert Figures 3A, B, C, and D here]

Figures 3A&B show the trend of firm-initiated tweets and 3C&D show the trend of follower's engagement for Fama-French 10 industries. The trend is broadly similar for all the industries with consumer durables industry having the highest volume of tweets and telephone and television transmission industry having the highest frequency of tweets per firm-quarter. These industries also lead in total and average engagement categories, respectively.

[27] There is only one firm which joined Twitter in 2006 but did not tweet during that year. Hence, there are zero tweets and engagement in 2006.
[28] One of the reasons could be that Twitter increased the number of characters which can be used in a tweet from 140 to 280 in November 2017.

Panel D of Table 1 shows the Pearson's correlation between key variables for the subsample of Tweet firm-quarters. The table shows that on a univariate basis, there is a signification positive correlation between *LOG (TWEETS)* and earnings, unexpected earnings, sales, and log (market value). Similarly, there is a significant correlation between *LOG (ENGAGEMENT)* and earnings, unexpected earnings, sales, and log (market value).*CHANGE_LOG (TWEETS)* and *CHANGE_LOG (ENGAGEMENT)* have a significant positive correlation with earnings, sales, sales growth, unexpected sales growth, and log (market value). Stock return is positively correlated with *CHANGE_LOG (TWEETS)* and *CHANGE_LOG (ENGAGEMENT)* but not with *LOG (TWEETS)* or *LOG (ENGAGEMENT).* This provides some initial support for my hypotheses.

5. Empirical Results

In this section, I discuss the results of testing my hypotheses using the three models discussed in Section 3.3.

5.1 Firm Value and Stock Returns

I present the results of testing Hypotheses 1A&B to demonstrate value relevance of the information content of firms' tweet volume and followers' engagement.

[Insert Table 2 here]

Table 2 shows the results of testing the association between tweet (engagement) volume/ change in tweet (engagement) volume and firm value using Model 1. The coefficients of *LOG (TWEETS)*, *LOG (ENGAGEMENT)* and *RESPONSE* are all positive and statistically significant at 1% level. Interestingly, the coefficient of *LOG (TWEETS)* becomes insignificant whereas that of *LOG (ENGAGEMENT)* remains positive and significant when both are included in the model in column 4; though F-test shows that the coefficients are not statistically different (p-value of 0.46). The results are also economically meaningful. One standard deviation increase in *LOG (TWEETS)* is associated with an average increase of approximately

3.25% ((e $^{(0.01*2.001)}$ – 1) = 0.0325)[29] in the firm-value. Similarly, one standard deviation increase in *LOG (ENGAGEMENT)* is associated with an average increase of approximately 6.13% ((e $^{(0.011*2.835)}$ – 1) = 0.0613) in the firm-value. Columns 5, 6 and 7 show the results when the change specification of tweets (engagement) is used. The coefficient of *CHANGE_LOG (TWEETS)* is insignificant in column 5 but the coefficient of (*CHANGE_LOG (ENGAGE))* is positive and statistically significant at 10% level in column 6. The coefficient of *CHANGE_LOG (ENGAGE)* remains positive and significant at 1% level whereas the coefficient of *CHANGE_LOG (TWEETS)* remains negative when both and are used together in column 7; F-test shows that the coefficients are also statistically different (p-value of 0.02). This suggests that engagement of followers is more informative about the firm-value than the volume of tweets by the firm. This provides evidence in support of Hypothesis 1A.

[Insert Tables 3A and B here]

Next, I examine whether the change in tweet (engagement) volume is informative to capital market participants in Table 3 using Model 2. The dependent variable is the monthly excess stock return which is the excess stock return over the 1-month Treasury bill rate and all independent variables are also monthly. The analysis employs Fama-MacBeth cross-sectional regressions to Fama-French three-factor and five-factor models. Asset pricing models (Fama and French, 1993; Cahart, 1997; Fama and French, 2014) state that differences in common factor betas can explain all the cross-sectional differences in stock returns and, hence, individual firm characteristics and idiosyncratic risk does not matter. I use prior 36 months data to, first, compute the common factor betas on a monthly rolling basis for each firm. This ensures that the monthly common factor betas for any month have been computed using the last 36 months data. Next, I perform cross-sectional regressions for each month after including the variables of interest and, finally, take the time-series average of the slopes in monthly regressions. The slopes reported

[29] In a log-log regression specification of the type Log(Y) = $\beta_0 + \beta_1$Log(X), if Log(X) increases by 1 unit then Y increases by ((e $^{\beta_1}$ – 1) times.

in Table 3, therefore, are the coefficients of the common factor betas and tweet(engagement) volume variables and t-statistics have been calculated using Newey-West corrected standard errors to account for any autocorrelation (two lags) in the error terms.

Panel A of Table 3 displays the results for FF three-factor model (Fama and French, 1993)[30] – Market return (MKTRF), Size (SMB) and Book-to-Market (HML). FF five-factor model (Fama and French, 2015)[31] introduced two new factors – operating profitability (RMW) and investment (CMA) – which I also include in Panel B. Jegadeesh and Titman (1993) provide evidence that past winners tend to outperform past losers in the following years and, therefore, I include Momentum factor in both the specifications. The full sample has been used in the first three columns and a subsample of only Tweet firm-months in columns 4, 5 and 6 in both the panels. The coefficients of *CHANGE_LOG (TWEETS) _MON* and *CHANGE_LOG (ENGAGE) _MON* are positive and statistically significant at the 1% level in columns 1&2 and columns 3&4 in both Panels A and B. This indicates that change in tweet (engagement) volume incrementally explains the cross-sectional differences in stock returns beyond the common factor betas. This also suggests that changes in tweet (engagement) volume are incrementally informative to the market participants and are priced by them in the same period. The coefficient of *CHANGE_LOG (ENGAGE) _MON* remains positive and significant whereas the coefficient of *CHANGE_LOG (TWEETS) _MON* becomes insignificant (except in Column 3 of Panel A) when I include both *CHANGE_LOG (TWEETS) _MON* and *CHANGE_LOG (ENGAGE) _MON* in the specification in columns 3 and 6.

Overall, I find strong evidence for my hypotheses 1A and B suggesting that the aggregate information in the firm's tweets (engagement) gets incorporated in the firm value during the same period,

---

[30] Fama and French (1993) established that three common risk factors – overall market factor, firm size and book-to-market equity – explain the average returns on stocks. This has been widely used in accounting and finance literature to test for the presence of anomalies such as PEAD Ball and Brown, 1968; Bernard and Thomas, 1990, accruals (Sloan, 1996), momentum (Jegadeesh and Titman, 1993) etc.

[31] Fama and French (2015) added two new common risk factors – operating profitability (RMW) and investment (CMA) – to the FF three factor models. This is known as the FF-five factor model.

beyond the concurrent information already contained in other known sources of information such as press, analyst forecasts, and voluntary disclosures. Another interpretation of the results could be that change in a firm's tweet (engagement) volume in a given period is a leading indicator of the expected stock returns during that period as this can be observed in real-time by managers, capital market participants, and financial intermediaries. The results also suggest that the market participants seemingly find a change in followers' engagement more informative than a change in firm's tweet volume.

5.2 Unexpected Earnings and Unexpected Sales Growth

Next, I test hypothesis 2 A & B using unexpected earnings (*UE_EARNINGS*) and unexpected sales growth (*UE_SALES_GRI*) as the dependent variables. I compute unexpected earnings as the difference between actual EPS for the quarter and the consensus analyst forecast EPS for the quarter scaled by previous quarter end's stock price and unexpected sales growth as the difference between actual quarterly sales and the consensus analyst sales forecast for the quarter divided by previous quarter's sales.

[Insert Tables 4 A and B here]

I create 5 portfolios by sorting on unexpected earnings, with portfolio 1 having the most positive *UE_EARNINGS*, portfolio 3 having *UE_EARNINGS* closest to zero and portfolio 5 having the most negative *UE_EARNINGS*. I, then, compute *LOG (TWEETS)* and *LOG (ENGAGEMENT)* for each portfolio. Panel A of Table 4 shows the mean value of *UE_EARNINGS* and *LOG (TWEETS)* and *LOG (ENGAGEMENT)* of the five portfolios. One can observe that the value of *LOG (TWEETS)* and *LOG (ENGAGEMENT)* increases and attains the maximum in portfolio 3 (which has the smallest absolute value of *UE_EARNINGS*) and then starts declining, with portfolio 5 having smaller values than portfolios 3 & 4. This suggests a non-linear relationship between the volume of tweets/engagement and unexpected earnings. It could also mean that the volume of tweets/engagement assists the analysts in making more accurate earnings forecasts. F-test rejects the null hypothesis that the value of *LOG (TWEETS) /LOG (ENGAGEMENT)* is the same for portfolios 1 & 3 and for portfolios 5 & 3. In Panel B, I create portfolios by first sorting on size and then on unexpected earnings. This allows me to observe the relationship between unexpected earnings and tweet/engagement volume controlling for size. Again, I find a similar pattern as in Panel A, and a non-linear relationship between *UE_EARNINGS* and *LOG (TWEETS/LOG (ENGAGEMENT)*. F-test, again, rejects the null hypothesis that the value of *LOG (TWEETS)/LOG (ENGAGEMENT)* is the same for portfolios 1 & 3 and for portfolios 5 & 3. In untabulated results, I find statistically similar results when I create triple sorted portfolios on size, ROA and unexpected earnings.

[Insert Tables 5 A and B here]

I examine the association between tweet/engagement volume and unexpected earnings in a multiple regression setting using Model 3.As unexpected earnings could be positive, negative or zero, I also define an indicator variable *NEG_UE_EARNINGS* which is equal to 1 if the unexpected earning is negative, and 0 otherwise. This helps me to test the association between the tweet volume and the signed unexpected earnings. Panel A of Table 5 has the results for the full sample. The coefficient of *LOG (TWEETS)* (*LOG (TWEETS)*NEG_UE*) is negative (positive) and significant in column 1. Similar results hold for *LOG (ENGAGEMENT)* and *LOG (ENGAGEMENT)* NEG_UE* in column 2. This is consistent with my prediction for Model 3 in Section 3.3. The positive coefficient of *LOG (TWEETS)*NEG_UE* and *LOG (ENGAGEMENT)*NEG_UE* means that the unexpected earnings becomes less negative or closer to zero since *NEG_UE_EARNINGS* represents negative earnings surprise. This implies that the unexpected earnings of firms which tweet is lower than that of non-Tweet firms - the reference group in the full sample is non-Tweet firm quarters. The effects are also economically significant. The mean previous quarter's share price used as the deflator of unexpected earnings is approximately $30. This indicates that a 50% increase in tweets is associated, on average, with a 3 cents reduction in unexpected earnings per share when the unexpected earnings is positive (-0.002% *50 *$30 =3 cents)[32] and 10.5 cents increase in unexpected earnings per share when the unexpected earnings is negative ((0.009 – 0.002) % *50 *$30 =10.5 cents) which is not trivial. Similarly, a 50% increase in engagement is associated, on average, with a 1.5 cents reduction in unexpected earnings per share when the unexpected earnings is positive (-0.001% *50 *$30 =1.5 cents) and 1.5 cents increase in unexpected earnings per share when the unexpected earnings is negative ((0.002 – 0.001) % *50 *$30 =1.5 cents).

The results for Hypothesis 2A hold in column 3 when I use *RESPONSE* which proxies for the level of engagement of followers per unit of tweeting by a firm. In untabulated results, I find that engagement volume is more informative to the analysts than tweet volume when I include both in the model.[33]

These results are consistent with Hypothesis 2A which suggests that there is incremental information in the volume of tweets (engagement) aggregated quarterly. The analysts also seem to be factoring in this incremental information, enabling them to make more accurate earnings forecast which reduces the unexpected earnings – relative to firms which do not have a Twitter account. The results hold for positive as well as negative unexpected earnings. However, I do not find support for Hypothesis 2B when I include *CHANGE_LOG (TWEETS)* in the model in column 4 and only weak support when I include *CHANGE_LOG (ENGAGE)* in the specification in column 5.

