5-2015

The Effect of Social Media on Trading Behavior: Evidence From Twitter

Tanya Paul
University of Pennsylvania

The Effect of Social Media on Trading Behavior: Evidence From Twitter

Abstract

Since 2008, many firms have begun to use Twitter as a form of communicating news to consumers and investors. Twitter enables the firm to manage the information content of the tweet because of its emphasis on the 140 characters. In addition, the marginal investor who may read this "managed" tweet and react accordingly is most likely unsophisticated since Twitter is an information-pushing platform rather than an information-pulling platform such as Bloomberg. As a result, I hypothesize that using Twitter to communicate with investors actually leads to asset mispricing. I document a negative relationship between the log of abnormal volume and firm-initiated announcement tweets for both product recalls and monthly sales announcements. I also document a negative relationship between the absolute value of abnormal equity returns and firm-initiated announcement tweets for monthly sales data. Further analysis shows that this relationship holds true only for monthly sales events with positive returns and not for those with negative returns. This suggests that it is because Twitter serves to lower information asymmetry, resulting in less positive returns than before. These results have implications for understanding how dissemination and unsophisticated trading can affect liquidity and market efficiency.

Keywords

disclosure, dissemination, disclosure quality, Twitter, social media, product recall, monthly sales announcements

Disciplines

Business

The Effect of Social Media on Trading Behavior: Evidence from Twitter

Tanya Paul

The Wharton School, University of Pennsylvania

Advisor: Prof. Catherine Schrand, Accounting

May 2015

Since 2008, many firms have begun to use Twitter as a form of communicating news to consumers and investors. Twitter enables the firm to manage the information content of the tweet because of its emphasis on the 140 characters. In addition, the marginal investor who may read this "managed" tweet and react accordingly is most likely unsophisticated since Twitter is an information-pushing platform rather than an information-pulling platform such as Bloomberg. As a result, I hypothesize that using Twitter to communicate with investors actually leads to asset mispricing. I document a negative relationship between the log of abnormal volume and firm-initiated announcement tweets for both product recalls and monthly sales announcements. I also document a negative relationship between the absolute value of abnormal equity returns and firm-initiated announcement tweets for monthly sales data. Further analysis shows that this relationship holds true only for monthly sales events with positive returns and not for those with negative returns. This suggests that it is because Twitter serves to lower information asymmetry, resulting in less positive returns than before. These results have implications for understanding how dissemination and unsophisticated trading can affect liquidity and market efficiency.

Keywords: disclosure, dissemination, disclosure quality, Twitter, social media, product recall, monthly sales announcements.[12]

[1] I would like to thank my advisor Professor Catherine Schrand, John C. Hower Professor of Accounting, for her continual support and guidance as well as insightful comments throughout the length of this project. I would also like to thank Professor Daniel Taylor, Professor Brian Bushee, Professor. Paul Shaman, and Professor. Richard Kihlstrom for their comments and support of my undergraduate research. I also gratefully acknowledge the helpful comments of fellow WRS scholars Ola Aboukhsaiwan and Kimberly Yao in addition to Vybhavi Bharadwaj who assisted in data collection. Finally, I am grateful to Dr. Utsav Schurmans and Bethany Schell for their continuous support.

I. Introduction

Many studies find that company disclosure through financial reports, management discussion, management forecasts, earnings calls, and other company announcements reduces information asymmetry, which is important for efficient capital markets and lowering the cost of capital for firms. However, perfectly accurate and unbiased disclosure can in fact be costly in many ways. For example, it could lead to an overreaction in capital markets or reveal competitive advantages to competitors or litigators. Additionally, even with a decision to disclose information, firms may not be successful in disseminating it to investors. Traditionally, firms releases announcements on its website and sent information to newswires such as PR Newswire and Businesswire. From there, the press selects which news to cover. However, there is some bias involved; the press cherry-picks the news it wants to cover. This usually means that press coverage focuses on more visible firms as the public would be more interested to read about them (Miller 2006).

With the advent of social media, the nature of corporate disclosure has changed. Firms can now directly access investors and consumers through sites such as Twitter, Facebook, LinkedIn, Instagram, etc. Followers of a social media account get push notifications about news; they can share news of interest with family and friends. In 2013, the SEC ruled that social media is a viable source of dissemination of corporate information.[3] The extent to which firms use social media as a mode of dissemination today greatly varies by the type and size of the firm as it is not required by the SEC. As a result, the increase of dissemination of important information through these online platforms will only increase moving forward, and the implications of this have not been heavily studied.

This study uses Twitter as its empirical setting for specific reasons that will be described below, as well as the fact that it is a representative microblogging site. It was founded in 2006 and enables its users to post and follow those who post "tweets", a 140-character message. In 2013, Twitter was one of the top 10 most visited websites in the world. Currently, it has 500 million users (up from 100 million in 2012), making it one of the fastest growing social media sites. [4] Additionally, anecdotal evidence suggests that many companies use Twitter or Facebook as one of its primary social media outlets to release information. Between the top social media sites such as Twitter and Facebook, Twitter was specifically chosen for this study because of its following three key characteristics:

1. *Information Content*. This is the main focus of this study. While studies have shown that disclosure does reduce information asymmetry, this would depend on the quality of the content provided in the Tweet. While firms are disseminating information, it may be highly biased in order to manage market reaction and minimize the cost of disseminating information. Twitter is a way for the firm to control the content that is disseminated in wide audiences. Before, firms were less able to do this as press coverage had a flavor of its own bias in the article. In the 140 characters allowed in a Tweet, a firm is able to prioritize the information of the announcement and influence what about the announcement the investor sees. In psychology, this appeals to the theory of belief perseverance, which is the idea that one sticks to his or her initial beliefs even if there is evidence contradicting this. Additionally, the language of a Tweet allows for more sentimental words rather than boilerplate and standard language that is usually found in a company announcement.

[3] In July of 2012, the CEO of Netflix posted on his personal Facebook that Netflix had consumed 1 billion hours that month for the first time. This sparked discussion on whether social media was an appropriate method of dissemination of company announcements. The SEC announced in 2013 that companies could use social media to announce information as long as investors were aware that social media was being used to disseminate.

[4] Taken from the Twitter website

2. *Quantity of Information.* Firms can choose to exceed the 140 character by posting several times about the same event. This not only increases the amount of information that will reach investors but also catches the attention of investors. This is particularly useful for small firms.
3. *Reach*. Compared to before, Twitter is a form of pushing information rather than pulling information. This means that in order to get news about a company before, investors had to seek it out. Here, the company pushes information to anyone who follows the company on Twitter. This is precisely how, as mentioned before, the reach of the news is more broadly disseminated. Additionally, those who receive the information can also participate in its dissemination; they can posts to share them with their own followers, leading to a network effect.

A study by Blankespoor, Miller, and White (2014) suggests that Twitter and other forms of social media help reduce information asymmetry amongst investors and allows for higher market liquidity as a result of broad dissemination. They measured bid-ask spreads as a proxy for liquidity and found that Twitter results in narrower bid-ask spreads. However, the relationship between liquidity and market efficiency is highly debated in research. Theory suggests that when liquidity increases, informed traders are able to trade against uninformed traders with lower transaction costs (Kyle 1985). This encourages informed traders to seek out information in order to take advantage of these arbitrage opportunities, making the market more efficient and reducing information asymmetry. On the contrary, liquidity may actually decrease market efficiency. This is based off the assumption that increased liquidity is actually a proxy for unsophisticated trading, which could drive the price of the security away from its fundamental value. The study by Blankespoor, Miller, and White (2014) does not specify whether the type of liquidity that results from dissemination through social media is efficiency-increasing or efficiency-decreasing. This leads to the development of the primary question being studied: how does the nature of information disseminated by companies affect trading behavior? This is motivated by understanding whether the characteristics of social media actually help provide valuable information dissemination and increase market efficiency or serve to increase unsophisticated trading, leading to pricing away from the fundamental security value.

To investigate this question, I document how the effect of companies disseminating information through Twitter affects the abnormal returns and trading volume of the stock in a (0, 2) trading window around product recalls and monthly automobile manufacturing sales announcements. I collect data from before the firm opened a Twitter account ("Pre-Twitter Period) and after the firm opened a Twitter account ("Post-Twitter Period") in order to observe what the effect of disseminating information on Twitter has on trading behavior. All of the companies in the study disseminated information using traditional methods of PR Newswire and announcements on the company website as discussed earlier in both the Pre-Twitter and the Post-Twitter Period. The main difference between the two periods was that these companies also tweet about the announcement in the Post-Twitter Period, on top of releasing a newswire about it.

The next section provides more background, related literature, and a hypothesis development. The section after that will discuss the nature of the data and the variables collected. Section 4 provides descriptive statistics as well as results. Section 5 talks about potential issues. Section 6 will conclude.

II. Background, Motivation, and Related Literature

Information Asymmetry and Disclosure

As mentioned earlier, the firm faces a tradeoff in the decision to disclose information. This discussion is relevant as the firm faces a decision on not only deciding to disclose at all on Twitter since it is not required by the SEC to do so but also to what extent it should disclose on Twitter. On one hand, the firm could increase liquidity, decrease its cost of capital, and reduce information asymmetry overall (Kim and Verrecchia 1994) with additional voluntary disclosure. On the other hand, there are costs that the firm could incur by disclosing. Berger and Hann (2007) suggest that agency costs and proprietary costs are reasons why managers may choose to withhold information. Supported by this is a study by Hayes and Lundholm (1996) which also documents that proprietary costs are a reason why managers are hesitant to disclose certain types of information. Another potential concern that a firm may have is increased vulnerability to litigation through voluntary disclosure; however, a study by Field, Lowry, and Shu (2005) did not find evidence that voluntary disclosure led to litigation. In fact, results seemed to point to the opposite idea. In addition to these factors, managers may also use voluntary disclosure to influence certain factors, such as analyst following as documented by Lang and Lundholm (1996). Their study found that increased voluntary disclosure can lead to increased investor following and reduced information asymmetry. Additionally, the results of the study by Bushee and Noe (2000) support the idea that firms with better disclosure practices have a higher institutional following; this leads to stock price volatility. To support this, Healy and Palepu (2001) suggest that increased voluntary disclosure is linked to better stock performance in conjunction with institutional ownership, analyst following, and stock liquidity.

Many of the studies above use the term "disclosure" slightly differently from this study. For example, the data collected by Healy and Palepu (2001) use firms that "have made large and sustained improvements in their disclosure quality" (Healy and Palepu 2001). Because there is an overall increase in the quality of the disclosure, it may be that information asymmetry decreases in this case because the firm and management are subsequently seen as more credible overall to capital markets. This is a significant distinction from my study; this type of disclosure is only true for the monthly sales data collected compared to the product recall data, as will be discussed later.

