Iowa Research Online
The University of Iowa's Institutional Repository

University of Iowa
Iowa Research Online

Theses and Dissertations

Summer 2016

Essays on social media and firm financial performance

Chanchal Bahadur Tamrakar
University of Iowa

Copyright 2016 Chanchal Bahadur Tamrakar

This dissertation is available at Iowa Research Online: http://ir.uiowa.edu/etd/2153

Recommended Citation

Tamrakar, Chanchal Bahadur. "Essays on social media and firm financial performance." PhD (Doctor of Philosophy) thesis, University of Iowa, 2016.
http://ir.uiowa.edu/etd/2153.

Follow this and additional works at: http://ir.uiowa.edu/etd

Part of the Business Administration, Management, and Operations Commons

ESSAYS ON SOCIAL MEDIA AND

FIRM FINANCIAL PERFORMANCE

by
Chanchal Bahadur Tamrakar

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Business Administration
in the Graduate College of
The University of Iowa

August 2016

Thesis Supervisor: Professor Thomas S. Gruca

Copyright by

Chanchal Bahadur Tamrakar

2016

All Rights Reserved

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Chanchal Bahadur Tamrakar

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Business Administration at the August 2016 graduation.

Thesis Committee: ___________________________________
Thomas S. Gruca, Thesis Supervisor

____
Gary J. Russell

____
Dhananjay Nayakankuppam

____
Sanghak Lee

____
Ashish Tiwari

ABSTRACT

Consumers are spending more time online and their involvement in social media is also growing. Furthermore, consumers truly trust the information they find online. Therefore, I expect that positive social media mentions of a given brand will influence a consumer's awareness, attitudes, affection, etc. towards that brand. The brand value chain model suggests that such a change in consumer mindset should translate into improved marketplace performance and, ultimately, better firm financial performance.

Previous researchers have studied the relationship between online user generated content and firm performance. They find that various metrics, (e.g. user ratings, comment volume or valence) impact firm performance. However, the extant research focuses on a single online platform (e.g., CNET), few methods of online posting (e.g., blog posts), or a narrow set of industries. In this study, I focus on social media sentiment expressed across multiple platforms for 180 monobrand firms spanning 10+ industries. I use total comments, total positive comments, total negative comments, proportion of positive comments, and proportion of negative comments as my measure of social media metrics.

First, I use the portfolio sort method to determine if firms with higher social media comment volume or higher positive (negative) comments generate higher (lower) abnormal returns, as determined by the Fama French 4 factor model (Fama, French, and Kenneth 1993; Fama and French 1996; Carhart 1997). Using monthly and daily returns data over a period of more than 2 years, I find no significance differences between the returns earned by the top and bottom 20% of the firms as ranked by various social media metrics. Contrary to prior research, this result suggests that social media sentiment is already fully priced into stock returns.

I then examine the possible relationship between social media metrics and firm financial performance by analyzing whether social media sentiment improves forecasts of a firm's quarterly cash flow. I modify the Lorek & Willinger (1996) multivariate time-series regression model to include social media comment volume and sentiment information to predict future cash flow. Using the Mean Absolute Percentage Error (MAPE) a guide to forecast accuracy, I find that utilizing social media information does not provide any improvement in the prediction of future quarterly cash flow forecast. I further analyze the relationship between social media comment metrics, and firm quarterly cash flow by utilizing a cross sectional regression model. I find no significant effect of social media metrics on the ability to predict future firm quarterly cash flow. Panel data estimation of both the cash flow model also does not find any significant effect of the social media metrics on quarterly cash flows.

PUBLIC ABSTRACT

Consumer involvement in social media is growing. Since consumers trust the information they find online, positive social media mentions of a given brand will influence a consumer's mindset towards that brand. The brand value chain model suggests that an improved consumer mindset translates into better marketplace performance and, ultimately, better firm financial performance.

Previous research on the relationship between online user generated content and firm performance examines how various metrics, (e.g. user ratings, comment volume or valence) impact firm performance. However, the extant research focuses on a small number of online platforms, a few types of online postings, or single industry. This study examines social media sentiment across multiple platforms for 180 monobrand firms spanning more than 10 industries.

The portfolio sort method is used to determine if firms with higher social media comment volume or higher positive (negative) comments generate higher (lower) abnormal returns. Monthly (and daily) returns data over a period of more than 2 years shows no significant difference between the top & bottom 20% of the firms as ranked by various social media metrics. Contrary to prior research, this study finds that that social media sentiment is already fully priced into stock returns.

A second study examined the effects of social media metrics on a firm's quarterly cash flow. Results from a multivariate time-series regression model and cross-sectional model (both from the accounting literature) show that social media metrics add no incremental explanatory power to known determinants of quarterly cash flow.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Consumer involvement in digital media is increasing and has overtaken traditional media in terms of average time spent consuming media (eMarketer 2014a). Consumers spend more time online than they do watching TV or listening to the radio (eMarketer 2014a). Even as the digital share of consumer time increases, the relative share dedicated to social media is becoming larger (eMarketer 2014b) and consumers now spend more time engaged in social media than in any other form of online activity (McCarthy 2014).

Online media provide a platform where consumers may share information about a company or product (Tirunillai and Tellis 2012). Consumers are now able to publicly share their product experiences, opinions, reviews, and feedback online (Wilde 2014; Bickart and Schindler 2001; Trusov et al. 2009). Most importantly, consumers tend to trust what other users have shared online (Nielsen 2012).

Prior research suggests that online information shapes consumer behavior (Bickart and Schindler 2001) and decision-making. In a survey by Dimensional Research, 88% of participants stated that reading a review about customer service had an impact on their buying decisions (Zendesk 2013). Since customers trust the information shared by others online, this information may impact consumer mindsets about a brand in multiple ways. First, online comments could increase brand awareness. For example, Sonnier, McAlister, and Rutz (2011) argue that positive, negative, and neutral information may help spread awareness about a firm's product. Earlier, Godes and Mayzlin (2004) suggested that increased online conversations would increase the likelihood of someone knowing about a product. Second, depending on the direction of the conversation, online comments by other consumers may create a positive or negative attitude toward a brand. Liu (2006) suggested that positive word of mouth would enhance, and negative word of mouth would reduce other consumers' attitudes toward a product. Third, online comments could spur brand-associated activity, such as further conversations or research about the brand.

Each of these - brand awareness, attitude, and activity – is a measure of the consumer mindset (Keller and Lehmann 2003; Mizik and Jacobson 2008) that influences purchase decisions. These decisions, in turn, affect firm sales, and ultimately firm financial performance (Keller and Lehmann 2003).

Popular business writers suggest that, since consumers spend so much time on social media, the conversation recorded there is a potential gold mine of information about how customers really feel about the products and brands they buy and use every day (Luo, Zhang, and Duan 2013; McKinsey and Company 2012; Microsoft Dynamics 2015). Thus managers should care about online user generated content because of its potential to benefit or harm their firms.

Positive consumer reviews, ratings, and word of mouth have the potential to provide benefits to a firm, while their negative counterparts can create problems for the company (Tirunillai and Tellis 2012). An incident involving a musician whose guitar was broken during a United Airlines flight is often presented as a clear example of how social media could have an impact on a firm. His video on YouTube about the broken guitar gathered more than 150,000 views within one day and had more than 4 million views within 1 month (huffingtonpost 2009). As of today, the video has more than 14 million views, with more than 77,000 Likes and 1,500 Dislikes. Press reports speculated that

United Airlines suffered financial losses because of the negative word of mouth generated due to this incident (huffingtonpost 2009).

The possible relationship between online user generated content (e.g. reviews, comments and conversations) and firm performance should be of interest of managers. Managers are encouraged to monitor various digital social media metrics in order to improve financial performance (Luo et al. 2013). However, anecdotal evidence suggests that managers are not convinced about the link between social media sentiment and firm performance. According to the CMO survey of 351 marketing leaders in 2014, social media spending was 9.4% of the total marketing budgets, and is expected to reach approximately 22% in the next five years (thecmosurvey 2014). While the managerial focus on online user generated content is increasing, the same report also suggests that demonstrating the effect of social media spending on businesses is considered challenging. Only 15% of marketing leaders say that they have quantitative proof that social media spending is justified.

While spending on social media is increasing, more research is needed for managers to understand its value. For example, "Measuring and Communicating the Value of Marketing Activities and Investments" is a top research priority (MSI 2014 - 2016). Within this topic, "How should the ROI of digital and social marketing activities be measured?" is a key area of interest. This study seeks to contribute to the existing literature in social media impact on firm performance by examining the relationship between social media sentiment and firm performance.

Other researchers have studied the impact of online user generated content, e.g. reviews, comments, ratings, etc. on firm performance (Luo et al. 2013; Duan, Gu, and

Whinston 2008; Tirunillai and Tellis 2012; etc.). Some of them find a positive relationship between comment volume (Sonnier et al. 2011; Liu 2006), product ratings (Luo et al. 2013) and firm performance. On the other hand, Duan et al. (2008) and Liu (2006) do not find any impact of consumer ratings on key marketing metrics, e.g., movie revenue and product sales. Recently, it has been suggested that negative consumer comments have a stronger negative impact on firm performance than a positive impact due to positive comments (Tirunillai and Tellis 2012).

These prior studies provide a good starting point for this study. However, this stream of research has some important potential limitations. Some studies focus only on product sales rather than overall firm performance (Chevalier and Mayzlin 2006; Liu 2006; Duan et al. 2008; Sonnier et al. 2011). Other studies limit their focus to a single firm (Sonnier et al. 2011), industry (Luo and Zhang 2013) or product category (Chevalier and Mayzlin 2006). The generalizability of other studies is potentially limited by their focus on a specific type of online information, such as customer reviews or ratings (e.g., Chevalier and Mayzlin 2006; Tirunillai and Tellis 2012). This research focuses on firm financial performance utilizing a data set that consists of some 180 companies belonging to more than 10 industries. In addition, the measures of social media sentiment span more than 100 outlets, including *Facebook, Twitter, Reddit, YouTube, Photobucket*, etc.

In the first essay, I analyze social media sentiment and stock returns for a sample of 180 monobrand (Mizik and Jacobson 2009) companies over a 2+ year period. Previous researchers have studied the impact of comment volume (Luo and Zhang 2013; Chevalier and Mayzlin 2006) and comment valance (Liu 2006; Sonnier et al. 2011) on firm financial performance. In this study, the range of social media sentiment metrics is expanded to include total comments, total positive comments, total negative comments, proportion of positive comments, and proportion of negative comments. These metrics capture a number of aspects of social media comment volume and valance.

Firm performance in this study is operationalized in terms of monthly abnormal stock market returns and daily abnormal stock market returns. Abnormal stock returns are calculated using the Fama-French-Carhart 4 factor model (Fama, French, and Kenneth 1993; Fama and French 1996; Carhart 1997). To model the effect of social media sentiment on stock market returns, the portfolio sort method (Jacobson and Mizik 2009; Luo et al. 2010) is employed.

The time frame for this study covers the period between June 2012 and July 2014. The sample is limited to monobrand firms. Mizik and Jacobson (2009) refer to monobrand firms as firms that use a branded house strategy. Therefore, the majority of their business and, presumably, customer conversations online, are aligned with a single brand name (Mizik and Jacobson 2009). A total of 180 publicly traded monobrand companies have complete stock market data on COMPUSTAT.

The portfolio sorting approach is an extensively used methodology in finance for exploring the relationship between firm characteristics and expected returns (Patton and Timmerman 2007). The first step is to rank all firms in a given time period – monthly or daily, based upon the social media metric under study. The entire sample of firms is then split into 5 equally sized portfolios (quintiles), based on their ranking. From the Fama-French-Carhart four factor model (Fama et al. 1993; Fama and French 1996; Carhart 1997), the average abnormal returns are computed for each firm for each time period in each of the five portfolios. While this approach was developed in finance, it has been increasingly used in marketing (e.g., Luo and Zhang 2013; Luo et al. 2013; Tirunillai and Tellis 2012; Tuli and Bharadwaj 2009).

Comparing the average abnormal returns across the various portfolios, mispricing can be detected. If a given social media metric provides information about firm financial performance, then one might expect that the top portfolio would perform significantly better than the bottom portfolio. The top quintile firms should have significantly positive or negative returns, depending on the metric under study.

Previous researchers who studied sentiment analysis have used different methodologies to obtain their data and categorize conversations based on their valence (positive or negative with respect to a given brand). For example, Luo et al. (2013) and Liu (2006) used graduate students to categorize contents as positive or negative. More recently, Tirunillai and Tellis (2012) used a naïve Bayesian classifier and Support Vector Machine classifier for the same task. For the current study, social media data were collected using a publicly available social media monitoring site known as *socialmention.com*. Some benefits of using a publicly available social media monitoring service are replicability and transparency.

A limitation of the first study is that the null results could be due to at least two possibilities. One possibility is that there is no impact of the social media metrics on financial markets. Another possibility is that if any social media information has an impact on stock abnormal returns, it is already priced into the stock prices, and no advantage can be gained by trading in stocks based on available social media information. By using the portfolio sort methodology, it can be conclusively stated that there is no mispricing. However, there still may be an effect of media sentiment on firm performance that is not reflected in stock returns, which are notoriously volatile. For this reason, my second essay analyzes whether social media sentiment provides any incremental value on predicting the future quarterly cash flows of the firm, or if any significant relationship exists between the two.

For my second study, I utilize the same data source but include additional data beyond June 2014. The final data set consists of data until December of 2015. Some firms that did not have the required quarterly cash flow data available in Compustat were not included in the second study which resulted in a slightly fewer number of firms. In this essay, I use the quarterly cash flow data, which has been considered a key measure of firm value by many marketing scholars (Gruca and Rego 2005; Anderson et al. 2004; Day and Fahey 1988; Srivastava, Shervani, and Fahey 1998), as the key variable of interest.

I utilize a multivariate time series regression model and a cross sectional regression model to analyze the relationship between social media and firm quarterly cash flow. The results indicate no significant relationship between social media metrics and future firm quarterly cash flows. Social media also does not seem to provide any additional benefits in terms of forecasts of future quarterly cash flow.

This dissertation utilizes well established methodologies in the finance and accounting literatures to study the impact of social media on firm performance. While marketing literature identifies abnormal returns and quarterly cash flow as an important variable of interest, most studies in this field have not utilized the methodologies that are well established in the literature from which those measures are adopted from, i.e. finance and accounting. In the first essay I utilize the portfolio sort method which is a dominant and popular approach in the finance literature to establish and test for expected stock returns of a firm based on various firm characteristics (Patton & Timmerman 2007). In my second study I utilize the multivariate time series regression model (Lorek & Willinger 1996) and a cross sectional regression model, similar to the ones used by Folsom et al. (2016) and Atwood et al. (2010), which are well established in the accounting literature to study periodic cash flows. Since the dataset comprises of observations for more than 150 firms over a span of multiple quarters I also utilize panel data estimation techniques to analyze whether any significant relationship between the social media metrics and quarterly cash flow exists.

The results of the first study suggest that there is no discernable relationship between portfolio abnormal returns and social media sentiment metrics, i.e. no significant abnormal return accrues to portfolios formed based on different social media metrics. I find that a portfolio of firms that have higher volume of comments, higher volume of positive comments, or higher proportion of positive comments do not provide higher abnormal returns than firms that have low volume of comments or lower proportion of positive comments. I also find that firms with higher proportion of negative comments or a larger number of negative comments do not provide lower abnormal returns than firms with fewer negative comments online. This result is contrary to the results found in prior literature in marketing that social media metrics based information could enable traders to earn significantly higher abnormal returns (Tirunillai and Tellis 2011). Thus, mispricing of social media information does not seem to be widespread. Any possible relationship between social media metrics and stock returns appears to be caused by the abnormal returns obtained by the Computer and Electronic manufacturing sector. Once this sector is isolated, I was unable to find any significant evidence of mispricing. This is an important finding since it suggests that the prior results are not necessarily generalizable.

The results of my second study suggest that utilization of social media information does not provide any significant benefits in terms of future cash flow prediction. Utilizing a well-established multivariate time series regression model (Lorek & Willinger, 1996), I find that social media comment volume and sentiments do not provide any improvements in future quarterly cash flow predictions. The results of a cross sectional regression analysis with firm, industry, and quarterly fixed effects also show that there is no significant relationship between the social media metrics and firm quarterly cash flow. The panel data analysis of the social media metrics and firm quarterly cash flow also fails to yield any significant results.

In summary, this dissertation has followed well established methodologies in finance and accounting to study the impact of various social media metrics on abnormal stock returns quarterly cash flows. This is the first study, we know of, that analyzes the relationship between various social media metrics and firm quarterly cash flow. Previous studies in marketing, that found that firms with better social media metrics have higher abnormal returns, suggest that the markets are not efficient. The first essay provided evidence that the markets are indeed efficient and investors cannot gain significant abnormal returns based on social media information. The second essay also finds no significant relationship between social media and firm quarterly cash flow.

CHAPTER 1: STUDYING MISPRICING OF SOCIAL MEDIA INFORMATION IN THE FINANCIAL MARKETS

1.1 LITERATURE REVIEW

1.1.1 User Generated Content and Social Media

Consumer involvement in social media has grown rapidly within the last decade. While approximately 8% of online adults used social networking sites in 2005, the figure has grown to 74% in 2013 (Pew Research Center 2014). Social networking sites such as Facebook and Twitter have seen rapid growth. The number of Facebook users grew to more than 1.1 billion in 2013 (www.facebook.com 2013)[1], and Twitter users are numbered at more than 250 million (www.twitter.com 2014)[2].

Consumer involvement in social media has not only grown in terms of the number of consumers, but also in terms of the time they spend on these websites. The average share of digital media time spent per day via mobile or online stands at more than 47% in 2014, compared to other major media. Television, for example, stands at around 36% (eMarketer 2014a). Consumers spend most of their online media time with social media. They spend an average of 37 minutes/day on social media, while spending 29 minutes using email and 23 minutes using search platforms (Richter 2013).

Not only do consumers spend a great deal of time expressing their opinions online, but they also trust the information provided by other users. For example,

[1] https://www.facebook.com/facebook/info?tab=milestone
[2] https://about.twitter.com/company

Nielsen's Global Trust in Advertising (2012) report surveyed more than 28,000 consumers in over 56 countries. They found that 70% of respondents trusted consumer opinions posted online, compared to less than 50% who trusted any form of advertising, and 33% who trusted online banner advertisements. Dellarocas and Wood (2008) state that consumers reveal feedback and recommendations online, and they do so with selfless intentions. Consumers may believe that social media content reveal less biased consumer sentiment (Hanson and Kalyanam 2007).

While most online user generated content may be from anonymous strangers, user generated content in the social media environment comes from people whom we know, or from people who are part of a social community. This is a key distinction, since Nielsen found that 92% of people surveyed said they trusted recommendations from friends/family. Since social media such as Facebook and Twitter involve the people we know (friends, family, or people that we decide to follow), consumers' trust of the opinions shared on social media about any product or company should also be higher. Consequently, one might expect that social media sentiment should have a greater impact on the financial performance of a company. Since prior studies have shown a positive impact of anonymous online recommendations and reviews on firm performance (Tirunillai and Tellis 2012; Luo and Zhang 2013; Luo et al. 2013), I expect a similarly strong effect from brand-based conversations on social media.

In the next section, I will discuss online user generated content and how it relates to traditional consumer word of mouth.

1.1.2 User Generated Content and Word of Mouth

The definition of "word of mouth" has changed. In the late 1960s, Arndt (1967) described word of mouth as "Oral, person-to-person communication between a receiver and a communicator whom the receiver perceives as non-commercial, concerning a brand, product, or a service." In the Internet age, Hennig-Thurau, Gwinner, Walsh, and Gremler (2004) defined online word of mouth as "any positive or negative statement made by potential, actual, or former customers, about a product or company, which is made available to a multitude of people and institutions via the Internet." Hence, while traditional word of mouth is face to face, online user generated content is not necessarily face to face, or even oral. It is generally written and experienced via a medium that does not require sender and receivers to be in direct contact.

Word of mouth can be a source of product awareness (Dhar and Chang 2007; Godes and Mayzlin 2004). Researchers have also established that word of mouth communication shapes consumers' attitudes and behaviors (Brown and Reingen 1987). Similarly, in the current world of the Internet, user generated content online provides information that could increase product awareness (Kiecker and Cowles 2002; Sonnier et al. 2011) and could affect attitudes toward products (Liu 2006). Consumer conversations online can provide signals relating to the awareness, attitudes, and affection of consumers toward a brand (Chevalier and Mayzlin 2006; Li and Hitt 2008). One advantage of online user generated content, such as star ratings, is that they have less of a chance of being misunderstood (King et al. 2014). As a written source of information, online user generated content has permanence, which is, in and of itself, an advantage over traditional WOM (Bickart and Schindler 2001).

In trying to understand how information online affects consumer behavior, researchers have created parallels between traditional WOM and online sources, including forums (Bickart and Schindler 2001), blogs, customer reviews (Chevalier and Mayzlin 2006; Luo and Zhang (2013), product ratings (Duan et al. 2008), social networking sites (Tirunillai and Tellis 2012), and any user generated content online (Ye et al. 2011). Posting on these sites by consumers is characterized as online WOM.

While previous studies have examined the relationship between various measures of online user generated content and firm value, none of this research is focused specifically on social media. Most of the research is focused on user generated content, which is created by anonymous strangers in the form of reviews, ratings, blog posts, etc. In the online environment, an individual may or may not understand the motivation of the user who posts a certain message (King et al. 2014). Self-interest on the part of a seller may negatively impact both credibility and informativeness (Resnick et al. 2000). Despite this, Nielsen (2012) finds that consumers trust the opinions of strangers that are posted online (70% trusted consumer opinions posted online). However, they trust recommendations from friends and families even more (92%). When receiving information, consumers often use source characteristics (identity descriptive information, etc.) as a heuristic device on which to base their decisions (Forman, Ghose, and Wiesenfeld 2008). Studies have also found that word of mouth from a person having strong ties with the receiver of the word of mouth will have a greater impact on their decision-making (Bansal and Voyer 2000; Brown and Reingen 1987). Therefore, while online user generated content, in general, may have an impact on consumer decisions, it is likely that user generated content on social media platforms should be more influential because of source credibility and because of the stronger relationship between the information source and the receiver.

While research on the impact of online user generated content is of interest to both managers and researchers, many of these studies have only used the volume of information available or the average consumer rating as the independent variable. More recently, advances in computer science have allowed researchers to measure consumer sentiment from online sources, such as social media platforms. In the next section, the role of sentiment analysis is discussed in reference to the user generated content literature.

1.1.3 User Generated Content and Sentiment Analysis

Online user generated content has changed the way business gets done (Farshid et al. 2011). Consumers are constantly creating and initiating conversations about products, brands, etc. to educate one another using a variety of online resources (Blackshaw and Nazzaro 2004). One of the ways to understand and interpret user generated content is by analyzing the customer sentiments being expressed there. User generated content in social networks, blogs, online forums, etc. generate potential sources of information (Thelwall, Wouters, and Fry 2008) and is also commercially exploited to extract customer opinions about brands and products (Thelwall, Buckley, and Paltoglou 2011).

Sentiment analysis of social media can be considered as the identification of opinions expressed online, whether they are positive, negative, or neutral toward a product, brand, or company. As overt and public actions, these conversations should provide us with an indication of consumer attitudes (Microsoft Dynamics, 2015).

Because consumers are sharing their product, brand, and firm experiences online and others are listening in on those experiences, managers need to be careful about what is being said. The "United breaks guitars" example is a case in point.