I repeat the analysis using a subsample of firms which have a Twitter account and have started tweeting. The results are shown in Panel B of Table 5 and are statistically similar to those for the full sample in Panel A for the level of tweets (engagement). However, I do not find results consistent with my

---

[32] In a level-log regression specification of the type $Y = \beta_0 + \beta_1 \text{Log}(X)$, for small changes in X, we can interpret β_1 as "if we change X by one percent, we'd expect Y to change by (β_1/100) units of Y".

[33]In a test of comparison of coefficients, I find that coefficient of *LOG (ENGAGEMENT)* is significantly different from the coefficient of *LOG (TWEETS)* (F-value of 7.30***) and coefficient of *LOG (ENGAGEMENT)*NEG_UE* is significantly different from the coefficient of *LOG (TWEETS) *NEG_UE (*F-value of 4.05**).

Hypothesis 2B when I use change in the level of tweets (engagement) in columns 4 and 5. This may be because analysts are paying closer attention to the level of tweeting and followers' engagement and not to the change in these variables, which might be more difficult for them to measure and incorporate in their forecasts. It is also worthwhile noting that in both panels the magnitude of the interaction term between tweet (engagement) volume and negative unexpected earnings is higher than the magnitude of tweet (engagement) volume. This may indicate that an increase in tweet (engagement) volume is more helpful to analysts to make more accurate forecasts when there is negative news; for the full sample, F-test rejects the null hypothesis that the magnitudes of *LOG (TWEETS)* and *LOG (TWEETS)*NEG_UE* are equal in Column 1 (p-value of 0.00) and that the magnitudes of *LOG (ENGAGEMENT)* and *LOG (ENGAGEMENT)*NEG_UE* are equal in Column 2(p-value of 0.00).

[Insert Table 6 here]

I examine the association between tweet/engagement volume and unexpected sales growth in a multiple regression setting using Model 3 for a subsample of firms which have a Twitter account and have started tweeting – only Tweet firm-quarters. As in the case of unexpected earnings, unexpected sales growth also could be positive, negative or zero. Therefore, I define an indicator variable *NEG_U_SALES_GR* which is equal to 1 if the unexpected sales growth is negative, and 0 otherwise. This helps me to test the association between the tweet volume/engagement (change in tweet volume/engagement) and the signed unexpected sales growth. Table 6 displays the results of the analysis. The coefficient of *TWEET_START* (*TWEET_START* NEG_UE_SALES_GR*) is negative (positive) and statistically significant at 1% level in column 1. The coefficients of *LOG (TWEETS)* and *LOG (TWEETS)*NEG_UE_SALES_GR* in column 2 are negative and positive, respectively, and statistically significant at 1% level. Similarly, the coefficients of *LOG (ENGAGEMENT)* and *LOG (ENGAGEMENT)* NEG_UE_SALES_GR* in column 2 are negative and positive, respectively, and statistically significant at 1% level. This is consistent with my prediction for Model 3. *NEG_UE_SALES_GR* represents negative sales growth surprise. Therefore, the positive coefficient of *LOG (TWEETS)*NEG_UE_SALES_GR* and *LOG (ENGAGEMENT)*NEG_UE_SALES_GR* implies that unexpected sales growth is becoming less negative or closer to zero. This suggests that the volume of tweets is providing additional information to the analysts which helps them make more accurate sales forecasts. The results for the change specification are in columns 4 and 5. I do not find support for Hypothesis 2B as the coefficients of all the variables of interest are statistically insignificant. The last column shows the results with *RESPONSE* as the variable of interest which provides some support for hypothesis 2A as the coefficient of *RESPONSE* is negative and significant but the coefficient of the interaction term with *NEG_UE_SALES_GR,* though positive is statistically insignificant.

Collectively, Tables 4, 5 and 6 provide strong support for hypotheses 2A, and some weak evidence in support of hypothesis 2B, especially in the full sample. I control for other known sources of public information and also include firm fixed-effects in all my regressions. In addition, the analysts also factor in all other public as well as private information, including voluntary disclosures by firms, even after the end of the current quarter (but before the earnings announcement) while making their earnings and sales forecast for the quarter. Many accounting studies consider analyst consensus forecast as representative of the market expectation. This suggests that analysts may be using the incremental information in the volume of a firm's tweets (engagement) to make more accurate earnings and sales forecasts which helps reduce the firm's unexpected component of earnings and sales growth.

Overall, I find strong evidence in support of my hypotheses 1A, 1B and 2A. Results of Tables 2 and 3 provide strong evidence that the aggregate information in firm-initiated tweets and followers' engagement in a given period is value relevant to the market participants and gets reflected in the stock prices during the same period. This means that the tweet (engagement) volume should be value relevant and contribute positively to the firm value. In addition, results of Tables 4, 5 and 6 suggest that analysts use the information in the volume of a firm's tweets and its followers' engagement to make more accurate forecasts which reduces the unexpected earnings and unexpected sales growth of the firm. Therefore, the volume of a firm's tweets (engagement) is incrementally informative to the market participants and financial intermediaries beyond other known sources of concurrent information such as press releases, newspaper coverage, and voluntary disclosures.

5.3 Additional Tests

The results in Sections 5.1 and 5.2 suggest that the aggregate information in the volume of Firm-initiated tweets (engagement) gets priced by the capital market concurrently and is also informative to the financial intermediaries. A key question to then ask is "What is the information in these tweets and engagement which the market participants and analysts find valuable?" In this section, I attempt to provide an answer to this question. Marketing studies have shown that firms use their presence on social media for brand building, marketing campaigns and sales promotions, in addition to their traditional marketing activities (Trusov, Bucklin and Pauwels, 2009; Erdogamus and Cicek, 2012). Initially, firms focused on acquiring more followers. But they soon realized that the response or buzz they are able to generate from the followers is a more important measure of the effectiveness of their social media marketing activities and started adopting new and innovative strategies and techniques to leverage social media for stimulating customer engagement and demand (Schniederjans, Cao and Schniedarjans, 2013; Rishika, Kumar, Janakiraman and Bezawada, 2013; Gong, Zhang, Zhao and Jiang, 2017; Lee, Hosanagar and Nair, 2018).

Therefore, the volume of a firm's tweets and the engagement level of the followers aggregated over a given period may convey incremental information about the firm's sales, sales growth, and earnings during that period; a high level of tweets and engagement is associated with higher earnings and sales. Similarly, a change in the level of tweets and engagement over a period may be a leading indicator of the sales growth to be expected during that period. However, it is inconclusive whether and how tweeting influences product demand and sales (Gong et al., 2017). Additionally, all firms may not have the same ability to harness the power of social media for increasing demand for their products and services or the followers may not be representative of the customer base of the firm. Therefore, ex-ante it is not clear whether the aggregate level of tweeting and engagement is informative about the likely business performance of the firm during the period – sales, sales growth, and earnings.

[Insert Table 7 here]

Table 7 shows the results of testing the association between earnings and tweet (engagement) volume. Columns 1-3 have the results for the full sample The coefficients of *LOG (TWEETS)*, *LOG (ENGAGEMENT)* and *RESPONSE* are positive and significant at 1% level in the first three columns and the magnitude is also 54%, 54% and 125% of the coefficient of *LOG (PRESSRELEASES)* and multiple times the coefficient of *LOG (NEWSPAPERS)*, both of which proxy for other information sources about the firm's performance. This suggests that use of Twitter by firms is indicative of the actual earnings during the quarter, as the reference group in the full sample is all non-Tweet firm-quarters (including firms which never join Twitter during the sample period). The coefficient of *LOG (NUM_ANALYSTS)* is insignificant in all three columns. The results are similar in Columns 4-6 which uses a subsample of only Tweet firm-quarters for the analysis. The only difference is that the coefficient of *RESPONSE* becomes insignificant in Column 6. This suggests that tweets and engagement aggregated at the quarterly level convey some information about the likely actual earnings during the quarter.

[Insert Tables 8 A and B here]

The strong association between volume of tweets/engagement and earnings could be because firms with a Twitter account may be using that platform to communicate and engage[34] with their followers

[34] The response by followers of a firm's Twitter account is only a lower bound of the interest shown by potential customers and other stakeholders. Indeed, one of the strong features of Twitter, and other social media platforms, is that when someone retweets, that person's followers can also view the 'shared' tweet and respond to it, thereby increasing the visibility of the tweet. It is very difficult to get the complete reach of a 'shared' tweet and, therefore,

about their products and services, sales promotions, new product offerings and also to respond to any customer service issues or queries which customers might raise (Lee et al., 2018; Gong et al., 2018). I next test the relationship between tweet/engagement volume, and sales and sales growth using the full sample in Table 8. *LOG (SALES)* is the dependent variable in Columns 1-3 and *SALES_GROWTH* in Columns 3-6. I include advertising expense scaled by assets, previous quarter's deferred revenue, previous quarter's sales and previous quarter's sales growth as additional controls (Tang, 2018) because these variables might also affect current quarter's sales and, hence, the sales growth. The coefficients of all the measures of our Variables of Interest are positive and significant at 1% level (except the coefficient of *RESPONSE* which is significant at 5% level). The results are qualitatively similar if I use a subsample of only Tweet firm-quarters for the test; the only difference being that the coefficient of *RESPONSE* becomes insignificant. This indicates a strong positive association between the volume of tweets (engagement) and sales and change in volume of tweets (engagement) and sales growth of the firm.

Collectively, Tables 7 and 8 provide strong evidence that the tweet (engagement) volume of a firm is informative about the likely earnings, sales and sales growth of the firm during the period. This implies that a firm's tweet (engagement) volume is a leading indicator of its business performance during that period. It is this predictive ability of the tweet (engagement) volume for earnings, sales and sales growth during that period which the capital market participants may be incorporating into stock prices contemporaneously.

5.4 Robustness Tests

In this section, I perform a series of robustness tests to check whether the results are sensitive to using different samples, specifications and alternative measures.

---

measuring engagement using the likes, retweets and replies of followers of the firm's Twitter account is only a lower bound on the complete reach and total response generated for a given tweet.

[Insert Table 9 A and B here]

First, I apply Model 2 to a subsample of the full sample which excludes firms which have never created a Twitter account. This allows me to observe the informational effect of tweets and engagement on the Tweet firm relative to when that firm did not have a Twitter account. Columns 1 and 2 in Panel A Table 9 show the results for FF three-factor model and Columns 3 and 4 for FF five-factor model. The coefficients of *CHANGE_LOG (TWEETS) _MON* and *CHANGE_LOG (ENGAGE) _MON* are positive and significant in all columns.

Thus far, I have only used firm-initiated tweets from the Primary Twitter account of the firm. However, some firms may, in addition to the Primary Twitter account, have other Twitter accounts catering to specific geographies, business segments or functions. Therefore, I repeat the analysis using tweets and engagement from all the Twitter – Primary and Secondary[35] – accounts of the firms. The results of testing Hypothesis 1B using this expanded tweet sample are shown in Panel B of Table 9. The positive and significant coefficients of *CHANGE_LOG (TWEETS) _MON* and *CHANGE_LOG (ENGAGE) _MON* again suggest that changes in tweet (engagement) volume are incrementally informative to the market participants and are priced by them in the same period.

[Insert Table 10 here]

There might be differences in tweet volume by firms across industries driven by the type of products sold, type of consumers as well as how other peer firms in the industry utilize Twitter. Therefore, I compute another measure of tweet and engagement volume after adjusting for the median SIC 2-digit tweet/engagement volume. The results for earnings and unexpected earnings are displayed in Table 10 for the subsample of only Tweet firm-quarters and are broadly consistent with the earlier results.