Liquidity and Unsophisticated Trading

The study mentioned earlier by Blankespoor, Miller, and White (2014) found that company-initiated Tweets increase liquidity. Liquidity is generally considered to be good in capital markets because it is generally linked with market efficiency. The Kyle model (1985) shows that when informed investors enter the market, bid-ask spreads widen. However, his model does not necessarily hold the opposite result for uninformed investors. When uninformed investors enter the market, the bid-ask spread could narrow but could also widen depending on the aggressiveness at which rational investors arbitrage. If the results of the aforementioned study could be explained by the entrance of uninformed investors in the market, then this begs the question whether social media as a method of dissemination leads to market efficiency because liquidity leads to lower trading cost (O'Hara 2005) or whether it leads to asset mispricing. This is further explored in this study.

Several empirical studies give evidence in favor of both theories. Wurgler and Zhuravskaya (2002) document that stocks without close substitutes experience a higher price jump when included in the S&P index. These stocks are more illiquid and hence experience asset mispricing. Sadka and Scherbina (2007) suggest that there exists a close relationship between liquidity and mispricing by using analyst disagreement as a proxy for firms with high information asymmetry and liquidity. However, other studies support the opposing view that increased liquidity actually decreases. Bloomfield, OHara, and Saar (2009) suggest in an experimental study that when uninformed traders enter the trading scene, they increase market liquidity but at the same time hurt efficiency because markets are unable to correct the asset mispricing. Flepp, Nuesch, and Franck (2014) demonstrate that the type of liquidity does indeed matter for market efficiency by showing that liquidity through the entrance of noise traders decreases market efficiency at the betting exchange on weekends but not weekdays.

Barber, Odean, and Zhu (2006) prove that in the presence of three conditions 1) misinterpretation of available information. 2) non-cancellation of uninformed trade and 3) limits on rational arbitrage, trading from not fully rational traders drive prices away from fundamental values for at least temporarily. Additionally, it has been empirically shown that when a firm's message board (Yahoo) was established, this lead to increase in daily trading volume, lower returns, and higher volatility (Jones 2006).

Social Media

Social media refers to "electronic communication through which users create online communities to share information, ideas, personal messages, and other content"[5]. Because of the sheer number of people on Twitter and the advent of Cashtags, where users can use #{Insert firm ticker here} to hashtag their Tweets. Because this is linked to financial information, now it is very simple for people to find financial information about a company on Twitter by searching by the firm's cashtag (example: General Motors's (GM) cashtag is $GM). As mentioned before, it has been shown that Twitter as a form of dissemination increases the liquidity of the stock, using bid-asks spreads as a proxy for liquidity. This is especially true for small visibility firms (Blankespoor, Miller, White 2014). Another study showed that using social media such as Twitter and Facebook to disseminate news about product recalls help lower negative market reactions because of the firm's ability to do "damage control" (Lee, Hutton, Shu 2015). Outside of firm-initiated Tweets, another study showed that sentiment of tweets (including non-company tweets) is associated with abnormal returns (Oh and Sheng 2011), and message volume can predict next-day trading volume (Wysocki 1998).

Product Recalls

There are four main reasons for using product recalls as one of the announcements of study. The first is that it is a "compelling" announcement to post for firms. This is because the number and types of important firm announcements put out on Twitter may be relatively low since (a) Twitter is a relatively new source of dissemination and (b) Investor Relations may be hesitant to provide information on the Twitter for fear that a mistake could blow the situation out of proportion. This is supported by the fact that the 2013 AON Risk Management survey cited social media as an upcoming area of risk for a company. However, due to chances of litigation that the company did not take enough measures to announce a product recall, companies post about recalls on Twitter since it directly affects consumers and would be an item of interest. The second reason is that product recalls is what one would consider a "negative" event – rational traders would trade in the same direction. This fulfils one of the required three conditions in the study by Barber, Odean, and Zhu (2006) mentioned earlier. The third reason is that product recalls are unexpected in nature in contrast to earnings announcements, which are anticipated. This means that investors have less preconceived notions about the event before they read about it in the news or on Twitter or have expectations. However, it should be noted that if a firm consistently has many product recalls, this advantage may not exist anymore. The fourth reason is that product recalls fits in with the design of the study well. The nature of product recalls has not changed significantly over the last 10 years, the period of the study.

However, using product recalls as one of the events in the study has certain disadvantages. One is it suffers from a possible endogeneity problem. The decision to disclose information about product recalls on Twitter is an endogenous one, evidenced from the fact that very few of the companies that are listed as

[5] Merriam Webster definition

having product recalls on the CPSC website actually disclose anything about a recall on their Twitter account. This is discussed in more detail in Section V. Additionally, the sample size is limited for this very reason.

Monthly Sales Announcement

There are three primary reasons why monthly sales announcement was chosen as one of the announcement types being collected in the data. The first is that, in contrast to product recalls, monthly sales are not necessarily a negative event. It could be either positive or negative. It may contain more content to interpret as negative or positive for unsophisticated investors or more room for interpretation. Related to this, a study found that investors underreact to textual information embedded in news stories (Tetlock, Saar-Tschansky 2008), which may be the case for a monthly sales announcement. Additionally, it has the benefit of being an expected announcement. Car companies release monthly sales data during the first week of every week. For this reason, investors may be able to tell how well a company did that month before it is released or have expectations about the event. The third reason is that the nature of the disclosure of the firm with respect to monthly sales is inherently different from product recalls; it builds and reflects a potential overall increase in disclosure. From the data, we very rarely observe firms choosing certain months to Tweet about monthly sales and refraining from doing so during certain months. This stems from the fact that the SEC does not require that announcements be posted on all platforms of dissemination at all times; it would be very possible for a firm to post only positive announcements on Twitter and never mention a negative announcement on Twitter (however, firms do not have a such option on PR Newswire).

III. Sample Selection and Variables

Data Collection

Overall methodology

Two different sets of data was collected for this analysis as mentioned above – product recalls and monthly car sales announcements. In order to understand how firm-initiated Tweets affect abnormal returns and abnormal trading volume, I manually collected roughly 90 data points on both recalls and announcements from before and after the company opened a Twitter account and began tweeting about company news. All firms included in the study did not change their disclosure behavior as a result of creating a Twitter account. This means that in addition to the channels they already disseminated through (ex: PR Newswire and Busineswire), these firms also began to use Twitter as an additional channel of dissemination. This distinction is important as it is critical for the empirical design; in 2010, Google went so far as to only release its earnings announcements on social media and its website only. Google would not have been a firm included in this study, as a result. To further illustrate, I provide an example: in 2006, Walmart had a product recall that it released a newswire about. In 2010, they had another product recall which they released a newswire about. The main difference is that in 2010, Walmart also tweeted about it on its Twitter to alert consumers and investors. Below, I discuss the specific procedures in obtaining the data for the product recall dataset and the monthly sales announcement dataset.

Product Recalls

To find a list of product recalls, I use the Consumer Product Safety Commission website. Next, I check whether this company is a) public and b) whether they have issued a newswire about it archived on Factiva. I also check to make sure that all the data points in the Pre-Twitter sample occur at a point in time before the company created a Twitter. Next, for the Post-Twitter sample, I use the CPSC list of product recalls and check whether the company in question a) has a Twitter and b) if so, whether this recall disclosed on Twitter. After this step, I have a set of companies that have product recalls before and after the company created a Twitter account and tweeted about the recall. Next, I find the PERMNO using CRSP and use Eventus find abnormal returns and abnormal trading volume of the stock compared to an estimation length of 255 46 days before the event. In addition to these two response variables, I also collect data on various characteristics of both the newswire and the Tweet the company posts. [6]

Monthly Sales Announcements

Data on General Motors and Honda monthly sales announcements were collected for this sample. This is because both General Motors and Honda began tweeting about their monthly sales after creating a Twitter account in 2008, which was imperative to the design of this study.[7] Next, I used Factiva to find the monthly sales newswires and recorded various characteristics of these announcements which are discussed in more detail in the next section. In addition to this, I collect information on the tweet(s) that the company posted for those announcements that occurred after the company opened a Twitter account. I did not include the month every year (generally January) in which the company announces annual sales as well as monthly sales since the returns and volume for that month would be an outlier from the other monthly sales announcements. Lastly, similar to above, I use Eventus to find abnormal returns and abnormal trading volume of the stock around a 2 day trading window, (0, +2).

Abnormal returns and volume can be calculated by the following equation:

$$A_{it} = R_{it} - \overline{R_i}.$$

Where A_{ji} is the arithmetic mean return of the common stock of the j[th] firm computed over the estimation period R_j from its return on day t.

Response variables and Regressors

Response Variables: The variables of interest that were measured were abnormal returns (*Return*) and abnormal log of trading volume (*LnVolume*), both of which Healy, Hutton, and Palepu (1999) found were affected in the presence of unsophisticated trading.

Regressors: PrePost is a binary variable and the primary regressor of interest that indicates whether this product recall was tweeted about (this occurs in the “Post-Twitter” period, or after the company establishes a Twitter account). If it falls in the pre-Twitter period (before the firms opens a Twitter account) then it equals 0. This would capture whether the firm tweeting about an announcement had an effect on abnormal returns or abnormal trading volume.

Because of the two types of events, I control for different things in the two types of events. Below are descriptions of the relevant control variables for each type of event. There are two types of variables: variables that characterize the underlying company announcement and ones characterizing the tweets that the companies also put out in the Post-Twitter period for cross-sectional predictions. I control for these in the regression to isolate the effect of whether tweeting about an event significantly impacts abnormal returns or abnormal log of trading volume.

[6] Here, the newswire is referred to as the newswire the company releases or the announcement of the company website. The tweet is the tweet the company posts about the event on its Twitter page.

[7] It must be noted that not all firms tweet about its company announcements; Urban Outfitters does not mention any company related news on its Twitter and keeps its Twitter as a way to advertise its new products and sales.

Product Recalls

Announcement Control Variables

The market cap *(market cap*) of the company the year of the announcement was recorded. This is a proxy for the "visibility" of the firm under the assumption that high market cap firms are better known.