The popular business press suggests that marketers need to monitor multiple online media for information pertaining to their brands (McKinsey&Company 2012; Harvard Business Review 2010; Proulx 2010; Wirthman (2013). Fragmented information sources where consumer opinions of products and services are posted mean that traditional methods of tracking product image may no longer work (Kim 2006 cited in Pang and Lee 2008). Many top managers consider "social listening" as being extremely important to their business (Microsoft Dynamics, 2015). This same report suggests that accurately tracking brand sentiment - how people feel about your brand - is very important to strategic marketing. Tracking online sentiment, for many businesses, has turned into a "virtual currency" that can make or break a product and could help businesses improve their bottom line and transform their online information searches (Wright 2009). Managers have always been interested in determining what consumers think about their brands, and they are scrambling to harness the power of technology in social networking and web applications to learn about the impact of online user generated content on firm performance (Luo and Zhang 2013).

The high level of participation in social media across the entire population means that there is a very large source of information pertaining to brands available online. The problem with using such a resource is that it is difficult to synthesize the vast amount of available information (Sonnier et al. 2011; Liu 2006). Many studies have tried to incorporate semantic information present in online user generated contents. They have used various approaches to facilitate the synthesis of online information. For example, Godes and Mayzlin (2004), Liu (2006), Luo and Zhang (2013) used a single source of information for their studies. Others have focused on a single metric, such as the volume of comments (Luo and Zhang 2013) or star ratings (Chevalier and Mayzlin 2006; Luo and Zhang 2013).

To summarize online user generated content in terms of sentiment, different authors have used different tools and techniques. For example, Liu (2006) used three judges to read 12,136 messages and assign them to positive, negative, mixed, neutral, and irrelevant categories. Sonnier et al. (2011) used a proprietary web crawler to collect the volume of positive, negative, and neutral online communications. Tirunillai and Tellis (2012) used web crawling technology to collect data and a naïve Bayesian classifier and Support Vector Machine Classifier to categorize the valance of the contents they found. Luo et al. 2013 used automated web crawling software to collect ratings data, and they employed two graduate students to categorize blog posts as reflecting positive or negative sentiment. In the case of human-based coding, various rules were used to determine valence. For example, Liu (2006) required 2 out of 3 judges to agree in order to assign a valence category.

With the advent of new tools and technology, sentiment mining has become more accessible. Sentiment analysis methods allow decision-makers to extract and characterize opinions from unstructured consumer generated online content (Pang and Lee 2008). Automated sentiment analysis reduces the need to have a human being (or three) read the large volume of information available online (Yu et al. 2013). It reduces the effort of having information hand coded and cross checked by multiple raters for accuracy. In this study, a publicly available and free social media monitoring site - www.*socialmention.com* - is used for the data collection. This website provides regularly updated information on various metrics, such as total positive comments, total negative comments, sources of comments, etc. It was queried on a daily basis using an automated web scraping, web harvesting and content extraction tool called "Visual Web Ripper."

In the next two sections, I will discuss the brand value chain model put forward by Keller and Lehmann (2003) and how online user generated content relates to the model. The brand value chain model focuses on how marketers use various means to impact consumer mindsets, and subsequently, firm financial performance. The subsequent section will examine how user generated content could influence the brand value chain, and thus impact firm financial performance.

1.1.4 Brand Value Chain

The brand value chain model, developed by Keller and Lehmann (2003), theoretically explains the factors and processes that impact a firm's financial performance. The model shows that brand-related investments ultimately affect a firm's financial value (e.g., stock price, market capitalization, etc.) through changes in the consumer mindset and the effects of these changes on the marketplace performance of the brand (e.g. market share, higher price premiums, profitability, etc.).

Firm-initiated programs include such activities as research and development, marketing communications, promotions, and employee training. These investments may influence the customer mindset, which consists of awareness, associations, attitudes, attachment and activities with regard to the brand (Keller and Lehmann 2003). Marketing communications and promotions can increase the consumer's ability to recognize and recall the brand that leads to higher brand awareness (Villarejo-Ramos and Sanchez-Franco 2005; Keller 1993). The product itself and its benefits should lead to positive brand associations. The performance of the brand and the level of satisfaction it provides create negative or positive attitudes toward the brand (Keller and Lehmann 2003).

High levels of brand awareness and a positive attitude toward the brand should increase the probability of brand choice. Over time, this would be reflected in higher levels of consumer loyalty (Keller 1993). The consumer's mindset thus shapes his or her behaviors toward the brand. These actions should then be reflected in terms of a willingness to pay a higher price premium, changes in price elasticity, increased market share or a higher share of requirements (Keller 1993).

The performance of a brand in the marketplace has an impact on firm performance through its impact on cash flow growth, acceleration and variability (Gruca and Rego 2005). In turn, improvements in a firm's cash flow will be reflected in the firm's overall financial standing with respect to stock price, price to earnings ratio, market capitalizations, etc. (Keller and Lehmann 2003).

The brand value chain model developed by Keller and Lehmann (2003) does not incorporate the influence of online user generated content. In contrast to the brand value chain view of the world, consumers are not only affected by marketing programs initiated by the firm, but also by other consumers. For example, Goh, Heng and Lin (2013) find that both marketer generated content (e.g. posted messages by firms on Facebook brand communities) and user generated content (e.g. posted messages by consumers on Facebook brand communities) affect consumer purchasing behavior. They find that while marketer generated content plays a persuasive role, user generated content plays both persuasive and informative roles (pp. 103). They also find that the marginal persuasive effect of information provided by other consumers is more than 22 times that of marketer generated content. In the same vein, Yu et al. (2013) analyzed user generated content (blogs, forums, news, and Twitter) and conventional media (newspaper, TV, magazines) for 824 publicly traded firms across 6 industries. They found that user generated content online had a stronger relationship with firm financial performance than conventional media. Scholz et al. (2013) found that user generated content (wall posts and comments on Facebook) and marketer generated content (wall posts and comments) differed in their impact on product awareness and conversion rates (the proportion of web shop visits that lead to a purchase). While both user and marketer generated content created awareness, only user generated content succeeded in creating higher conversion rates.

In the next section, I will review the previous literature focusing on the link between online user generated content and the various stages of the brand value chain model.

1.1.5 User Generated Contents and the Brand Value Chain

Evidence of online user generated content having an impact on the various aspects of the brand value chain has been established by previous research. Table A1 provides a list of the major studies in this field.

In this section, I will review previous literature that suggests a relationship between user generated content and the different stages of the brand value chain model.

1.1.5.1 User generated content and the consumer mindset

Keller and Lehmann (2003) suggest that customers' mindsets include thoughts, feelings, experiences, images, perceptions, attitudes, etc. that exist in consumers' minds with respect to a given brand. Researchers have argued that online user generated content relating to a product or a brand could impact other consumers' mindsets. For example, Sonnier et al. (2011) suggest that information contained in user generated content may help spread awareness about a firm's product, while Godes and Mayzlin (2004) also argue that increased conversations about a product would lead to an increased likelihood of someone knowing about the product. Hutter et al. (2013) find that active engagement with the social media activity of a brand increases awareness of the brand. Scholz et al. (2013) find that both marketer and user generated content are effective in creating awareness among online social network users.

Awareness is only the first step. Before taking action, the consumer must have interest and desire (Strong 1925). These important aspects of brand attitude may be influenced by information that consumers find online. For example, Liu (2006) argues that positive word of mouth enhances other consumers' attitudes toward a product, while negative word of mouth reduces it. Bruhn, Schoenmueller, and Schafer (2012) find that user generated content has a major influence on non-attribute based brand image, but no influence in terms of the attribute based brand image.

1.1.5.2 User Generated Content and Brand Performance

In a study by McKinsey (King et al. 2014), it is suggested that the traditional consumer decision-making journey has changed. They suggest that instead of a linear funnel process where consumers narrow down their choices, today's consumers add and delete brands based upon information from online sources. As such, brands that were not a part of the original consideration set may enter the consumer's final consideration set if there is positive online word of mouth with respect to that brand. This study suggests that online user generated content can impact the marketplace performance of brands. However, this is not really news. Chevalier and Mayzlin (2006) found that customer reviews posted on Retail websites affected other consumers' purchasing behavior on those sites. Liu (2006) found that the volume of posted messages on online websites explains both aggregate and weekly box office revenues of movies. Sonnier et al. (2011) found that online comment volume was positively correlated with firm sales. They also found that positive and negative comments had a larger effect than neutral comments.

It must be said that the evidence of the relationship between user generated content and market performance is not conclusive. For example, Liu (2006) found that the valence of messages on Yahoo message board postings, as measured by the percentage of positive and negative messages, did not have any impact on revenues. Scholz et al. (2013) found no significant effect of positive or negative user generated content on differences in conversion rate, while neutral comments had a positive effect.

1.1.5.3 User Generated Content and Firm Performance

Multiple studies have skipped the intermediate steps in the brand value chain to focus directly on the impact of user generated content involving firm financial performance. Tirunillai and Tellis (2012) found that user generated content in online platforms help predict stock returns and trading volume. Luo and Zhang (2012) found that consumers' reviews posted online contributed to firm value. Luo et al. (2013) determined consumer ratings, search and traffic volume to the website under study, and blog sentiments as a leading indicator of firm equity value.

The current study extends this last stream of research by focusing on the impact of social media sentiment content on firm financial performance. The approach here is an analysis of possible mispricing regarding social media information in the financial markets. Mispricing studies have been used in finance, as well as in the marketing literature to study the relationship between a variable of interest and the value of the firm, as determined by the stock market. In the next section, I will briefly discuss mispricing and the possible mispricing of social media information in the financial markets.

1.1.6 Mispricing

The ability to predict firm financial performance has always been of interest to researchers. Over time, the stock prices and earnings of firms in the US have become less correlated, and this phenomenon may be attributed to the inability of financial markets to account for intangible assets while determining firm value (Lev and Zarowin 1999). Firms that have better marketing assets (e.g., higher levels of satisfaction or stronger brands) should have higher levels of performance in the financial markets (Gruca and Rego 2005; Mortanges and Riel 2003; Kerin and Sethuraman 1998). One intangible asset of current interest includes consumer online WOM and online user generated content.

Many studies in the marketing literature (Fornell et al. 2006; Jacobson and Mizik 2009; Tirunillai and Tellis 2012) have been conducted to understand the relationship between a marketing variable of interest and firm performance. Mispricing of marketing performance indicators, such as customer satisfaction and brand equity, is considered as an important element in arguments for the relationship between those indicators and firm value (Bell, Ledoit, and Wolf 2013). To study the possible relationship between social media information and firm value, I will use the portfolio sort method, which is a popular method in the finance literature to study mispricing (Patton and Timmerman 2007).

The portfolio sort approach is a way to test the hypotheses of a relationship between firm valuation and a given marketing variable (Patton and Timmerman 2007). The Efficient Market hypothesis suggests that the stock price reflects all relevant information and provides a fair valuation of the stock (Fama 1970). If an indicator under investigation (e.g., social media sentiment) were to provide additional information relating to firm performance that is not already incorporated into the stock price, then an investor could benefit by trading in undervalued stocks. The resulting returns should be higher than the overall market return. An example of this approach can be found in the customer satisfaction literature, where Fornell et al. (2006) found evidence that investors could beat the market by using the American Customer Satisfaction Index (ACSI)-based trading strategies. Evidence of mispricing, where the average return between the portfolios being compared is different, could indicate a relationship between such an indicator and firm value (Bell et al. 2013).

Researchers have used the presence or absence of mispricing to study the relationship between various marketing activities and firm financial performance. For example, Jacobson and Mizik (2009), and Fornell et al. (2006) studied the market mispricing of customer satisfaction to see whether a portfolio of firms with higher customer satisfaction performed better than a portfolio of firms with lower customer satisfaction. In the case of online user generated content and its impact on firm financial performance, Tirunillai and Tellis (2011) found that there was evidence of mispricing. They found that a return of approximately 8% over normal market returns could be made by trading stocks based on user generated content in online platforms (consumer reviews). It is important to note that there is the possibility of inter-industry differences when it comes to mispricing, as Jacobson and Mizik (2009) found that mispricing of customer satisfaction existed for computer/Internet-based firms, but not for firms in other sectors.

If social media sentiment does have an impact on firm performance and the financial markets do not properly impound this information into stock prices, then the stocks of firms with better metrics would be undervalued. Similarly, stocks with worse metrics would be overvalued. Therefore, an investor buying or selling stocks based on this information would be able to outperform the market by holding the stocks and waiting for the social media information to impact other marketing metrics, which subsequently impact the stock prices. If the impact of social media is correctly embedded in stock prices, there would be no benefit to an investor using the social media information for trading purposes.

This study differs in several ways from prior research on online user generated content. Some of these expected improvements are presented in Table A2.

First, this study focuses on social media comments. Previous studies have focused on online user generated content, such as product reviews and ratings (Luo and Zhang 2013; Chevalier and Mayzlin 2006; Tirunillari and Tellis 2012), online message boards (Liu 2006), or online consumer ratings and blog posts (Luo et al. 2013). This paper focuses on messages posted on various social media sites such as *Facebook, Twitter, Reddit, YouTube*, etc. This is important, since much more consumer time is spent on these platforms, and many companies have moved their spending in the direction of social media outlets. Also, comments on social media may have a greater impact because of their source credibility, etc. as discussed earlier.

Second, this study utilizes multiple online platforms for information collection. While Luo and Zhang (2013) used *CNET.com*; Chevalier and Mayzlin (2006) used *Amazon.com* and *bn.com*; and Liu (2006) used *Yahoo*, the data source employed in this study gathered sentiment information from more than 100 different social media sites.

Third, the sample includes a larger number of firms spanning multiple industries. Many of the previous studies were conducted using only one firm or one product or one industry. The sample analyzed here consists of 180 firms across more than 10 different sectors and industries (based on the NAICS industry classification system).

As in other studies, the social media data presented here are characterized in terms of positive, neutral and negative sentiments. Consistent with the finance literature, firm performance is measured using abnormal returns (as determined by the Fama-French model). However, this study uses the portfolio sorting method from finance to ascertain the relationship between social media sentiment and firm performance.

1.2 DATA AND MEASURES

1.2.1 Social Media Data and Measures

The raw data on social media sentiment were collected from a social media monitoring site: *www.socialmention.com*. A review by Brandwatch (2013) concluded that *socialmention.com* is one of the best free online social media monitoring tools available. Every day, a web-crawling agent ("Visual Web Ripper") queried the site for a set of 250 brands associated with monobrand companies. These monobrand companies are the same ones used by Mizik and Jacobson (2009).

By using publicly available data, this study can be more easily extended or replicated.

According to its website, *"Social Mention is a social media search and analysis platform that aggregates user generated content from across the universe into a single stream of information. It allows you to easily track and measure what people are saying about you, your company, a new product, or any topic across the web's social media landscape in real-time. Social Mention monitors 100+ social media properties directly including: Twitter, Facebook, FriendFeed, YouTube, Digg, Google etc."* (http://www.*socialmention.com*/about/).

Daily data were collected from www.*socialmention.com* for a period of over 2 years beginning on May 25th, 2012 and ending on July 31st, 2014. *Socialmention.com* provides various sentiment metrics, such as positive/neutral/negative comments. It also provides a listing of sites on which the comments were made. The social media sites that *socialmention.com* monitors include *Facebook, Twitter, YouTube, Reddit,* etc. The web-crawling agent visited the site daily and collected those daily data for a 24-hour period at a predetermined time every day.

Data were collected for over 250 monobrand companies, which are publicly traded, belonging to over 10 different sectors and multiple industries. Monobrands are firms in which a single brand represents the majority of the firm's business and firms that use a branded house strategy (Mizik and Jacobson 2008, 2009). This approach was intended to capitalize on the singular focus of these firms regarding their corporate brands. For a firm with a branded house strategy, any positive, neutral or negative sentiment expressed online about the brand could potentially impact the firm's financial market performance. This impact on stock returns might be less for a firm possessing a diverse portfolio of brands because any given brand would only constitute a small proportion of the firm's sales. Some of the monobrand firms in the sample were *Canon, FedEx , Oracle, Starbucks , Walmart, Amazon.com, Chevron, Comcast, Nike, Southwest Airlines*, etc. These firms belong to various NAICS Sectors and Subsectors, details about which are presented in Table A3.

There were 5 different social media sentiment metrics constructed for this research. They are listed next.

1. Total Positive Comments – This is the total number of positive comments made about the monobrand for a given time period.

2. Total Negative Comments – This is the total number of negative comments made about the monobrand for a given time period.

3. Total Comments – This is the total number of comments made about the monobrand for a given time period. This is calculated by summing the total positive comments, total negative comments and total neutral comments for that period.

4. Proportion of Positive Comments – This is calculated by dividing the total positive comments by the total comments in a given time period.

5. Proportion of Negative comments – This is calculated by dividing the total negative comments by the total comments in a given time period.

1.2.2 Financial Performance Data and Measures

The daily and monthly financial performance data were obtained from the COMPUSTAT and CRSP data sets using the Wharton Research and Data Services. Firm financial performance was measured using a firm's abnormal or excess return obtained beyond the expected average return in the stock market performance, based on the Fama-French model (Fama et al. 1993; Fama and French 1996; Carhart 1997). The Fama-French model to calculate abnormal returns has been used by Luo and Zhang (2013), Luo et al. (2013), Tirunillai and Tellis (2011), Tuli and Bharadwaj (2009) and various other researchers in marketing to measure firm performance, since it controls for important variations in stock returns identified in the finance literature.

The Fama-French-Carhart model measures excess returns as a function of 4 factors, which are: market returns, a small minus big market capitalization factor (size effect), a high minus low market to book ratio factor (value effect), and a momentum factor. The Fama-French- Carhart model is represented as follows, where the return of a firm is

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{mkt,i}(R_{mkt,t} - R_{f,t}) + \beta_{smb,i} SMB_t + \beta_{hml,i} HML_t + \beta_{mom,i} MOM_t + e_{i,t} \qquad (1)$$

Where

$R_{i,t}$ = returns for firm i on time t

$R_{mkt,t}$ = average market returns on time t

$R_{f,t}$ = risk free rate on time t

SMB_t = Small minus Big (Size effect factor)

Represents the return on a portfolio invested in a group of small cap vs. large cap stocks.

HML_t = High minus Low (Value effect factor)

Represents the premium that stocks with high book-to-market ratio enjoy over stocks with low book-to-market value.

MOM = Momentum

Represents the empirically observed momentum phenomenon in stock returns

Using a rolling regression method to get the estimated factor coefficients of the Fama French 4 factor model, the abnormal return for the time period t+1 is calculated as

$$AR_{i(t+1)} = (R_{i(t+1)} - R_{f(t+1)}) - \{\hat{\alpha}_{i,t} + \hat{\beta}_{mkt,i,t}(R_{mkt,(t+1)} - R_{f,(t+1)}) + \hat{\beta}_{smb,i,t} SMB_{(t+1)} + \hat{\beta}_{hml,i,t} HML_{(t+1)} + \hat{\beta}_{mom,i,t} MOM_{(t+1)}\} \qquad (2)$$

The beta values in Equation 2 are estimated using a rolling 36-month window for the monthly level analysis and 250-day window for the daily level analysis.

For a few companies present in the initial data set, the abnormal returns could not be calculated throughout the period of the study because of various reasons such as mergers, acquisitions, bankruptcy, etc. Those companies were removed from the data set.

I matched the social media data from social mention with the firm's stock abnormal returns, as calculated above. In terms of monthly data, the final data consisted of 180 monobrand companies. Some companies were also removed because of data collection errors during the web-crawling process. Thus, the final monthly data set contained 4,680 pooled time-series cross-sectional observations of the 180 firms for a 26-month period. This is the largest sample of firms to be studied among published research that has examined the relationship between social media sentiment and firm value. In terms of daily data a total of 175 firms were included in the data set with each firm providing 486 business days of data, from 4^{th} June 2012 to 27^{th} June 2014. The final data set contained 85.050 pooled time-series cross-sectional observations for over two years.

1.3 METHODOLOGY AND RESULTS

The portfolio sort method is a popular approach in finance, which explores the relationship between expected stock returns and some observable firm characteristic (Patton and Timmermann 2007). In this approach, the stocks are sorted based on the variable of interest at a beginning date. First, the stocks are grouped into multiple portfolios (3, 5, 10 or more). After a certain holding period, the average abnormal returns for each of the portfolios, from top to bottom, are calculated. Whether or not there is a significant relationship is then determined by finding out if the top and bottom portfolios have significantly different average abnormal returns (Patton and Timmermann 2007). Other researchers test to see whether the abnormal returns are different from zero (Jacobson and Mizik 2009).

Patton and Timmermen (2007) mentioned that portfolio sorts are widely used for the following reasons. Firstly, companies that drop out or enter the sample are easier to handle. Secondly, no assumption of a linear relationship between the variable of interest and returns is required. Third, any differences between the returns from the top and bottom portfolios could be interpreted as profits accrued from a trading strategy.

Some examples in the marketing literature using a similar methodology are Jacobson and Mizik (2009), O'Sullivan et al. (2009), and Aksoy et al. (2008). They use the portfolio-based study to examine the possible mispricing of customer satisfaction in financial markets. Jacobson and Mizik (2009) detail various methodological approaches to assess financial market mispricing using various portfolio approaches. In this paper, I will use a firm-specific risk model with a rolling data window, where the firm variable of interest will be the social media information available to us.

First, I sorted all available stocks in the data based on the total comments for each of the time period, daily or monthly, so as to create portfolios based upon their ranking. Then, for each time period, all of the available stocks were ranked from 1st to 180th for the monthly data and 175th for the daily data. Five groups were then created, with 20% of the firms in each group based upon their ranking, from 1st to last. Since the number of firms available was 180 for the monthly data, each of the 5 portfolios consisted of 36 firms each. The first portfolio consisted of the 1st to 36th ranked firm. The second portfolio consisted of the 37th to 72nd ranked firms, and so on. This process was repeated for each month. Similarly 5 portfolios were created based on the daily data as well. In the portfolio sort methodology a firm's ranking based on the social media metrics could be different across different time period. Therefore, a firm could move from one portfolio to another across periods. For the occasions when there was a tie in the social media metrics, all of the firms in the tie were given the mean value of the ranking, and the quintiles were decided based on the mean. This could give rise to an uneven number of observations in each portfolio.

For robustness check I also conducted the portfolio analysis without the α value in the Fama-French model in equation 2. The result of the analysis were similar whit or without α.

1.3.1 Monthly Portfolio Analysis (Contemporaneous)

The descriptive statistics of the various social media metrics by month are presented in Table A4. The total number of comments for all firms at the end of June 2012 was 245,367, which included 61,825 positive, 171,142 neutral, and 12,400 negative comments. From the monthly summary statistics in Table A4, it can be seen that a majority of the comments were generally neutral. The average proportion of positive comments for the same month was 0.212, and the average proportion of negative comments was 0.047. Generally, the proportion of negative comments was in the 0.05 range, and the proportion of positive comments was around the 0.2 range.

Figure B3 shows the trend for the total comments, total positive comments, and total negative comments over time, aggregated to the monthly level. The positive and negative total comments trends seemed to be fairly stable, while the total comments trend seemed to vary by month, with an increase in the later months of the data collection period. A breakdown of this by industry sector showed a similar overall trend.

Figure B4 shows the time series plot of the abnormal returns for portfolio 1 and portfolio 5 based on total comments. Figure B5 through figure B8 shows the time series plot of the abnormal returns based on total positive comments, total negative comments, proportion positive comments, and total negative comments respectively.

Figure B9 shows the time series plot of total comments, total positive comments, and total negative comments for a few companies present in the data set.

Table A6 shows the summary statistics of the various social media metrics by sector for the monthly data. The total comments across all sectors for the entire duration of the data collection were 6.9 million, with an average of 1,477 comments per firm per month. The proportion of positive comments per observation across all sectors was around 0.20, and the proportion of negative comments was around 0.04. In terms of the number of comments per observation, the Information sector had the highest total comment volume, with over two thousand comments per firm per month, compared to others, which had around 1,300 to 1,700 comments.

For the portfolio based on the total number of comments across 26 months for all of the 180 firms, the top portfolio had a total of 2.76 million comments, while the bottom portfolio had approximately 117,000 total comments (See Table A7). Table A7 further breaks down the various social media metrics by sector and portfolio.