[35] 195 firms in my sample have 1,209 Secondary accounts in addition to having a Primary account. All Twitter accounts – Primary and Secondary – have approximately 29 million tweets, 223 million likes, 140 million retweets and 21 million replies.

I repeat the analysis for unexpected earnings and unexpected sales growth using industry fixed-effects (Fama-French 48 or SIC 2-digit) instead of firm fixed-effects and find even stronger (untabulated) results than before. I also use an alternative measure of unexpected earnings computed as the difference between current and previous quarter's earnings scaled by the share price at the end of the previous quarter and find qualitatively similar results. This shows that the findings are not sensitive to different samples, specifications and alternative measures of variables of interest or dependent variables.

6. Conclusion

Twitter has, arguably, emerged as the most popular social media for dissemination of information by firms. In this paper, I study the aggregate information in firms' tweets and followers' engagement on their official Twitter accounts. Specifically, I examine whether the volume of a firm's tweets and its followers' engagement is informative to capital market participants and financial intermediaries, namely investors and analysts. I find results broadly consistent with my hypotheses which suggests that the firm's tweet/engagement volume and a change in the firm's tweet/engagement volume convey incremental informative to investors and analysts over and above the information contained in other known sources of information such as press releases, newspaper coverage, analyst forecast and voluntary disclosures. The results also suggest that the firm's tweet (engagement) volume helps predict the firm's earnings, sales, and sales growth in a given period. It is conceivable that the investors find this information valuable which they price into stock prices concurrently. Similarly, analysts may be factoring in this predictive ability of the information in the volume of a firm's tweets (engagement) to make more accurate forecasts. This is an important finding which should be useful for managers, investors as well as financial analysts. However, I don't make any claims of causation as there may be an unobservable omitted correlated variable influencing both the tweet volume and the firm's financial performance. Also, Twitter may only represent a subset of the dissemination effort by a firm on social media - most firms have a presence on other social media too such as Facebook, YouTube, and Instagram, etc.

Appendix A

Variables Description

Dependent Variables	
EPS	Diluted earnings per share excluding extraordinary items at the end of the current quarter
LOG(SALES)	Natural log of firm's sales in the current quarter
MON_EXCESS_RETURN	Excess of the firm's monthly stock return over minus the 1-month T-bill rate ($R_{i,t}$ – Rf_t)
SALES_GROWTH	(Sales for the current quarter / Sales for the previous) -1
TOBIN'S Q	Market value of assets/book value of assets=(Book value of assets + Market value of Common Stock - Book value of Common Stock)/Book Value of Assets
UE_EARNINGS	Actual EPS for the quarter (reported in I/B/E/S) minus the most recent consensus analyst forecast EPS for the current quarter, divided by the previous quarter-end stock price
UE_SALES_GR	Actual quarterly sales (reported in I/B/E/S) minus the most recent consensus analyst sales forecast for the current quarter, divided by the previous quarter's sales for each firm-quarter.
Variables of Interest	
CHANGE_LOG(ENGAGEMENT)	Log(Engagement) of the current quarter minus Log(Engagement) of the previous quarter
CHANGE_LOG(ENGAGEMENT)_MON	Log(Engagement) of the current month minus Log(Engagement) of the previous month
CHANGE_LOG(TWEETS)	Log(Tweets) of the current quarter minus Log(Tweets) of the previous quarter
CHANGE_LOG(TWEETS)_MON	Log(Tweets) of the current month minus Log(Tweets) of the previous month
LOG(ENGAGEMENT)	Natural log of the sum of total retweets, total likes and total replies by followers of a firm's Twitter account in the current quarter
LOG(TWEETS)	Natural log of the sum of all tweets by a firm in the current quarter
LOG(ENGAGEMENT)_MED_IND_ADJ	Log(Engagement) of the current quarter for a firm adjusted for the same quarter median value of the SIC 2-digit industry to which the firm belongs
LOG(TWEETS)_MED_IND_ADJ	Log(Tweets) of the current quarter for a firm adjusted for the same quarter median value of the SIC 2-digit industry to which the firm belongs

RESPONSE	Log(Engagement) divided by Log(Tweets) for each Tweet firm-quarter
Control Variables	
ACQUISITION	1 if the firm makes an acquisition during the quarter , and 0 otherwise
ADV_EXP_QTR	Annual advertising expense divided equally over the four quarters and scaled by average total assets of the quarter
CMA	Conservative Minus Aggressive (CMA from Fama-French Factors – Monthly Frequency)
HML	High minus low. (HML from Fama-French Factors – Monthly Frequency)
LEVERAGE	Sum of long-term debt and debt in current liabilities scaled by total assets of the firm at the end of the current quarter
LOG(ASSETS)	Natural log of the firm’s total assets at the end of the current quarter
LOG(PRESSRELEASES)	Natural Log of one plus the number of press releases issued by the firm and distributed via a news provider during the quarter.
LOG(NEWSPAPERS)	Natural Log of one plus the number of news articles written about a firm during the quarter.
LOG(NUM_ANALYSTS)	Natural Log of one plus number of analysts following (from IBES database) during the quarter
LOSS	1 if net income for the quarter is negative, and 0 otherwise
MKT_RET_QTR	Equally-weighted market return for the current quarter
MKT RF	Excess return on the market (MKTRF from Fama-French Factors – Monthly Frequency)
MOM	Up minus down. (MOM from Fama-French Factors – Monthly Frequency)
MTB	Ratio of market value of equity to book value of equity at the end of the current quarter
NEG_UE_EARNINGS	1 if unexpected earnings (UE) for the quarter is negative, and 0 otherwise
NEG_UE_SALES_GR	1 if unexpected sales growth (*U_SALES_GR*) for the quarter is negative, and 0 otherwise
RMW	Robust Minus Weak (RMW from Fama-French Factors – Monthly Frequency)
ROA	Net Income in the current quarter scaled by average assets of the firm at the end of the current and previous quarters
SMB	Small minus big. (SMB from Fama-French Factors – Monthly Frequency)

References

Amir, E., and B. Lev.1996. Value-Relevance of Nonfinancial Information: The Wireless Communications Industry. *Journal of Accounting and Economics* 22: 3–30.

Ball, R., and P. Brown. 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6(2): 159-178.

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

Bernard, V. L., and J. Thomas. 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics* 13(4): 305-340.

Blankespoor, E., G.S. Miller, and H.D. White. 2014. The role of dissemination in market liquidity: Evidence from firms' use of Twitter. *The Accounting Review* 89 (1): 79–112.

Bloomfield, R. 2002. The incomplete revelation hypothesis and financial reporting. *Accounting Horizons* 16 (3): 233–243.

Bollen, J., H. Mao, and X. Zheng. 2011. Twitter mood predicts the stock market. *Journal of Computational Science* 2: 1–8.

Bowen, R., Davis, A., & Matsumoto, D. 2002. Do conference calls affect analysts' forecasts? *The Accounting Review* 77 (2):285−316.

Bradshaw, M. 2004. How do analysts use their earnings forecasts in generating stock recommendations? *The Accounting Review* 79(1): 25−50

Brown, L. 1993. Earnings forecasting research: its implications for capital markets research. *International Journal of Forecasting* 9(3): 295−320.

Brown L., A. Call, M. Clement, and A. Sharp. 2015. Inside the "Black Box" of Sell-Side Financial Analysts. *Journal of Accounting Research* 53(1):1-47.

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

Chen, X., Q. Cheng, and K. Lo. 2010. On the relationship between analyst reports and corporate disclosures: Exploring the roles of information discovery and interpretation. *Journal of Accounting and Economics* 49 (3): 206–226.

M. Coller, and T. Yohn.1997.Management forecasts and information asymmetry: an examination of bid-ask spreads. Journal of Accounting Research 35 (2): 181-191.

Crowley, R., W. Huang, and H. Lu. 2018. Discretionary Dissemination on Twitter. *Working paper* (SSRN).

Curtis, A., V. Richardson, and R. Schmardebeck. 2016. Investor attention and the pricing of earnings news. *Handbook of Sentiment Analysis in Finance* (Chapter 8): 212–232.

Deng, Z, B. Lev, and F. Narin. 1999. Science and Technology as Predictors of Stock Performance. *Financial Analysts Journal* 55(3): 20–32.

Erdogamus, I.E., and M. Cicek. 2012. The impact of social media marketing on brand loyalty. *Social and Behavioral Sciences* 58(2): 1353-1360.

Fama, E. 1970. Efficient capital markets: A review of theory and empirical work. *Journal of Finance* 25: 383–-417.

Fama, E., and K. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33: 3–56.

Fama, E., and K. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116(1): 1 – 22.

Fama, E., and K. French.2015. Dissecting anomalies with a five-factor model. *Review of Financial Studies* 29(1): 69 – 103.

Fama, E., and J. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81(3): 607–636.

Francis, J., K. Schipper, and L. Vincent. 2002. Earnings announcements and competing information. *Journal of Accounting and Economics* 33 (3): 313–343.

Francis, J., and L. Soffer. 1997. The relative informativeness of analysts' stock recommendations and earnings forecast revisions. *Journal of Accounting Research* 35 (2): 193–212.

Frankel, R., S. Kothari, and J. Weber. 2006. Determinants of the informativeness of analyst research. *Journal of Accounting and Economics* 41(1): 29−54.

Givoly, D. and J. Lakonishok. 1979. The information content of financial analysts forecasts of earnings: some evidence on semi-strong efficiency. *Journal of Accounting and Economics* 1: 165–186.

Gong, S., J. Zhang, and P. Zhao. 2017. Tweeting as a Marketing Tool: A Field Experiment in the TV Industry. *Journal of Marketing Research* 54(6): 833-850.

Hales, J.2007. Directional Preferences, Information Processing, and Investors' Forecasts of Earnings. *Journal of Accounting Research* 45 (3): 607–628.

Hayes, R., & Levine, C. 2000. An approach to adjusting analysts' consensus forecasts for selection bias. *Contemporary Accounting Research* 17(1): 61−83.

Healy, P.M., and K. G. Palepu.2001. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics* 31(1-3): 405–440.

Hilary, G., and C. Hsu. 2013. Analyst Forecast Consistency. *Journal of Finance* 68(1): 271-297.

Hirshleifer, D., S.S. Lim, S.H. Teoh.2002. Limited Investor Attention and Stock Market Misreactions to Accounting Information. *The Review of Asset Pricing Studies* 1(1): 35-73.

Hirshleifer, D., S.S. Lim, S.H. Teoh.2003. Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics* 36: 337-386.

Hodder, L., P. E. Hopkins, and D. A. Wood. 2008. The Effects of Financial Statement and Informational Complexity on Analysts' Cash Flow Forecasts. *The Accounting Review* 83(4): 915-956.

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

Hughes, K. E.2000. The Value Relevance of Nonfinancial Measures of Air Pollution in the Electric Utility Industry. *The Accounting Review* 75 (2): 209–28.

Imhoff, E. A., and G. J. Lobo. 1992. The effect of ex-ante earnings uncertainty on earnings response coefficients. *The Accounting Review* 67 (2): 427–439.

Ittner, C., and D. Larcker. 1998. Are Nonfinancial Measures Leading Indicators of Financial Performance? An Analysis of Customer Satisfaction. *Journal of Accounting Research* 36: 1–35.

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

Jung, M., J. Naughton, A. Tahoun, and C. Wang.2018. Do firms strategically disseminate? Evidence from corporate use of social media. *The Accounting Review* 93(4): 225-252.

Lee, D., K. Hosanagar, and H. Nair. 2018. Advertising content and consumer engagement on social media: Evidence from Facebook. *Management Science* 64(11): 5105-5131.