Tweet Variables

As discussed before, there are 3 distinct characteristics of tweets that I use to test for the effects of tweeting: Information Content, Reach, and Quantity. Here, quantity does not apply as no company in the sample set Tweeted multiple times about the same recall within the (0, +2) period. One is the number of Retweets of a company Tweet (*Retweet*). This continuous variable measures the reach in terms of it reaching additional social networks as people retweet Tweets they read. I predict that as the reach increases, the effect on abnormal returns and volume should increase. Additionally, *Link,* a binary variable, controls for whether the company included a hyperlink to the announcement in the Tweet, which is another variable controlling for reach. The second characteristic of a tweet is information content. *Problem.or.No* measures whether the company mentions the cause of the recall in the tweet, which can be seen as a form of information content manipulation because an unsophisticated investor may be more likely to react if the reason for the recall is included. For example, Walmart in 2013 tweeted "Product Recall Alert: Wahl Total Care due to burn hazard." Here, they mention that burn hazard is the reason for the recall. In 2014, Walmart tweeted "#ProductRecall Alert: Comfort Research Bean Bag Chairs", not mentioning why the product was being recalled. *Numerical* is also a variable that measures the information content; it measures whether the company discloses the number of units recalled in the tweet. Additionally, *Impact* controls for the impact of the recall, measured as number of units recalled X price of each unit. This is because this data set contains high impact recalls such as car recalls in addition to smaller product recalls such as Bean Bag chairs.

Monthly Sales Announcements

Announcement Control Variables

GM, a binary variable, is 1 if the company is GM, 0 if the company is Honda. *Beta* controls for the estimated beta of the firm in the estimation period of 200 days 43 days before the event. This is a relevant variable for the monthly sales announcements and not product recalls because of the seasonality of the data. *Event* controls for whether the event was "positive" or "negative". This was estimated by looking at whether monthly sales for that month increased or decreased compared to the same period in the prior year. This is a standard auto sales measure used in announcements. *Up_Net* captures the manipulation of the information by the firm. It is the difference between the number of positive words that appear in the announcement – the number of negative words that appear in the announcement. Positive words include: Gain, increase, up, +, positive, crown, top, rising. Negative words include: down, decrease, declining, loss, -. *Salesper* measures the percentage increase or decrease in total sales that month compared to the same month the previous year. *SalesNum* measures the gross sales of cars that month.

Tweet Variables

Retweet and *Link* are the same as what was discussed above. *Number of Tweets* measures the number of times a company Tweeted about the monthly sales. This is a measure of quantity, one of the 3 characteristics of a tweet that did not apply to the product recall dataset but applies here. *NetSent* measures positive words that appear in the Tweets minus the negative words that appear in the Tweets. This is a measure of the quality of the information provided in the tweet. For example, a company may manipulate information during a negative sales event by including many positive words to counteract the negative overall results. I predict that the higher the NetSent in the Post-Twitter period compared to the Pre-Twitter period, the abnormal returns and trading volume should be higher. *Number*, a binary variable, measures whether the firm mentions the gross amount of sales in the Tweet(s). *Per,* also a binary variable, measures whether the firm mentions the % increase or decrease of sales in the Tweet(s). Per and Number also measure the quality of the information provided in the tweet; the most informative tweets would include the % increase and gross number of sales of that month in the tweet.

Model

To estimate the effect of Twitter on the response variables abnormal return and abnormal log of trading Volume, I use the model given below. It should be noted that my design is not a difference-in-differences. That design would not have allowed me to measure various attributes of the tweet.

Where Return $$Return = B_0 + B_1 PrePost + B_2 AnmtControls + B_3 TweetControls \times PrePost$$ represents either Abnormal Returns or log (Abnormal Trading Volume). Here, I interact all the Twitter-specific variables such as *Link* and *Number of Retwe*ets with the *PrePos*t variable. This is so that the effect of dissemination through Twitter on abnormal returns can be written as

$$\frac{\partial Return}{\partial PrePost} = B_{1+} B_3 TweetControls$$

Due to the limited sample size, I use model selection using AIC (Akaike Information Criterion) in order to find which variables add the most information to the regression as assessed by the likelihood function.

Results

This section examines whether the dissemination through Twitter has an effect on Abnormal Returns and log(Abnormal Trading Volume).

Using AIC variable selection criterion, the following regressors were selected for product recall.

$$Volume = B_0 + B_1 Prepost + B_2\ Retweet{*}Prepost + B_3 Impact + B_4 Marketcap$$

$$Return = B_0 + B_1 Impact$$

$$AbsReturn = B_0 + B_1 Retweet + B_2\ Retweet{*}Prepost + B_4 Marketcap$$

Using AIC variable selection criterion, the following regressors were selected for monthly sales announcements.

$$Volume = B_0 + B_1 salesper + B_2\ salenum$$

$$Return = B_0 + B_1 salesper + B_2\ salenum$$

$$AbsReturn = B_0 + B_1 PrePost + B_2\ UpNet + B_4 salesper$$

IV. Regression Results

Because Twitter allows companies to manage the content released to investors, I predict that disseminating corporate news through Twitter will result in less negative equity returns over the two day trading window. Because tweets are a form of “pushing” information rather than pull and because companies have the ability to manage the content of the Tweet, this will result in unsophisticated investors following these tweets to trade in a positive direction, resulting in a less negative overall return for product recalls and bad sales events and to higher abnormal returns for positive sales events. This is exacerbated by the fact that monthly sales announcements do not make top headlines like big recalls, so tweeting about them will cause more unsophisticated traders to enter than product recalls. I also predict that trading volume will also increase as a result of dissemination to a much broader audience. If the model yields a significant coefficient for the variable *PrePost*, then we may infer that, all else equal, the use of Twitter to disseminate information did indeed affect the response variable in question, either abnormal returns (*Return*) or abnormal trading volume (*Volume*).

[Insert Table 7 here]

However, results suggest that for both product recalls and sales, the advent of Twitter is insignificant when comparing the abnormal equity returns. Hence, it appears as if the dissemination of information through Twitter does not have a significant effect on the abnormal equity returns. However, for log(abnormal volume), *PrePost* is significant for both product recalls and monthly sales announcements, with a p- value of .05 and a coefficient of -1. For the absolute value of abnormal equity returns, PrePost was significant for monthly sales announcements but not product recalls. For product recall, the only significant regressor on return was impact of the product recall. For monthly sales, it was salesper and salesnum. [8]

Because it is impossible to know what the market expectations were before those events, I also regress the absolute value of abnormal returns on the collected variables. For the absolute value of abnormal returns of product recall, the main effects were the number of Retweets and market cap of the company. It should also be noted here that the constant was very significant, indicating that there is some variable that accounts for abnormal returns which I did not collect data for. For the absolute return of monthly sales, PrePost, UpNet, and the constant were significant. For the trading volume for product recalls, Prepost, Retweet, and ProblemInfo were significant with p-values of .05. For trading volume for monthly sales, significant regressors were Prepost, salesnum, link, and beta, all with p-values of less than .05.

This means that disseminating through Twitter does have an effect on the trading volume. While the coefficient on this is negative, this is possible because the data is in terms of log volume and not sheer magnitude. A negative coefficient could imply that while the trading volume still increased, the rate at which trading volume increased actually decreased. This makes sense given the time period; while I hypothesized that trading would increase as a result of unsophisticated trading, I did not account for the fact that Twitter took off in popularity in 2008, around the time of the financial crisis, lowering the number of unsophisticated investors that would enter and trade. Even for events that seemingly would have trades in one direction like a product recall, it is possible unsophisticated traders were misinterpreting signals in the opposite direction. This led to cancellation of the direction of their trades, making it appear like no significant difference in equity returns while trading volume increased.

To further understand the results and the fundamental question of Twitter on equity returns, I separated the Product Recall data into 3 groups: positive abnormal returns and negative abnormal returns.

[Insert Table 8 here]

Product Recalls

For the positive subgroup, as the impact of the recall increased, the returns became less and less positive (Figure 1). Perhaps this means that either (1) there is something about the low impact recalls that lets the firm manage the information content or (2) consumers panic less because it is a seemingly less alarming issue and/or it isn't a company they know much about. In order to understand this, I look at the information

---

[8] Salesper – the increase or decrease in monthly sales compared to the prior period; Salesnum – the gross number of car sales that month

content released by the firm both through its announcement and through its tweet. Figure 3 we see that the mean abnormal returns is lower when the firm includes the reason for the recall in the tweet. Perhaps that means that a lot of the positive returns are generated because the market cap of the firm is big, so the firm has more to lose by tweeting about this event. Thus, firms manage information on the tweet. It is also possible that firms also manage the information content depending on how large the impact or how much it believes unsophisticated investors will react to the news. In Panel B of Table 2, we see that returns become less positive as the number of retweets increase. This may be that as dissemination increases, more unsophisticated investors trade. However, this is unsupported by the data as there is no seeming relation between Retweets and volume of Trade. Figure 4 does show management of information by market cap, which supports my earlier hypothesis.

For negative abnormal returns, there seems to be less negative returns in the Post-Twitter than before. This seems to support the original theory that for recalls that are likely to be more negative, Twitter does make a difference; the abnormal returns are lower. This is consistent with the result in the product recall study mentioned earlier (Lee, Hutton, Shu 2015). We see that as market cap increases, there is less negative returns. Perhaps because high market cap companies manage information content. It could also be because unsophisticated traders react more to companies they know about (GM, Ford, etc) and they misunderstand the signals give off.

Additionally, in Figure 9, we do see information content management here as market cap increases. Even with the divisions into subgroups, AIC does not select PrePost as a regressor at all.

[Insert Table 4 and Table 9 here]

Monthly Sales

In Table 4, we see that the net positive words is much higher for positive sales events than negative sales events. This makes sense – no clear sense of information management. The puzzling thing about the data is the fact that there are negative sales events that actually have positive abnormal returns. This begs the question of whether for negative events (Events =1), do the net positive words in the announcement change? For negative events, Net_Pos =11.63. For positive events, it is Net_Pos = 19.33. This does indicate that the firm's management of content in the announcement leads to positive events.

For positive returns, the return in the post period is actually lower than in the pre-period (Figure 9). I conclude that for a bad event (Event =1), Twitter actually helps disseminate information and bring it to back to fundamentals. This result does not hold true for the negative return subset. This is likely because the market has already fairly priced it, so further dissemination of a bad event with hard-to-misinterpret signals will not change abnormal returns by very much. This conclusion is supported by the regression results in Table 9. In Table 7, I find that PrePost lowers the absolute value of abnormal equity returns for monthly sales announcements. Because the data contains both positive and negative, it hard to understand what is going on in the results. The regressions in Table 9 provide some insight. As mentioned above, PrePost is only significant for the positive returns in monthly sales. Here, the results suggest that the advent of Twitter actually lowers the abnormal returns because it helps disseminate information. We do not see a similar in Panel A of Table 7 for the negative subset, supporting the idea that for negative returns in the Post-Twitter period, the market either has likely already priced it and the firm does not further manage content due to fear of litigation costs. Hence, unsophisticated traders do not trade in a different direction from sophisticated investors.