The various subsectors that are analyzed for the monthly data study were Information sector, manufacturing sector, which was broken down into Computer and Electronic manufacturing and all other manufacturing, and all sectors, except for the manufacturing and Information sector. The Information sector consists of the NAICS sector 51 and includes firms such as *Adobe Systems Inc., Cablevision, Comcast Corp., Direct TV, eBay, Microsoft Corp., Verizon Communications Inc., Viacom, Vonage,* etc. In terms of the manufacturing sector and its sub groups, the first group consists of the NAICS subsector 334 (Computer and Electronic Product Manufacturing), and the second group consists of all other manufacturing sectors under the NAICS 31, 32, and 33 heading. The Computer and Electronic Product Manufacturing sector consists of 19 firms, which includes *Advanced Micro Devices Inc*., *Hewlett Packard Co*., *Canon Inc*., *Nokia Corp*., *Sony Corp*., *Texas Instruments Inc*., *Intel Corp*., etc. The other manufacturing sector consists of 64 firms that include *3M Co*., *Boeing Co*., *Conagra Inc*., *Du Pont*, *Exxon Mobil Corp*., *Nike Inc*., *Pepsi Co. Inc*., *Tyson Foods Inc*., *Kimberly-Clark Corp*., etc.

It can be seen that for the overall sector, the mean total comments for firms in portfolio 1 was around 2,964 and portfolio 5 was 125. This mean comment volume per firm for each month seemed low. This might be because *socialmention.com* monitors social media sites where privacy concerns may be high. For example, a person can choose who is able to see his or her posting on Facebook. If the posting is only made available to friends, then anyone outside of the friend's circle will be unable to see the posting. However, if a posting is made public, then everyone on Facebook and *socialmention.com* will be able to see the posting.

Considering the total comments as the social media metric, a total of 103 out of the 180 firms made it at least once into our top portfolio. However, only 34 firms made it to the top portfolio more than 10 times in the 26-month window, and only 16 firms made it to the top portfolio more than 20 times. Six firms, *Amazon, Canon, eBay, Microsoft, Nike, Nokia*, and *Sony* were in the top portfolio for all of the 26 months.

In terms of the bottom portfolio, there were 52 firms out of the 180 that were present in the bottom portfolio at least once. Twenty-nine firms appeared in the bottom portfolio at least 20 times in the 26-month period.

Considering the total negative comments as the social media metric, a total of 93 firms were in the top portfolio at least once in the 26-month period. This means they had the highest volume of negative comments during that month. Thirteen firms were in the top of the total negative comment portfolio for more than 20 of the 26 months. Firms that were consistently in the top portfolios in terms of the highest volume of total negative comments were *DirecTV, Clorox, Kroger, Monsanto, Verizon, Walmart, Comcast*, and *Microsoft*.

Looking at the proportion of negative comments with respect to the total comments, we can see that 34 firms were in the top portfolio on more than 10 occasions, and 12 firms on more than 20 occasions. The firms with the highest proportion of negative comments were *Clorox, Comcast, Direct TV, Kroger, Pep Boys, FedEx, Bank of America,* etc. Notice that although *Microsoft* and *Walmart* were in the top portfolio in terms of the volume of negative comments, they were not on the list for the top portfolio when I used the proportion of negative comments.

In terms of the proportion of positive comments, 29 firms appeared in the top portfolio more than 10 times, and 6 firms more than 20 times. The top firms in terms of appearance in the top portfolio were *Best Buy, Capital One, Revlon*, and *Allstate*. More details on the portfolios and their top ten firms in terms of appearance in that portfolio are provided in Table A8, and details for the bottom portfolio and the ten firms with the most appearances in the bottom portfolio are provided in Table A9.

For the monthly portfolio analysis, for each of the firms in the data set, the factor coefficients of the 4 factor Fama-French model were estimated using a linear regression with a 36-month rolling window (3 years) data set prior to the target month. In other words, for the monthly data, the Fama-French 4 factor model was estimated for each firm and each month to obtain firm-specific monthly factor loadings by using the 36-month data set prior to that month. In this way, the predicted returns were calculated for each firm for each month from June 2012 to July 2014. The abnormal returns were then calculated by subtracting the monthly stock returns, as reported in the financial data, and the monthly predicted returns, calculated above, as explained in Equation 2.

This entire process was then repeated for each of the remaining social media metrics: total positive comments, total negative comments, proportion of positive comments, and proportion of negative comments.

Figures 5 through 9 show the comparison of the top and bottom portfolio for the different social media metrics. For all of the social media metrics under consideration, if there were any differences between the top and bottom portfolio, then we would expect to see a general separation between the two trends. No such separation seemed apparent in these data. Looking at Figure B8 (Abnormal returns by the proportion of negative comments), it can be seen that the abnormal returns for portfolio 5 are generally above zero. However, for all other metrics, the estimated abnormal returns hovered around 0, and the top and bottom portfolio estimated abnormal returns crossed each other numerous times.

Table A10 shows the estimates of the average abnormal returns for all of the different portfolios under consideration for all five social media metrics. The first row provides the estimates for the abnormal returns for the five portfolios, based on the total comments about the firm in the social media. The estimated abnormal return for Portfolio 1 (i.e., firms with the highest volume of total comments) was 0.0007, with a *t*-statistic of 0.29. This was not significant at the 0.5% level. This result suggests that the abnormal return was not different from zero. Similarly, the estimated abnormal return for Portfolio 5 (i.e., firms with the lowest volume of total comments) was -0.0011, with a *t*-statistic of -0.403, which was also not significant. Therefore, there is no convincing evidence of mispricing. It can be seen that Portfolio 5 showed a lower value for abnormal returns than Portfolio 1.

The second row of Table A10 shows the estimates for the average abnormal returns for portfolios based on the total positive comments. The estimate for Portfolio 1 (i.e., firms with the highest volume of total positive comments) was 0.007, with a *t*-statistic of 0.303 and for Portfolio 5 was -0.0004, with a *t*-statistic of -0.153. None of the expected abnormal returns were significantly different from zero, which is suggestive of no market mispricing based on the total positive comments. Row 3 shows estimates for the average abnormal returns for portfolios based on the total negative comments. None of the estimated abnormal returns were significant in this case, either. The fourth row shows the estimates for the average abnormal returns for portfolios based on the proportion of positive comments, where the proportion of positive comments is the ratio of the total positive comments to the total comments for the firms. The estimate of abnormal returns for Portfolio 1 (i.e., firms with the highest proportion of positive comments with respect to total comments) was -0.0013, with a *t*-statistic of -0.49, which was also not significant and suggests no mispricing in the market. Considering the proportion of negative comments metric, the estimate of abnormal returns for Portfolio 1 was -0.0021, with a *t*-statistic of -0.856, which was not significant and thus suggests no mispricing. However, looking at the estimates of average abnormal returns in terms of the proportion of negative comments (row 5 of Table A10), Portfolio 5 (i.e., firms with the lowest proportion of negative comments) generate significant positive returns of 0.0053, with a *t*-statistic of 2.081. This suggests that firms with the lowest proportion of negative comments provide some significant abnormal returns.

Identifying the possible source of the mispricing observed could be possible if the overall data set is broken down by different industry sectors. This will also help us assess whether industry-specific differences are present in the mispricing. I break down the overall data set into the Information sector and the manufacturing sector, which is further broken down into Computer and Electronics manufacturing and all other manufacturing. I also study all sectors, except for the manufacturing and Information sector. The NAICS classification uses two digits of 31, 32, and 33 to identify the manufacturing sector and 51 to identify the Information sector. Portfolio analysis of the Information sector (Table A11) and all other sectors except for the information and manufacturing sector (Table A12) show no mispricing of any of the social media metrics in the top and bottom portfolio. This could indicate that the source or mispricing for the proportion of negative comments is not the Information sector or other sectors.

An analysis of the manufacturing sector (Table A13) abnormal returns shows that there is no mispricing of the total comments, total positive comments, total negative comments and proportion of positive comments metrics across the top and bottom portfolios. While considering the abnormal returns in terms of the proportion of negative comments (row 5 of Table A13), Portfolio 5 (i.e., manufacturing firms with the lowest proportion of negative comments) generates significant positive returns of 0.0075. This indicates that the manufacturing sector could be the source of the non-zero abnormal returns.

I further analyze the manufacturing sector by breaking it down into Computer and Electronics manufacturing and all other manufacturing. Portfolio analyses of all other manufacturing sectors show that there is no mispricing in the top and bottom portfolio for all of the social media metrics (Table A14).

However, a look at the Computer and Electronics product manufacturing sector (Table A15) shows that this sector could be the possible source of the mispricing that was previously observed in the overall data. It can be seen that Portfolio 5 of the proportion of negative comments metric in this sector shows significant returns.

It is also seen that Portfolio 1 for the total comments and total negative comments are also significant for the Computer and Electronics product manufacturing sector, along with Portfolio 5 of the proportion of negative comments metric. However this result seems to be generated by an extreme outlier. Data for the company *Nokia* for the month of September 2013 has almost 3 times the return compared to the next highest value for the month. This jump in stock returns seems to be because of *Microsoft's* announcement that it was going to purchase *Nokia* in September 2013. Analysis with this outlier removed shows no significant abnormal returns for the total comments, total negative comments. However this did not change the significance of the proportion of negative comments.

The significance of abnormal returns for the different portfolios in the Computer and Electronic sector could indicate that this sector may have characteristics that are different from other sectors. Thus mispricing may be the result of some other factors impacting the sector. Therefore it cannot be concluded that it is because of the social media metrics being discussed in the paper.

1.3.2 Monthly Portfolio Analysis (Lagged)

To understand lagged effect of social media metrics on firm abnormal returns I analyze how the different social media metrics impact a firms next month abnormal returns. The methodology used in this analysis is similar to the one used in the previous section. Table A17 through table A21 shows the results of this study. Table A17 shows the estimates of the average abnormal returns for all of the different portfolios under consideration for all five lagged social media metrics. Similar to the previous analysis the first row provides the estimates for the abnormal returns for the five portfolios, based on the total comments. The 2^{nd} through 5^{th} row provides estimates of abnormal returns for the next month based on total positive, total negative, proportion of positive, and proportion of negative comments respectively. Similar to the results of the contemporaneous portfolio analysis, none of the portfolio across the 5 measures of social media provides significant abnormal returns except for the portfolio 5 of the proportion of negative comments. Portfolio 5 which includes firms that have the lowest proportion of negative social media comments shows a small but significant positive abnormal return of 0.006 with a *t*-statistic of 2.486. This again suggests that firms with the lowest proportion of negative comments provide some significant abnormal returns, thus indicating mispricing.

I further analyze the results by breaking down the data set by different industries. Table A19 shows the result of portfolio analysis of the manufacturing sector. This result is similar to the previous contemporaneous analysis in that most of the portfolio shows a lack of significant abnormal returns. Portfolio 1, which includes the firms with the highest proportion of negative comments, provides a small but significant abnormal positive return of 0.8%, which is surprising and contrary to the expectations that negative comments would be detrimental to the firms. In terms of the Computer and Electronics manufacturing sector (Table A20) the portfolio that included firms that had high volumes of total comments provided a significant abnormal returns of 3%. This result is similar to the results explained in the previous section. From the results it seems that there is some evidence of mispricing in the Computer and Electronics manufacturing sector.

Similar to the results in the previous section no significant abnormal returns were found in the Information sector and Manufacturing sector (excluding the Computer and Electronics manufacturing). Thus it can be concluded that there is no mispricing in the Information and Manufacturing sectors.

Thus from the analysis of lagged social media metrics it can be concluded that there is no mispricing in the financial markets. There is no evidence that firms with better social media indicators (i.e. higher total comments, higher volume of positive comments, or higher proportion of positive comments) provide higher abnormal returns than firms that are worse in those indicators, in subsequent time periods. There is also no evidence that firms with negative social media indicators (i.e. higher volume of negative comments, or higher proportion of negative comments) provide lower abnormal returns than firms that perform better in those indicators.

1.3.3 Daily Portfolio Analysis (Lagged)

Portfolio analysis of the various social media metrics utilizing lagged daily data was also conducted. Table A5 shows the daily data summary for each firm-day observation. It can be seen that the average total comment for a firm is around 50 comments per day with an average of 2 negative comments and 10 positive comments with the majority of the comments being neutral. Firms in the data set have generally higher average proportion of positive comments, 17%, compared to the average proportion of negative comments of 3%. One reason that the daily comment volume totals are lower than the monthly comment volume is that the daily returns data does not include weekend and holidays data while the monthly data includes those comments.

The factor coefficients of the 4 factor Fama-French model were estimated using a linear regression with a 150-day rolling window data set prior to the target month. In this way, the predicted returns were calculated for each firm for each day from June 6th 2012 to June 27th 2014. The abnormal returns were then calculated by subtracting the daily stock returns, as reported in the financial data, and the monthly predicted returns, calculated above, as explained in Equation 2.

Only overall industry analysis was conducted for the daily data. It was not possible to conduct industry level analysis since industry specific breakdown yielded too few observations in terms of social media data. Table A23 shows the results of the daily level portfolio analysis. None of the portfolio provides any significant returns for any of the social media metrics which indicates a lack of mispricing at the daily level.

1.4 POWER TEST

This section reports the results of a power experiment to investigate whether the test has reasonable power to detect significant abnormal returns. I incorporate a monthly abnormal return of 5 basis point (μ=0.0005) and increase it by a multiple of c, where c is an integer scaler, such that a unit increase in c translates into an increment of 5 basis points in the monthly abnormal returns. In finance literature, monthly return of 10 basis points (c=2) is considered small while a monthly return of 50 basis points (c=10) is considered large (Fama and French 1996; Ray, Savin, & Tiwari 2009). The steps for the power calculations are described as follows.

S1) Run the following regression using the original data spanning an estimation window of 36 months from June 2009 to May 2012 (the same pre event, i.e., the time period prior to the creation and recording of user generated content, estimation period used in equation 2). Where $R_{i,t}$ = actual return for firm *i* during the month *t*.

$$R_{i,t} = \alpha_i + \beta_{m,i,t}\left(R_{m,i,t} - R_{f,t}\right) + \beta_{SMB,i,t}SMB_t + \beta_{HML,i,t}HML_t + \beta_{MOM,i,t}MOM_t + \varepsilon_{i,t}$$

S2) Calculate the error standard deviation for each firm *i* (using the 36 residuals for each firm from the above equation) for each firm.

S3) Simulate a time series of error terms for the various firms using $e_{i,t} = N(0, \sigma(\varepsilon_i))$. The length of the time series is equal to 26 months (the period with social media data) plus 36 months (estimation period as discussed in step 1).

S4) Using the estimated β values from Step (1), and using the random error terms generated from step (3) above, calculate the simulated excess returns (SR^{*}_{it}) for each of the 36 plus 26 (=62) months in the following manner:

a. For the initial 36 months (the pre-event window) simulate returns for each firm according to the following equation:

$$S_{pre}R^{*}_{i,t} = 0 + \beta_{m,i,t}\left(R_{m,i,t} - R_{f,t}\right) + \beta_{SMB,i,t}SMB_t + \beta_{HML,i,t}HML_t + \beta_{MOM,i,t}MOM_t + N\left(0, \sigma\left(\varepsilon_i\right)\right)$$

(Note: the abnormal returns generated above for the initial 36 months equals to zero)

b. For the 26 months event window, simulate returns for each firm according to the following equation

$$S_{event}R^{*}_{i,t} = \mu + \beta_{m,i,t}\left(R_{m,i,t} - R_{f,t}\right) + \beta_{SMB,i,t}SMB_t + \beta_{HML,i,t}HML_t + \beta_{MOM,i,t}MOM_t + N\left(0, \sigma\left(\varepsilon_i\right)\right)$$

(Note that the 'abnormal' return during the last 26 month period is non-zero, by design, since we are interested in studying the power of the test. I use different values of μ, starting with $\mu = 0.0005$ (c=1) for all firms for the first set of results, followed by $\mu = 0.0010$ (c=2) and so on.

S5) Using the pre event (36 months) of simulated excess return calculated above ($S_{pre}R^{*}_{i,t}$) run firm specific regressions to get the new estimates, $\hat{\alpha}_i$, $\hat{\beta}_i$, which are then used to calculate the residuals or 'abnormal' returns during event window (26 month) of the time series of simulated returns (these 'abnormal' returns are then used for portfolio analysis):

$$AR^{*}_{i,t} = S_{event}R^{*}_{i,t} - \left[\hat{\alpha}_{i,t} + \hat{\beta}_{m,i,t}\left(R_{m,i,t} - R_{f,t}\right) + \hat{\beta}_{SMB,i,t}SMB_t + \hat{\beta}_{HML,i,t}HML_t + \hat{\beta}_{MOM,i,t}MOM_t\right]$$

S6) Carry out the portfolio based analysis with total comment volume as the factor, as done in the original test. This time using the above abnormal returns (estimated in step 4).

S7) I repeated steps 3 to 5 a total of 10,000 times (for each case involving a particular value of μ used in Step 3b. above) and stored the results. I checked the proportion of times the test rejects (at the 0.05 level) the null hypothesis, namely, that the 'abnormal' portfolio return equals zero.

1.4.1 Results

Table A24 shows the result of the power calculation. At $\mu = 0.0005$ (abnormal return of 5 basis points) around 10% of the 10,000 simulations rejected the null hypothesis of abnormal return equals to zero. At c=10, i.e. μ=0.005 (abnormal return of 50 basis points) the rejection rate was around 50% while at c=20 (abnormal return of 100 basis points) the rejection rate was over 90%. This simulation results suggest that the test has reasonable power for economically meaningful portfolio abnormal returns in excess of 50 basis points per month. For abnormal returns that are lower than this threshold, there is insufficient power.

In light of the above results, the original abnormal returns that were documented are quite insignificant in economic terms (and hence, statistically insignificant). One important information to consider is that the abnormal returns calculated for the different portfolios in the original analysis were sometimes not of a consistent sign as well as were not of a sign as expected (for e.g. in Table A17 Portfolio 3 showed an abnormal return that was negative while all other portfolios 2 and 4 were positive for total positive comments. Similarly for the portfolio with the firms that had the highest proportion of positive comments showed an abnormal return which was negative while all other portfolios had positive abnormal returns).

The lack of rejection (of the null hypothesis) based on the original results suggests one of two possibilities

(a) That there is no significant, post event, measurable impact of social media content on firm stock returns (i.e., the market reacts efficiently to the publicly available social media content),

(b) That there is a significant, post event, impact of social media content on firm stock returns, (i.e., the market reacts inefficiently to the public social media content), but the test lacks sufficient power to detect this impact.

Based upon the results of the power calculation conducted above it can be concluded that the test has sufficient power to detect economically meaningful abnormal returns that are in excess of 50 basis points (one-half percent) per month. This provides evidence in support of the first possibility, namely, that there is no (economically) significant measurable impact of social media content on firm stock returns, after the public dissemination of such content (i.e., post-event). Since social media content is public information, the market incorporates that information and reacts efficiently to the social media content.

However, the study's findings should not be interpreted to mean that social media content is not important for firms. The findings only imply that the firm stock prices react efficiently to the social media content which is generally public information.

1.5 CONCLUSIONS

1.5.1 Summary, Limitations, and Future Research

As consumer and firm involvement in social media grows, there is increasing interest amongst researchers and managers to figure out the impact social media have on a brand or a firm's financial performance. This study seeks to contribute to the literature in this field and tries to ascertain whether firm performance can be affected by social media comment volume and the valance of those comments. Five different social media metrics (total comments, total positive comments, total negative comments, proportion of positive comments, and proportion of negative comments) were used to see if any of those metrics could provide important clues to managers and investors in figuring out firm financial performance. However, the Efficient Market Hypothesis would suggest that the price of any stock reflects all available information in the marketplace and provides an accurate estimate relating to the value of a firm. As such, using publicly available information, such as social media sentiment, should not provide any benefit to any investors because such information would already be priced into the stock market.

In this study, I demonstrated that the markets seem to be efficient when it comes to the benefits of using social media information to gain advantage. Most of the social media metrics that I used for the study show no possibility of gaining abnormal returns that are significantly different from zero. In terms of the daily analysis the results suggest that there is no evidence of mispricing and social media metrics does not have any impact on abnormal returns of any portfolios. The result from the monthly analysis also suggests that financial market mispricing due to information available in social media is mostly absent, and when present is limited to the Computer and Electronics manufacturing sector for the same month period. Why this is the case could be part of a future research study.

It should also be noted that the possible source of mispricing may not be the result of social media metrics, but of some anomaly in the Computer and Electronics manufacturing sector during the data collection period of 2012 to 2014. It could also be due to the financial market, which is slow to incorporate the impact of social media information in this sector. When the lagged social media metrics in the Computer and Electronics manufacturing sector is considered I find no significant abnormal returns for those portfolios. This could indicate that the market takes some time to adjust in the Computer and Electronics manufacturing sector.

A limitation of this study is the inability to differentiate between marketer generated content in the social media versus consumer generated content. However, the sheer volume of the number of consumer generated content on social media and the limited number of the same by companies means that the volume of marketer generated content might be significantly less compared to user generated content.

Another limitation is that this method of data collection does not have access to all word of mouth communication that might be available through other means. Facebook and other social media sites might have access to private conversation data pertaining to firms that we do not have access to. This means that the study is only able to capture a fraction of the actual conversations happening in social media.

The use of *socialmention.com* as a data source also has an impact on the study. The exact mechanism of data collection and classification of the comments is also unknown. The lack of this knowledge means that there is no explanation for the spike in the total comments volume observed in the ending time periods of this data set.

Finally there is possibility of measurement error. This is because sentiment classification algorithms used by the website is unlikely to be error free and the mechanism of classification is unknown.

This study indicates that the market accurately values social media information in most cases, if such value exists. This may be because the markets react quickly and accurately to any such available information. Also, the absence of a significant result in our analysis does not necessarily mean that there is no impact of social media on firm financial performance. The results in this study only show us that the markets react efficiently with regards to information contained in the social media, which is public knowledge. In the next chapter, I look into the relationship between the various social media metrics on a firm’s cash flow to see whether social media has an impact on firm financial performance.

CHAPTER 2: SOCIAL MEDIA AND FIRM QUARETERLY CASH FLOWS

The brand value chain model (Keller and Lehmann 2003) suggests that consumers' mindset play an important role in consumer decisions on whether to buy a brand and how much to pay. User generated content, and sentiments expressed in those contents, play an important role in developing or changing consumer mindset about a product or a brand. The decision on purchase of the product and how much should be paid, then impacts how the brand performs in the marketplace. While I analyzed the direct link between user generated content/social media and shareholder value in the first chapter, in this chapter I move further up the brand value chain and investigate the relationship between social media and firm performance as measured by firm quarterly cash flow.

Only a decade ago, most interpersonal communication happened in person or over the phone. Today, billions of people converse with each other via Internet-based social media (Sonnier et al. 2011). While traditional communication was one to one, or one to a few, online communication is one-to-many, or many-to-many, which facilitates the diffusion of online user generated content (Godes et al. 2005). Consumers trust in online user generated content is high, and their trust on recommendations from friends and family is even higher (Nielsen 2012). This should mean that consumer trust on user generated content in social media should also be higher. This is because social media sites are generally built around a community of people, including friends and family, whom we know. As such information available through social media should have higher credibility than anonymous online user generate content.

Researchers have found that source credibility, and strong ties between information source and receiver, should have a high impact on consumer decision making (Bansal and Voyer 2000; Brown and Reingen 1987; Forman et al. 2008). Godes et al. (2005) also argue that consumers' choice is affected by other people's actions and these actions include recommendations from a friend. A survey by Dimensional Research found that 88% of participants stated that reading an online review about customer service had an impact on their purchase decisions (Zendesk 2013). Mayzlin (2006) also found that consumers' online reviews impacted consumer purchase decisions. Hence information shared in social media regarding a brand or a product should have a higher impact on consumer decision making and choice, which should lead to improved sales, which then should lead to improved cash flow (Srivastava et al. 1998).