Lee, L.F., A.P. Hutton, and S. Shu. 2015. The role of social media in the capital market: Evidence from consumer product recalls. *Journal of Accounting Research* 53 (2): 367–404.

Lehavy, R., F. Li, and K. Merkley. 2011. The Effect of Annual Report Readability on Analyst Following and the Properties of Their Earnings Forecasts. *The Accounting Review* 86(3): 1087-1115.

Lobo, G., M. Song, M.H. Stanford.2017. The Effect of Analyst Forecasts during Earnings Announcements on Investor Responses to Reported. *The Accounting Review* 92(3): 239-263.

Luo, X., and J. Zhang. 2013. How Do Consumer Buzz and Traffic in Social Media Marketing Predict the Value of the Firm? *Journal of Management Information Systems* 30(2): 213-238.

Luo, X., J. Zhang, and W. Duan. 2013. Social Media and Firm Equity Value. *Information Systems Research* 24(1): 146-163.

Lys, T., and S. Sohn, 1990. The association between revisions of financial analysts' earnings forecasts and security price changes. *Journal of Accounting and Economics* 13(4): 341–364.

Mao, Y., W. Wei, B. Wang, and B. Liu. 2012. Correlating S&P 500 stocks with Twitter data. *Proceedings of the First ACM International Workshop on Hot Topics on Interdisciplinary Social Networks*: 69–72.

Matsumoto, D. A., M. Pronk, and E. Roelofsen.2011.What Makes Conference Calls Useful? The Information Content of Managers' Presentations and Analysts' Discussion Sessions. *The Accounting Review* 86 (4): 1383–414.

Miller, G., and D. Skinner.2015. The Evolving Disclosure Landscape: How Changes in Technology, the Media, and Capital Markets Are Affecting Disclosure. *Journal of Accounting Research* 53(2): 221-239.

Newey, W., and K. West. 1987. A simple positive-definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703–708.

O'Brien, P. C. 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics* 10 (1): 53–83.

Plumlee, M. A. 2003. The effect of information complexity on analysts' use of that information. *The Accounting Review* 78(1): 275-296.

Rajgopal, S., T. Shevlin, and M. Venkatachalam. 2003. Does the stock market fully appreciate the implications of leading indicators for future earnings? Evidence from order backlog. *Review of Accounting Studies* 8(4): 461-492.

Rajgopal, S., M. Venkatachalam, and S. Kotha. 2003. The Value Relevance of Network Advantages: The Case of E-commerce Firms. *Journal of Accounting Research* 41(1): 135–162.

Rishika, R., A. Kumar, R. Janakiraman, and R. Bezawada.2013. The Effect of Customers' Social Media Participation on Customer Visit Frequency and Profitability: An Empirical Investigation. *Information Systems Research* 24(1): 108-127.

Rogers, J.L., D.J. Skinner, and S.L.C. Zechman. 2016. The role of the media in disseminating insider-trading news. *Review of Accounting Studies* 21 (3): 711-739.

Rui, H., Y. Liu, and A. Whinston.2013. Whose and what chatter matters? The effect of tweets on movie sales. *Decision Support Systems* 55 (4): 863-870.

Schniederjans, D., E.S. Cao, and M. Schniederjans. 2013. Enhancing financial performance with social media: An impression management perspective. *Decision Support Systems* 55(4): 911-918.

Sloan, R. G. 1996. Do stock prices fully reflect the information in accruals and cash flows about future earnings? *The Accounting Review* 71(3): 289–315.

Tang, V.W.2018. Wisdom of Crowds: Cross-Sectional Variation in the Informativeness of Third-Party-Generated Product Information on Twitter. *Journal of Accounting Research* 56 (3): 989–1034.

Trueman, B., F. Wong and X. Zhang. 2001. The Eyeballs Have It: Searching for the Value in Internet Stocks. *Journal of Accounting Research* 38: 137–162.

Trusov, M., R.E. Bucklin, and K. Pauwels.2009. Effects of Word-of-Mouth Versus Traditional Marketing: Findings from an Internet Social Networking Site. *Journal of Marketing* 73(5): 90-102.

Xie, H. 2001. The mispricing of abnormal accruals. *The Accounting Review* 76(3): 357–374.

Yu, Y., W. Duan, and Q. Cao.2013.The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision Support Systems* 55(4):919-926.

Zhang, Y. 2008. Analyst responsiveness and the post-earnings-announcement drift. *Journal of Accounting and Economics* 46 (1): 201–215

Figure 1: Aggregation of Information: Firm's Tweets and Engagement of Followers

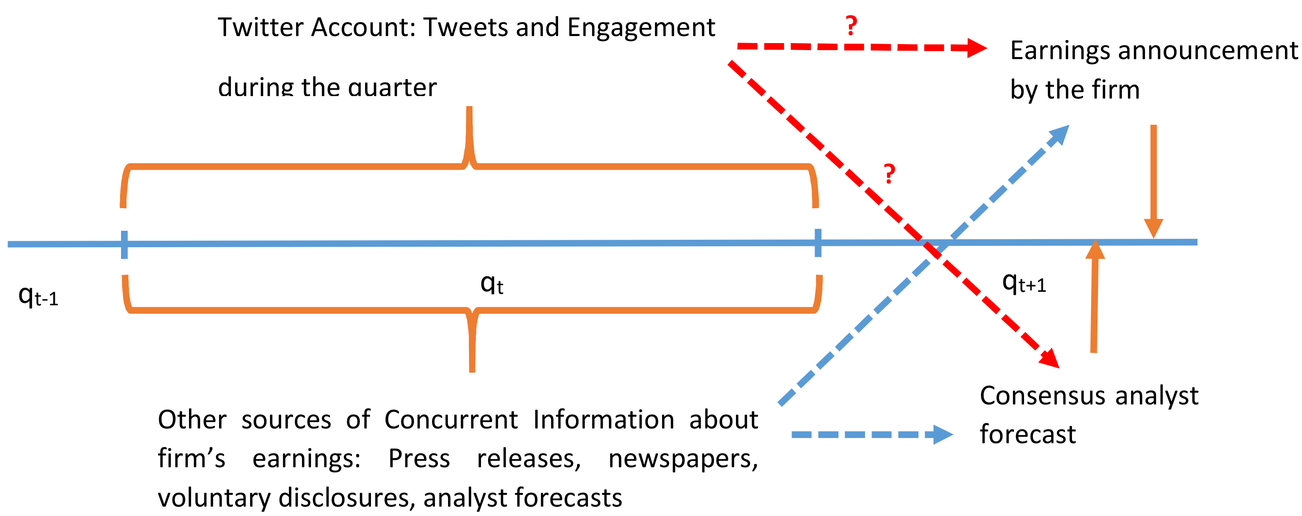

Figure 2: Time Trend of Proportion of Firms which Tweet

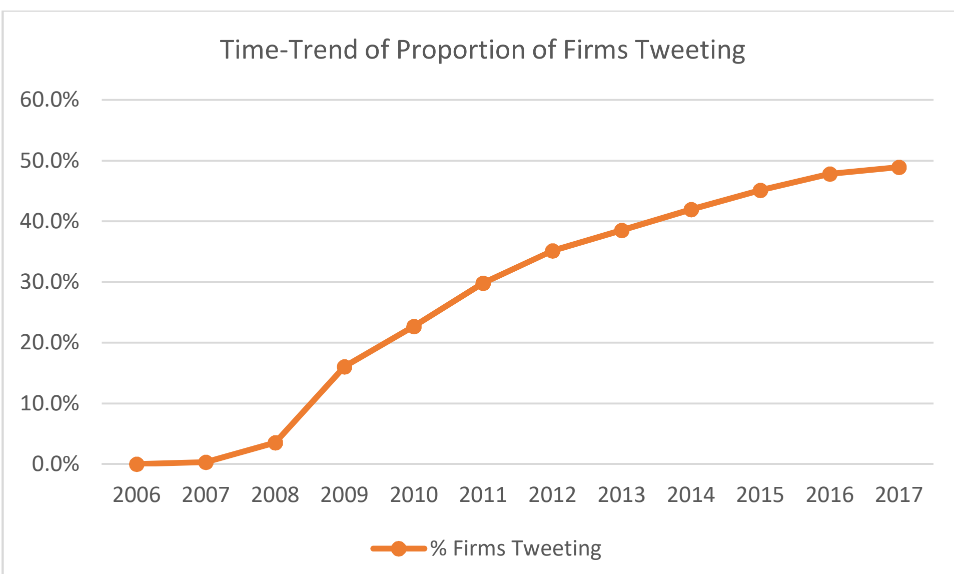

Figure 3: Time Trend of Tweets and Engagement: Fama-French Ten Industry Classification