In Figure 12, we see as the number of positive words in the tweets increase, we see an increase in abnormal returns. Then, as the number of tweets increase, we again see a drop in returns. Most likely this is because for tweets that have high content, the return is near 0 because it should have actually have been negative and it was “managed” upwards to reach 0.

Figure 13 shows that returns were less negative in the Post-Twitter Period than in the Pre-Twitter Period. It is possible this is because either the Post-Twitter Period contains more bad news overall or else this can be attributed to information management of firms in the Post-Twitter Period through their tweets. The Post-Twitter period here has 25% true “negative” events while the Pre-Twitter Period had 62% negative events. This may be a big reason as to why we see a difference in returns – so there is no evidence that firms manage the information content. Graphically, we also see that the amount of announcement manipulation went up in the Twitter period. This is probably because there were more positive events in the Pre Period, which leads to a higher Up_Net overall. Regression results show that for the negative subset, including the % increase in the Tweet led to higher returns.

Further analysis that there is indeed a difference between the two types of events. For product recalls, we see that Prepost and all Twitter variables are insignificant; the only relevant variable is impact. For monthly sales with positive returns, disseminating through Twitter actually lowers the returns. For negative product recalls, disseminating through Twitter actually makes returns less negative through the managing information (i.e. including positive sales increases in certain divisions in the Tweet). The cancellation of the two types of effects is why overall we saw that Twitter does not seem to overall affect price volatility or volume.

[Insert Table 10 here]

Influential Points

In order to see whether there were any influential points in the dataset, I plotted the Cook’s distance for the model for returns for both product recalls and monthly sales announcements. For product recalls, 3 points were removed. For monthly sales announcements, 6 significantly influential points were removed.

Table 10 exhibits the regression results once the influential points were removed.
Panel B. Before the influential points were removed, the regressors selected through model selection were Salesper and Salesnum[9]. After the influential points were removed, the selected regressors were beta, tone, UpNet, Salesper, Salesnum, Retweeet, per, and Prepost. Out of those, beta, UpNet, and Salesnum were significant. Hence, the removal of influential points did not change my conclusion about the role of Prepost. For log volume, the regressors selected after removing influential points also changed. Before, Prepost, salesnum, and beta were significant. After removing influential points, link also became a significant variable. This meant that while a company-initiated Tweet actually decreases the amount of trading, including a link increases the amount of trading volume. For the absolute value of abnormal returns, PrePost was significant before influential points were removed. After removal, no variable was significant – only the intercept was significant. This hints that must be other variables that prove significant for returns.

Panel A. For product recalls, there was no significant regressor for returns. However, the removal of influential points led to ProblemInfo (whether the company discloses the reason for the recall on Twitter) being significant. The negative coefficient suggests that as firms disclose the reason for the recall on Twitter, the abnormal returns are negative. This is consistent with the original hypothesis. Before the points were removed, the market cap and the number of times a company tweet is retweeted by the public were significant. Afterwards, there was no significant regressor. The significant intercept here also hints that there

[9] Salesper- the % increase or decrease of monthly sales compared to the same month the previous year
Salesnum- the gross number of sales for that month

is an explanation for the abnormal returns outside of the given regressors. It is likely to be unrelated to the company-initiated tweet. For volume, Prepost and Retweet were significant. However, the removal of influential points led nothing to be significant after.

Whether these points should have been removed or not is controversial. For one, removing so many points from an already limited data-set would be removing valuable information. Secondly, none of the removed points had a Cook's distance of more than 1, indicating that it could still perform fine for the regression. Another reason why the results above for log(volume) may not have been valid is because when I plotted the Cook's distance for return, there were highly influential points. Removing them may have made for a better analysis. However, when I plotted Cook's distance for the former model for log(volume), there were very few influential points. Hence, removing the points that affected the model for return would possibly improve the model for return, but it would certainly hurt the model for volume. Perhaps that is why we saw the volume for both product recalls and monthly sales announcements become insignificant after the removal of the points.

V. Further Analysis and Potential Issues

One significant issue with the analysis above is that variable PrePost, the variable measuring voluntary disclosure, is potentially endogenous variable. This is especially true for product recalls where the ability to manage information within both the announcement and the Tweet is limited because of potential litigation consequences. This can be seen from the data; firms tended to limit the management of the information for disclosures, even within Tweets. Because the sample data contains only product recalls that were tweeted about post 2008, I face a bias because those firms chose to disclosure for reasons that may or may not be controlled for in the data. Out of all the firms that had product recalls after 2008, very few of them actually mention it on their Twitter. Perhaps this is because firms are well aware of the fact that this could have detrimental consequences on price volatility and trading volume, so they purposely do not disseminate through Twitter. Perhaps this is an explanation as to why PrePost was not significant in the product recall data for both returns and the absolute value of the returns; the firms who disclose on Twitter product recalls in the first place are under the belief that the market will not significantly react to it. Hence, we do not see a significant change in price volatility and only a minor, albeit significant, increase in trading volume. This is especially true as out of all the product recalls that car companies such as GM, Nissan, Honda, and Ford faced after 2008, only a small portion of them were discussed on Twitter. The ones that were disclosed were most likely disclosed because the impact was very high, and the company feared litigation risk and/or believed that the cost of disseminating to the consumer was lower than the cost of an overreaction to the product recall from investors. An example would be the recent GM ignition recall in which GM was fire for. However, some of the less significant recalls may not be disclosed on Twitter due to fear that it would cause unnecessary price volatility from unsophisticated investors. On top of having a selection bias with the events, there is a selection bias with the companies selected as well. Only a small percentage of all firms on Twitter tweet about financial information on their page. For example, firms like Dunkin' Donuts and Urban Outfitters limit their tweets to information about new products and sales.

Another potential issue is market cap. Companies with a small market cap and a smaller investor following may abstain from disclosing on Twitter as the reach of their Tweets would not be to a broad enough audience; this would be a direct violation of Regulation FD. Because smaller cap companies tend to be firms with lower visibility, these would be the firms that actually might experience higher price volatility and trading volume as a result of disclosing information on Twitter. This may be another reason why I did not see an significant change in abnormal returns in the post-Twitter period; the events featured may have been high profile recalls, and unsophisticated traders were likely to find out about this recall anyway through the news, etc. This is especially true for car recalls. Additionally, we may have seen less controlling of information about product recalls on the Tweets because of the language; one of the advantages of Tweeting is to be able to use words that appeal more to emotion (e.g, "Ford took the crown in sales this month!"). This is hard to do with a product recall as firms are limited in their ability to use positive words in connection with a product recall.

A potential way to fix this problem would be through collecting data on firms who had product recalls but did not disclose on Twitter. While there is a bias here too as it is impossible to observe what the firm avoided by not disclosing on Twitter, we can still observe the disclosure decision on Twitter with regards to both the market cap of the firm and the size of the impact. Perhaps controlling for the number of followers in the period would also allow us to look into the hypothesis about Regulation FD given above.

The monthly sales data faces less of an endogeneity problem than the product recalls data. This is because out of the two firms observed in the 4 year period, GM and Honda, both disclosed on Twitter every month regardless of the type of event. A potential reason for that could be that the potential to manipulate the information content of both announcements as well as Tweets is much higher for monthly sales. This is likely because there is less litigation risk from doing so and firms also have more "material" to work with – both GM and Honda have several divisions and types of cars; even when it is an overall negative or positive monthly sales events, there is bound to be at least one division or car that did do well in sales. This can be emphasized in both the announcement and in the car. However, the problem that monthly sales data faced most likely lay in the frequency of the event. Perhaps unsophisticated investors might see that that GM or Ford consistently is very positive about sales on their Twitter page, and they see this every month. This removes the shock factor in the news; they may not trade on this information as their expectations were already set from the month before. A way to get around this may be to control for the nature of the monthly sales data one or two months before. This would mean to use lagged EVENTS on t-1 and t-2 as a proxy for the "expectation" of the unsophisticated investor.

Perhaps the problem in the results lay in my belief about the unsophisticated investor. Unsophisticated investors may know that they are unsophisticated; hence, they may rely on opinions of more sophisticated investors to make their trade. This is not a hard task as the availability of information and opinions and articles online have increased significantly after 2008. Perhaps the reason we see no change in price volatility is because after a unsophisticated investor, who is likely to be risk averse, reads about a product recall or a monthly sales announcement on Twitter, may go to other sources to corroborate or find out more based on what they read. After they read news articles from WSJ or other news sources, their interpretation of that event may converge to what the market believes overall. Hence, they trade in the same direction (or, the distribution of the direction of their trade is the same as sophisticated investors). Perhaps this can be further researched by looking at the quartiles of the abnormal returns and how many fall in each quartile for both the post and the pre period. Hence, we would observe no change in abnormal returns cross-sectionally but an increase in trading volume.

VI. Conclusion

This study empirically examines whether firms' use of Twitter as a viable mode of dissemination of corporate information affects trading behavior. Twitter enables the firm to manage the information content of the Tweet because of its 140 characters. In addition, the marginal investor who may read this "managed" tweet and react accordingly is most likely unsophisticated since Twitter is an information-pushing platform rather than an information-pulling platform such as Bloomberg. As a result, I hypothesize that dissemination through Twitter leads to less negative abnormal returns for product recalls and bad monthly sales events. Furthermore, I hypothesize that the abnormal trading volume will increase after the advent of Twitter due to the entrance of unsophisticated traders.

To do so, I look at abnormal returns and abnormal trading volume in a 2 day period after the announcement of a product recall and a monthly sales announcement. I compare differences between the abnormal returns and trading volume that tweet about the event on their company Twitter in addition to sending out a newswire as opposed to when a firm would only post a newswire before the firm created a company Twitter account.

I document a relationship between the log of abnormal volume and firm-initiated announcement tweets for both product recalls and monthly sales announcements. However, the relationship is actually a negative one suggesting that either trading volume actually decreased or that the rate of increase of trading volume decreased. The second one is the most likely explanation due to the fact that the sample takes place shortly after the financial crisis where many retail investors were not in a place to casually trade on information, so the rate at which trading volume increases took a hit. I also document a negative relationship between the absolute value of abnormal equity returns and firm-initiated announcement tweets for monthly sales data. Further analysis shows that this relationship holds true only for monthly sales events with positive returns and not for those with negative returns. This suggests that it is because Twitter serves to lower information asymmetry, resulting in less positive returns than before. This also means that either firms fail to successfully manage information content or else do not attempt to through Twitter. Hence, I conclude that Twitter as a mode of dissemination does lead to higher market efficiency as opposed to asset mispricing as a result of unsophisticated traders trading away from the fundamental value.