Some researchers have argued that user generated content/social media has an impact on firm performance and have examined the direct link between the two. For example, Anderson et al. (2004) suggest that, in terms of consumer word of mouth and user generated content, positive word of mouth and recommendations from satisfied customers should influence shareholder value. Establishing a link between user generated content and firm's cash flow, Fornell (1992) and Anderson (1998) suggest that positive word of mouth should lower acquisition costs for new customers which should increase net cash flow. Anderson et al. (2004) further suggest that positive word of mouth should help a firm penetrate new and existing markets which should lead to accelerated cash flow. Other researchers (Luo 2009; Anderson 1996; Singh and Pattanayak 2014) also suggest that consumer word of mouth has an impact on firm intangible assets such as customers repurchase intention, defection rates, new customer acquisition, and lower acquisition cost, all of which should have an impact on future cash flows of the firm. Pertaining to this relationship between user generated content and firm cash flow, Luo (2009) found that a high level of negative word of mouth led to more shortfalls in a firm's future cash flow and a lower level of negative word of mouth led to fewer shortfalls in a firm's future cash flow in the airline industry. Tirunillai and Tellis (2012) suggest that consumers consult user generated content online to make purchase decisions, which translates into sales and therefore, into higher future cash flows.

In summary, comments and sentiments expressed in social media reflects peoples positive or negative experiences, which in turn influences other consumers and their decision to engage with the product or the firm. So I expect that any social media comments in the current time period would influence other consumers purchase decisions and subsequently the firms' performance in the subsequent time period. Hence, in this essay (Essay 2), I examine the relationship between social media metrics and firm financial performance by examining the following two questions.

1. Does social media comment volume or sentiment provide any improvements in a firm's quarterly cash flow forecast?
2. Does a significant relationship exists between social media comment volume, social media sentiment and firm quarterly cash flow?

2.1 DATA AND MEASURES

Quarterly cash flow data and other financial data was collected from COMPUSTAT and CRSP using the Wharton Research Data Services. Data from the 3nd quarter of 2012 to the final quarter of 2015 was obtained which is a total of 14 quarters of financial data. However, most firms' quarterly financial data for the final quarter of 2015 was not yet available in Compustat at the time of this analysis hence the final data includes 13 quarters of financial data.

2.1.1 Cash Flow as a Measure of Firm Performance in Marketing Literature

This study will use cash flow from operations as a measure of firm performance. Gruca & Rego (2005) suggest that "modeling future cash flow is consistent with the current theory (and practice) of firm valuation". Previous researchers who suggest using cash flow as a measure of firm value include Rappaport (1986), Srivastava et al. (1998), Gurca & Rego (2005), and Rao & Bharadwaj (2008). Rappaport (1986) stresses that marketing actions must be linked with cash flow as it is the ultimate determinant of firm value and wealth created for shareholders. Cash flow is considered to be a key determinant of a firm's value (Day and Fahey 1988; Srivastava et al. 1998; Gruca and Rego 2005). Ambler and Roberts (2006, p. 4) state "at the end of the day, marketing is the creation of cash flows". Ideally, a firm's current market value is determined by its future cash flows (Srivastava et al. 1998; Mizik and Jacobson 2009). Moreover, Srivastava et al. (1998) suggest that the value of any marketing strategy is driven by its ability to 1) accelerate cash flow, 2) increase the level of cash flow, 3) reduce the risk associated with cash flow, and 4) increase the residual value of cash flows. Thus, utilizing cash flow in studies of marketing assets, activities or outcomes should help researchers better understand the impact of marketing on firm value (Srivastava et al., 1998). Rust et al (2004) considered that marketing activities generate a reservoir of cash flow that has not yet translated into revenues. Gruca and Rego (2005) showed that customer satisfaction increased future cash flow while reducing cash flow variability.

Studying the impact of brand related assets on cash flow, Larkin (2013) found that brand stature reduced the forward-looking volatility of cash flows. Rust et al (2004) further suggest that the value of brand equity is due to the incremental discounted cash flow generated from sales. Kim, Mahajan, and Srivastava (1995) demonstrated a strong relationship between growth in customer base and cash flow in the cellular industry. Tim (1997) mentions customers' and influencers' attitude impact brand equity which subsequently impact cash flow. To this end, previous research has focused on cash flow as a measure of brand and firm performance and, therefore, value.

However, there are few studies examining the ability of metrics associated with user generated content to improve forecasts of firm cash flow. Outside the realm of online user generated content and social media Luo (2009) studied negative word of mouth and its impact on firm's future cash flow in the airlines industry.

2.1.2 Variables Used for Cash Flow Analysis

The variables that were used for the quarterly cash flow analysis, and how they were calculated, are listed next.

- CFO: - Cash flow from operations. Quarterly cash flow values were obtained by using the yearly cash flow data available in Compustat with the code *oancfy*. Quarterly cash flow data was calculated by subtracting the previous quarter's *oancfy* value from the target quarter except for the 1st quarter.
- OIBD: - Operating income before depreciation obtained using COMPUSTAT with code *oibdpq*
- REC: - Accounts receivable obtained using Compustat with code *rectq*

- INV: - Inventory in stock obtained using Compustat with code *invtq*
- PAY: - Accounts payable obtained using Compustat with code *apq*
- CFOA :- Cash flow from operations (*CFO*) divided by total assets (*atq*)
- OIBDA: - Operating income before depreciation obtained using COMPUSTAT with code *oibdpq* and divided by total assets
- RECA: - Accounts receivable obtained using Compustat with code *rectq* and divided by total assets
- INVA: - Inventory in stock obtained using Compustat with code *invtq* and divided by total assets
- PAYA: - Accounts payable obtained using Compustat with code *apq* and divided by total assets.
- ASSETS: - Total assets/liabilities of the firm on that quarter. Obtained using Compustat with code *atq.*
- EARN: - Earnings reflect earnings before extraordinary items (*ibq*) divided by total assets (*atq*).
- ACCRUAL: - Calculated by subtracting *CFOA* from *EARN*.
- SIZE :- Size of the firm as calculated by taking the log of sales (log (*saleq*))
- BTM: - Book to Market ratio of the firm at the end of the quarter. Calculated as (*ceqq/prccq*cshoq*). Where *ceqq* is the Total Common/Ordinary Equity, *prccq* is the closing price of the quarter, and *cshoq* is the common shares outstanding.
- LEVERAGE:- Ratio of long term debt to market value of equity. Calculated as (*dlttq/prccq*cshoq*). Where *dlttq* is the Total long-term debt, *prccq* is the closing price of the quarter, and *cshoq* is the common shares outstanding.

- AGE: - Age of the firm in years based on initial pricing data availability on the CRSP monthly file.
- DIV: - Indicator variable equal to one if the firm had dividend in that quarter and zero otherwise.
- BUSSEG: - Number of operating segments reported by firm on the Compustat segments file.
- GEOSEG:- Number of geographic segments reported by firm on the Compustat segments file.

Table A26 reports the summary statistics for the 165 firms that were included in this analysis. The table summarizes statistics for a total of 15 industry sectors as per the NAICS classification. The highest concentrations of firms are in the manufacturing sector. The Computer and Electronic manufacturing sector contains 17 firms while the smallest sector in the dataset is accommodation and food sector with a total of 3 firms. The average number of total comments per quarter per firm was 5,365 with average positive comments of 860 comments and average negative comments of 171. In terms of proportion of negative and positive comments the average values were 0.035 and 0.168 respectively. The Information sector which includes broadcasting and publishing industry were the sector with the highest volumes of comments while the beverage and tobacco manufacturers generated the least amount of social media comments. The fast food and broadcasting sector was the most negatively commented upon with around more than 5% of the comments being negative while the fast food sector had the highest proportion of positive comments with around 20% of the comments being positive.

Table A25 reports the summary statistics for different variables included in this quarterly cash flow analysis. The average age of the firms with available data was 39 years with a Book to market ratio of 0.362 and leverage of 0.364. The firms had average total assets of $112 billion (median $19.7 billion) and average operating cash flow of $1.3 billion (median $375 million).

2.2 METHODOLOGY AND RESULTS

Quarterly cash flow prediction models are a topic of very high interest to both accounting and finance researchers (Lorek 2014). Therefore, a number of excellent prediction models are available (Lorek 2014). Some of these models are cross-sectional models such as those used by Wilson (1986, 1987), and Bernard and Stober (1989). Other models such as the multivariate time series regression model (Lorek & Willinger 1996, 2011) is also widely used in accounting literature. Univariate time series models such as a model attributed to Brown and Rozeff (1979) and Griffin (1977) are also used as quarterly cash flow prediction models, but will not be considered for this study because of the lack of lengthy time series data.

For my analysis, I use a cross sectional regression model based on Folsom, Hribar, Mergenthaler, and Peterson (2016) as well as a multivariate time series regression model of quarterly cash flow popularized by Lorek and Willinger (1996, 2011).

2.2.1 Multivariate Time-Series Regression Model (MULT)

The multivariate time-series regression model approach was popularized by Lorek and Willinger (1996). This model has strong empirical support in the literature based on its short term (one step ahead) predictive power (Lorek and Willinger 2011). The model is specified below

$$\begin{aligned} CFO_t = \alpha + \beta_1(CFO_{t-1}) + \beta_2(CFO_{t-4}) + \beta_3(OIBD_{t-1}) + \beta_4(OIBD_{t-4}) + \beta_5(REC_{t-1}) \\ + \beta_6(INV_{t-1}) + \beta_7(PAY_{t-1}) + e_t \end{aligned} \tag{5}$$

where, CFO_t is the operating cash flow at time t, $OIBD_{t-1}$ is operating income before depreciation at time t-1, REC_{t-1} is the accounts receivable at time t-1, INV_{t-1} is inventory at time t-1, PAY_{t-1} is the accounts payable at time t-1, and e_t is the current error term. The inclusion of lagged values of the dependent variable CFO at (t-1) and (t-4) is intuitively consistent with ARIMA modeling procedures which rely on past values to predict future values (Lorek and Willinger 1996). CFO_{t-1} is included to capture adjacent effects of cash flow while CFO_{t-4} is included to capture seasonal effects of cash flow. OIBD, REC, INV, & PAY are values relating to accrual based earnings. Lorek and Willinger (1996) suggest that OIBD provides a better proxy for accrual based earnings and also a better descriptive fit and results in better cash flow predictions than when net income is used. Similar to CFO an adjacent and seasonal lag is used as proxy for earnings. REC, INV, and PAY are variables whose use is consistent with previous cash flow prediction models used by Lorek and Willinger as well as Wilsons (1986) cross-sectional cash flow regression model. Lorek and Willinger (1996) state that this multivariate time series regression model allows firm-specific parameter estimations unlike a cross sectional regression model and also includes parsimonious set of accrual accounting variables, which should be better than ARIMA models which only uses past values of cash flow series.

Lorke and Willinger (1996) suggest that the selection of independent variables in this models were based on the following considerations –

- A desire to create a model which was more parsimonious than previous cash flow prediction models.
- Inclusion of lagged values of the cash flow at t-1 and t-4 are consistent with ARIMA models that utilize past dependent variable values to predict future values.
- CFO_{t-1} should be able to capture the adjacent effects of cash flow and CFO_{t-4} should be able to capture the seasonal effects of cash flow.
- OIBD is included as a proxy for accrual-based earnings. First lag and fourth lag are included to capture adjacent and seasonal effect similar to cash flow from operations mentioned above.
- REC, PAY, INV use is similar to previous multivariate cross sectional regression model used by Wilson (1986). Lagged variables of these variables are used to invoke a random walk assumption.

Apart from the variables included in the MULT model I incorporate the social media measures in equation 5 which provides firm specific estimates of the coefficients.

$$\begin{aligned} \mathrm{CFO_t} = \alpha &+ \beta_1(\mathrm{CFO_{t-1}}) + \beta_2(\mathrm{CFO_{t-4}}) + \beta_3(\mathrm{OIBD_{t-1}}) + \beta_4(\mathrm{OIBD_{t-4}}) + \beta_5(\mathrm{REC_{t-1}}) \\ &+ \beta_6(\mathrm{INV_{t-1}}) + \beta_7(\mathrm{PAY_{t-1}}) + \beta\ \mathrm{SMM}_{t-1} + e_t \end{aligned} \tag{6}$$

The one step ahead quarterly cash flow prediction was generated in an ex ante fashion for the MULT model using 12 quarters of data from 3rd quarter of 2012 to 2nd quarter of 2015. First I used equation 5, not including social media metrics variables, to get the coefficients for each individual firms in the data set by utilizing ordinary least squares estimation technique. Then I used those estimates to generate the 13th quarter's cash flow prediction. I then use equation 6, which includes social media metrics variable, to generate the coefficients of all variables, for each individual firm, which is then used to calculate the predicted value of the 13th quarter cash flow. I then use the actual cash flow value for the 13th quarter, obtained using Compustat, and compare it with both predicted values to determine which model provides more accurate forecasts for each company.

I use the same methodology used by Lorek and Willinger (2011) and use Mean absolute percentage error (MAPE) as the primary error metric. The MAPE is calculated as

$$\text{MAPE} = \frac{\sum \left| {}^{(\text{A-F})}\!/_{\text{A}} \right|}{\text{N}} \tag{7}$$

where, N is the number of predictions, A is the actual quarterly CFOs, and F is the forecasted quarterly CFOs. Consistent with Lorek & Willinger (1996, 2011), all forecast errors greater than 100% are truncated to 100%.

Comparison of the two models, i.e., MAPE of MULT without the social media metrics (equation 5) and MULT with social media metrics (equation 6), provides us with an indication of whether the use of social media metric provide any benefit in terms of cash flow predictions.

I also analyze the multivariate time-series regression model by using contemporaneous measure of social media metrics.

$$\begin{aligned} \text{CFO}_{\text{t}} = \alpha + \beta_1(\text{CFO}_{\text{t-1}}) + \beta_2(\text{CFO}_{\text{t-4}}) + \beta_3(\text{OIBD}_{\text{t-1}}) + \beta_4(\text{OIBD}_{\text{t-4}}) + \beta_5(\text{REC}_{\text{t-1}}) \\ + \beta_6(\text{INV}_{\text{t-1}}) + \beta_7(\text{PAY}_{\text{t-1}}) + \beta\ \text{SMM}_t + e_t \end{aligned} \tag{8}$$

One reason for using contemporaneous measure is that social media information flow is fast and any positive or negative conversation online may have an impact on the firm within a matter of days. This could mean that social media conversations may affect the firm cash flow in the same quarter. I utilize the same approach of comparing MAPE to analyze whether contemporaneous social media information provides any benefits for cash flow predictions.

2.2.1.1 Results - Multivariate time-series regression model (MULT)

Table A26 and 27 shows the MAPE values for the MULT models without and with lagged social media metrics. Forecast errors that were greater than 100% were truncated to 100 percent which is similar to the procedure used by Lorek and Willinger (1996). The contemporaneous MULT model without social media metrics had a truncation frequency of 29% which was lower than the truncation frequency of all other models that utilized some form of social media metrics. The lagged MULT model without social media metrics had a truncation frequency of 32% which was also lower than all other models with lagged social media metrics.

The best performing MULT model, amongst both contemporaneous and lagged, was the prediction model based on no social media information. The MAPE metrics of the model without social media metrics was 50.5% for the contemporaneous model and 55% for the lagged model. These figures are very similar to the ones obtained by Lorek and Willinger (1996) in their MULT quarterly cash flow prediction model. Compared to this model that does not incorporate any social media information, all others models that incorporate various social media metrics have a higher value of MAPE. The MAPE values for the contemporaneous model ranged from 53.45% for total negative comments to 55.46% for proportion of negative comments. In terms of lagged social media metrics the MAPE values ranged from 56.63% for proportion of negative comments to 60.53% for total positive comments. This suggests that inclusion of those social media metrics does not provide any additional value to the prediction of future quarterly cash flow.

Considering the changes in adjusted R^2 values of the entire 165 individual firm specific regression 62 firms saw increases in the value of adjusted R_2, when lagged total comment was included in the MULT model, while 85 firms saw decreases. The remaining 17 had adjusted R^2 values that were negative. Without the lagged total comment in the model the average adjusted R^2 value for the 142 firms, excluding the 23 firms with negative values, was 0.666. When the lagged total comment is included in the model the adjusted R2 value was 0.695. For lagged positive comments, 60 firms saw increases and 91 firms saw decreases in adjusted R^2 values after including the variable. The average adjusted R^2 value went from 0.666 to 0.691 when positive comment was included. For lagged negative comments, 52 firms saw increases while 95 saw decreases with the average adjusted R^2 going from 0.666 to 0.699. For lagged proportion of positive comments, 55 firms saw increases while 92 saw decreases with the average adjusted R^2 going from 0.666 to 0.691. Similarly, for lagged proportion of negative comments, 62 firms saw increases while 84 saw decreases with the average adjusted R^2 going from 0.666 to 0.691.

Also when considering individual firm specific regression only 14 out of the 164 firms had the lagged total comment value significant at 0.05 level of significance. The number of firms that had significant lagged social media variables were 16 for total positive comments, 14 for total negative comments, 8 for proportion of positive comments, and 14 for proportion of negative comments respectively.

It should be kept in mind, while interpreting the above results, that the number of observations for each of the OLS regression is only 12 which may be considered to be too few.

2.2.2 Cross-Sectional Regression Model

A cross-sectional regression model is utilized to examine the association between current social media metrics and future cash flows. The empirical specification for this model is based on prior literature (e.g. Folsom et al. 2016, Atwood et al. 2010) in accounting.

$$\begin{aligned}CFOA_{it+1} = {} & \beta_0 + \beta\, SMM_{it} + \beta_1(EARN_{it}) + \beta_2(SIZE_{it}) + \beta_3(BTM_{it}) \\ & + \beta_4(LEVERAGE_{it}) + \beta_5(DIV_{it}) + \beta_6(AGE_{it}) + \beta_7(BUSSEG_{it}) \\ & + \beta_8(GEOSEG_{it}) + \beta_9(EARN_{it}\ X\ SIZE_{it}) + \beta_{10}(EARN_{it}\ X\ BTM_{it}) \\ & + \beta_{11}(EARN_{it}\ X\ LEVERAGE_{it}) + \beta_{12}(EARN_{it}\ X\ DIV_{it}) \\ & + \beta_{13}(EARN_{it}\ X\ AGE_{it}) + \beta_{14}(EARN_{it}\ X\ BUSSEG_{it}) \\ & + \beta_{15}(EARN_{it}\ X\ GEOSEG_{it}) + e_{it}\end{aligned} \tag{3}$$

Subscript i is firm and t is the year. This model is a yearly cash flow model and will be the basis of my quarterly cash flow model. Previous quarterly cash flow models have used a 4th quarter lagged as it has been shown to be the best predictor of current quarterly cash flow because of seasonality. All other variables are as described in the previous sections. Folsom et al. (2016) suggest that inclusion of firm size, number of business segments, number of geographical segments, and industry fixed effects as control variables help to control firm level complexity in the model. Two different specifications of the main regression model (equation 3) is presented in this paper. In the first specification, industry and time fixed effects is included to control for correlations across industries and time. The second specification consists of firm and year fixed effects. Firm and time fixed effect is also included to hold the economics of the firm constant. While this is an established methodology that has been used by Hribar et al. (2015), Atwood et al. (2009), etc. I analyze the appropriateness of this model for this current data. Detailed explanation of the model and model selection process is detailed next.

The data collection in this study was done in a time series of cross sectional observations. The final data consists of a cross section of 164 firms for which data had been collected daily for over 13 quarters. Generally with a cross sectional time series data, panel data estimation techniques are used. This is because ordinary least squares estimation assumes errors to be independent and identically distributed. This assumption may be violated in panel data. Also using pooled OLS would not utilize the firm specific and time specific dimensions available to us. In panel data analysis there are a few estimation methods that are popular, the fixed effects and the random effects model. In a fixed effects model a regressor and individual or time specific error are not independent, while in the random effects model the individual specific errors are not correlated with the regressor. Pooled OLS would be the most efficient estimator if the individual component was missing completely.

I first estimate the pooled, fixed, and random effect model for the data set and utilize the Breusch-Pagan Lagrange multiplier test to accesses whether there is any panel effects present in the data or not, i.e. whether the data can be treated as pooled data that can be estimated by OLS or whether any panel data estimation techniques are required. The Breusch-Pagan Lagrange multiplier test was significant at the 0.05% level with a p value of 0.008. This implies that panel effects are present in the data and panel data estimation is preferable to OLS techniques.

After identifying that panel data estimation is preferable I used the Hausman's specification test to choose between the fixed effect and the random effect model. The null hypothesis for the Hausman's test is that the random effects model is preferred vs. the alternative hypothesis of fixed effects model being preferred (Green, 2008, chapter 9). The p value of the test is 0.0028 which suggests that the null hypothesis of random effects is rejected and hence a fixed effects model is the preferred model. Furthermore I also test for whether time-fixed effects are present in the panel data and found that there is a need to use time fixed effects based on the Breusch-Pagan Lagrange Multiplier test and a F test for time specific effects.

The results of the above tests indicate that a fixed effects model incorporating both firm and time fixed effects is an appropriate model for the data. The established approach utilized by Folsom et al. (2016), Atwood et al. (2010) also incorporate both firm and time fixed effects in their model of cash flow prediction. The firm fixed effects will help to control for unobserved factors that differ between firms but are constant over time for each firm, while the time fixed effects will help control for unobserved factors that are shared by all firms at a given point in time.

$$
\begin{aligned}
CFOA_{it+1} = \alpha_0 &+ \beta_0 SMM_{it} + \beta_1(EARN_{it}) + \beta_2(SIZE_{it}) + \beta_3(BTM_{it}) + \beta_4(LEVERAGE_{it}) \\
&+ \beta_5(DIV_{it}) + \beta_6(AGE_{it}) + \beta_7(BUSSEG_{it}) + \beta_8(GEOSEG_{it}) \\
&+ \beta_9(EARN_{it} \text{ X } SIZE_{it}) + \beta_{10}(EARN_{it} \text{ X } BTM_{it}) \\
&+ \beta_{11}(EARN_{it} \text{ X } LEVERAGE_{it}) + \beta_{12}(EARN_{it} \text{ X } DIV_{it}) \\
&+ \beta_{13}(EARN_{it} \text{ X } AGE_{it}) + \beta_{14}(EARN_{it} \text{ X } BUSSEG_{it}) \\
&+ \beta_{15}(EARN_{it} \text{ X } GEOSEG_{it}) + \beta_{16} CFOA_{it} + \beta_{17} CFOA_{it-3} + e_{it}
\end{aligned} \tag{4}
$$

$$e_{it} = \mu_i + \lambda_t + \nu_{it}$$

Where μ_i is the firm effects and λ_t is the quarter effects and ν_{it} is the error term.

I also test for serial correlation in the data using Wooldridge's test for FE panels, which is considered a good test for panels with large N and short T and is also robust to general heteroscedasticity. The Wooldridge's test for serial correlation in fixed effect panels shows a p value of 0.809 with the null hypothesis being that there is no serial correlation. Hence serial correlation does not seem to be a problem for this study.

To check for unit root I use the Augmented Dickey-Fuller test which suggests that there is no unit root present in the cash flow series during the duration of this study and hence the time series is stationary.

The Breusch Pagen test for heteroscedasticity suggests that there is presence of heteroscedasticity. The presence of heteroscedasticity should not impact the estimates themselves but only their standard errors. This problem can be solved by estimating heteroscedasticity robust standard errors while estimating the fixed effects model. The results in terms of significance of the social media metrics did not change even after using heteroscedasticity consistent standard errors.

2.2.2.1 Results – Cross Sectional Regression Model

Table A30 and A31 presents regression estimates of future cash flows on current social media metrics and current earnings and its interactions with other control variables to determine whether the various social media metrics have any effect on cash flow. Both the firm and time fixed effects are included in this estimation but is not shown in the table. The coefficient of the social media metrics, β, reflects how that metric relates to the next quarter's cash flow from operations. A significant positive coefficient would suggest a positive relationship between total comment volume and future quarterly cash flow while a significant negative coefficient would suggest a negative relationship.