Figure 3A: Time- Trend of Total Tweets

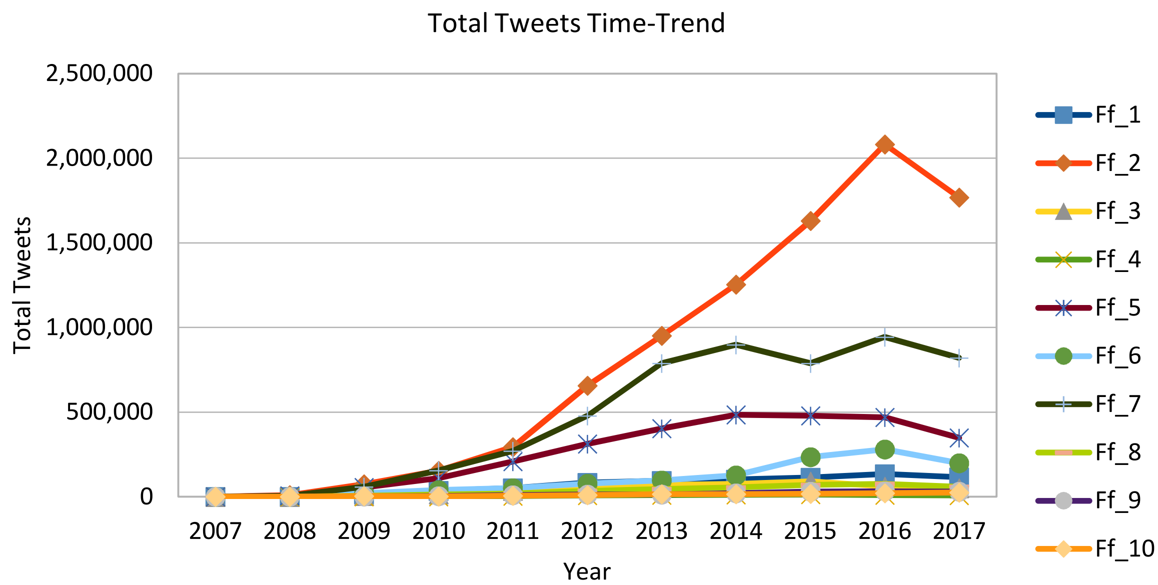

Figure 3B: Time- Trend of Average Tweets

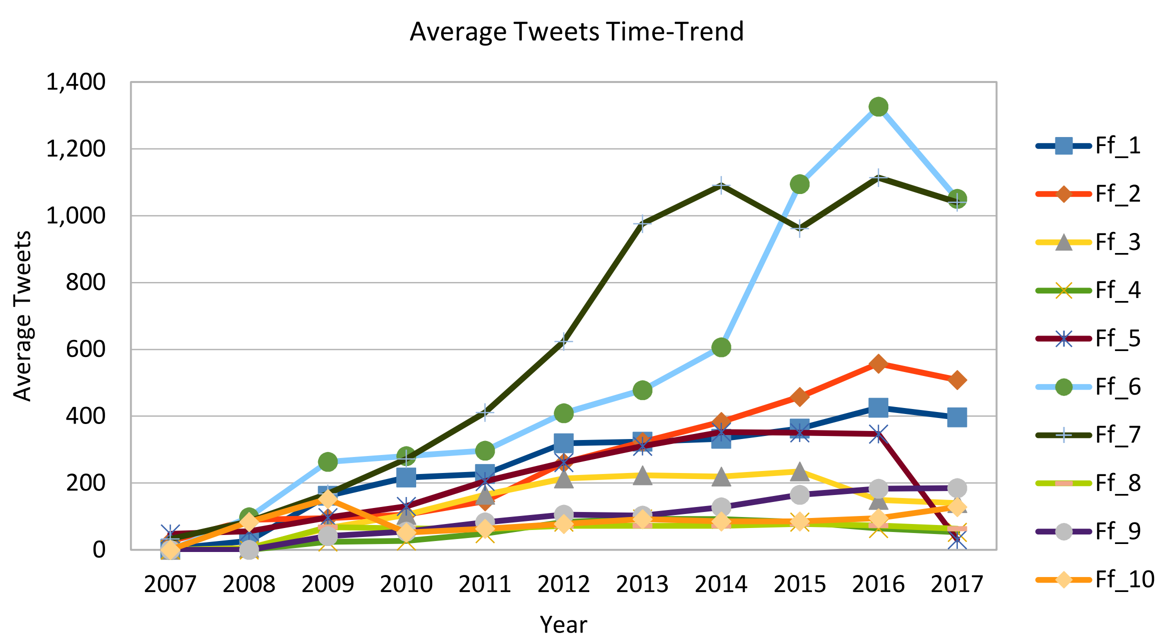

Figure 3C: Time- Trend of Total Engagement

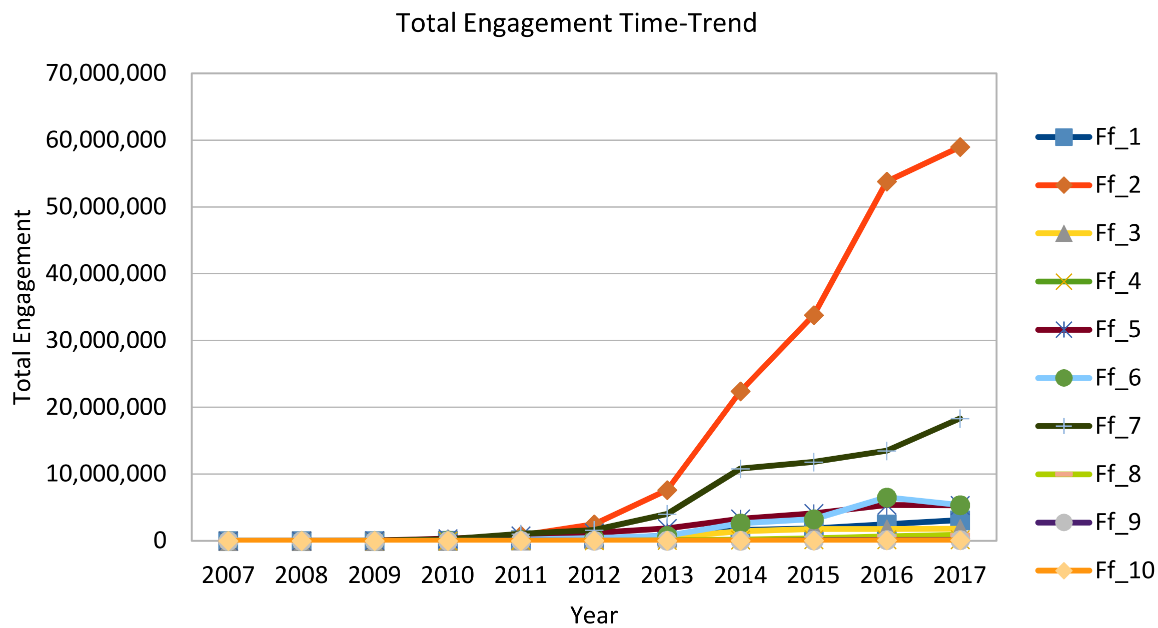

Figure 3D: Time Trend of Average Engagement

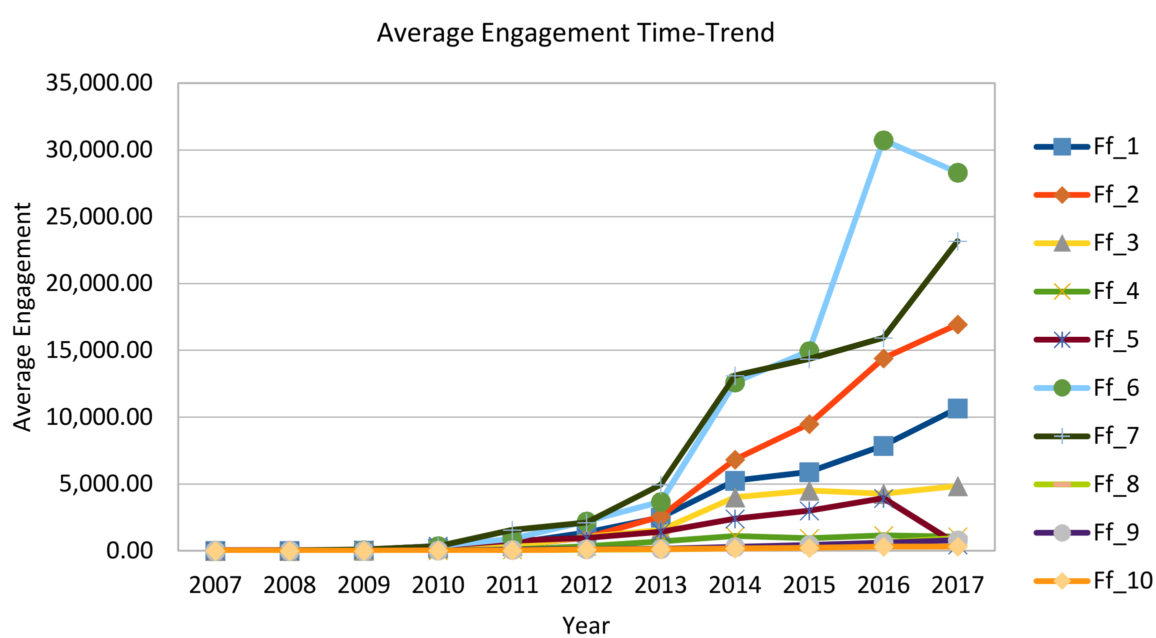

Table 1: Descriptive Statistics of Key Variables

Panel A: All Firm-quarters

Variables	**# Firm-quarters**	**Mean**	**Median**	**Std. Dev.**	**P25**	**P75**
ASSETS	178,236	5760.193	678.501	20118.570	144.803	2971.019
MTB	178,236	3.050	1.984	5.718	1.159	3.569
LEVERAGE	178,236	0.643	0.230	2.091	0.000	0.795
NUMBER_ANALYSTS	178,236	6.182	3.000	7.761	0.000	10.000
EPS	178,236	0.212	0.130	0.704	-0.060	0.450
UE_EARNINGS	128,903	0.000	0.010	0.267	-0.030	0.056
SALES	178,236	764.535	112.205	2191.382	21.771	476.562
SALES_GROWTH	172,040	0.051	0.019	0.322	-0.055	0.100
UE_SALES_GR	121,163	0.007	0.005	0.123	-0.023	0.037
MARKET VALUE	178,236	4005.689	654.268	10644.220	155.915	2598.457
NEWSPAPERS	178,236	5.609	0.000	23.032	0.000	3.000
PRESS_RELEASES	178,236	30.254	1.000	205.662	0.000	8.000
STOCK_RET_QTR	121,421	0.028	0.015	0.247	-0.112	0.143
MARKET_RET_QTR	178,236	0.023	0.037	0.078	-0.007	0.062

Panel B: Only Tweet Firm-quarters

Variables	**# Tweet Firm-quarters**	**Mean**	**Median**	**Std. Dev.**	**P25**	**P75**
TWEETS	46,559	374.735	71.000	2555.022	16.000	231.000
ENGAGEMENT	46,559	6149.960	82.000	8772.134	9.000	534.000
LOG(TWEETS)	46,559	4.009	4.263	2.001	2.773	5.442
LOG(ENGAGEMENT)	46,559	4.343	4.407	2.835	2.197	6.280
CHANGE_TWEETS	46,559	8.170	0.000	230.214	-18.000	23.000
CHANGE_ENGAGEMENT	46,559	138.223	2.000	3383.682	-13.000	57.000
CHANGE_LOG(TWEETS)	46,559	0.045	0.000	0.705	-0.234	0.284
CHANGE_LOG(ENGAGEMENT)	46,559	0.130	0.000	0.775	-0.176	0.437
RESPONSE	46,559	0.962	1.042	0.555	0.707	1.284
NUMBER_ANALYSTS	46,559	9.293	7.000	9.207	1.000	15.000
EPS	46,559	0.288	0.180	0.722	-0.040	0.570
UE_EARNINGS	34,211	0.000	0.001	0.027	-0.001	0.002
SALES	46,559	1341.393	240.407	3062.878	51.446	997.207
SALES_GROWTH	45,505	0.044	0.020	0.266	-0.043	0.090
UE_SALES_GR	37,972	0.007	0.005	0.100	-0.016	0.031
MARKET VALUE	46,559	7255.058	1329.581	15394.990	283.764	5603.612
NEWSPAPERS	46,559	13.808	2.000	38.592	0.000	12.000
PRESS_RELEASES	46,559	57.675	3.000	297.968	0.000	17.000
STOCK_RET_QTR	35,064	0.038	0.028	0.220	-0.084	0.142
TOBIN'S Q	46,559	2.272	1.664	1.825	1.223	2.568

Panel C: Time-Trend of Tweets and Engagement

Year	# Tweet Firm-quarters	Total Tweets	Average Tweets	Total Engagement	Average Engagement
2006	0	0	0	0	0
2007	26	1,075	41.35	795	30.58
2008	296	22,344	75.49	9,190	31.05
2009	1,765	214,696	121.64	116,921	66.24
2010	3,083	442,882	143.65	736,757	238.97
2011	4,008	780,994	194.86	2,620,917	653.92
2012	4,864	1,424,146	292.79	5,471,190	1,124.83
2013	5,484	1,998,715	364.46	13,902,839	2,535.16
2014	6,205	2,619,061	422.09	39,947,510	6,447.24
2015	6,896	3,082,487	447.00	54,617,596	7,920.84
2016	7,352	3,767,680	512.47	79,750,502	10,849.06
2017	6,580	3,093,196	470.09	89,067,354	13,539.18
Total	**46,559**	**17,447,276**	**374.73**	**286,241,571**	**6,149.96**

Panel D: Pearson Correlation between Key Variables for Tweet Firm-quarters

	LOG(TWEETS)	*LOG(ENGAGEMENT)*	*CHANGE_LOG(TWEETS)*	*CHANGE_LOG(ENGAGE)*	*RESPONSE*	*LOG(PRESS RELEASES)*	*LOG(NEWSPAPERS)*
LOG(TWEETS)	1						
LOG(ENGAGEMENT)	0.839***	1					
CHANGE_LOG(TWEETS)	0.166***	0.050***	1				
CHANGE_LOG(ENGAGE)	0.127***	0.125***	0.571***	1			
RESPONSE	0.413***	0.714***	-0.043***	0.107***	1		
LOG(PRESSRELEASES)	0.167***	0.340***	-0.052***	-0.047***	0.329***	1	
LOG(NEWSPAPERS)	0.291***	0.374***	0.021***	0.042***	0.233***	0.224***	1
EPS	0.127***	0.169***	0.0123*	0.017**	0.117***	0.119***	0.289***
UE_EARNINGS	0.047***	0.056***	0.002	0	0.034***	0.036***	0.055***
LOG(SALES)	0.275***	0.344***	0.017**	0.027***	0.227***	0.211***	0.610***
SALES_GROWTH	-0.006	-0.002	0.040***	0.038***	0.005	-0.009	-0.038***
UE_SALES_GR	0.011	0.005	0.013*	0.012*	-0.009	-0.001	-0.001
LOG(MARKET_VALUE)	0.272***	0.372***	0.014*	0.028***	0.291***	0.300***	0.565***
STOCK_RET_QTR	-0.016***	-0.033***	0.022***	0.028***	-0.027***	0.025***	0.020***

	EPS	*UE_EARNINGS*	*LOG(SALES)*	*SALES_GROWTH*	*UE_SALES_GR*	*LOG(MARKET_VALUE)*	*STOCK_RET_QTR*
EPS	1						
UE_EARNINGS	0.287***	1					
LOG(SALES)	0.481***	0.097***	1				
SALES_GROWTH	0.095***	0.067***	-0.038***	1			
UE_SALES_GR	0.050***	0.202***	0.035***	0.314***	1		
LOG(MARKET_VALUE)	0.444***	0.115***	0.792***	-0.032***	0.034***	1	
STOCK_RET_QTR	0.046***	0.067***	0.030***	0.001	0.086***	0.092***	1

Panel A shows the descriptive statistics of key variables for the full sample comprising of Tweet as well as all Non-tweet firm-quarters

Panel B shows the descriptive statistics of key variables for the sub-sample comprising of Tweet firm-quarters only (firms which have a Twitter account and have started tweeting)

Panel C shows the time- trend of Tweet Firm-quarters, Total Tweets, Average Tweets, Total Engagement, and Average Engagement from 2006 to 2017.