Table 1: Summary Statistics using all data

Panel A provides descriptive statistics for product recall announcements and Panel B provides descriptive statistics for monthly sales announcements. The mean return for monthly sales announcements is higher than product recalls, as expected. The absolute return between the two types is very similar, which is surprising given that monthly sales announcements have positive events, so we might expect a significantly abnormal higher return. The volume of trade around a product recall is much higher than a monthly sales announcement. This could potentially be explained by the entrance of unsophisticated traders. Other common variables they share are link – 33.7% of all product recall data contained a link. 43% of all monthly sales announcements contain a link.

Panel A: Summary Statistics for Product Recalls

Statistic	N	Mean	St. Dev.	Min	Max
Return	83	0.0001	0.033	-0.118	0.111
AbsReturn	83	0.022	0.024	0.0002	0.118
volume	83	0.139	1.487	-3.020	5.000
Prepost	83	0.518	0.503	0	1
Impact	83	10,855.530	20,654.300	0	107,000
Link	83	0.337	0.476	0	1
Retweet	83	6.193	10.159	0	47
Tone	83	0.133	0.341	0	1
ProblemInfo	83	0.253	0.437	0	1
Favorites	83	1.000	2.513	0	12
beta	83	1.000	0.390	0.330	1.940
marketcap	83	73.596	66.391	1.190	244.180

Panel B: Summary Statistics for Monthly Sales Announcements

Statistic	N	Mean	St. Dev.	Min	Max
Return	93	0.003	0.038	-0.115	0.151
AbsReturn	93	0.028	0.027	0.001	0.151
Volume	93	0.004	0.038	-0.120	0.150
Prepost	93	0.591	0.494	0	1
beta	93	1.196	0.235	0.780	1.790
tone	93	0.882	0.325	0	1
UpNet	93	16.409	9.744	-3	39
salesper	93	0.040	0.153	-0.350	0.600
salesnum	93	199,879.600	105,129.600	45,948	739,578
Retweet	93	25.258	36.143	0	179
Link	93	0.430	0.498	0	1
NetSent	93	2.473	3.516	0	17
TweetNum	93	2.108	2.838	0	15
Per	93	0.430	0.498	0	1
Num	93	0.312	0.466	0	1

Product Recalls

Table 2: Summary Statistics by Equity Returns for Product Recalls

Panel A provides descriptive statistics for the negative returns in the sample. Panel B provides descriptive statistics for the positive returns in the sample. It is interesting that the mean of the negative sample almost perfectly cancels out the mean of the positive sample. Positive returns had slightly lower average ln(volume). The mean market cap of firms with negative returns around a product recall is 80 M compared to 67 M for positive returns. We also see that firms with positive returns also have a higher number of retweets and only 57% of them contain a link compared to the 70% of negative returns that contain a link. For negative returns, 58.3% of firms had provided the reason for the negative recall compared to the 36.8% of firms with positive returns that had provided it. The most interesting statistic in this dataset is the fact that the mean impact for the positive returns is actually 12 B compared to 9 B for negative returns.

Panel A: Summary Statistics for Negative Returns

Statistic	N	Mean	St. Dev.	Min	Max
Return	40	-0.023	0.025	-0.118	-0.001
AbsReturn	40	0.023	0.025	0.001	0.118
LnVolume	40	0.158	1.592	-3.020	4.550
PrePost	40	0.600	0.496	0	1
marketcap	40	80.262	66.304	1.720	244.180

Panel B: Summary Statistics for Positive Returns

Statistic	N	Mean	St. Dev.	Min	Max
Return	43	0.022	0.024	0.0002	0.111
AbsReturn	43	0.022	0.024	0.0002	0.111
LnVolume	43	0.121	1.400	-2.790	5.000
marketcap	43	67.396	66.647	1.190	244.180
PrePost	43	0.442	0.502	0	1

Panel C: Summary Statistics for Twitter Variables - Negative Returns

Statistic	N	Mean	St. Dev.	Min	Max
Retweet	24	10.750	9.918	0	41
Link	24	0.708	0.464	0	1
neg.ProblemInfo	24	0.583	0.504	0	1
Impact2	24	9,500.750	16,116.000	0	54,464

Panel D: Summary Statistics for Twitter Variables - Positive Returns

Statistic	N	Mean	St. Dev.	Min	Max
Retweet	19	13.474	13.251	0	47
Link	19	0.579	0.507	0	1
pt.ProblemInfo	19	0.368	0.496	0	1
Impact2	19	12,410.630	28,845.160	0	107,000

Table3: Summary Statistics by Time Period for Product Recall

Panel A provides descriptive statistics for the Pre-Twitter subsample. Panel B provides descriptive statistics for the Post-Twitter subsample. It is interesting to note that the average return for the post-Twitter sample is -.002 while the average return for the Pre-Twitter sample is .003. Additionally, it actually appears as though the mean trading volume has actually decreased in the Post-Twitter sample which directly contradicts the hypothesis. The fact that the average mean impact is a similar number in both periods is good to know; this implies that there is no cross-sectional difference between the "seriousness" of the recall. Additionally, the average market cap of the Post Twitter period is substantially higher than in the Pre Twitter period. This may imply that there is substantial bias in the data collection; only large companies may Tweet about recalls on Twitter.

Panel A: Summary Statistics for Pre-Twitter Subsample

Statistic	N	Mean	St. Dev.	Min	Max
Return	40	0.003	0.035	-0.096	0.111
AbsReturn	40	0.023	0.026	0.0002	0.111
LnVolume	40	0.248	1.411	-3.020	4.550
InfoContent	40	0.000	0.000	0	0
Impact2	40	10,929.730	18,906.320	0	106,000
marketcap	40	45.647	45.869	2.460	194.220

Panel B: Summary Statistics for Post-Twitter Subsample

Statistic	N	Mean	St. Dev.	Min	Max
Return	43	-0.002	0.032	-0.118	0.082
AbsReturn	43	0.022	0.023	0.001	0.118
LnVolume	43	0.037	1.563	-2.790	5.000
Retweet	43	11.953	11.445	0	47
Link	43	0.651	0.482	0	1
Tone	43	0.256	0.441	0	1
InfoContent	43	0.488	0.506	0	1
Impact2	43	10,786.510	22,382.100	0	107,000
marketcap	43	99.596	72.222	1.190	244.180

Monthly Sales Announcements

Table 4: Summary Statistics by Returns for Monthly Sales

Panel A provides descriptive statistics for monthly sales announcements that resulted in negative abnormal returns. Panel B provides descriptive statistics for monthly sales announcements that resulted in positive returns. Panel C gives summary statistics for the Twitter variables for negative returns, and Panel D provides summary Statistics for the Twitter for positive returns. This was broken out this way because only around half the data points have data for the Twitter Variables because they belong in the Post-Period results.

We note that just as for product recalls, the average return for negative returns versus positive returns almost perfectly cancels each other out. Positive returns have a higher log volume than negative returns. For positive returns, the salesnum is substantially higher than for negative returns. The number of retweets of the company tweet is also higher for positive returns. Additionally, the positive return events have a higher number of tweets posted by the company than the negative.

Panel A: Summary Statistics for Negative Returns

Statistic	N	Mean	St. Dev.	Min	Max
Return	43	-0.027	0.025	-0.115	-0.002
AbsReturn	43	0.027	0.025	0.002	0.115
Volume	43	-0.024	0.026	-0.120	0.010
Prepost	43	0.465	0.505	0	1
beta	43	1.176	0.239	0.780	1.790
tone	43	0.860	0.351	0	1
UpNet	43	16.860	10.020	-3	39
salesper	43	0.013	0.152	-0.350	0.450
salesnum	43	184,196.000	100,047.800	45,948	413,473

Panel B: Summary Statistics for Postive Returns

Statistic	N	Mean	St. Dev.	Min	Max
Return	50	0.028	0.029	0.001	0.151
AbsReturn	50	0.028	0.029	0.001	0.151
Volume	50	0.027	0.029	-0.020	0.150
Prepost	50	0.700	0.463	0	1
beta	50	1.213	0.233	0.820	1.640
tone	50	0.900	0.303	0	1
UpNet	50	16.020	9.584	-2	38
salesper	50	0.063	0.151	-0.310	0.600
salesnum	50	213,367.600	108,489.700	71,031	739,578

Panel C: Summary Statistics for Twitter Variables - Negative Returns

Statistic	N	Mean	St. Dev.	Min	Max
Retweet	20	32.650	30.538	4	114
Link	20	0.650	0.489	0	1
NetSent	20	3.900	4.266	1	17
TweetNum	20	3.150	3.249	1	15
Per	20	0.650	0.489	0	1
Num	20	0.450	0.510	0	1

Panel D: Summary Statistics for Twitter Variables - Positive Returns

Statistic	N	Mean	St. Dev.	Min	Max
Retweet	35	48.457	41.405	7	179
Link	35	0.771	0.426	0	1
NetSent	35	4.343	3.412	0	13
TweetNum	35	3.800	2.709	1	11
Per	35	0.771	0.426	0	1
Num	35	0.571	0.502	0	1

Table 5: Summary Statistics by Period – Monthly Sales

Panel A provides descriptive summary statistics for the Pre-Twitter period. Panel summarizes key statistics in the Post-Twitter period. We see that the mean return for monthly sales in the Post-Twiter period is higher than in the Pre-Period. Could this be explained by disemmination through Twitter? Trading volume in the Post-Period is higher than in the Pre-Period? We also note that tone has changed substantially in the Post-Twitter period – companies rarely post a neutral sales announcement now. We also note that mean gross sales has decreased in the Post-Twitter period. Additionally, 75% of all Tweets contain links to the original corporate announcement and 73% contain the percentage of sales increase compared to the prior year in the Tweet. 53% of all tweets contain information about the gross number of sales that month.