Table A30 shows the results of the regression examining the relationship between social media and one quarter ahead cash flow with firm and time fixed effects while table A31 shows the same with industry and time fixed effects. The β coefficient, representing coefficient of total comments, is not statistically significant suggesting that there is no relationship between total comments and next quarter's cash flow. Similar analysis of the other social media metrics also suggest no significant relationship between those social media metrics and quarterly cash flow except for positive comments. The coefficient for positive comment has a value of -0.0000026 and is statistically significant at the 0.05 level.

Comparing the adjusted R^2 value between the models that included the social media metrics and did not include those metrics it can be seen that inclusion of social media metrics did not change the adjusted R^2 values which indicates that there was no benefits accrued by adding the social media information to the model.

2.2.3 Panel Data Analysis of Lorek and Willinger (1996) Model

The Lorek and Willinger (1996) model with the addition of social media metrics is specified below. The multivariate time series regression model conducted in the previous section has a limitation as the time series available in the data is not extensive.

Only 13 quarters of data for each firm is available which means that only 12 quarters of data were used to estimate the OLS model in the previous section. Since the data is of a time series and cross sectional nature, it is possible to use panel data estimation techniques with the same variables utilized by Lorek and Willinger (1996), which will be done in this section.

Table A32 – 36 shows the results of the pooled OLS, fixed effects, random effects, and first differenced model. However, the first step in this section is to identify which panel data estimation should be the most appropriate for the given data set. I first use the Breusch-Pagan Lagrange multiplier test to accesses whether there is any panel effects present in the data or not, i.e. whether there is evidence of significant differences across companies leading to a simple OLS regression not being preferred. The Breusch-Pagan Lagrange multiplier test was significant at the 0.05% level. F test of the joint significance of the fixed effects intercept, where the null hypothesis is that all fixed effect intercepts are zero, is also conducted. In this test the null hypothesis is rejected at 0.05 level indicating that a fixed effects model is appropriate instead of the pooled OLS regression model.

After identifying that fixed effects panel data estimation is preferable I used the Hausman's specification test to choose between the fixed effect and the random effect model. The null hypothesis for the Hausman's test is that the random effects model is preferred vs. the alternative hypothesis of fixed effects model being preferred. The p value of the test is significantly lower than zero which suggests that the null hypothesis be rejected and hence a fixed effects model is again the preferred model. When testing for whether time-fixed effects should be included in the model the Breusch-Pagan Lagrange Multiplier test and a F test suggest that time specific effects be included in the model.

To check for unit root I use the Augmented Dickey-Fuller test which suggests that there is no unit root present in the cash flow series during the duration of this study and hence the time series is stationary. This suggests that there is need to take the first difference of the variables.

The Breusch Pagen test for heteroscedasticity suggests that there is presence of heteroscedasticity. As mentioned earlier the presence of heteroscedasticity should not impact the estimates themselves but only their standard errors. This problem is solved by estimating heteroscedasticity robust standard errors while estimating the fixed effects model. The results did not change even after using heteroscedasticity consistent standard errors.

I also test for serial correlation in the data using Wooldridge's test for FE panels, which shows a p value of 0.735 with the null hypothesis being that there is no serial correlation. The Wooldridge's test is suitable for "short" panels with small T and large n. This indicates that serial correlation does not seem to be a problem for this study.

I further tested if a first differenced model is better than the fixed effect model by using Wooldridge's first-difference-based test. Wooldridge (2002) proposes using a serial correlation test, which can also be considered a specification test, to choose between the fixed effects model and the first difference model. This test is done for both the fixed effects and first differenced model. The fixed effects model has a null hypothesis of no serial correlation in original errors while the first differenced model has a null hypothesis of no serial correlation in differenced error. The p value for Wooldridge's first difference based test for the fixed effects was significant at 0.05 level and the p value for the same test for the first differenced model was also significant at the 0.05 level. This indicates that there is serial correlation present in errors and hence whichever model is used will have issue with serial correlation.

Based on the above analysis the final model chosen was the fixed effects model that includes time fixed effects as well.

$$\begin{aligned} \mathrm{CFOA}_{\mathrm{it}} = {} & \beta_1\mathrm{CFOA}_{\mathrm{it-1}} + \beta_2\mathrm{CFOA}_{\mathrm{it-4}} + \beta_3\mathrm{OIBDA}_{\mathrm{it-1}} + \beta_4\mathrm{OIBDA}_{\mathrm{it-4}} \\ & + \beta_5\mathrm{RECA}_{\mathrm{it-1}} + \beta_6\mathrm{INVA}_{\mathrm{it-1}} + \beta_7\mathrm{PAYA}_{\mathrm{it-1}} + \beta_7\mathrm{SMM}_{\mathrm{it-1}} \\ & + e_{it} \end{aligned} \tag{9}$$

Where

$$e_{it} = \mu_{\mathrm{i}} + \lambda_{\mathrm{t}} + \nu_{\mathrm{it}}$$

The Arellano method (1987) is widely used to estimate the heteroscedasticity and serial correlation consistent covariance estimator for fixed effects model. I use the method to estimate the fixed effects model. Utilizing the method did not change any of the results.

I repeated the above analysis for all the social media metrics.

2.2.3.1 Results – Panel data estimation

Table A32 to A36 presents the results of the panel data estimation for all the social media metrics under consideration. The total number of observations was 1968 for each of the models. As discussed in the previous section the fixed effects with time fixed effects estimation technique is the preferred estimation technique for analysis of the data.

Table A32 shows the results of all panel data model with total comment volume as a social media metric. For the fixed effects with time fixed effect model, while the other variables that were used by Lorek and Willinger (1996) such as OIBD, Account Receivable, Inventory, Accounts Payable are significant the total comment volume is not significant.

Analyzing the results with total positive comments, total negative comments, proportion of positive comments, and proportion of negative comments I fail to find any significance of those variables in the model.

2.3 CONCLUSION

2.3.1 Summary, Limitations, and Future Research

Social media and consumer user generated content is a hot button topic in marketing. As consumer involvement in social media grows, there are concerns on how this impacts firm performance. This second essay contributes towards identifying the impact of the various social media metrics on firm financial performance. While many research studies examine the relationship between user generated content and marketing metrics (such as sales) and financial performance metrics (such as stock prices and stock abnormal returns) there is little research on the impact on firm cash flow. This study attempts to bridge this gap and provides us with a clearer picture on the relationship between social media and firm performance. Five different social media metrics (total comments, total positive comments, total negative comments, proportion of positive comments, and proportion of negative comments) were used in this study to see if any of those metrics had a significant effect on firm quarterly cash flows.

The results suggest that no such effect of social media on cash flow exists. Analysis of future cash flow prediction model utilizing the multivariate time series regression model showed no significant benefit in using social media data while predicting future quarterly cash flows. The cross sectional regression analysis failed to show significant effect of any of the social media metrics on cash flow as well.

The limitation of this study is similar to the limitation of my previous study in chapter 1. In this study I am unable to differentiate between marketer generated content in the social media versus consumer generated content. Also this data set does not incorporate all social media data that are present online because of privacy settings. Valuable data sources in terms of social media such as Facebook do not allow insights into private conversations which limits the data availability from such sources. This means that not all social media conversation is being captured by our data source. Another limitation is the lack of knowledge regarding how *socialmention.com* collects and classifies data.

CHAPTER 3: GENERAL DISCUSSION AND CONCLUSIONS

In recent years the use of digital media among the general population has become ubiquitous. Consumers are spending more time using various firms of digital medium to gain information as well as communicate. Amongst the various forms of digital media one form is of particular interest to marketers, namely social media. Consumers are now not only able to learn about product and services from their friends, families, and strangers but are now also able to share about their own experiences with the world. This has created interest amongst marketing managers as they grapple to fully understand how social media content has an impact on their bottom lines.

The relationship between social media and firm performance is based on the idea that social media comments and sentiments expressed in those comments should have an impact on other consumers' mindsets about the given firm. Whether a consumers' mindset is positive or negative should then influence their decision about purchasing the firm's product or services. This affects how the firm's products performs in the marketplace, which ultimately impacts the firm's financial performance in terms of cash flow generations or stock market returns.

Previous research - in the context of online user generated content and its impact on firm performance - focused on specific types of user generated content, such as user reviews, blog posts, ratings, etc., on firm performance. However, a big part of the user generated content that is present online is in the form of social media communications which has not been studied. My dissertation seeks to bridge this gap by examining the relationship between a specific type of user generated content, i.e. social media activities of a company, and firm performance. I study whether different types of social media metrics, such as positive comment volume, negative comment volume, and proportion of positive or negative comment are important in determining firm performance.

The results of my study suggest that it does not. My dissertation finds that no investor can gain significant abnormal returns based on the publically available social media information. This could indicate that either social media is not an important indicator of financial performance or that social media is important but any information obtained through social media is already priced into the financial markets. Prior studies in marketing that have investigated the impact of various social media metrics on firm abnormal returns have utilized methodologies which are not standard in the finance literature. When utilizing the most popular methodology in finance (where the measure of abnormal returns originate) to study abnormal returns, i.e. the portfolio sort method, I do not find any evidence that firms that have higher values of social media comment volume or higher positive sentiments perform any better than firms that have lower values of those metrics.

To get a clearer understanding between social media and firm performance my second essay studied the relationship between the various social media metrics and firm quarterly cash flow. Again in this study I did not find any evidence that social media comment volume or sentiments had any impact on firm quarterly cash flows. Previous studies in marketing that investigate the relationship between various user generated content and firm performance do not utilize quarterly cash flow as a measure of firm performance. To my knowledge this is the first study to analyze the impact of online social media content on firm quarterly cash flow. I use quarterly cash flow models that are established in the accounting literature to analyze the relationship between social media and quarterly cash flow. This study does not find any evidence that social media content has any impact on firm quarterly cash flows.

The impact of social media on firm performance is not well established. There is doubt amongst marketing managers that social media content is not as effective as other sources when it comes to its impact on their bottom lines. Surveys by various organizations have found that managers are not convinced about the benefits of social media. For example, in the cmosurvey (2016), more than 60% of the managers stated that social media did not contribute or contributed very little to their firms' performance. Only 3.4% of the managers reported that social media contributed very highly to their firm performances. Also in the same survey, around 90% of the firms stated that social media contribution to firm performance is about average or below average. Similarly, Custora (2013) also shows that social media such as twitter and Facebook lag far behind organic search and email in terms of customer acquisition. The survey found that email is almost 40 times more effective than Facebook and Twitter combined in terms of acquiring customers. Also, the survey found that the customers acquired through Facebook was worth only 1.31% more than average while customer acquired through twitter was worth 23% less than average in terms of customer lifetime value.

While the ability of social media to improve firm performance is debatable, companies are spending more on social media. Firms worldwide social network advertisement spending increased from $11.36 billion in 2013 to $23.68 Billion in 2015 (emarketer 2015). In 2016 firms spent a little more than 10% of their marketing budget on social media and this is expected to double within the next 5 years (cmosurvey 2016). This level of spending in social media indicates that social media is not free. Firms want consumers to engage with the product or brand and talk about them in social media, which requires time, effort, and money on part of the firms.

While expense in social media might translate into higher volume of consumer conversations leading to better brand impression and even purchase intent, the benefits might not outweigh the costs. Higher sales does not necessarily mean higher profit if the expense required to generate a sale is greater than the benefit accruing from that sale. This could be one possible reason as to why this study did not find a relationship between social media and quarterly cash flows and stock abnormal returns.

This paper contributes to the current marketing literature in the following ways. First, I use social media information as the focus of my study. Prior literature has mainly focused on other forms of user generated contents such as reviews, ratings, etc. Second, this study incorporates a wide range of firms. While previous studies use data from a single firm or a handful of firms this study utilizes data from over 160 firms. This should provide evidence as to whether the results from the previous literature are generalizable or not. Third, I use established methodology in financing and accounting to study the impact of social media on abnormal returns and quarterly cash flows. While prior literature in marketing used methodologies popular in marketing I utilize well established methodology in the finance literature to analyze the impact of social media on abnormal stock returns, which is a financial measure of firm performance. Similarly, while cash flow has been used as a measure of firm performance by previous marketing researchers I utilize well established methodology in accounting to analyze the impact of social media on quarterly cash flow.

APPENDIX A: TABLES

Table A1. Previous Research on Impact of User Generated Content on Customer Mindset, Market Place Performance, and Financial Performance

Authors (year)	**Findings**
Sonnier et al. (2011)	Online comments volume is positively correlated with firm sales. Positive and negative comments have larger effect than neutral comments. Effect size of positive comments is greater than of negative comments.
Luo et al. (2013)	Buzz rating highly predict firm returns and risks. Blogs and consumer ratings are leading indicator of firm equity value
Liu (2006)	Buzz volume but not buzz valance had a positive impact on revenue. Buzz volume has in informational effect on awareness.
Tellis and Johnson (2007)	Ratings about product quality influences firm stock prices
Chevalier and Mayzlin (2006)	Consumer ratings have a positive impact on firm sales
Morgan and Rego (2006)	Word of mouth positively associated with market share
Tirunillai and Tellis (2012)	Online chatter predicts stock returns and trading volumes. Negative ratings have stronger impact on stock returns than positive ratings
Duan et al. (2008)	Consumer rating do not have an impact on movie revenue but online posting volume has a significant effect
Luo and Zhang (2013)	Buzz and traffic explains a large portion of total variance of firm value

Table A2. Relevant Literature on Online User Generated Content

	Source	**Number of sources for data collection**	**Online Metric**	**Sentiment Analysis**	**Performance Metric**	**Industries**
Chevalier and Mayzlin (2006)	Online UGC	2 – amazon.com and bn.com	Volume and rating	No	Sales	1 - Book
Liu (2006)	Online UGC	1 – Yahoo Movies	**Volume and valance**	Yes	Box office Revenue	1 - Movie
Sonnier et al. (2011)	**Social Media**	**Multiple social media sites**	**Volume and valance**	**Yes**	Sales Revenue	1 - Durable goods firm
Tirunillai and Tellis (2012)	Online UGC	3 – amazon.com, epinions.com, Yahoo Shopping	**Volume and Valance**	**Yes (Reviews)**	**Stock Abnormal Return**	**Multiple Industry** - 15 firms
Luo and Zhang (2013)	Online UGC	1- CNET	Volume and rating	No	Stock Abnormal Return	1- Computer hardware and software - 9 Firms
Luo et al. (2013)	Online UGC	Multiple sources - (CNET, Blogs, etc.)	Volume and rating	**Yes (blogs)**	Stock Abnormal Return	1- Computer hardware and software - 9 Firms
This Study	**Social Media**	**Multiple social media sites**	**Volume and valance**	**Yes**	**Stock Abnormal Return**	**Multiple Industry** - 180 Firms

Table A3. NAICS Classification of Firms

NAICS Sector	NAICS Subsector
Agriculture, Forestry, Fishing and Hunting	Crop Production
Manufacturing	Food manufacturing; Beverage and Tobacco Product Manufacturing; Apparel Manufacturing; Computer and Electronic Product Manufacturing; Paper Manufacturing; Electrical Equipment, Appliance, and Component Manufacturing; Chemical Manufacturing; Transportation Equipment Manufacturing; Petroleum and Coal Products Manufacturing; Leather and Allied Product Manufacturing; Plastics and Rubber Products Manufacturing etc.
Retail Trade	Motor Vehicle and Parts Dealers; Electronics and Appliance Stores; Food and Beverage Stores; Clothing and Clothing Accessories Stores; General Merchandise Stores etc.
Transportation and Warehousing	Air Transportation; Water Transportation; Couriers and Messengers
Information	Publishing Industries (except Internet); Motion Picture and Sound Recording Industries; Broadcasting (except Internet); Telecommunications etc.
Finance and Insurance	Credit Intermediation and Related Activities; Securities, Commodity Contracts, and Other Financial Investments and Related Activities; Insurance Carriers and Related Activities
Real Estate and Rental and Leasing	Real Estate; Rental and Leasing Services
Professional, Scientific, and Technical Services	Professional, Scientific, and Technical Services
Administrative and Support and Waste Management and Remediation Services	Administrative and Support Services
Accommodation and Food Services	Accommodation; Food Services and Drinking Places
Other Services (except Public Administration)	Personal and Laundry Services

Table A4. Summary of Volume of Social Media Comments by Month

YYYYMM	Total comments	Total positive comments	Total neutral comments	Total negative comments	Average proportion of negative comments	Average proportion of positive comments
201206	245367	61825	171142	12400	0.047	0.212
201207	290066	75279	198957	15830	0.050	0.221
201208	316178	80965	218028	17185	0.047	0.218
201209	239921	59191	168445	12285	0.046	0.204
201210	201671	49257	141406	11008	0.052	0.209
201211	119620	24380	89097	6143	0.044	0.190
201212	163752	34812	120310	8630	0.056	0.216
201301	156060	35313	111835	8912	0.061	0.234
201302	136210	30768	98601	6841	0.048	0.212
201303	200406	47526	141838	11042	0.049	0.219
201304	274075	58858	199572	15645	0.051	0.208
201305	222076	46913	163126	12037	0.051	0.204
201306	143898	31851	104895	7152	0.049	0.209
201307	232789	49208	170911	12670	0.048	0.190
201308	241396	49563	179917	11916	0.045	0.193
201309	227453	42737	175191	9525	0.038	0.176
201310	153762	29682	117420	6660	0.039	0.175
201311	298980	62127	223922	12931	0.042	0.189
201312	267341	53148	203075	11118	0.038	0.187
201401	225986	49731	167450	8805	0.040	0.200
201402	218595	49548	160441	8606	0.040	0.223
201403	315602	69731	231672	14199	0.047	0.214
201404	367721	72107	282087	13527	0.035	0.196
201405	706510	103908	584174	18428	0.028	0.156
201406	546383	76209	455697	14477	0.025	0.151
201407	452764	61518	378917	12329	0.025	0.124

Table A5. Summary of Daily Data

	N	Minimum	Maximum	Sum	Mean	Std. Deviation
Total Comment	85050	0	452	4291332	50.46	62.83
Positive Comments	85050	0	271	885042	10.41	15.79
Neutral Comments	85050	0	398	3225115	37.92	50.20
Negative Comments	85050	0	96	181175	2.13	4.14
Percent Positive comments	85050	0	1.00		0.17	0.21
Percent Negative Comments	85050	0	1.00		0.03	0.07

Table A6 Summary Statistics of Social Media Comments by Sectors - Monthly

Sectors		N	Sum	Mean
All Sector – 180 Firms	Total Comment	4680	6915712	1477.72
	Total Positive Comments	4680	1397668	298.65
	Total Negative Comments	4680	296871	63.43
	Proportion of Positive Comments	4680		0.20
	Proportion of Negative Comments	4680		0.04
Information Sector – 21 Firms	Total Comment	546	1156698	2118.49
	Total Positive Comments	546	245219	449.12
	Total Negative Comments	546	60699	111.17
	Proportion of Positive Comments	546		0.21
	Proportion of Negative Comments	546		0.06
All Manufacturing Sector – 83 Firms	Total Comment	2158	2993111	1386.98
	Total Positive Comments	2158	537612	249.13
	Total Negative Comments	2158	102196	47.36
	Proportion of Positive Comments	2158		0.18
	Proportion of Negative Comments	2158		0.03
Computer and Electronics Manufacturing Sector – 19 Firms	Total Comment	494	836792	1693.91
	Total Positive Comments	494	152469	308.64
	Total Negative Comments	494	21562	43.65
	Proportion of Positive Comments	494		0.18
	Proportion of Negative Comments	494		0.03
Manufacturing Sector (except computer and electronic products) – 64 Firms	Total Comment	1664	2156319	1295.86
	Total Positive Comments	1664	385143	231.46
	Total Negative Comments	1664	80634	48.46
	Proportion of Positive Comments	1664		0.18
	Proportion of Negative Comments	1664		0.04
All Sectors (except manufacturing and Information) - 76 Firms	Total Comment	1976	2765903	1399.75
	Total Positive Comments	1976	614837	311.15
	Total Negative Comments	1976	133976	67.80
	Proportion of Positive Comments	1976		0.21
	Proportion of Negative Comments	1976		0.05

Table A7. Summary of Social Media Comments by Sector and Portfolio - Monthly

Sectors	Portfolio	Total Comments		Total Positive Comments		Total Negative Comments		Proportion of Positive Comments	Proportion of Negative Comments
		Sum	Mean	Sum	Mean	Sum	Mean	Mean	Mean
All Sector - ***180 Firms***	1	2768042	2963.6	682525	730.0	150549	161.2	.35	.10
	2	1899098	2026.8	359812	384.8	80415	86.7	.23	.05
	3	1422189	1517.8	235647	251.0	46797	49.3	.19	.04
	4	708940	757.4	103542	110.7	16959	18.0	.14	.02
	5	117443	125.5	16142	17.2	2151	2.3	.07	.01
Information Sector -***21 Firms***	1	396883	3816.2	103640	996.5	26391	256.2	0.33	0.13
	2	284855	2739.0	57512	553.0	14741	137.8	0.24	0.07
	3	261307	2010.1	49192	381.3	11924	93.2	0.20	0.05
	4	150675	1448.8	25360	241.5	5727	55.1	0.17	0.03
	5	62978	605.6	9515	91.5	1916	18.4	0.12	0.02
All manufacturing Sector - ***83 Firms***	1	1215172	2907.1	250971	601.8	49603	117.0	0.30	0.07
	2	863099	1957.1	150647	340.8	29333	67.0	0.22	0.04
	3	588579	1334.6	91620	209.7	16401	38.1	0.17	0.03
	4	277006	626.7	37886	85.5	6060	13.7	0.13	0.02
	5	49255	118.4	6488	15.5	799	1.9	0.07	0.01
Computer and Electronic Manufacturing Sector - ***19 Firms***	1	292990	3756.2	64390	825.5	8513	107.7	.31	.06
	2	237812	2286.6	43746	416.6	6514	62.04	.22	.03
	3	183736	1766.6	27428	263.7	4140	39.81	.18	.03
	4	98841	950.3	13198	128.1	1979	19.03	.14	.02
	5	23413	225.1	3707	35.64	416	4.08	.09	.01
Manufacturing Sector (except computer and electronic products) - ***64 Firms***	1	846395	2712.8	172586	549.6	38084	120.9	.31	.08
	2	638134	1888.0	112488	331.8	24314	71.3	.22	.05
	3	425701	1259.5	68108	202.1	13138	39.3	.17	.03
	4	210652	623.2	27668	81.9	4540	13.5	.13	.02
	5	35437	104.8	4293	12.8	558	1.7	.07	.01
All Sectors (except manufacturing and Information) - ***76 Firms***	1	1046680	2683.8	301688	777.5	64855	167.6	0.41	0.11
	2	778515	1996.2	157776	401.5	39885	100.2	0.24	0.06
	3	641240	1541.4	108490	262.1	21763	53.3	0.20	0.04
	4	265659	681.2	42100	107.1	6816	17.2	0.15	0.02
	5	33809	86.7	4783	12.3	657	1.7	0.07	0.01