Panel D shows the Pearson Coefficient between the Dependent Variables and Variables of Interest for the sub-sample comprising of Tweet firm-quarters only (firms which have a Twitter account and have started tweeting) All variables are as defined in Appendix A

*** p<0.01, ** p<0.05, * p<0.1

Table 2: Association between Firm Value and Tweet/Engagement Volume – Only Tweet Firm-quarters

	Dependent Variable = *TOBIN'S'Q*						
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LOG(TWEETS)	**0.016***			0			
	(1.697)			(0.029)			
LOG(ENGAGEMENT)		**0.021****		**0.021***			
		(2.072)		**(1.833)**			
RESPONSE			**0.028***				
			(1.826)				
CHANGE_LOG(TWEETS)					0		-0.008
					(0.05)		(-1.547)
CHANGE_LOG(ENGAGE)						**0.009***	**0.013*****
						(1.833)	**(3.303)**
LOG(ASSETS)	-0.867***	-0.869***	-0.863***	-0.869***	-0.862***	-0.862***	-0.862***
	(-7.231)	(-7.186)	(-7.268)	(-7.183)	(-7.289)	(-7.294)	(-7.293)
ROA	-1.465	-1.465	-1.47	-1.465	-1.468	-1.47	-1.468
	(-1.506)	(-1.506)	(-1.511)	(-1.506)	(-1.511)	(-1.511)	(-1.509)
ROA_1	0.547	0.545	0.546	0.545	0.548	0.549	0.548
	(0.87)	(0.866)	(0.87)	(0.865)	(0.873)	(0.875)	(0.874)
LOG(PRESSRELEASES)	0.170***	0.168***	0.170***	0.168***	0.171***	0.171***	0.171***
	(4.047)	(4.082)	(4.025)	(4.085)	(4.037)	(4.044)	(4.045)
LOG(NEWSPAPERS)	0.395***	0.393***	0.396***	0.393***	0.396***	0.396***	0.396***
	(11.356)	(11.361)	(11.417)	(11.357)	(11.416)	(11.412)	(11.398)
LOSS	-0.263***	-0.263***	-0.264***	-0.263***	-0.264***	-0.263***	-0.263***
	(-4.512)	(-4.515)	(-4.523)	(-4.508)	(-4.518)	(-4.516)	(-4.517)
ACQUISITION	-0.052**	-0.052**	-0.053**	-0.052**	-0.053**	-0.053**	-0.053**
	(-2.329)	(-2.327)	(-2.369)	(-2.329)	(-2.371)	(-2.378)	(-2.374)
CONSTANT	8.033***	8.105***	8.014***	8.105***	8.001***	8.003***	8.004***
	(9.591)	(9.367)	(9.641)	(9.378)	(9.671)	(9.68)	(9.679)
Observations	46,559	46, 559	46, 559	46, 559	46, 559	46, 559	46, 559
R-squared	0.766	0.766	0.766	0.766	0.766	0.766	0.766
Year-qtr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustering of Errors	Industry	Industry	Industry	Industry	Industry	Industry	Industry

Table 2 shows the results of the association between tweet/engagement volume and the firm's market value.

Both panels display the results of OLS regression for a sub-sample of only tweet firm-quarters (includes only firms which have created a Twitter account and have started tweeting) for the sample period 2006 - 2017 using Model 1: $TOBIN'SQ_{i,t} = \beta_0 + \beta_1 TWEET\_VOLUME/CHANGE\_TWEET\_VOLUME_{i,t} + \Sigma\beta_j CONTROLS_{i,t} + YEAR\_QTR\ FIXED\ EFFECTS + FIRM\ FIXED\text{-}EFFECTS + \varepsilon_{i,t}$

Robust t statistics are in parentheses, *** p<0.01, ** p<0.05, * p<0.1; All variables are defined in Appendix A.

Table 3: Association between Stock Returns and Change in Tweet/Engagement Volume

Panel A: Stock Returns and Change in Tweet/Engagement Volume – Fama-French Three Factors

	Dependent Variable = *MON_EXCESS_RETURN*					
	All Firm-Months			**Only Tweet Firm-Months**		
	(1)	(2)	(3)	(4)	(5)	(6)
CHANGE_LOG(TWEETS)_MON	**0.004*** **		**0.001***	**0.003*** **		-0.001
	(4.94)		**(1.711)**	**(5.223)**		(-1.465)
CHANGE_LOG(ENGAGE)_MON		**0.004*** **	**0.003*** **		**0.002*** **	**0.002** **
		(4.279)	**(2.755)**		**(4.832)**	**(2.518)**
MKTRF	0.533**	0.532**	0.537**	0.524***	0.524***	0.534***
	(2.605)	(2.604)	(2.627)	(3.098)	(3.095)	(3.159)
SMB	0.021	0.022	0.02	0.184*	0.187*	0.184*
	(0.149)	(0.154)	(0.14)	(1.746)	(1.77)	(1.745)
HML	-0.175	-0.175	0.172	0.04	0.041	0.042
	(-0.898)	(-0.896)	(0.887)	(0.294)	(0.298)	(0.031)
MOM	-0.694*	-0.694*	-0.697*	-0.404*	-0.404*	-0.406*
	(-1.916)	(-1.916)	(-1.917)	(-1.716)	(-1.718)	(-1.709)
CONSTANT	-0.048***	-0.048***	-0.048***	-0.011***	-0.011***	-0.011***
	(-3.732)	(-3.733)	(-3.736)	(-3.543)	(-3.565)	(-3.6)
Observations	498,695	498,695	498,695	141,507	141,507	141,507
R-squared	0.089	0.089	0.089	0.072	0.072	0.074

Panel B: Stock Returns and Change in Tweet/Engagement Volume – Fama-French Five Factors

	Dependent Variable = *MON_EXCESS_RETURN*					
	All Firm-Months			Only Tweet Firm-Months		
	(1)	(2)	(3)	(4)	(5)	(6)
CHANGE_LOG(TWEETS)_MON	**0.004*** **		0.001	**0.002*** **		0.001
	(4.978)		(1.37)	**(4.957)**		(1.317)
CHANGE_LOG(ENGAGE)_MON		**0.004*** **	**0.003*** **		**0.002*** **	**0.002** **
		(4.362)	**(2.771)**		**(4.605)**	**(2.181)**
MKTRF	0.539**	0.539**	-0.545	0.468***	0.469***	0.474***
	(2.566)	(2.566)	(-2.592)	(2.777)	(2.779)	(2.825)
SMB	-0.003	-0.003	-0.004	0.185*	0.187*	0.187*
	(-0.021)	(-0.018)	(-0.028)	(1.77)	(1.788)	(1.803)
HML	-0.267	-0.267	-0.264	-0.054	-0.054	-0.05
	(-1.451)	(-1.449)	(-1.44)	(-0.408)	(-0.403)	(-0.373)
MOM	-0.473	-0.472	-0.478	-0.168	-0.167	-0.172
	(-1.522)	(-1.522)	(-1.532)	(-0.783)	(-0.781)	(-0.794)
RMW	-0.089	-0.089	-0.088	-0.049	-0.05	-0.052
	(-0.985)	(-0.986)	(-0.971)	(-0.609)	(-0.618)	(-0.651)
CMA	-0.122	-0.122	-0.12	-0.127	-0.128	-0.12
	(-1.212)	(-1.211)	(-1.19)	(-1.606)	(-1.608)	(-1.525)
CONSTANT	-0.045***	-0.045***	-0.045***	-0.011***	-0.011***	-0.011***
	(-3.736)	(-3.738)	(-3.739)	(-3.504)	(-3.525)	(-3.553)
Observations	486,539	486,539	486,539	140,363	140,363	140,363
R-squared	0.128	0.128	0.128	0.103	0.103	0.104

Panel A shows the results of the Fama-French three-factor model for testing the association between change in tweet (engagement) volume and the monthly excess stock returns.

Panel B shows the results of the Fama-French five-factor model for testing the association between change in tweet (engagement) volume and the monthly excess stock returns.

Both panels incorporate Momentum factor as well and display the results using Fama-MacBeth monthly cross-sectional regressions with Newey-West corrected standard errors for autocorrelation (two lags) used for calculating t-statistics. The reported slopes are computed as the time-series average of the slopes in monthly regressions of excess stock returns on the explanatory variables for the sample period 2006 - 2017 using Model 2: $(R_{i,t} - Rf_t) = \beta_0 + \beta_1 CHANGE\_LOG\ (TWEET)_{i,t}/CHANGE\_LOG\ (ENGAGE)_{i,t} + \beta_2 MOM_t + \Sigma\beta_J FAMA\text{-}FRENCH\_FACTORS_t + \varepsilon_{i,\ t}$

*** p<0.01, ** p<0.05, * p<0.1; All variables are defined in Appendix A.

Table 4: Unexpected Earnings and Tweet/Engagement Volume

Panel A: Association between Unexpected Earnings and Tweet/Engagement Volume - Portfolio Sorting on Unexpected Earnings

PORTFOLIO	# TWEET FIRM-QUARTERS	UE	LOG(TWEETS)	LOG(ENGAGEMENT)
1	6,843	-0.021	3.702	3.831
2	6,842	0.000	4.385	4.895
3	6,842	0.001	4.638	5.387
4	6,842	0.002	4.346	4.792
5	6,842	0.017	3.836	3.896
Total	34,211	0.000	4.181	4.560

Difference between LOG (TWEETS):
Portfolio (1) – (3): F-Value 817.01***

Portfolio (5) – (3): F-Value 599.53***

Difference between LOG (ENGAGEMENT):
Portfolio (1) – (3): F-Value 1106.5***

Portfolio (5) – (3): F-Value 1016.51***

Panel B: Association between Unexpected Earnings and Tweet/Engagement Volume - Portfolio Sorting on Size and Unexpected Earnings

PORTFOLIO	# TWEET FIRM-QUARTERS	LOG(ASSETS)	UE	LOG(TWEETS)	LOG(ENGAGEMENT)
1	7,049	7.223	-0.020	3.943	4.205
2	7,061	7.134	-0.001	4.219	4.621
3	6,517	7.380	0.001	4.442	4.930
4	6,827	7.277	0.002	4.296	4.721
5	6,757	7.214	0.016	4.022	4.348
Total	34,211	7.244	0.000	4.181	4.560

Difference between LOG (TWEETS):
Portfolio (1) – (3): F-Value 224.45***

Portfolio (5) – (3): F-Value 155.73***

Difference between LOG (ENGAGEMENT):
Portfolio (1) – (3): F-Value 228.82***

Portfolio (5) – (3): F-Value 144.52***

Panel A shows the relationship between unexpected earnings and tweet/engagement volume for five portfolios created by sorting on unexpected earnings.