Panel A: Summary Statistics fo the Pre-Twitter Period

Statistic	N	Mean	St. Dev.	Min	Max
Return	38	-0.005	0.051	-0.115	0.151
AbsReturn	38	0.037	0.034	0.001	0.151
Volume	38	-0.004	0.049	-0.120	0.150
Prepost	38	0.000	0.000	0	0
beta	38	1.140	0.232	0.790	1.790
tone	38	0.737	0.446	0	1
UpNet	38	13.079	8.744	-3	39
salesper	38	-0.042	0.142	-0.350	0.220
salesnum	38	211,448.300	117,699.700	45,948	413,473

Panel B: Summary Statistics for Post-Twitter Period

Statistic	N	Mean	St. Dev.	Min	Max
Return	55	0.008	0.026	-0.052	0.070
AbsReturn	55	0.021	0.018	0.002	0.070
Volume	55	0.009	0.026	-0.050	0.070
Prepost	55	1.000	0.000	1	1
beta	55	1.234	0.231	0.780	1.630
tone	55	0.982	0.135	0	1
UpNet	55	18.709	9.807	-2	39
salesper	55	0.097	0.134	-0.100	0.600
salesnum	55	191,886.700	95,804.340	91,531	739,578
Link	55	0.727	0.449	0	1
NetSent	55	4.182	3.712	0	17
TweetNum	55	3.564	2.904	1	15
Per	55	0.727	0.449	0	1
Num	55	0.527	0.504	0	1

Table 6: Summary Statistics by Type of Event for Monthly Sales

Panel A provides descriptive statistics for negative monthly sales. This means that compared to the same month the previous year, total sales decreased. Panel B provides descriptive statistics for positive monthly sales or months that did better compared to the same month one year before. We note that that for negative monthly sales, the tone of the headline of the announcement was neutral 25% of the time compared to for positive monthly sales where it was neutral only 5% of the time. This makes sense – during negative sales months, companies would not want to have the headline of the announcement by negative. Instead, they opt for a neutral headline such as "GM releases May sales". Additionally, we note that for negative monthly sales announcement, a link is provided in the tweet 80% of the time, which is interesting. Additionally, I expected the NetSent to be higher in the negative monthly sales than for the positive monthly sales, which would point to firms managing information content. In order words, they distract the investor from the negative sales event by overemphasizing the positive aspects of sales that month, although overall sales decreased.

Panel A: Summary Statistics for Negative Monthly Sales

Statistic	N	Mean	St. Dev.	Min	Max
Return	33	-0.006	0.051	-0.115	0.151
AbsReturn	33	0.037	0.035	0.003	0.151
Volume	33	-0.006	0.052	-0.120	0.150
Prepost	33	0.303	0.467	0	1
GM	33	0.455	0.506	0	1
Beta	33	1.171	0.228	0.790	1.790
Tone	33	0.758	0.435	0	1
Up_Net	33	11.636	9.842	-3	39
Salesper	33	-0.082	0.138	-0.350	0.390
SalesNum	33	192,460.400	112,129.000	45,948	410,332

Panel B: Summary Statistics for Positive Monthly Sales

Statistic	N	Mean	St. Dev.	Min	Max
Return	60	0.008	0.028	-0.052	0.073
AbsReturn	60	0.022	0.019	0.001	0.073
Volume	60	0.009	0.027	-0.050	0.070
Prepost	60	0.750	0.437	0	1
GM	60	0.517	0.504	0	1
Beta	60	1.209	0.240	0.780	1.630
Tone	60	0.950	0.220	0	1
Up_Net	60	19.033	8.704	-2	39
Salesper	60	0.108	0.114	-0.060	0.600
SalesNum	60	203,960.200	101,818.900	93,626	739,578

Panel C: Summary Statistics for Twitter Variables - Negative Monthly Sales

Statistic	N	Mean	St. Dev.	Min	Max
Link	10	0.800	0.422	0	1
NetSent	10	2.000	1.414	0	4
TweetNum	10	2.300	1.160	1	4
Per	10	0.600	0.516	0	1
Num	10	0.500	0.527	0	1

Panel D: Summary Statistics for Twitter Variables - Positive Monthly Sales

Statistic	N	Mean	St. Dev.	Min	Max
Link	45	0.711	0.458	0	1
NetSent	45	4.667	3.896	0	17
TweetNum	45	3.844	3.104	1	15
Per	45	0.756	0.435	0	1
Num	45	0.533	0.505	0	1

Table 7: Regression results using Model Selection using all data.

Panel A shows the regression results for product recalls. Here, the full sample of 84 recalls was used. Panel B shows regression results for monthly sales. The full sample of 93 monthly sales announcements was used. For volume for monthly sales and product recalls, PrePost was a significant regressor. This means that all else held constant, firm-initiated tweets made a difference in the trading volume. However, the constant is very significant for absolute value of the returns for both product recalls and the monthly sales.

Regression Results for Product Recalls

	Dependent variable:		
	volume	Return	AbsReturn
	(1)	(2)	(3)
Prepost	-1.000**		
	(0.489)		
Retweet	0.037*		0.0005*
	(0.020)		(0.0002)
ProblemInfo	0.708		
	(0.461)		
Impact		-0.00000	
		(0.00000)	
marketcap			-0.0001**
			(0.00004)
Constant	0.252	0.001	0.024***
	(0.233)	(0.004)	(0.004)
Observations	83	83	83
R^2	0.060	0.025	0.078
Adjusted R^2	0.024	0.013	0.055
Residual Std. Error	1.471 (df = 79)	0.030 (df = 81)	0.022 (df = 80)
F Statistic	1.683 (df = 3; 79)	2.041 (df = 1; 81)	3.387** (df = 2; 80)

Regression Results for Monthly Sales

	Dependent variable:		
	Return	AbsReturn	volume
	(1)	(2)	(3)
Prepost		-0.010*	-1.507***
		(0.006)	(0.436)
UpNet		-0.001**	
		(0.0003)	
salesper	0.051*	-0.024	
	(0.026)	(0.020)	
salesnum	0.00000**		-0.00000**
	(0.00000)		(0.00000)
Link			0.837*
			(0.433)
beta		0.010	2.508***
		(0.012)	(0.685)
Constant	-0.014	0.033**	-1.157
	(0.009)	(0.015)	(0.761)
Observations	93	93	93
R^2	0.074	0.178	0.210
Adjusted R^2	0.053	0.140	0.174
Residual Std. Error	0.037 (df = 90)	0.025 (df = 88)	1.401 (df = 88)

Table 8: Regression results by using negative returns and positive returns for product recalls

Panel A shows regression results for the negative subset of product recalls. This means that I stratified the recalls into two groups: those that had negative abnormal returns and those with positive abnormal returns. Under the assumption that those with positive returns exceeded expectation and the ones with negative abnormal returns did not meet expectation, I look at how company-initiated tweets may have affected abnormal returns and volume.

Panel A: Regression Results for Product Recalls - Negative Returns

	Dependent variable:	
	Return	LnVolume
	(1)	(2)
Impact	-0.000^{*}	
	(0.000)	
nRetweet		0.044
		(0.034)
PrePost		-0.422
		(0.630)
Constant	-0.019^{***}	0.126
	(0.004)	(0.399)
Observations	40	40
R^2	0.080	0.045
Adjusted R^2	0.056	-0.007
Residual Std. Error	0.024 (df = 38)	1.598 (df = 37)
F Statistic	3.316^{*} (df = 1; 38)	0.868 (df = 2; 37)

Panel B: Regression Results for Product Recalls - Positive Returns

	Dependent variable:	
	Return	LnVolume
	(1)	(2)
marketcap	-0.0001^{*}	
	(0.0001)	
pLink		-1.320^{*}
		(0.664)
PrePost	0.012	-0.272
	(0.009)	(0.546)
pProblemInfo	-0.016	
	(0.013)	
Impact2	0.00000	
	(0.00000)	
pRetweet		0.048^{*}
		(0.027)
Constant	0.026^{***}	0.329
	(0.006)	(0.274)
Observations	43	43
R^2	0.113	0.143
Adjusted R^2	0.020	0.078
Residual Std. Error	0.023 (df = 38)	1.345 (df = 39)
F Statistic	1.213 (df = 4; 38)	2.177 (df = 3; 39)

Table 9: Regression results for Positive and Negative Returns for Monthly Sales

Panel A shows regression results for the negative subset of monthly sales. This means that I stratified the monthly sales into two groups: those that had negative abnormal returns and those with positive abnormal returns. Under the assumption that those with positive returns exceeded expectation and the ones with negative abnormal returns did not meet expectation, I look at how company-initiated tweets may have affected abnormal returns and volume.

Panel A: Regression Results for Monthly Sales - Negative Subset

	Dependent variable:	
	Return	Volume
	(1)	(2)
UpNet	0.001*	
	(0.0004)	
salesper	0.067**	-0.072
	(0.026)	(1.742)
salesnum	0.00000**	-0.00001**
	(0.00000)	(0.00000)
Link	-0.015	
	(0.010)	
Per	0.018*	
	(0.010)	
Num		-1.726**
		(0.677)
GM	-0.017*	0.987
	(0.010)	(0.741)
Prepost	0.006	-0.079
	(0.011)	(0.647)
Constant	-0.055***	2.220***
	(0.009)	(0.650)
Observations	43	43
R^2	0.483	0.243
Adjusted R^2	0.380	0.141
Residual Std. Error	0.020 (df = 35)	1.473 (df = 37)
F Statistic	4.680*** (df = 7; 35)	2.379* (df = 5; 37)

Panel B: Regression Results for Monthly Sales - Positive Subset

	Dependent variable:	
	Return	Volume
	(1)	(2)
Prepost	-0.023**	
	(0.009)	
beta	0.034*	
	(0.017)	
salesper		-0.056**
		(0.027)
Constant	0.003	0.031***
	(0.020)	(0.004)
Observations	50	50
R^2	0.157	0.082
Adjusted R^2	0.121	0.063
Residual Std. Error	0.027 (df = 47)	0.028 (df = 48)
F Statistic	4.377** (df = 2; 47)	4.267** (df = 1; 48)

Table 10: Regression Results after Removing Influential Points

Panel A provides regression results after removing influential points for product recalls (6 points with a large Cook's distance were removed). Panel B provides regression results after removing influential points for monthly sales announcements. The new regression results are somewhat different from the previous regression. Prepost is still only significant for volume, and it still shows that company initiated tweets actually reduces the abnormal log (volume). However, including a link to the original announcement actually increases volume.