Table A8. Firm in Top 10 of Top Portfolio for Each Metric

Appearance count out of 26 months				
Total Comments	**Total Positive Comment**	**Total Negative Comment**	**Proportion of positive comments**	**Proportion of Negative Comments**
AMAZON COM INC (26)	AMAZON COM INC(26)	DIRECTV(26)	BEST BUY COMPANY INC(26)	CLOROX CO(26)
CANON INC(26)	BEST BUY COMPANY INC(26)	CLOROX CO(25)	CAPITAL ONE FINANCIAL CORP(26)	COMCAST CORP NEW(26)
EBAY INC(26)	CAPITAL ONE FINANCIAL CORP(26)	KROGER COMPANY(25)	REVLON INC(25)	DIRECTV(26)
MICROSOFT CORP(26)	STARBUCKS CORP(26)	MONSANTO CO NEW(25)	ALLSTATE CORP(25)	KROGER COMPANY(25)
NIKE INC(26)	CANON INC(25)	VERIZON COMMUNICATIONS INC(25)	KRISPY KREME DOUGHNUTS INC(21)	PEP BOYS MANNY MOE and JACK(25)
NOKIA CORP(26)	NIKE INC(25)	WAL MART STORES INC(25)	NORDSTROM INC(21)	FEDEX CORP(24)
SONY CORP(26)	NORDSTROM INC(25)	COMCAST CORP NEW(24)	AMAZON COM INC(20)	BANK OF AMERICA CORP(23)
PEPSICO INC(25)	DISNEY WALT CO(24)	SAFEWAY INC(24)	PRICELINE COM INC(20)	SAFEWAY INC(23)
STARBUCKS CORP(25)	EBAY INC(24)	FEDEX CORP(23)	AVON PRODUCTS INC(20)	AUTOZONE INC(23)
YAHOO INC(25)	MICROSOFT CORP(24)	BANK OF AMERICA CORP(21)	WEIGHT WATCHERS INTL INC NEW(26)	URBAN OUTFITTERS INC(22)

Table A9. Firm in Bottom 10 of Bottom Portfolio for Each Metric

Appearance count out of 26 months				
Total Comments	**Total Positive Comment**	**Total Negative Comment**	**Proportion of positive comments**	**Proportion of Negative Comments**
T D AMERITRADE HOLDING CORP(26)	T D AMERITRADE HOLDING CORP(26)	T D AMERITRADE HOLDING CORP(26)	GENERAL DYNAMICS CORP(26)	B P PLC(23)
ANN INC(26)	B P PLC(26)	B P PLC(26)	DENNYS CORP(23)	DENNYS CORP(23)
B P PLC(26)	CIGNA CORP(26)	DENNYS CORP(26)	BRISTOL MYERS SQUIBB CO(20)	ROCKWELL AUTOMATION INC(22)
CIGNA CORP(26)	DENNYS CORP(26)	ETHAN ALLEN INTERIORS INC(26)	B P PLC(19)	CAPITAL ONE FINANCIAL CORP(21)
CHIQUITA BRANDS INTL INC(26)	ENERGIZER HOLDINGS INC(26)	I T T CORP(26)	ADOBE SYSTEMS INC(19)	ENERGIZER HOLDINGS INC(21)
DENNYS CORP(26)	ETHAN ALLEN INTERIORS INC(26)	OWENS ILL INC(26)	UNISYS CORP(19)	I T T CORP(20)
ENERGIZER HOLDINGS INC(26)	I T T CORP(26)	SKECHERS U S A INC(26)	LOWES COMPANIES INC(18)	NEWELL RUBBERMAID INC(20)
ETHAN ALLEN INTERIORS INC(26)	OWENS ILL INC(26)	ANN INC(25)	NEWELL RUBBERMAID INC(18)	T D AMERITRADE HOLDING CORP(19)
I T T CORP(26)	SEARS HOLDINGS CORP(26)	CIGNA CORP(25)	BOEING CO(17)	KEYCORP NEW(19)
M G M RESORTS INTERNATIONAL (26)	SKECHERS U S A INC(26)	SEARS HOLDINGS CORP(25)	NOVARTIS A G(17)	ADOBE SYSTEMS INC(19)

Table A10. Monthly - Estimates of Abnormal Returns for All Firms (Contemporaneous)

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Total Comments	0.0007 [0.290]	0.0007 [0.309]	0.0000 [0.009]	0.0014 [0.593]	- 0.0011 [-0.403]
Total Positive Comments	0.0007 [0.303]	-0.0002 [-0.065]	0.0032 [1.319]	-0.0015 [-0.695]	-0.0004 [-0.153]
Total Negative Comments	-0.0035 [-1.526]	0.0022 [0.761]	0.0013 [0.628]	0.0000 [0.018]	0.0018 [0.673]
Proportion of Positive Comments	-0.0013 [-0.492]	-0.0003 [-0.157]	-0.0000 [-0.009]	0.0018 [0.770]	0.0017 [0.681]
Proportion of Negative Comments	-0.0021 [-0.856]	-0.0039 [-1.417]	-0.0004 [-0.1980]	0.0031 [1.354]	**0.0053 [2.081]**
Net Proportion of Positive Comments	-0.0025 [-0.956]	0.0034 [1.301]	-0.0031 [-1.375]	**0.0097 [3.800]**	**-0.0057 [-2.328]**

(t-statistics in parenthesis)

Table A11. Monthly -Estimates of Abnormal Returns for Information Sector (Contemporaneous)

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Total Comments	0.0032 [0.510]	-0.0041 [-0.683]	0.0020 [0.348]	0.0109 [1.302]	-0.0001 [-0.018]
Total Positive Comments	0.0054 [0.794]	-0.0077 [-1.465]	0.0044 [0.652]	0.0090 [1.235]	0.0002 [0.040]
Total Negative Comments	-0.0004 [-0.071]	-0.0010 [-0.149]	0.0038 [0.554]	0.0063 [1.001]	0.0030 [0.448]
Proportion of Positive Comments	0.0047 [0.679]	-0.0034 [-0.441]	-0.0044 [-0.805]	0.0035 [0.559]	0.0131 [1.796]
Proportion of Negative Comments	0.0029 [0.334]	-0.0014 [-0.200]	0.0062 [1.145]	-0.0046 [-0.726]	0.0079 [1.324]
Net Proportion of Positive Comments	-0.0014 [-0.265]	0.0045 [0.486]	0.0074 [-1.334]	0.0074 [1.100]	0.1136 [1.788]

(t-statistics in parenthesis)

Table A12. Monthly -Estimates of Abnormal Returns for All Sectors Except Manufacturing and Information (Contemporaneous)

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Total Comments	-0.0029 [-0.819]	0.0009 [0.206]	-0.0003 [-0.076]	0.0001 [0.043]	-0.0027 [-0.587]
Total Positive Comments	-0.0008 [-0.230]	-0.0055 [-1.207]	0.0066 [1.587]	-0.0034 [-1.004]	-0.0020 [-0.424]
Total Negative Comments	-0.0043 [-1.178]	-0.0043 [-0.846]	0.0038 [1.117]	-0.0004 [-0.110]	0.0001 [0.028]
Proportion of Positive Comments	-0.0037 [-0.859]	-0.0012 [-0.282]	-0.0009 [-0.238]	-0.0002 [-0.055]	0.0011 [0.273]
Proportion of Negative Comments	-0.0072 [-1.833]	-0.0030 [-0.685]	-0.0018 [-0.405]	0.0033 [1.093]	0.0039 [0.875]
Net Proportion of Positive Comments	-0.0001 [-0.027]	-0.0036 [-0.823]	-0.0014 [-0.384]	**0.0109 [2.482]**	**-0.0106 [-2.566]**

(t-statistics in parenthesis)

Table A13. Monthly -Estimates of Abnormal Returns for Manufacturing Sector (Contemporaneous)

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Total Comments	0.0059 [1.456]	0.0007 [0.240]	-0.0008 [0.286]	-0.00148 [0.428]	0.0014 [0.371]
Total Positive Comments	0.0057 [1.338]	0.0026 [0.907]	-0.0036 [-1.212]	0.0003 [0.093]	0.0006 [0.161]
Total Negative Comments	0.0002 [0.065]	**0.0070 [2.076]**	-0.0023 [-0.699]	-0.0026 [-0.812]	0.0032 [0.863]
Proportion of Positive Comments	-0.0014 [-0.392]	0.0056 [1.619]	-0.0028 [-0.951]	0.0037 [1.085]	0.0002 [0.060]
Proportion of Negative Comments	-0.0047 [-1.412]	0.0011 [0.353]	0.0010 [0.302]	0.0005 [0.140]	**0.0075 [2.092]**
Net Proportion of Positive Comments	-0.0027 [0.735]	**0.0073 [2.070]**	-0.0016 [-0.543]	**0.0072 [2.063]**	-0.0052 [-1.511]

(t-statistics in parenthesis)

Table A14. Monthly - Estimates Abnormal Returns for Manufacturing Firms (Except Computer and Electronic Product Manufacturing) (Contemporaneous)

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Total Comments	-0.0020 [-0.541]	-0.0007 [-0.213]	-0.0013 [-0.470]	-0.0017 [-0.503]	-0.0031 [-0.766]
Total Positive Comments	-0.0018 [-0.421]	-0.0006 [-0.213]	-0.0021 [-0.737]	-0.0027 [-0.803]	-0.0015 [-0.411]
Total Negative Comments	-0.0034 [-1.054]	0.0047 [1.240]	-0.0042 [-1.334]	-0.0050 [-1.514]	-0.0011 [-0.290]
Proportion of Positive Comments	-0.0028 [-0.647]	0.0034 [0.906]	**-0.0081** **[-2.763]**	-0.0016 [-0.568]	0.0002 [0.078]
Proportion of Negative Comments	-0.0009 [-0.265]	-0.0063 [-1.689]	0.0019 [0.538]	-0.0040 [-1.299]	0.0004 [0.137]
Net Proportion of Positive Comments	-0.0010 [-0.224]	-0.0030 [-0.813]	-0.0027 [-0.894]	0.0024 [0.838]	-0.0046 [-1.341]

(t-statistics in parenthesis)

Table A15. Monthly - Estimates of Abnormal Returns for Manufacturing Sector (Computer and Electronic Product Manufacturing) (Contemporaneous)

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Total Comments	**0.0314** **[2.101]**	0.0077 [1.152]	0.0033 [0.450]	0.0081 [0.877]	0.0087 [0.951]
Total Positive Comments	0.0146 [1.341]	**0.0275** **[2.891]**	-0.0017 [0.226]	0.0037 [0.478]	0.0108 [1.007]
Total Negative Comments	**0.0300** **[2.138]**	0.0063 [0.832]	-0.0000 [-0.003]	-0.0071 [0.817]	0.0154 [1.636]
Proportion of Positive Comments	-0.0026 [-0.334]	0.0087 [0.972]	0.0087 [1.230]	**0.0223** **[2.084]**	0.0137 [1.284]
Proportion of Negative Comments	-0.0016 [-0.157]	-0.0002 [-0.032]	**0.0300** **[2.855]**	-0.0049 [0.597]	**0.0181** **[2.038]**
Net Proportion of Positive Comments	0.0001 [0.009]	0.0113 [1.314]	0.0153 [1.934]	0.0113 [1.385]	0.0135 [1.066]

(t-statistics in parenthesis)

Table A16. Monthly - Estimates of Abnormal Returns for Manufacturing Sector (Computer and Electronic Product Manufacturing) – Extreme outlier Removed (Nokia September 2013) (Contemporaneous)

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Total Comments	0.022 [1.84]	0.0077 [1.152]	0.0033 [0.450]	0.0081 [0.877]	0.0087 [0.951]
Total Positive Comments	0.0146 [1.341]	**0.0210** **[2.981]**	-0.0017 [0.226]	0.0037 [0.478]	0.0108 [1.007]
Total Negative Comments	0.0215 [1.904]	0.0063 [0.832]	-0.0000 [-0.003]	-0.0071 [0.817]	0.0154 [1.636]
Proportion of Positive Comments	-0.0026 [-0.334]	0.0087 [0.972]	0.0087 [1.230]	0.0157 [1.844]	0.0137 [1.284]
Proportion of Negative Comments	-0.0016 [-0.157]	-0.0002 [-0.032]	**0.0235** **[2.8175]**	-0.0049 [0.597]	**0.0181** **[2.038]**

(t-statistics in parenthesis)

Table A17. Monthly - Estimates of Abnormal Returns for All Firms (Lagged)

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Total Comments	0.0024 [1.027]	0.0035 [1.332]	-0.0018 [-0.694]	0.0023 [0.94]	-0.0000 [-0.01]
Total Positive Comments	0.0022 [0.829]	0.0011 [0.398]	0.0014 [0.599]	0.0029 [1.258]	-0.0012 [-0.448]
Total Negative Comments	-0.0019 [-0.762]	0.0029 [1.039]	0.0032 [1.446]	0.0013 [0.542]	0.0009 [0.328]
Proportion of Positive Comments	-0.0017 [-0.629]	0.0012 [0.417]	0.0001 [0.028]	0.0032 [1.334]	0.0036 [1.49]
Proportion of Negative Comments	-0.0025 [-0.883]	-0.0007 [-0.287]	0.0010 [0.43]	0.0023 [0.922]	**0.0062** **[2.486]**
Net Proportion of Positive Comments	0.0008 [0.294]	-0.0023 [-0.870]	0.0045 [1.742]	0.0012 [0.512]	0.0022 [0.881]

(t-statistics in parenthesis)

Table A18. Monthly - Estimates of Abnormal Returns for Information Sector (Lagged)

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Total Comments	-0.0020 [-0.377]	-0.0010 [-0.189]	0.0111 [1.628]	-0.0070 [-0.948]	0.0042 [0.474]
Total Positive Comments	0.0047 [0.671]	-0.0055 [-0.945]	0.0039 [0.715]	0.0037 [0.382]	0.0002 [0.030]
Total Negative Comments	0.0010 [0.158]	-0.0051 [-0.663]	0.0026 [0.378]	0.0034 [0.546]	0.0054 [0.737]
Proportion of Positive Comments	0.0058 [0.746]	-0.0005 [-0.061]	-0.0055 [-1.109]	0.0083 [1.269]	0.0012 [0.175]
Proportion of Negative Comments	0.0018 [0.227]	-0.0027 [-0.336]	0.0011 [0.177]	-0.0002 [-0.028]	0.0077 [1.255]
Net Proportion of Positive Comments	0.0069 0.929	-0.0002 -0.020	-0.0019 -0.380	0.0004 0.063	0.0032 0.549

(t-statistics in parenthesis)

Table A19. Monthly - Estimates of Abnormal Returns for Manufacturing Sector (Lagged)

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Total Comments	0.0005 [0.131]	0.0001 [0.029]	0.0000 [0.004]	0.0023 [0.743]	0.0078 [1.866]
Total Positive Comments	-0.0003 [-0.067]	0.0043 [1.275]	0.0010 [0.365]	-0.0012 [-0.392]	0.0067 [1.511]
Total Negative Comments	0.0031 [0.832]	-0.0026 [-0.751]	0.0043 [1.285]	-0.0016 [-0.494]	**0.0078** **[2.005]**
Proportion of Positive Comments	0.0013 [0.377]	0.0039 [1.067]	**0.0061** **[1.994]**	-0.0026 [-0.739]	0.0017 [0.445]
Proportion of Negative Comments	**0.0080** **[2.346]**	0.0028 [0.774]	-0.0019 [-0.466]	0.0033 [1.098]	-0.0017 [-0.515]
Net Proportion of Positive Comments	0.0041 [1.057]	-0.0046 [-1.319]	0.0086 [2.404]	0.0022 [0.760]	0.0000 [0.016]

(t-statistics in parenthesis)

Table A20. Monthly - Estimates of Abnormal Returns for Manufacturing Sector (Computer and Electronic Product Manufacturing) (Lagged)

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Total Comments	**0.0302** [**2.029**]	0.0102 [1.485]	0.0039 [0.592]	0.0048 [0.443]	0.0095 [0.992]
Total Positive Comments	0.0283 [1.978]	0.0072 [0.934]	0.0103 [1.609]	0.0073 [0.769]	0.0033 [0.285]
Total Negative Comments	0.0241 [1.758]	0.0062 [0.822]	0.0098 [1.407]	0.0041 [0.420]	0.0145 [1.267]
Proportion of Positive Comments	0.0056 [0.625]	0.0145 [1.579]	0.0044 [0.564]	**0.0264** [**2.373**]	-0.0012 [-0.106]
Proportion of Negative Comments	0.0079 [0.800]	0.0069 [0.878]	0.0164 [1.513]	0.0074 [0.768]	0.0160 [1.581]
Net Proportion of Positive Comments	0.0060 [0.676]	0.0046 [0.540]	**0.0193** [**2.040**]	**0.0227** [**2.306**]	-0.0034 [-0.299]

(t-statistics in parenthesis, number of observation in italics)

Table A21. Monthly - Estimates of Abnormal Returns for Manufacturing Firms (Except Computer and Electronic Product Manufacturing) (Lagged)

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Total Comments	-0.0057 [-1.605]	0.0056 [1.496]	0.0007 [0.198]	-0.0036 [-1.074]	0.0016 [0.401]
Total Positive Comments	-0.0021 [-0.525]	-0.0009 [-0.268]	0.0016 [0.514]	-0.0019 [-0.549]	0.0018 [0.455]
Total Negative Comments	0.0020 [0.638]	-0.0023 [-0.586]	0.0013 [0.382]	-0.0054 [-1.507]	0.0032 [0.850]
Proportion of Positive Comments	0.0074 [1.864]	-0.0041 [-1.069]	-0.0036 [-1.147]	-0.0022 [-0.679]	0.0008 [0.234]
Proportion of Negative Comments	0.0041 [1.189]	-0.0013 [-0.378]	-0.0032 [-0.805]	-0.0039 [-1.057]	0.0027 [0.818]
Net Proportion of Positive Comments	0.0034 [0.803]	-0.0089 [-2.461]	0.0007 [0.224]	0.0010 [0.316]	0.0022 [0.636]

(t-statistics in parenthesis)

Table A22. Daily - Estimates of Abnormal Returns for All sectors (Lagged)

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Total Comments	-0.00004	-0.00004	0.00013	-0.00005	0.00004
	[-0.342]	[-0.326]	[1.141]	[-0.457]	[0.318]
Total Positive Comments	-0.00002	0.00007	0.00003	-0.00013	0.00008
	[-0.140]	[0.583]	[0.298]	[-1.068]	[0.620]
Total Negative Comments	-0.00017	0.00003	0.00013	0.00008	0.00002
	[-1.555]	[0.260]	[0.982]	[0.676]	[0.154]
Proportion of Positive Comments	0.00002	-0.00003	0.00005	-0.00013	0.00018
	[0.179]	[-0.272]	[0.455]	[-1.135]	[1.253]
Proportion of Negative Comments	-0.00010	-0.00001	0.00005	0.00003	0.00007
	[-0.867]	[-0.082]	[0.460]	[0.272]	[0.542]

(t-statistics in parenthesis)

Table A23: Power Test

	c = 0	c = 1	c = 3	c = 5	c = 10	c = 15	c = 20
Portfolio 1	9.31%	9.72%	13.77%	21.49%	52.42%	81.90%	96.26%
Portfolio 2	8.12%	8.73%	12.30%	19.49%	48.65%	81.50%	94.41%
Portfolio 3	8.35%	8.37%	11.84%	18.00%	43.70%	73.40%	90.75%
Portfolio 4	8.82%	8.65%	12.38%	19.79%	50.70%	79.60%	94.79%
Portfolio 5	11.81%	12.07%	14.29%	19.01%	38.69%	65.20%	83.58%

Table A24. Social Media Metrics Breakdown by Industries

Industry	No of Firms (Obs.)	Mean				
		Positive Comments	Negative Comments	Total Comments	% Negative Comments	% Positive Comments
Bev & Tobacco Manufacturing	5 (65)	237	59	2,008	0.031	0.129
Chemical Manufacturing	15 (195)	728	153	4,868	0.034	0.150
Comp & Elec Manufacturing	17 (221)	870	116	5,948	0.023	0.153
Food manufacturing	10 (130)	782	168	5,020	0.039	0.173
Transport Equipment Manufacturing	6 (78)	551	129	5,321	0.028	0.115
Other Manufacturing	26 (338)	751	124	4,965	0.025	0.156
Broadcasting	4 (52)	1,201	356	7,870	0.051	0.157
Publishing	7 (91)	1,416	247	8,038	0.031	0.170
Other Information	7 (91)	1,315	254	7,392	0.039	0.198
Accommodation & Food	3 (39)	863	85	5,972	0.017	0.171
Fast Food	7 (91)	669	153	3,253	0.072	0.219
Finance	20 (260)	972	203	5,565	0.036	0.195
Retail	21 (273)	1,115	256	5,832	0.048	0.188
Transportation	6 (78)	662	168	5,024	0.031	0.133
Others	11 (143)	570	120	4,149	0.032	0.168
Total	165 (2145)	860	171	5,365	0.035	0.168

Table A25. Descriptive Statistics for Quarterly Cash Flow Analysis

	Count	Missing	Mean	Standard Deviation	Percentile 25	Median	Percentile 75
Total Comments	2145	0	5365	4538	1576	4462	7834
Total Positive Comments	2145	0	860	864	191	665	1233
Total Neutral Comments	2145	0	4334	3836	1222	3323	6335
Total Negative Comments	2145	0	171	170	31	128	261
Percent of Positive Comments	2145	0	0.168	0.106	0.10	0.15	0.217
Percent of Negative Comments	2145	0	0.035	0.030	0.015	0.03	0.049
Total Assets	2145	0	112210.44	336159.05	6115.20	19701.00	56265.80
Size	2145	2	3.50	0.66	3.077	3.57	3.997
Cash Flow From Operations (CFO)	2145	1	1347.88	3548.57	74.80	375.24	1400.50
CFO/Total Assets	2145	0	0.026	0.024	0.01	0.04	0.035
Net Sales/Turnover	2145	2	8593.59	15065.32	1195.00	3748.00	9936.222
Accounts Payable	2145	16	36414.50	171143.13	284.02	1193.00	4881.00
Cash	2145	52	4408.47	10274.15	349.00	1200.90	4282.00

Table A25. Continued

	Count	Missing	Mean	Standard Deviation	Percentile 25	Median	Percentile 75
Inventory	2145	38	9191.94	41629.21	164.00	998.93	3607.00
OIBD	2145	119	1550.80	2678.96	165.02	494.84	1616.90
Accounts Receivable	2145	44	32071.98	129511.21	422.90	1539.20	6826.00
Age	2145	247	39	26	17	28	54
Earnings	2145	0	0.015	0.026	0.005	0.014	0.025
Book-to-Market (BTM)	2145	2	0.362	0.435	0.16	0.30	0.516
Leverage	2145	2	0.364	0.533	0.08	0.17	0.425
Dividend	0	2054					
	1	91					
Accrual	2145	1	1354.32	778.075	682.99	1351.45	2025.98
GeoSeg	2145	382	6	6	2	4	8
BusSeg	2145	382	5	5	1	3	8

Table A26. MAPE (Contemporaneous)

	MAPE without Truncation	MAPE without Extreme Outliers (>1000%)	MAPE with Truncation at 100%	Truncation frequency
Without Social Media Metrics	**273.16**	**101.77**	**50.53**	**29.13**
Total Comment	289.92	119.82	53.64	32.28
Positive Comment	266.94	113.20	54.42	29.92
Negative Comment	332.1	125.27	53.45	29.13
Proportion of Positive Comments	268.13	114.92	53.49	33.07
Proportion of Negative Comments	321.71	134.29	55.46	37.00

Table A27 MAPE (Lagged Social Media Metrics)

	MAPE without Truncation	MAPE without Extreme Outliers (>1000%)	MAPE with Truncation at 100%	Truncation frequency
Without Social Media Metrics	**305.71**	**117.55**	**55.01**	**31.74**
Total Comment	388.64	185.46	59.43	36.50
Positive Comment	343.75	144.07	60.53	33.33
Negative Comment	369.41	175.14	59.79	34.92
Proportion of Positive Comments	314.32	153.44	58.55	35.71
Proportion of Negative Comments	352.98	126.93	56.63	35.71

Table A28. MAPE - Industry Breakdown (Contemporaneous)