Panel B shows the relationship between unexpected earnings and tweet volume/engagement for five portfolios created by sorting first on the size of the firm and then on unexpected earnings.

The sample for both panels comprises of only tweet firm-quarters during the period 2006-2017 (includes only firms which have created a Twitter account and have started tweeting).

*** p<0.01, ** p<0.05, * p<0.1; All variables are defined in Appendix A.

Table 5: Association between Unexpected Earnings and Tweet/Engagement Volume

Panel A: Unexpected Earnings and Tweet/Engagement Volume - All Firm-quarters

ALL FIRMS	**Dependent Variable = *UE_EARNINGS***				
	(1)	(2)	(3)	(4)	(5)
LOG(TWEETS)	**-0.001****				
	(-2.253)				
*LOG(TWEETS)*NEG_UE*	**0.002*****				
	(2.870)				
LOG(ENGAGEMENT)		**-0.001*****			
		(-2.667)			
*LOG(ENGAGEMENT)*NEG_UE*		**0.002*****			
		(3.513)			
RESPONSE			**-0.003*****		
			(-3.265)		
*RESPONSE*NEG_UE*			**0.008*****		
			(4.349)		
CHANGE_LOG(TWEETS)				0	
				(1.308)	
*CHANGE_LOG(TWEETS)*NEG_UE*				0.001	
				(1.014)	
CHANGE_LOG(ENGAGE)					**-0.001***
					(-1.691)
*CHANGE_LOG(ENGAGE)*NEG_UE*					0.001
					(1.223)
NEG_UE_EARNINGS	-0.028***	-0.028***	-0.028***	-0.025***	-0.025***
	(-5.533)	(-5.71)	(-5.85)	(-5.855)	(-5.845)
LOG(ASSETS)	-0.001*	-0.001*	-0.001*	-0.001**	-0.001**
	(-1.752)	(-1.681)	(-1.741)	(-2.032)	(-2.029)
MTB	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
	(-2.456)	(-2.442)	(-2.469)	(-2.476)	(-2.479)
LEVERAGE	0.000**	0.000**	0.000**	0.000**	0.000**
	(2.215)	(2.201)	(2.212)	(2.194)	(2.198)
LOG(NUM_ANALYSTS)	0	0	0	0	0
	(0.216)	(0.206)	(0.193)	(0.202)	(0.201)
LOG(PRESSRELEASES)	0	0	0	0	0
	(0.219)	(0.303)	(0.495)	(0.082)	(0.073)
LOG(NEWSPAPERS)	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	(-3.11)	(-3.085)	(-3.085)	(-3.155)	(-3.154)
UE_EARNINGS_1	0.062**	0.062**	0.062**	0.062**	0.062**
	(2.328)	(2.323)	(2.328)	(2.313)	(2.312)
CONSTANT	0.015***	0.015***	0.015***	0.015***	0.015***

	(4.321)	(4.262)	(4.215)	(4.457)	(4.457)
Observations	100,917	100,917	100,917	100,917	100,917
R-squared	0.256	0.257	0.255	0.252	0.252
Year-qtr FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Clustering of Errors	Industry	Industry	Industry	Industry	Industry

Panel B: Unexpected Earnings and Tweet/Engagement Volume - Only Tweet Firm-quarters

	Dependent Variable = *UE_EARNINGS*				
	(1)	(2)	(3)	(4)	(5)
LOG(TWEETS)	**-0.001***				
	(-2.695)				
*LOG(TWEETS)*NEG_UE*	**0.003****				
	(2.209)				
LOG(ENGAGEMENT)		**-0.001***			
		(-3.562)			
*LOG(ENGAGEMENT)*NEG_UE*		**0.002***			
		(3.620)			
RESPONSE			**-0.004***		
			(-4.916)		
*RESPONSE*NEG_UE*			**0.007***		
			(4.878)		
CHANGE_LOG(TWEETS)				0	
				(0.90)	
*CHANGE_LOG(TWEETS)*NEG_UE*				0	
				(0.39)	
CHANGE_LOG(ENGAGE)					0
					(0.97)
*CHANGE_LOG(ENGAGE)*NEG_UE*					0
					(0.46)
NEG_UE_EARNINGS	-0.030***	-0.030***	-0.026***	-0.019***	-0.019***
	(-4.269)	(-5.787)	(-7.946)	(-7.76)	(-7.72)
LOG(ASSETS)	0	0	0	-0.001	-0.001
	(0.451)	(0.177)	(0.474)	(0.71)	(0.70)
MTB	0	0	0	0	0
	(0.138)	(0.081)	(0.135)	(0.18)	(0.18)
LEVERAGE	0	0	0	0	0
	(0.673)	(0.594)	(0.658)	(0.71)	(0.71)
LOG(NUM_ANALYSTS)	-0.002*	-0.002*	-0.002*	-0.002*	-0.002*
	(-1.738)	(-1.792)	(-1.785)	(-1.68)	(-1.69)
LOG(PRESSRELEASES)	0	0	0	0	0
	(0.077)	(0.273)	(0.292)	(0.11)	(0.10)
LOG(NEWSPAPERS)	-0.001	-0.001	-0.001	-0.001	-0.001
	(-1.184)	(-1.09)	(-1.156)	(1.24)	(1.22)
UE_EARNINGS_1	0.056	0.055	0.055	0.056	0.056
	(1.550)	(1.515)	(1.535)	(1.53)	(1.53)
CONSTANT	0.022**	0.018*	0.021**	0.022**	0.021**
	(2.194)	(1.878)	(2.046)	(2.12)	(2.119)

Observations	32,899	32,899	32,899	32,899	32,899
R-squared	0.267	0.272	0.264	0.26	0.26
Year-qtr FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Clustering of Errors	Industry	Industry	Industry	Industry	Industry

Panel A shows the results of testing the association between tweet/engagement volume and the unexpected earnings for the full sample (includes both Tweet as well as all non-Tweet firm- quarters).

Panel B shows the results of testing the association between tweet/engagement volume and the unexpected earnings for a sub-sample of only tweet firm-quarters (includes only firms which have created a Twitter account and have started tweeting).

Both panels display the results of OLS regression for the sample period 2006 to 2017 using Model 3:
$UE\_EARNINGS_{i,t}/UE\_SALES\_GR_{i,t} = \beta_0 + \beta_1 TWEET\_VOLUME/CHANGE\_TWEET\_VOLUME_{i,t} + \beta_2(TWEET\_VOLUME/CHANGE\_TWEET\_VOLUME_{i,t}) *(NEG\_UE/NEG\_U\_SALES\_GR) + \beta_3 NEG\_UE/NEG\_U\_SALES\_GR + \Sigma\beta_j CONTROLS_{i,t} + YEAR\_QTR\ FIXED\ EFFECTS + FIRM\ FIXED\text{-}EFFECTS + \varepsilon_{i,t}$

Robust t statistics are in parentheses, *** p<0.01, ** p<0.05, * p<0.1; All variables are as defined in Appendix A.

Table 6: Association between Unexpected Sales Growth and Tweet/Engagement - Only Tweet Firm-quarters

TWEET_START==1	**Dependent Variable = *UE_SALES_GR***				
	(1)	(2)	(3)	(4)	(5)
LOG(TWEETS)	**-0.005*** **				
	(-2.887)				
*LOG(TWEETS)*NEG_UE_SALES_GR*	**0.014*** **				
	(3.073)				
LOG(ENGAGEMENT)		**-0.003*** **			
		(-2.972)			
*LOG(ENGAGEMENT)*NEG_UE_SALES_GR*		**0.008*** **			
		(3.504)			
CHANGE_LOG(TWEETS)			0		
			(0.765)		
*CHANGE_LOG(TWEETS)*NEG_UE_SALES_GR*			0.002		
			(1.052)		
CHANGE_LOG(ENGAGE)				0.001	
				(0.92)	
*CHANGE_LOG(ENGAGE)*NEG_UE_SALES_GR*				0	
				(0.055)	
RESPONSE					**-0.008*** **
					(-3.142)
*RESPONSE*NEG_UE_SALES_GR*					0.008
					(1.565)
NEG_UE_SALES_GR	-0.153***	-0.133***	-0.095***	-0.095***	-0.103***
	(-5.311)	(-6.158)	(-7.836)	(-7.83)	(-8.565)
LOG(ASSETS)	0.014***	0.014***	0.013***	0.013***	0.013***
	(4.015)	(4.12)	(3.824)	(3.835)	(3.867)
MTB	0	0	0	0	0
	(0.369)	(0.525)	(0.361)	(0.358)	(0.437)
LOG(NUM_ANALYSTS)	-0.005	-0.004	-0.004	-0.004	-0.004
	(-1.549)	(-1.386)	(-1.291)	(-1.281)	(-1.271)
LOG(PRESSRELEASES)	-0.001	-0.001	-0.001	-0.001	-0.001
	(-0.748)	(-0.682)	(-0.823)	(-0.841)	(-0.748)
LOG(NEWSPAPERS)	-0.001	-0.001	-0.001	-0.001	-0.001
	(-0.864)	(-0.848)	(-0.973)	(-0.971)	(-0.911)
LOG(SALES)_1	-0.017***	-0.017***	-0.018***	-0.018***	-0.018***
	(-5.259)	(-5.167)	(-5.394)	(-5.389)	(-5.344)
UE_SALES_GR_1	0.046**	0.047**	0.049**	0.049**	0.049**
	(2.128)	(2.194)	(2.199)	(2.22)	(2.227)
ADV_EXP_QTR	0.693	0.684	0.693	0.694	0.712
	(1.441)	(1.421)	(1.531)	(1.533)	(1.554)
CONSTANT	0.083***	0.071***	0.073***	0.072***	0.071***

	(3.281)	(2.748)	(3.114)	(3.106)	(3.006)
Observations	37,299	37,299	37,299	37,299	37,299
R-squared	0.416	0.412	0.4	0.4	0.401
Year-qtr FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Clustering of Errors	Industry	Industry	Industry	Industry	Industry

Table 6 shows the results of testing the association between tweet/engagement volume and the unexpected sales growth for a sub-sample of firms which have a Twitter account in the sample period 2006-2017 (excludes firms which have not created a Twitter account during the sample period) using Model 3: $UE\_EARNINGS_{i,t}/UE\_SALES\_GR_{i,t} = \beta_0 + \beta_1 TWEET\_VOLUME/CHANGE\_TWEET\_VOLUME_{i,t} + \beta_2(TWEET\_VOLUME/\ CHANGE\_TWEET\_VOLUME_{i,t}) *(NEG\_UE/NEG\_UE\_SALES\_GR) + \beta_3 NEG\_UE/NEG\_UE\_SALES\_GR + \Sigma\beta_j CONTROLS_{i,t} + YEAR\_QTR\ FIXED\ EFFECTS + FIRM\ FIXED\text{-}EFFECTS + \varepsilon_{i,t}$

Robust t statistics are in parentheses, *** p<0.01, ** p<0.05, * p<0.1; All variables are defined in Appendix A.