Panel A: Regression Results after Removing Outliers for Product Recalls

	Dependent variable:		
	Return	AbsReturn	volume
	(1)	(2)	(3)
ProblemInfo	-0.014*		0.723
	(0.007)		(0.469)
marketcap		-0.0001	
		(0.00005)	
Retweet		0.0004	
		(0.0004)	
Tone		-0.008	
		(0.011)	
Prepost		-0.006	-0.669
		(0.009)	(0.413)
Link		0.006	
		(0.009)	
Favorites		-0.0003	
		(0.001)	
Constant	0.004	0.026***	0.248
	(0.004)	(0.004)	(0.234)
Observations	80	80	80
R^2	0.045	0.089	0.039
Adjusted R^2	0.033	0.014	0.014
Residual Std. Error	0.029 (df = 78)	0.021 (df = 73)	1.481 (df = 77)
F Statistic	3.682* (df = 1; 78)	1.187 (df = 6; 73)	1.569 (df = 2; 77)

Panel B: Regression Results after Removing Influential Points for Monthly Sales Announcements

	Dependent variable:		
	Return	AbsReturn	volume
	(1)	(2)	(3)
Beta	-0.038**		2.696***
	(0.019)		(0.751)
tone	0.026**		
	(0.012)		
UpNet	-0.001**	-0.0002	
	(0.0004)	(0.0002)	
Salesper	0.039		
	(0.025)		
SalesNum	0.00000***		-0.00001**
	(0.00000)		(0.00000)
Retweet	0.0001		
	(0.0001)		
Per	0.009		
	(0.010)		
beta		0.007	
		(0.011)	
salesnum		-0.000	
		(0.00000)	
Link			0.940**
			(0.432)
Prepost	0.006	-0.007	-1.502***
	(0.010)	(0.005)	(0.465)
Constant	-0.003	0.024*	-1.220
	(0.021)	(0.012)	(0.767)
Observations	87	87	87
R^2	0.228	0.044	0.185
Adjusted R^2	0.149	-0.003	0.145
Residual Std. Error	0.028 (df = 78)	0.020 (df = 82)	1.379 (df = 82)
F Statistic	2.879*** (df = 8; 78)	0.943 (df = 4; 82)	4.650*** (df = 4; 82)

Positive Equity Returns – Product Recalls

Figure 1: Return vs. Impact

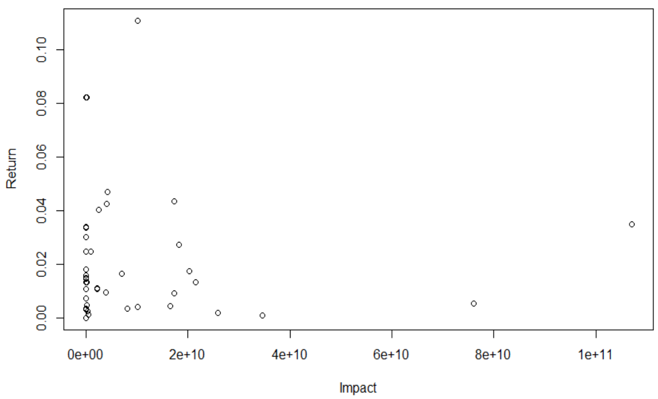


Figure 2: Return vs. Info Content

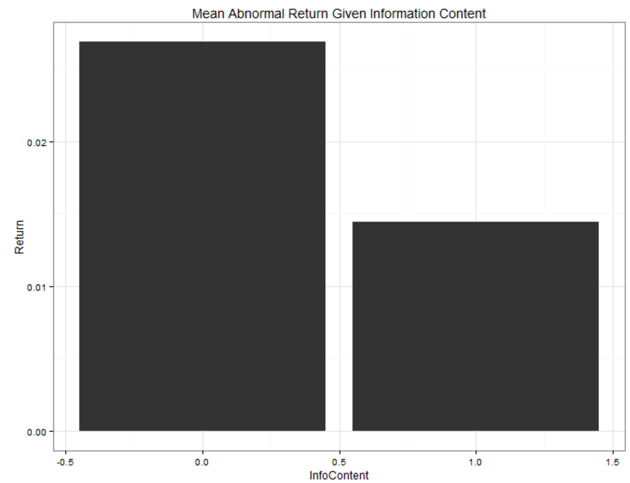


Figure 3: Return vs. Number of Retweets

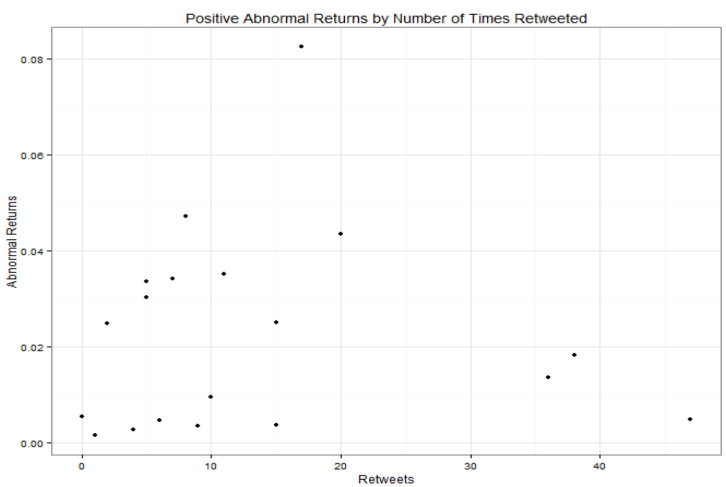
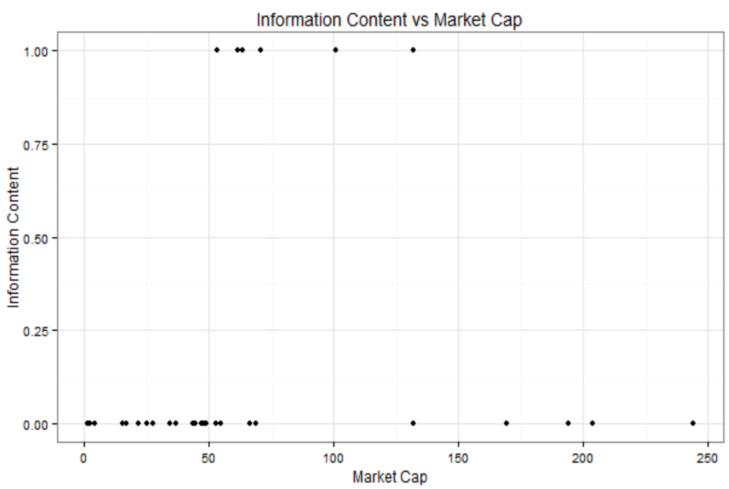