Industry		MAPE
MAPE (ALL)	Without Social Media Metrics	50.14
	Total Comment	53.64
	Positive Comment	54.44
	Negative Comment	53.45
	Proportion of Positive Comments	53.49
	Proportion of Negative Comments	55.46
Chemical Manufacturing	Without Social Media Metrics	59.14
	Total Comment	64.74
	Positive Comment	58.48
	Negative Comment	58.23
	Proportion of Positive Comments	66.86
	Proportion of Negative Comments	62.22
Comp and electronics manufacturing	Without Social Media Metrics	45.24
	Total Comment	47.96
	Positive Comment	49.47
	Negative Comment	41.19
	Proportion of Positive Comments	35.25
	Proportion of Negative Comments	48.33
Food manufacturing	Without Social Media Metrics	44.09
	Total Comment	35.42
	Positive Comment	42.89
	Negative Comment	53.38
	Proportion of Positive Comments	41.24
	Proportion of Negative Comments	45.16
Retail	Without Social Media Metrics	45.19
	Total Comment	53.91
	Positive Comment	61.36
	Negative Comment	58.56
	Proportion of Positive Comments	58.20
	Proportion of Negative Comments	41.31

Table A29. Correlations

	Earnings	CFO/ Assets	Size	BTM	Leverage	Bus Seg	Geo Seg	Positive Comm-ents	Negati-ve Comm-ents	Total Comm-ents	Percent Neg. Comm-ents	Percent Pos. Comm-ents
Earning	1.000											
CFO/Assets	**0.339**	1										
Size	**0.046**	-0.022	1									
BTM	**-0.222**	**-0.215**	**0.192**	1								
Leverage	**-0.274**	**-0.180**	**0.091**	**0.068**	1							
BusSeg	-0.031	-0.030	**0.202**	**0.181**	0.017	1						
GeoSeg	-0.027	-0.002	**0.204**	**0.123**	-0.023	**0.424**	1					
Positive	0.033	**0.046**	**0.118**	**0.046**	**-0.113**	0.046	-0.030	1.000				
Negative	0.026	0.041	**0.148**	0.036	-0.040	0.005	-0.027	**0.651**	1			
Total Comment	0.036	0.034	**0.144**	0.016	**-0.107**	-0.006	**-0.057**	**0.745**	**0.649**	1		
Percent Neg. Comment	0.041	**0.073**	-0.038	-0.027	**0.075**	-0.025	-0.010	0.029	**0.417**	**-0.132**	1	
Percent Pos. Comment	0.007	0.022	**-0.066**	0.041	0.037	0.032	-0.012	**0.436**	**0.128**	**-0.087**	**0.257**	1

Figures in Bold are significant at 0.05%

Table A30: Regression Test Using Overall Relation Between Earnings and Future Cash Flow – Firm and Quarter Fixed effects

	Estimates							
EARN	**0.611**	**0.633**	**0.389**	**0.389**	**0.395**	**0.388**	**0.391**	**0.390**
SIZE	**0.144**	**0.163**	**0.042**	**0.042**	**0.042**	**0.042**	**0.042**	**0.041**
BTM	-0.008	-0.009	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003
LEVERAGE	-0.001	-0.002	0.003	0.003	0.003	0.003	0.003	0.003
DIV	-0.014	-0.014	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005
GEOSEG	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BUSSEG	0.000	0.000	0.000 5	0.001	0.001	0.001	0.000 5	0.000 5
AGE	0.002	-0.007	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
EARN X SIZE	**-0.129**	**-0.114**	-0.033	-0.033	-0.032	-0.033	-0.033	-0.033
EARN X BTM	-0.027	-0.083	**-0.227**	**-0.227**	**-0.224**	**-0.228**	**-0.229**	**-0.227**
EARN X LEVERAGE	-0.151	-0.163	-0.098	-0.098	-0.106	-0.097	-0.101	-0.098
EARN X DIV	0.264	0.291	0.053	0.051	0.038	0.051	0.047	0.054
EARN X GEOSEG	0.000	-0.010	0.000	0.000	-0.001	0.000	0.000	0.000
EARN X BUSSEG	**-0.020**	**-0.018**	**-0.019**	**-0.019**	**-0.019**	**-0.019**	**-0.019**	**-0.019**
EARN X AGE	0.003	0.004	0.003	0.003	0.002	0.003	0.002	0.002
CFOAit		**-0.292**	**-0.090**	**-0.090**	**-0.089**	**-0.090**	**-0.089**	**-0.089**
CFOAit-3			0.710	0.711	0.709	0.711	0.711	0.711
Total Comments				0.000				
Positive Comments					**-0.000**			
Negative Comments						0.000		
Proportion of Positive Comments							-0.008	
Proportion of Negative Comments								-0.016
N	1,555	1,555	1,555	1,555	1,555	1,555	1,555	1,555
R2	0.617	0.651	0.830	0.830	0.831	0.831	0.831	0.831
Adjusted R2	0.570	0.608	0.809	0.809	0.810	0.809	0.809	0.809

Bold – Significant at 0.05 level

Table A31: Regression Test Using Overall Relation Between Earnings and Future Cash Flow – Industry and Quarter Fixed effects

	Estimates							
EARN	**0.745**	**0.787**	**0.299**	**0.299**	**0.299**	**0.299**	**0.299**	**0.301**
SIZE	0.002	0.002	0.000	0.000	-0.000	-0.000	-0.000	-0.000
BTM	**-0.007**	**-0.008**	0.001	0.001	0.001	0.001	0.001	0.001
LEVERAGE	-0.002	-0.003	0.001	0.001	0.001	0.001	0.001	0.001
DIV	-0.010	**-0.010**	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
GEOSEG	0.000	0.000	0.000	0.000	-0.000	-0.000	-0.000	-0.000
BUSSEG	**0.0005**	**0.0005**	**0.0003**	**0.0003**	**0.0003**	**0.0003**	**0.0003**	**0.0003**
AGE	**-0.000**	**-0.000**	**-0.000**	**-0.000**	**-0.000**	**-0.000**	**-0.000**	**-0.000**
EARN X SIZE	-0.006	0.011	0.009	0.008	0.010	0.008	0.009	0.010
EARN X BTM	**-0.399**	**-0.468**	**-0.219**	**-0.218**	**-0.218**	**-0.219**	**-0.219**	**-0.221**
EARN X LEVERAGE	**-0.349**	**-0.384**	**-0.147**	**-0.145**	**-0.147**	**-0.145**	**-0.146**	**-0.146**
EARN X DIV	-0.077	-0.102	-0.006	-0.007	-0.006	-0.005	-0.007	-0.004
EARN X GEOSEG	-0.004	-0.008	-0.004	-0.003	-0.004	-0.003	-0.003	-0.003
EARN X BUSSEG	**-0.028**	**-0.029**	**-0.015**	**-0.015**	**-0.015**	**-0.015**	**-0.015**	**-0.015**
EARN X AGE	**0.004**	**0.004**	**0.003**	**0.003**	**0.003**	**0.003**	**0.003**	**0.003**
CFOAit		**-0.130**	**-0.042**	**-0.042**	**-0.042**	**-0.042**	**-0.042**	**-0.042**
CFOAit-3			**0.758**	**0.758**	**0.758**	**0.758**	**0.758**	**0.758**
Total Comments				0.000				
Positive Comments					0.000			
Negative Comments						0.000		
Proportion of Positive Comments							-0.001	
Proportion of Negative Comments								-0.008
N	1,555	1,555	1,555	1,555	1,555	1,555	1,555	1,555
R2	0.542	0.55	0.817	0.818	0.817	0.818	0.817	0.818
Adjusted R2	0.532	0.54	0.814	0.813	0.813	0.813	0.813	0.813

Bold – Significant at 0.05 level

Table A32: Total Comments. Dependent Variable is $CFOA_t$

	OLS		Fixed		Fixed with time FE		Random Effects		First Difference	
	Coef.	St. Error	Coef.	St. Error	Coef.	St. Error	Coef.	St. Error	Coef.	St. Error
(Intercept)	-0.003**	-0.002					-0.003**	0.002		
$CFOA_{t-1}$	-0.062***	-0.017	-0.061***	-0.018	-0.054***	-0.019	-0.062***	0.017	-0.169***	-0.017
$CFOA_{t-4}$	0.755***	-0.017	0.712***	-0.019	0.695***	-0.02	0.754***	0.017	0.682***	-0.018
$OIBDA_{t-1}$	0.284***	-0.032	0.204***	-0.044	0.213***	-0.044	0.283***	0.032	0.079*	-0.047
$OIBDA_{t-4}$	-0.035	-0.028	-0.011	-0.041	-0.009	-0.041	-0.0350	0.028	0.011	-0.044
$RECA_{t-1}$	0.013**	-0.006	0.204***	-0.029	0.207***	-0.029	0.013**	0.006	0.287***	-0.039
$INVA_{t-1}$	0.010*	-0.005	0.102***	-0.038	0.092**	-0.039	0.010*	0.005	0.057	-0.054
$PAYA_{t-1}$	-0.007	-0.005	-0.133***	-0.039	-0.136***	-0.039	-0.007	0.005	-0.076*	-0.04
Total $Comment_{t-1}$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Observations	1,476		1,476		1,476		1,476		1,353	
Adjusted R2	0.697		0.8113		0.8134		0.792		0.723	

*p<0.1; **p<0.05; ***p<0.01

Table A33: Positive Comments. Dependent Variable is $CFOA_t$

	OLS		Fixed		Fixed with time FE		Random Effects		First Difference	
	Coef.	St. Error	Coef.	St. Error	Coef.	St. Error	Coef.	St. Error	Coef.	St. Error
(Intercept)	-0.003**	-0.001					-0.003**	0.001		
$CFOA_{t-1}$	-0.063***	-0.017	-0.061***	-0.018	-0.054***	-0.019	-0.062***	0.017	-0.169***	-0.017
$CFOA_{t-4}$	0.755***	-0.017	0.712***	-0.019	0.695***	-0.02	0.754***	0.017	0.683***	-0.018
$OIBDA_{t-1}$	0.282***	-0.032	0.205***	-0.044	0.213***	-0.044	0.281***	0.032	0.078*	-0.047
$OIBDA_{t-4}$	-0.035	-0.028	-0.009	-0.041	-0.007	-0.041	-0.035	0.028	0.011	-0.043
$RECA_{t-1}$	0.014**	-0.006	0.203***	-0.029	0.207***	-0.029	0.013**	0.006	0.287***	-0.039
$INVA_{t-1}$	0.010*	-0.005	0.102***	-0.038	0.092**	-0.039	0.010*	0.005	0.056	-0.054
$PAYA_{t-1}$	-0.007	-0.005	-0.131***	-0.039	-0.135***	-0.039	-0.00732	0.005	-0.075*	-0.04
Positive $Comment_{t-1}$	0.00000*	0.000	0.000	0.000	0.000	0.000	0.000*	0.000	0.000	0.000
Observations	1,476		1,476		1,476		1,476		1,353	
Adjusted R2	0.697		0.8112		0.8131		0.797		0.723	

*p<0.1; **p<0.05; ***p<0.01

Table A34: Negative Comments. Dependent Variable is $CFOA_t$

	OLS		Fixed		Fixed with time FE		Random Effects		First Difference	
	Coef.	St. Error	Coef.	St. Error	Coef.	St. Error	Coef.	St Error	Coef.	St. Error
(Intercept)	-0.003**	-0.001					-0.003**	0.001		
$CFOA_{t-1}$	-0.062***	-0.017	-0.061***	-0.018	-0.054***	-0.019	-0.061***	0.017	-0.170***	-0.017
$CFOA_{t-4}$	0.755***	-0.017	0.712***	-0.019	0.695***	-0.02	0.755***	0.017	0.683***	-0.018
$OIBDA_{t-1}$	0.281***	-0.032	0.205***	-0.044	0.213***	-0.044	0.281***	0.032	0.079*	-0.047
$OIBDA_{t-4}$	-0.035	-0.028	-0.009	-0.041	-0.007	-0.041	-0.034	0.028	0.011	-0.044
$RECA_{t-1}$	0.014**	-0.006	0.203***	-0.029	0.207***	-0.029	0.013**	0.006	0.287***	-0.039
$INVA_{t-1}$	0.010*	-0.005	0.102***	-0.038	0.093**	-0.039	0.010*	0.005	0.055	-0.054
$PAYA_{t-1}$	-0.008*	-0.005	-0.132***	-0.039	-0.135***	-0.039	-0.007*	0.005	-0.076*	-0.04
Negative $Comment_{t-1}$	0.000	0.000	0.000	-0.000	0.000	-0.000	0.000	0.000	0.000	-0.000
Observations	1,476		1,476		1,476		1,476		1,353	
Adjusted R2	0.697		0.8112		0.8131		0.802		0.723	

*p<0.1; **p<0.05; ***p<0.01

Table A35: Proportion of Positive Comments. Dependent variable is $CFOA_t$

	OLS		Fixed		Fixed with time FE		Random Effects		First Difference	
	Coef.	St. Error	Coef.	St. Error	Coef.	St. Error	Coef.	St. Error	Coef.	St. Error
(Intercept)	-0.003*	-0.002					-0.003*	0.002		
$CFOA_{t-1}$	-0.061***	-0.017	-0.060***	-0.018	-0.054***	-0.019	-0.060***	0.017	-0.170***	-0.017
$CFOA_{t-4}$	0.756***	-0.017	0.712***	-0.019	0.695***	-0.02	0.755***	0.017	0.682***	-0.018
$OIBDA_{t-1}$	0.281***	-0.032	0.204***	-0.044	0.213***	-0.044	0.281***	0.032	0.078*	-0.047
$OIBDA_{t-4}$	-0.035	-0.028	-0.009	-0.041	-0.007	-0.041	-0.034	0.028	0.014	-0.044
$RECA_{t-1}$	0.014**	-0.006	0.205***	-0.029	0.207***	-0.029	0.014**	0.006	0.287***	-0.039
$INVA_{t-1}$	0.011**	-0.005	0.102***	-0.038	0.092**	-0.039	0.010**	0.005	0.052	-0.054
$PAYA_{t-1}$	-0.008	-0.005	-0.132***	-0.039	-0.135***	-0.039	-0.00762	0.005	-0.076*	-0.04
Proportion of Positive $Comment_{t-1}$	0.004	-0.006	-0.004	-0.007	-0.003	-0.008	0.003	0.006	-0.007	-0.009
Observations	1,476		1,476		1,476		1,476		1,353	
Adjusted R2	0.696		0.8113		0.8131		0.81		0.723	

*p<0.1; **p<0.05; ***p<0.01

Table A36: Proportion of Negative Comments. Dependent variable is $CFOA_t$

	OLS		Fixed		Fixed with time FE		Random Effects		First Difference	
	Coef.	St. Error	Coef.	St. Error	Coef.	St. Error	Coef.	St. Error	Coef.	St. Error
(Intercept)	-0.002*	-0.001					-0.002*	0.001		
$CFOA_{t-1}$	-0.061***	-0.017	-0.060***	-0.018	-0.054***	-0.019	-0.060***	0.017	-0.170***	-0.017
$CFOA_{t-4}$	0.756***	-0.018	0.712***	-0.019	0.695***	-0.02	0.755***	0.018	0.683***	-0.018
$OIBDA_{t-1}$	0.282***	-0.032	0.203***	-0.044	0.212***	-0.044	0.282***	0.032	0.076	-0.046
$OIBDA_{t-4}$	-0.034	-0.029	-0.011	-0.041	-0.007	-0.041	-0.03371	0.029	0.013	-0.043
$RECA_{t-1}$	0.014**	-0.006	0.205***	-0.029	0.207***	-0.029	0.013**	0.006	0.287***	-0.039
$INVA_{t-1}$	0.011**	-0.005	0.104***	-0.038	0.093**	-0.039	0.010**	0.005	0.05	-0.054
$PAYA_{t-1}$	-0.008	-0.005	-0.133***	-0.039	-0.135***	-0.039	-0.00752	0.005	-0.078*	-0.04
Proportion of Negative $Comment_{t-1}$	-0.005	-0.017	-0.02	-0.021	-0.022	-0.023	-0.00478	0.017	-0.045*	-0.024
Observations	1,476		1,476		1,476		1,476		1,353	
Adjusted R2	0.696		0.8114		0.8131		0.81		0.723	

*p<0.1; **p<0.05; ***p<0.01

APPENDIX B: FIGURES

Figure B1. Brand Value Chain (Keller and Lehmann (2003)

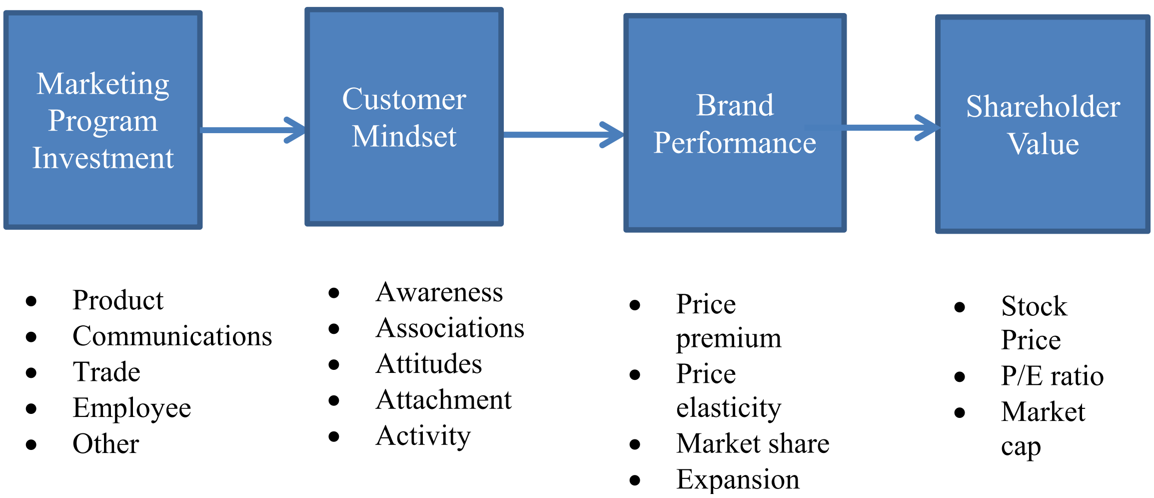


Figure B2. Brand Value Chain with Social Media

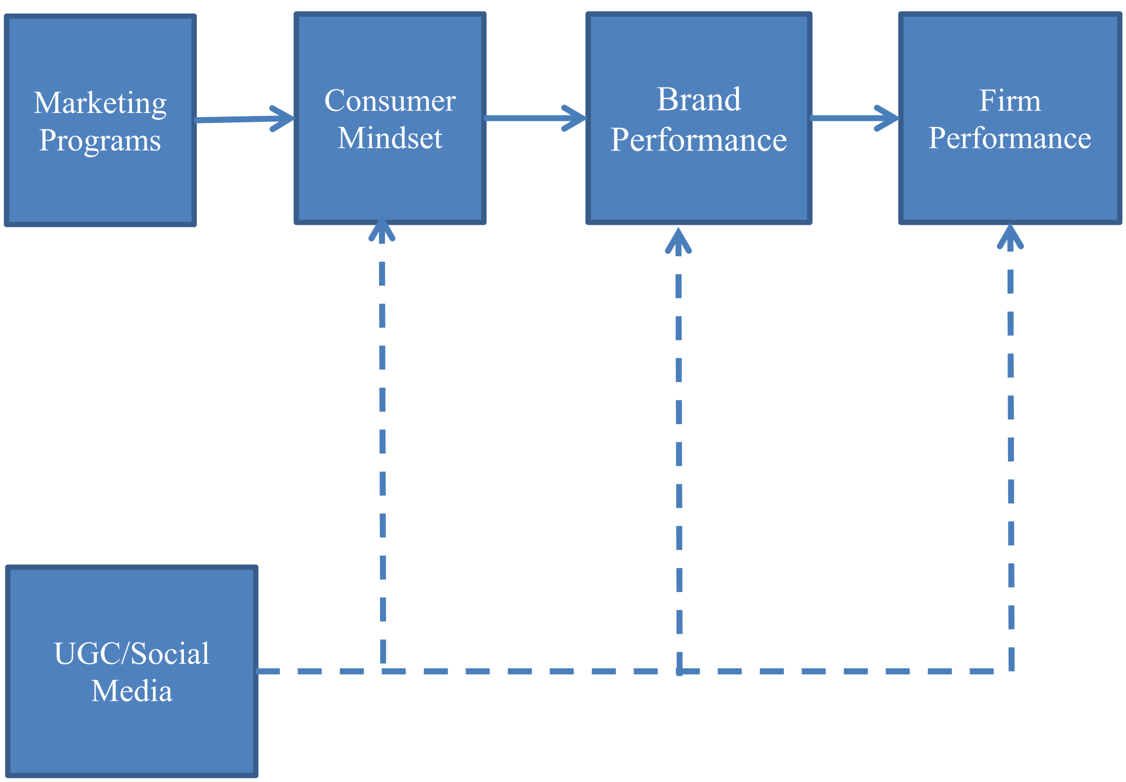


Figure B3. Time Series Plot of Selected Social Media Metrics

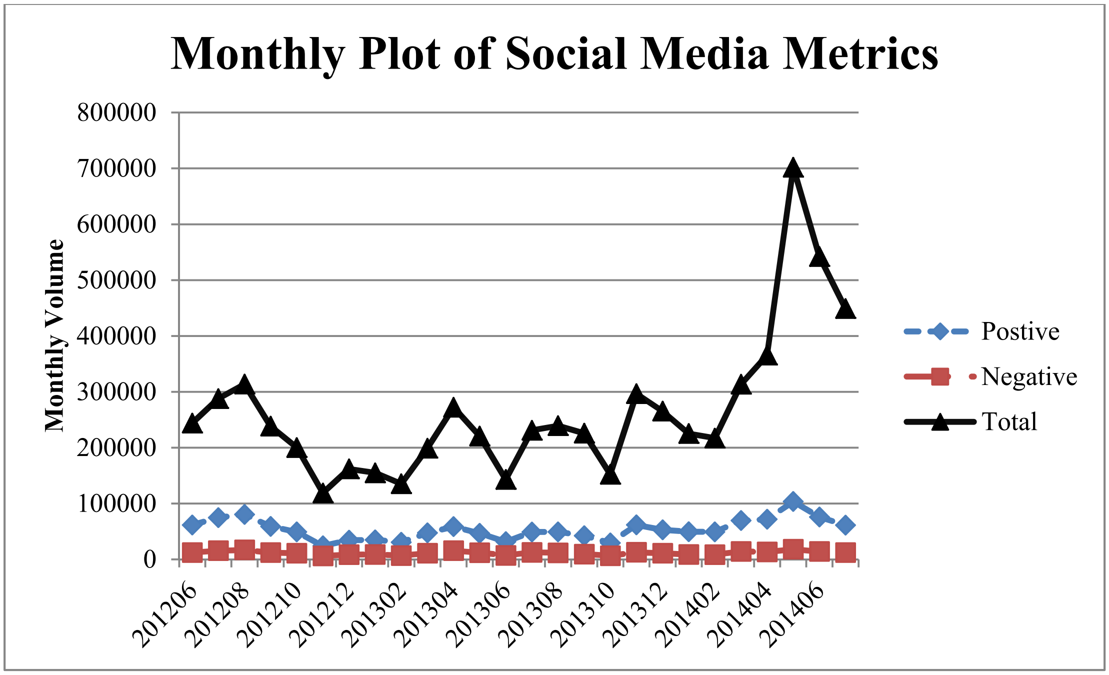


Figure B4. Abnormal Returns for top and Bottom Portfolio by Total Comments

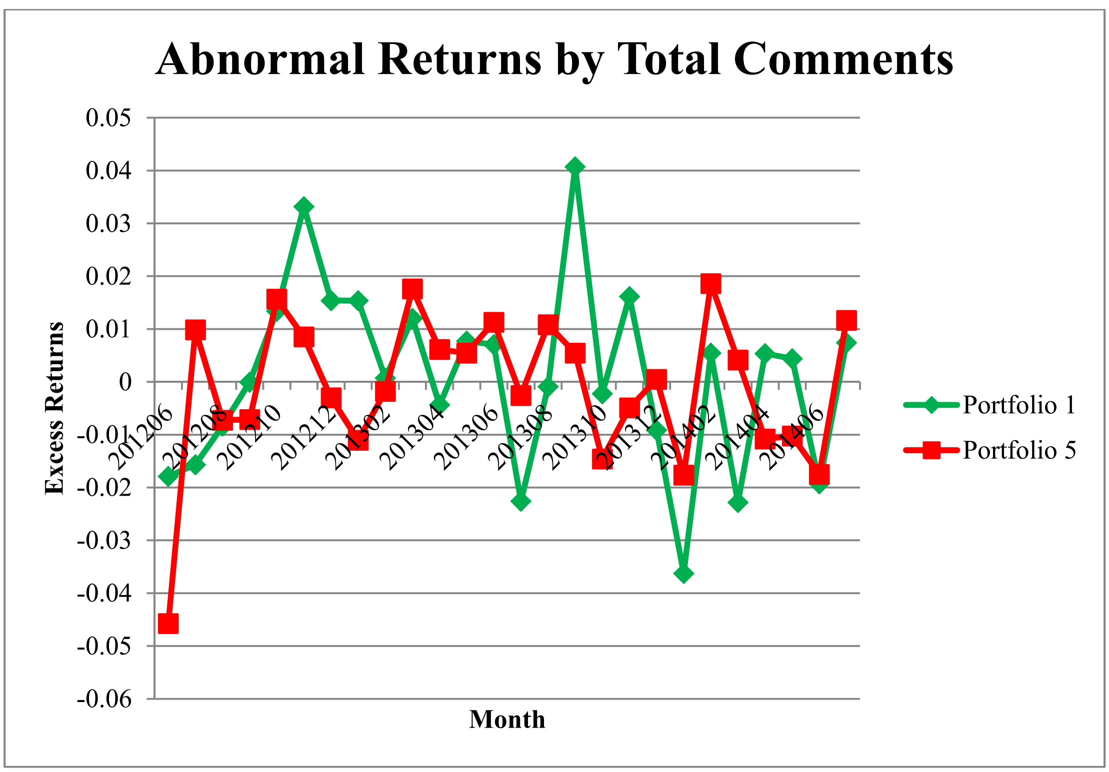


Figure B5. Abnormal Returns for top and Bottom Portfolio by Total Positive Comments