Table 7: Association between Earnings and Tweet/Engagement Volume

	Dependent Variable = *EPS*					
	All Firm-Quarters			**Only Tweet Firm-Quarters**		
	(1)	(2)	(3)	(4)	(5)	(6)
LOG(TWEETS)	**0.014***** **(4.269)**			**0.008**** **(2.392)**		
LOG(ENGAGEMENT)		**0.014***** **(4.730)**			**0.009***** **(3.414)**	
RESPONSE			**0.035***** **(3.533)**			-0.002 (0.266)
LOG(ASSETS)	0.039*** (3.455)	0.039*** (3.473)	0.039*** (3.435)	0.049*** (3.122)	0.048*** (3.065)	0.052*** (3.264)
MTB	0.005*** (3.632)	0.005*** (3.624)	0.005*** (3.647)	0.005*** (3.813)	0.005*** (3.806)	0.006*** (3.847)
LEVERAGE	-0.012*** (-4.447)	-0.012*** (-4.417)	-0.012*** (-4.456)	-0.014*** (4.244)	-0.014*** (4.211)	-0.014*** (4.239)
LOG(NUM_ANALYSTS)	-0.002 (-0.306)	-0.003 (-0.338)	-0.002 (-0.284)	0.006 (0.443)	0.006 (0.465)	0.007 (0.503)
LOG(PRESSRELEASES)	0.026*** (3.364)	0.023*** (2.923)	0.028*** (3.531)	0.011 (1.031)	0.01 (0.977)	0.011 (1.069)
LOG(NEWSPAPERS)	0.002 (0.208)	0.002 (0.202)	0.002 (0.214)	0.006 (0.381)	0.006 (0.358)	0.007 (0.425)
EPS_1	0.284*** (16.167)	0.284*** (16.200)	0.285*** (16.084)	0.201*** (7.084)	0.200*** (7.058)	0.201*** (7.097)
CONSTANT	-0.08 (1.348)	-0.078 (1.340)	-0.081 (1.345)	-0.472* (1.692)	-0.444 (1.581)	-0.490* (1.739)
Observations	178,236	178,236	178,236	46,559	46,559	46,559
R-squared	0.463	0.463	0.416	0.54	0.54	0.54
Year-qtr FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Clustering of Errors	Industry	Industry	Industry	Industry	Industry	Industry

Table 7 shows the results of testing the association between tweet/engagement volume and earnings for the full sample (includes both Tweet as well as all non-Tweet firm- quarters) in Columns (1) – (3) and for a sub-sample of only tweet firm-quarters (includes only firms which have created a Twitter account and have started tweeting) in Columns (4) – (6).

The Table displays the results of OLS regression for the sample period 2006 - 2017 using the following Model: $EPS_{i,t} = \beta_0 + \beta_1 TWEET\_VOLUME/CHANGE\_TWEET\_VOLUME_{i,t} + \Sigma\beta_j CONTROLS_{i,t} + YEAR\_QTR\ FIXED\ EFFECTS + FIRM\ FIXED\text{-}EFFECTS + \varepsilon_{i,t}$

Robust t statistics are in parentheses, *** p<0.01, ** p<0.05, * p<0.1; All variables are as defined in Appendix A.

Table 8: Association between Sales, Sales Growth and Tweet/Engagement Volume

Panel A: Sales, Sales Growth and Tweet/Engagement Volume - All Firm-quarters

	Dependent Variable = *LOG(SALES)*			**Dependent Variable = *SALES_GROWTH***		
	(1)	(2)	(3)	(4)	(5)	(6)
LOG(TWEETS)	**0.003***** **(3.213)**					
LOG(ENGAGEMENT)		**0.002***** **(2.745)**				
RESPONSE			**0.007**** **(2.309)**	**0.007**** **(2.456)**		
CHANGE_LOG(TWEETS)					**0.009***** **(2.958)**	
CHANGE_LOG(ENGAGE)						**0.010***** **(3.353)**
Observations	173,159	173,159	173,159	172,798	172,798	172,798
R-squared	0.982	0.982	0.982	0.241	0.241	0.241
Constant, Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year-qtr FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Clustering of Errors	Industry	Industry	Industry	Industry	Industry	Industry

Table 8 shows the results of testing the association between tweet/engagement volume and log (sales)/sales growth for the full sample (includes both Tweet as well as all non-Tweet firm- quarters).

Panel B shows the results of testing the association between tweet/engagement volume and log (sales)/sales growth for a sub-sample of only tweet firm-quarters (includes only firms which have created a Twitter account and have started tweeting).

Both panels display the results of OLS regression for the sample period 2006 to 2017 using the following Model: *LOG (SALES)* $_{i,t}$ *SALES_GROWTH*$_{i,t}$ = $\beta_0 + \beta_1$*TWEET_VOLUME/CHANGE_TWEET_VOLUME*$_{i,t}$ + $\Sigma\beta_J$*CONTROLS*$_{i,t}$ + *YEAR_QTR FIXED EFFECTS + FIRM FIXED-EFFECTS* + $\varepsilon_{i,t}$

Robust t statistics are in parentheses, *** p<0.01, ** p<0.05, * p<0.1; All variables are as defined in Appendix A.

Table 9: Stock Returns and Tweet/Engagement Volume – Additional Analysis

Panel A: Stock Returns and Tweet/Engagement Volume – Excluding Firms which never create a Twitter account

	Dependent Variable = *MON_EXCESS_RETURN*			
	FF Three Factors		FF Five Factors	
	(1)	(2)	(3)	(4)
CHANGE_LOG(TWEETS)_MON	**0.004***** **(5.019)**		**0.004***** **(5.179)**	
CHANGE_LOG(ENGAGE)_MON		**0.004***** **(4.3)**		**0.004***** **(4.395)**
MKTRF	0.646*** (3.04)	0.646*** (3.038)	0.677*** (3.074)	0.677*** (3.072)
SMB	0.059 (0.454)	0.06 (0.463)	0.012 (0.09)	0.013 (0.097)
HML	-0.193 (-0.946)	-0.192 (-0.942)	-0.274 (-1.382)	-0.273 (-1.38)
MOM	-0.703* (-1.834)	-0.703* (-1.835)	-0.501 (-1.511)	-0.501 (-1.51)
RMW			-0.106 (-1.093)	-0.106 (-1.095)
CMA			-0.179* (-1.691)	-0.179* (-1.69)
CONSTANT	-0.049*** (-3.832)	-0.049*** (-3.835)	-0.046*** (-3.866)	-0.046*** (-3.869)
Observations	252,412	252,412	247,351	247,351
R-squared	0.095	0.095	0.134	0.134

Panel B: Stock Returns and Tweet/Engagement Volume - All Tweets from Primary and Secondary Twitter Accounts

	Dependent Variable = *MON_EXCESS_RETURN*			
	All Firm-Months		**Only Tweet Firm-Months**	
	(1)	(2)	(3)	(4)
CHANGE_LOG(TWEETS)_MON	**0.003*****		**0.002*****	
	(5.158)		**(4.202)**	
CHANGE_LOG(ENGAGE)_MON		**0.004*****		**0.002*****
		(4.33)		**(4.446)**
MKTRF	0.544***	0.544***	0.477***	0.488***
	(2.59)	(2.589)	(2.819)	(2.823)
SMB	-0.005	-0.004	0.186*	0.189*
	(-0.032)	(-0.028)	(1.785)	(1.806)
HML	-0.268	-0.267	-0.054	-0.052
	(-1.457)	(-1.453)	(-0.407)	(-0.388)
MOM	-0.471	-0.47	-0.164	-0.163
	(-1.515)	(-1.514)	(-0.759)	(-0.756)
RMW	-0.089	-0.089	-0.048	-0.049
	(-0.982)	(-0.983)	(-0.596)	(-0.611)
CMA	-0.122	-0.122	-0.125	-0.125
	(-1.214)	(-1.21)	(-1.581)	(-1.571)
CONSTANT	-0.045***	-0.045***	-0.011***	-0.011***
	(-3.74)	(-3.741)	(-3.532)	(-3.555)
Observations	486,525	486,525	140,294	140,294
R-squared	0.128	0.128	0.103	0.103

Panel A shows the results of testing the association between change in tweet (engagement) volume and the monthly excess stock returns for a sub-sample of firms which have a Twitter account in the sample period (excludes firms which have not created a Twitter account during the sample period). Columns (1) – (3) are for Fama-French three-factor model and columns (4) – (6) are for the Fama-French five-factor model.

Panel B shows the results of the Fama-French five-factor model for testing the association between change in tweet (engagement) volume and the monthly excess stock returns. The tweets are from both the Primary as well as Secondary Twitter accounts of a firm. Columns (1) – (3) are for All Firm-Months and columns (4) – (6) are only Tweet Firm-Months.

Both panels incorporate Momentum factor as well and display the results using Fama-MacBeth monthly cross-sectional regressions with Newey-West corrected standard errors for autocorrelation (two lags) used for calculating t-statistics. The reported slopes are computed as the time-series average of the slopes in monthly regressions of excess stock returns on the explanatory variables for the sample period 2006 - 2017 using Model 2: $(R_{i,t} - Rf_t) = \beta_0 + \beta_1 CHANGE\_LOG\ (TWEET)_{i,t}/CHANGE\_LOG\ (ENGAGE)_{i,t} + \beta_2 MOM_t + \Sigma\beta_j FAMA\text{-}FRENCH\_FACTORS_t + \varepsilon_{i,t}$

*** p<0.01, ** p<0.05, * p<0.1; All variables are defined in Appendix A.

Table 10: Earnings, Unexpected Earnings and Industry-Adjusted Tweet/Engagement Variables

	Dependent Variable = *EPS*		Dependent Variable = *UE_EARNINGS*	
	(1)	(2)	(3)	(4)
LOG(TWEETS)_MED_IND_ADJ	0.003		**-0.001*** **	
	(1.027)		**(-2.775)**	
*LOG(TWEETS)_MED_IND_ADJ*NEG_UE*			**0.002*** **	
			(3.994)	
LOG(ENGAGEMENT)_MED_IND_ADJ		**0.005***		**-0.001** **
		(1.707)		**(-2.227)**
*LOG(ENGAGEMENT)_MED_IND_ADJ*NEG_UE*				**0.003** **
				(2.407)
NEG_UE			-0.019***	-0.018***
			(-7.78)	(-7.97)
Observations	46,559	46,559	32,899	32,899
R-squared	0.54	0.54	0.266	0.264
Controls	Yes	Yes	Yes	Yes
Year-qtr FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Clustering of Errors	Industry	Industry	Industry	Industry

The first two columns of Table 10 show the results of association between median industry-adjusted tweet/engagement volume and the earnings using Model 1: $EPS_{i,t} = \beta_0 + \beta_1 TWEET\_VOLUME/CHANGE\_TWEET\_VOLUME_{i,t} + \Sigma\beta_J CONTROLS_{i,t} + YEAR\_QTR\ FIXED\ EFFECTS + FIRM\ FIXED\text{-}EFFECTS + \varepsilon_{i,t}$

Columns 3 & 4 show the results of association between median industry-adjusted tweet/engagement volume and the unexpected earnings using Model 2: $UE\_EARNINGS_{i,t} = \beta_0 + \beta_1 TWEET\_VOLUME/CHANGE\_TWEET\_VOLUME_{i,t} + \beta_2(TWEET\_VOLUME/\ CHANGE\_TWEET\_VOLUME_{i,t}) *(NEG\_UE/NEG\_U\_SALES\_GR) + \beta_3 NEG\_UE/NEG\_U\_SALES\_GR + \Sigma\beta_J CONTROLS_{i,t} + YEAR\_QTR\ FIXED\ EFFECTS + FIRM\ FIXED\text{-}EFFECTS + \varepsilon_{i,t}$

The analysis is for a sub-sample of only tweet firm-quarters (includes only firms which have created a Twitter account and have started tweeting).

Robust t statistics are in parentheses, *** p<0.01, ** p<0.05, * p<0.1; All variables are as defined in Appendix A.