Figure 4: InfoContent vs Market Cap



Negative Equity Returns – Product Recalls

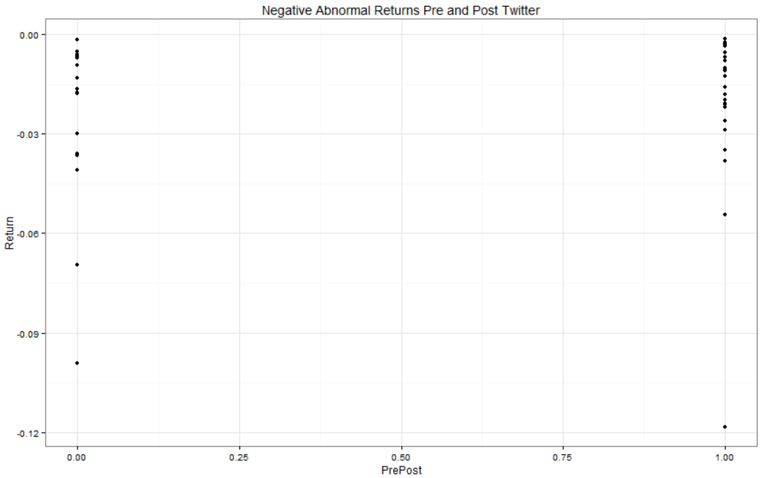


Figure 5: Return vs. Pre and Post Twitter

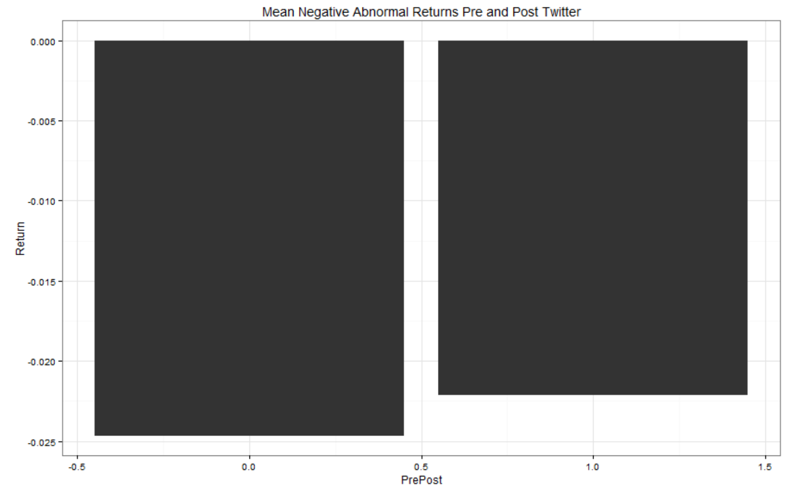


Figure 6: Average Return per Period

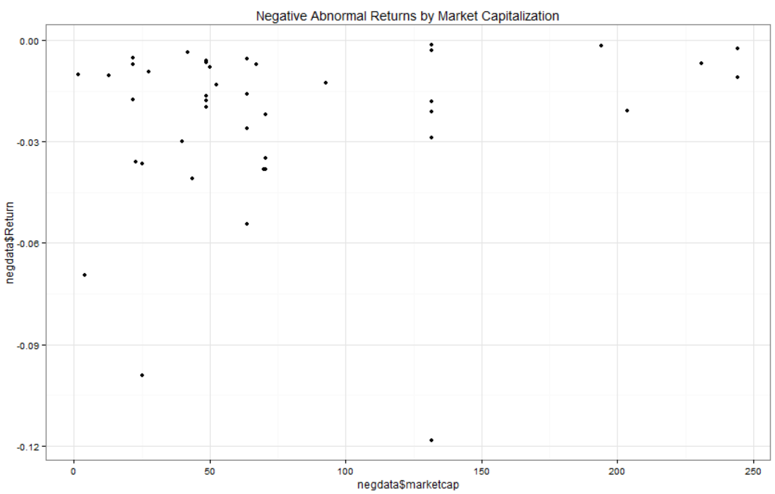


Figure 7: Return vs. Market Cap

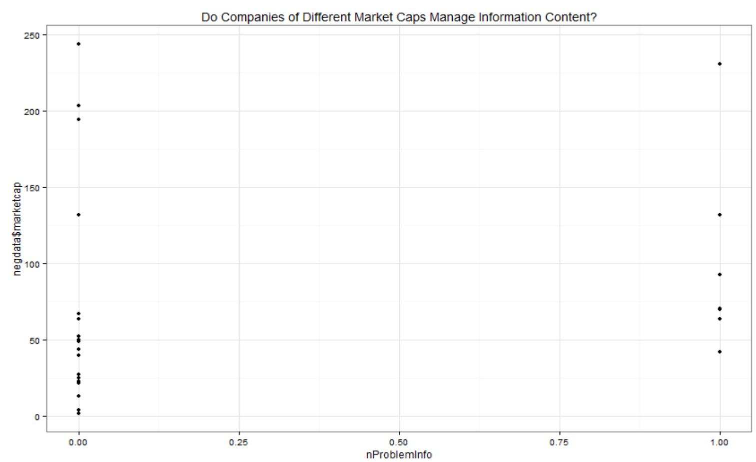


Figure 8: Market Cap vs Info Content

Positive Equity Returns – Monthly Sales

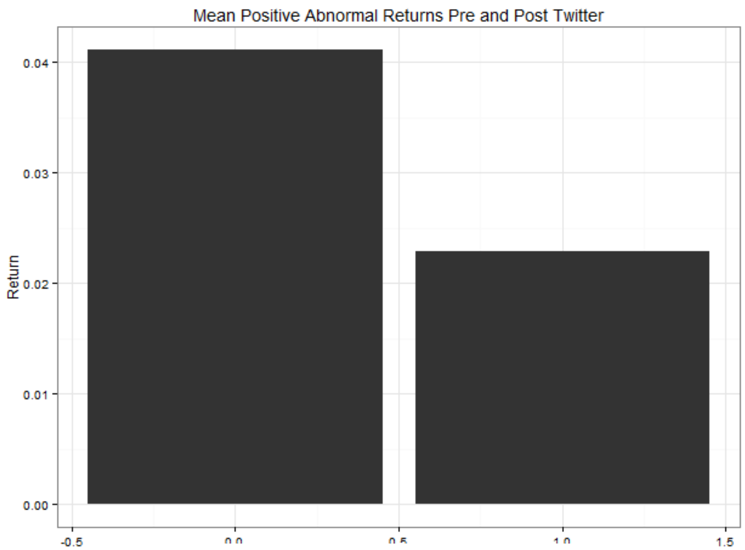


Figure 9: Mean Return vs Time Period

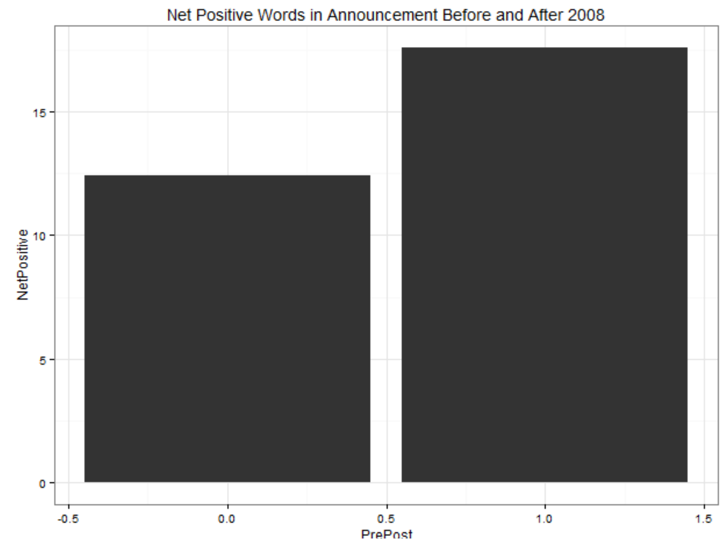


Figure 10: Net Positive Words vs Time Period

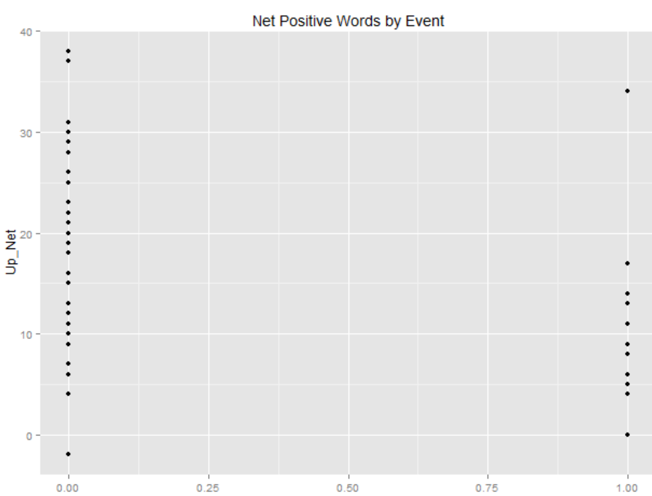


Figure 11: Net Positive Words by Event

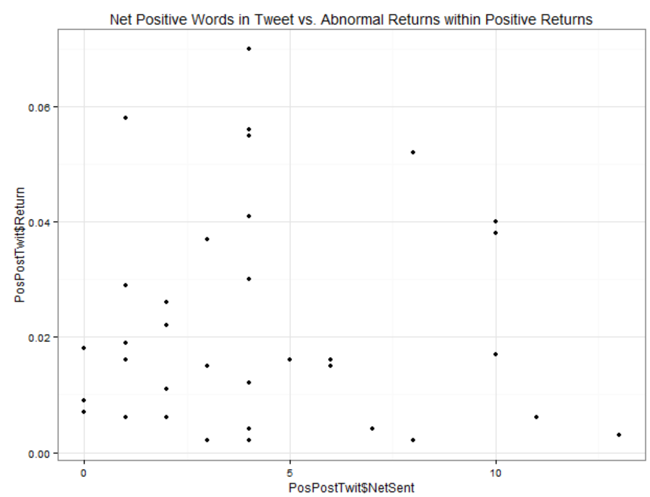


Figure 12: Returns vs Net Positive Words

Negative Equity Return – Monthly Sales

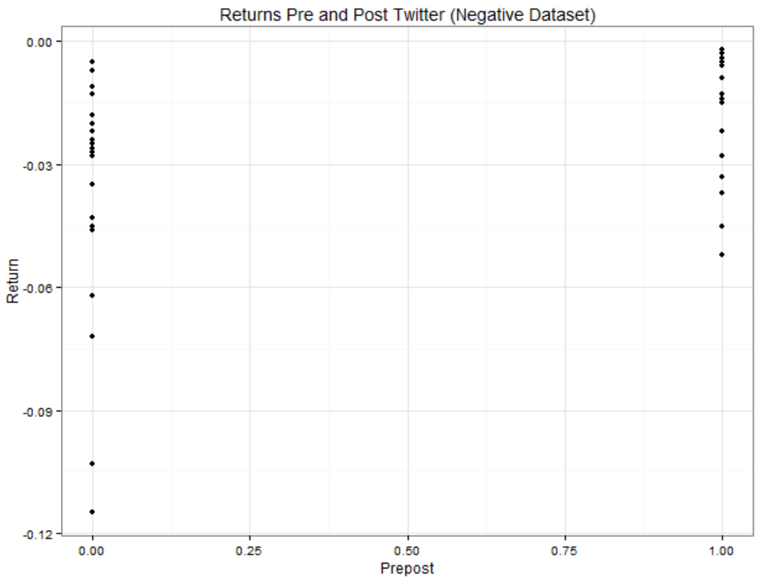


Figure 13: Returns vs Time Period

References

Antweiler, W. and M. Frank. 2004. Is all that talk just noise? The information content of stock message boards. *The Journal of Finance* 59 (3): 1259-1294.

Barber, B., T. Odean, and N. Zhu. 2009. Do retail trades move markets? *The Review of Financial Studies* 22 (1): 151-186.

Bensinger, G. "Netflix Cues up Facebook, Twitter for Disclosure." *The Wall Street Journal* April 10, 2013.

Berger, P. and R. Hann. 2007. Segment profitability and the proprietary and agency costs of disclosure. *The Accounting Review* 82 (4): 869-906.

Bloomfield, R., OHara, M., and Saar, G. (2009). How noise trading affects markets: An experimental analysis: Review of Financial Studies, 22, 2275-2302.

Bushee, B., J. Core, W. Guay, and S. Hamm. 2010. The role of the business press as an information intermediary. *Journal of Accounting Research* 48 (1): 1-19.

Bushee, B. and G. Miller. 2012. Investor relations, firm visibility, and investor following. *The Accounting Review* 87(3): 867-897

Bushee, B. and C. Noe. 2000. Corporate disclosure practices, institutional investors, and stock return volatility. *Journal of Accounting Research* 38 (Supplement): 171-202.

Core, John E. A Review of the Empirical Disclosure Literature: Discussion. Journal of Accounting & Economics. Vol. 31, Nos 1-3. September 2001.

Diamond, D. W. and R. E. Verrecchia. 1991. Disclosure, liquidity, and the cost of capital. *The Journal of Finance* 46 (4): 1325-1359

Drake, M., D. Roulstone, and J. Thornock. 2012. Investor information demand: Evidence from Google searches around earnings announcements. *Journal of Accounting Research* 50(4): 1001-104

Field, L., M. Lowry, and S. Shu. 2005. Does disclosure deter or trigger litigation? *Journal of Accounting and Economics* 39 (3): 487-507.

Hayes, R. and R. Lundholm. 1996. Segment reporting to the capital market in the presence of a competitor. *Journal of Accounting Research* 34 (2): 261-279.

Healy, P., A. Hutton, and K. Palepu. 1999. Stock performance and intermediation changes surrounding sustained increases in disclosure. *Contemporary Accounting Research* 16 (3): 485-520.

Hutton, A., G. S. Miller, and D. Skinner. 2003. The role of supplementary statements with management earnings forecasts. *Journal of Accounting Research* 41 (5): 886-890

Kim, O. and R. E. Verrecchia. 1994. Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics* 17(1): 41-67.

Kyle, A.S. 1985. Continuous auctions and insider trading. *Econometrica* 53(6): 1315-1336.

Lang, M. and R. Lundholm. 1996. Corporate disclosure policy and analyst behavior. *The Accounting Review* 71 (4): 467-492.

Lee, Lian Fen, and Hutton, Amy P. and Shu, Susan. The Role of Social Media in the Capital Market: Evidence from Consumer Product Recalls (January 28, 2015). Journal of Accounting Research.

Leuz, C. and R. E. Verrecchia. 2000. The economic consequences of increased disclosure. *Journal of Accounting Research* 38 (Supplement): 91-124.

O'Hara, M. (1995). Market Microstructure Theory. Blackwell Publishers Ltd.

Sadka, R., & Scherbina, A. (2007). Analyst disagreement, mispricing, and liquidity. Journal of Finance, 62, 2367-2403

Schrand, C. and B. Walther. 2000. "Strategic benchmarks in earnings announcements: the selective disclosure of prior-period earnings components." *The Accounting Review* 75 (2): 151-177.

Skinner, D. 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32(1): 38-60.

Tetlock, P.C, SAAR-Tsechansky, M and Macskassy, S (2008). More than Words: Quantifying Language to Measure Firms' Fundamentals. The Journal of Finance, 63: 1437-1467

Wysocki, P. 1998. Cheap talk on the Web: The determinants of postings on stock message boards. Working paper, University of Michigan.

Wurgler, J., & Zhuravskaya, E. (2002). Does arbitrage flatten demand curves for stocks? Journal of Business, 75, 583-608.

Wooldrige, J. "Econometrics Analysis of Cross Section and Panel Data." 2nd edition. The MIT Press (2010)..