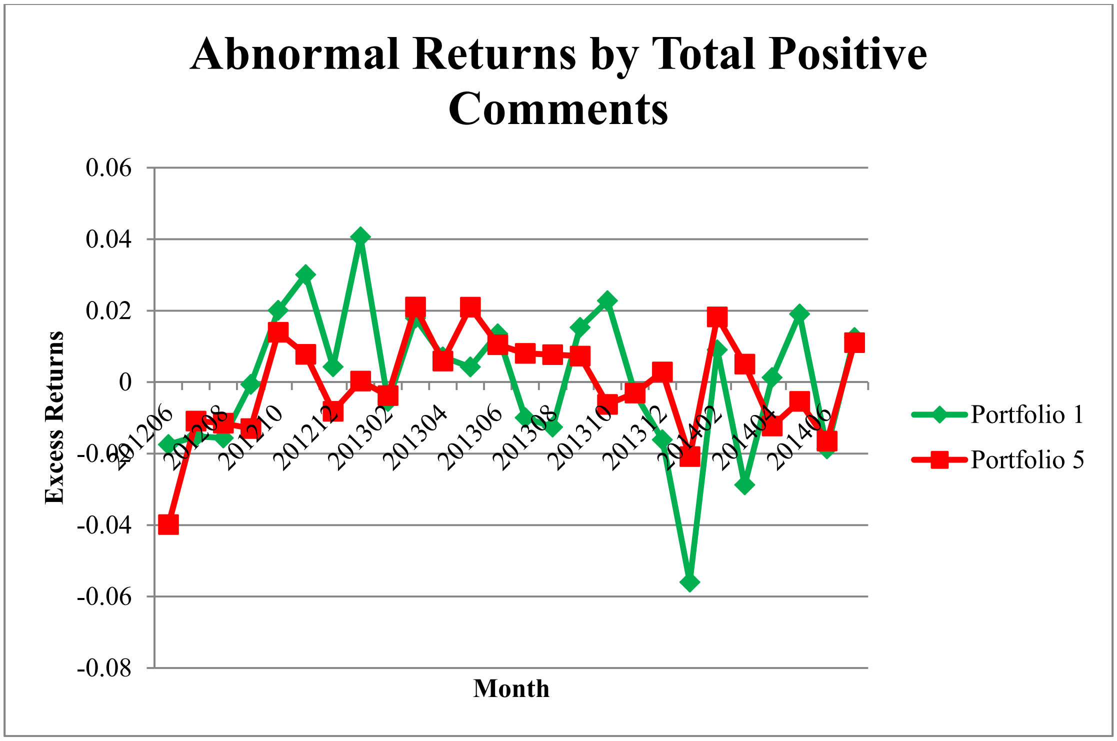


Figure B6. Abnormal Returns for top and Bottom Portfolio by Total Negative Comments

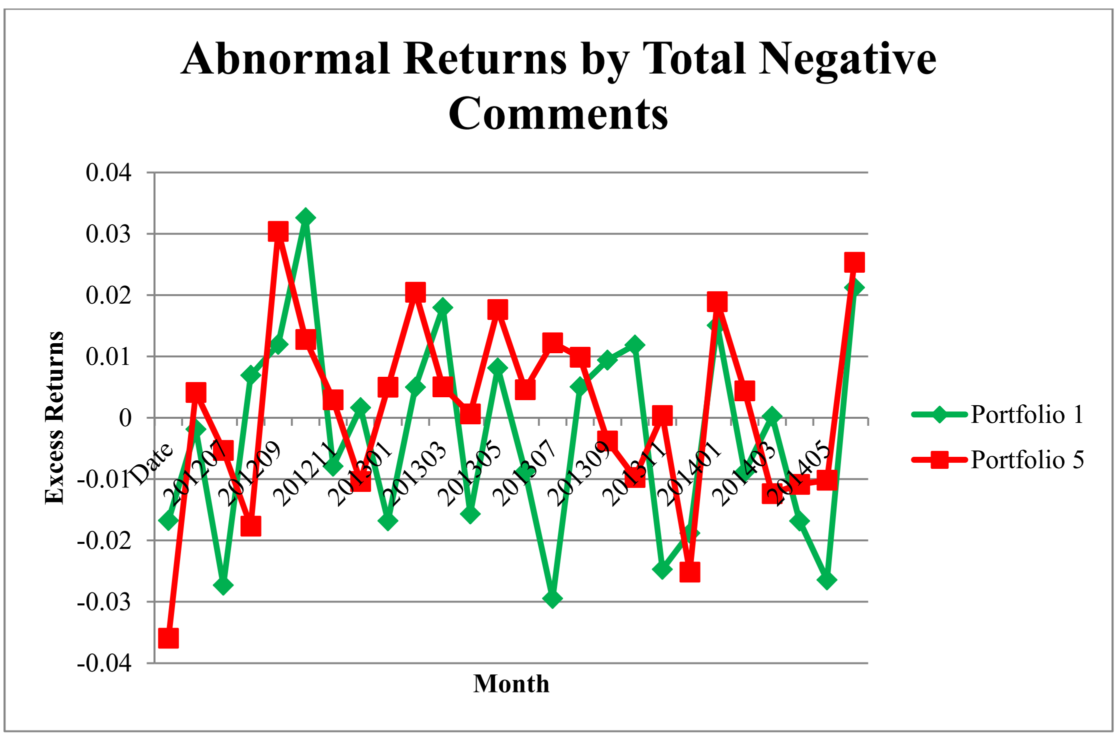


Figure B7. Abnormal Returns for top and Bottom Portfolio by Proportion of Positive Comments

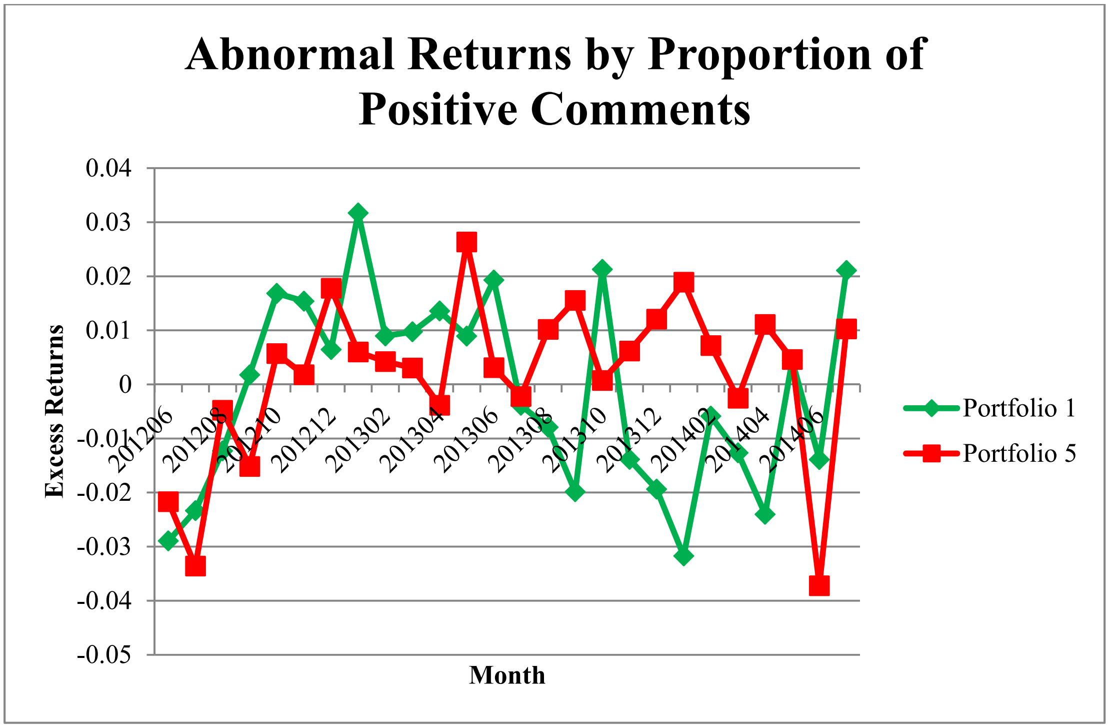


Figure B8. Abnormal Returns for top and Bottom Portfolio by Proportion of Negative Comments

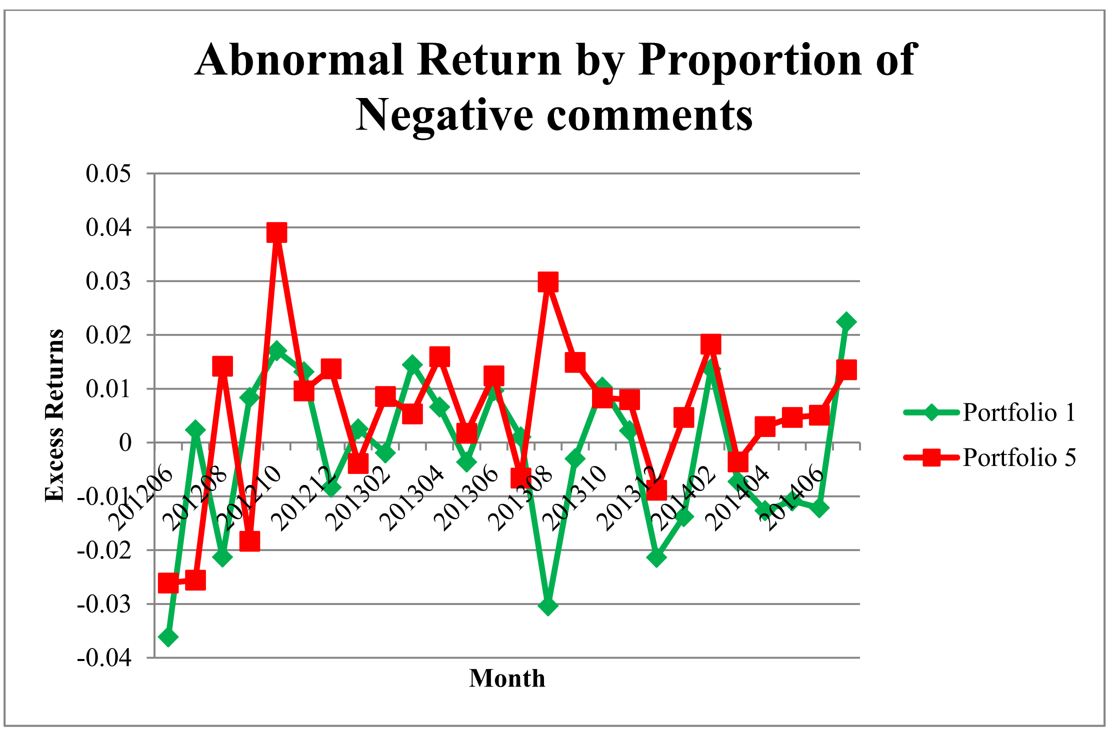
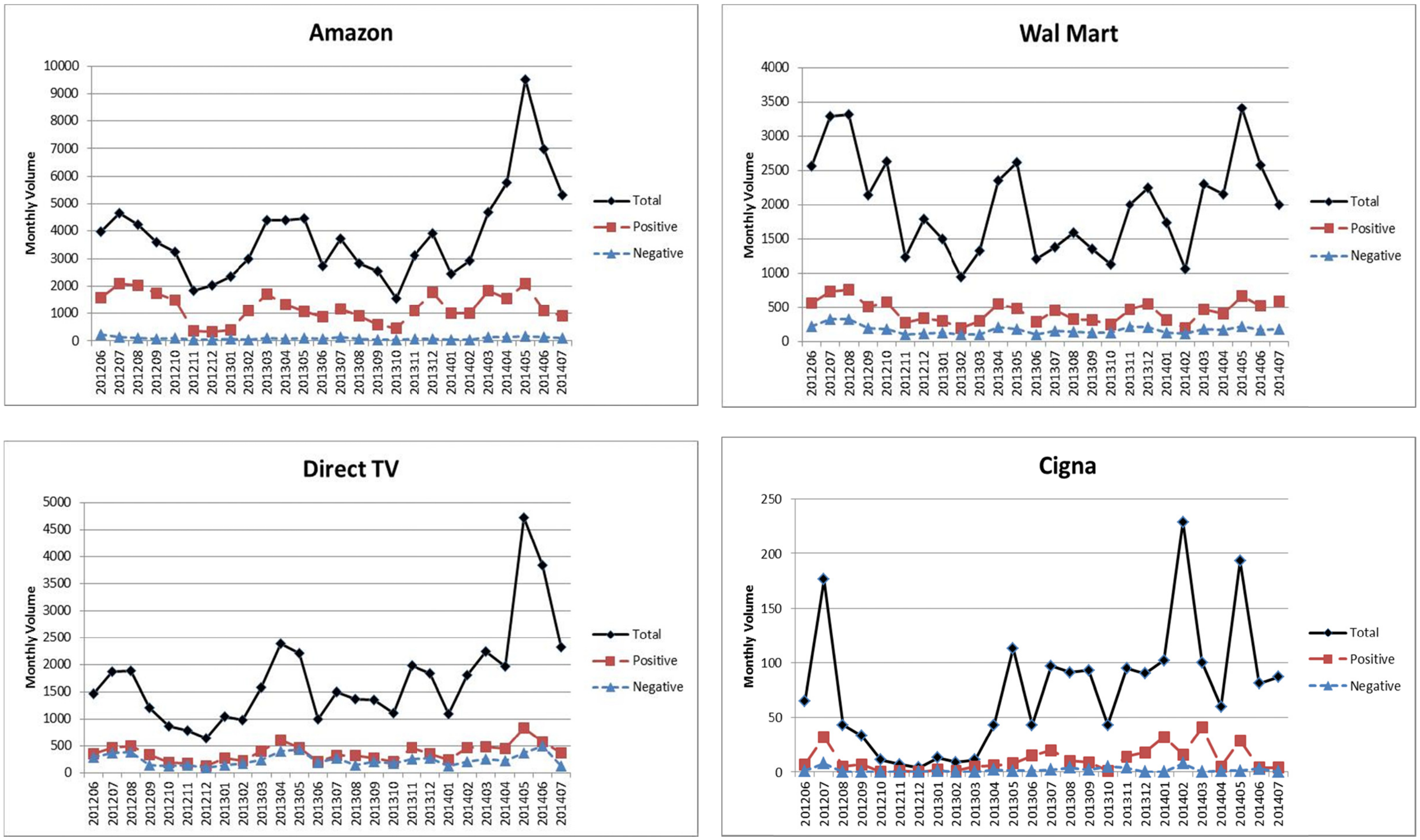


Figure B9. Time Series Plot of Select Social Media Metrics of Select Firms



REFERENCES

Aaker, D. A., and Jacobson, R., (1994) "The Financial Information Content of Perceived Quality." *Journal of Marketing Research,* 31 (Spring): 191-201.

Aaker, D. A., and Jacobson, R. (2001), "The Value Relevance of Brand Attitude in High-Technology Markets," *Journal of Marketing Research*, 38(4), 485-93

Aggarwal. R., Gopal, R., Sankaranarayanan, R., Singh. R. V., (2012) "Blog, blogger, and the firm: Can negative employee posts lead to positive outcomes?" *Information Systems Research* 23(2):306–322

Aksoy, L., Cooil, B., Groening, C., Keiningham, T. L.,Yalcın, A., (2008) "The long-term stock market valuation of customer satisfaction." Journal of Marketing, 72(4) 105–122.

Atwood, T.J., M.S. Drake, L.A. Myers. (2010). "Book-tax conformity, earnings persistence and the association between earnings and future cash flows." *Journal of Accounting and Economics* 50: 111-125.

Ambler, Tim and John Roberts (2006), "Beware the Silver Metric: Marketing Performance Measurement Has to Be Multidimensional," Report No. 06-113, *Marketing Science Institute*.

Anderson, E. W., (1998), "Customer Satisfaction and Word-of-Mouth," *Journal of Service Research, 1 (1), 1-14*

Anderson, E. W., Fornell, C., and Mazvancheryl, S. K., (2004) "Customer Satisfaction and Shareholder Value," *Journal of Marketing,* Vol. 68, 4, pp. 172-185

Arellano, M. (1987) "Computing robust standard errors for within group estimators", *Oxford Bulletin of Economics and Statistics*, **49**, pp. 431--434.

Arellano, M., Bond, S. R., (1991), "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *The Review of Economic Studies* 58 (2) 277–297.

Arndt, J. (1967). "Word-of-mouth advertising: a review of the literature." *New York: Advertising Research Foundation Inc.*

Atwood, T.J., M.S. Drake, L.A. Myers. 2010. "Book-tax conformity, earnings persistence and the association between earnings and future cash flows." *Journal of Accounting and Economics* 50: 111-125.

Bansal, H. S., Voyer, P. A. (2000) "Word-of-Mouth Processes Within a Service Purchase Decision Context," *Journal of Service Research,* Vol. 3, No. 2, 166-177

Barth, M. E., Clement, M. B., Foster, G., and Kaszkik. R., (1998) "Brand Values and Capital Market Valuation," *Review of Accounting Studies* 3:41-68.

Bathke, A. W., Jr., Lorek. K. S., (1984), “The relationship between time-series models and the security market’s expectation of quarterly earnings,” *The Accounting Review* 59: 163–176.

Baum, C. F., Lecture on Dynamic Panel Data Estimators, Personal Collection of Christopher F Baum, Boston College, Boston, MA.

Bell, D. R., Ledoit, O., Wolf, M., (2013) “A New Portfolio Formation Approach to Mispricing of Marketing Performance Indicators with an Application to Customer Satisfaction,” *Working Paper, University of Zurich*

Bernard, V., and Stober, T. L., (1989). “The timing, amount, and nature of information reflected in cash flows and accruals”. *The Accounting Review* 64 (October): 624-652.

Bickart, B., Schindler, R. M., (2001) “Internet forums as influential sources of consumer information,” *Journal of Interactive Marketing* 15 (3) 31–40

Blackshaw, P., Nazzaro, M., (2004) “Consumer Generated Media (CGM) 101” retrieved Feb 18 from *http://www.brandchannel.com/images/papers/222_cgm.pdf*

Brandwatch (2013), “Marketing: Top 15 Free Social Media Monitoring Tools”, retrieved Feb 20 from *https://www.brandwatch.com/2013/08/top-10-free-social-media-monitoring-tools/*

Brown, J. J., Reingen. P., (1987). “Social ties and word-of-mouth referral behavior.” *Journal of Consumer Research*, 14, 350–362.

Brown, L. D., & Rozeff, M. S. (1979), “Univariate time-series models of quarterly accounting earnings per-share: A proposal,” *Journal of Accounting Research*,17, 179–189

Bruhn, M., Schoenmueller, V., Schäfer, D.B., (2012) "Are social media replacing traditional media in terms of brand equity creation?", *Management Research Review*, Vol. 35 Iss: 9, pp.770 - 790

Brunswick Survey (2012). “2012 Brunswick investor use of digital and social media survey”, retrieved Feb 18th, 2015 from http://www.brunswickgroup.com/publications/surveys/2012-brunswick-investor-use-of-digital-and-social-media-survey/

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

Chevalier, J.D., and Mayzlin, D., (2006), " The effect of word of mouth on sales: Online book reviews," *Journal of Marketing Research*, 43(3), 345–354

Chen Y, Xie J (2008) "Online consumer reviews: A new element of marketing communications mix," *Management Science*, 54(3):477–491.

cmosurvery (2014), "Tweet this: Social media important to company performance but difficult to prove", retrieved Feb 18th , 2015 from http://www.cmosurvey.org/blog/tweet-this-social-media-important-to-company-performance-but-difficult-to-prove/

Conchar, M. P., Crask, M. R., and Zinkhan, G. M., (2005) "Market Valuation Models of the Effect of Advertising and Promotional Spending: A Review and Meta-Analysis?' *Journal of the Academy of Marketing Science* 33 (4): 445-460.

Custora., (2013) "E-Commerce Customer Acquisition Snapshot", New York, NY

Da, Z., Engelberg, J., Gao, P., (2011) "In search of attention," *Journal of Finance,* 66(5):1461–1499.

Das, S., and Chen, M., (2007). "Yahoo! for Amazon: Sentiment extraction from small talk on the web." *Management Science* , 53, 1375.

Day, G. S., and Fahey, L., (1988) "Valuing market strategies," *Journal of Marketing*, 52(3), 45–57.

Dellarocas, C., Wood, C., (2008) "The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias," *Management Science*, 54(3):460–476.

Dhar, V., Chang, E. A., (2009) "Does Chatter Matter? The Impact of User-Generated Content on Music Sales," *Journal of Interactive marketing*, 23, 300 - 307

Divol R, Edelman D, Sarrazin H (2012) "Demystifying social media". *McKinsey Quart*erly. (April).
http://www.mckinsey.com/insights/marketing_sales/demystifying_social_media

Duan, W., Gu, B., and Whinston, A.B., (2008) "Do online reviews matter? An empirical investigation of panel data'" *Decision Support Systems,* 45 1007 - 1016

Duan, W., Gu, B., and Whinston, A.B., (2009) "Informational cascades and software adoption on the Internet: An empirical investigation," *MIS Quarterly*, 33, 1, 23–48.

East, R., Hammond, K., Lomax, W., and Robinson, H. (2005), "What Is the Effect of a Recommendation?" *Marketing Review*, 5 (2), 145–57.

eMarketer (2014a), "Mobile continues to Steal Share of US Adults' Daily Time Spent with Media", retrieved Feb 18, 2015 from http://www.emarketer.com/Article/Mobile-Continues-Steal-Share-of-US-Adults-Daily-Time-Spent-with-Media/1010782

eMarketer (2014b), "Emarketer Roundup:Global Media Trends", retrieved Feb 18, 2015 from http://www.festivalofmedia.com/media/10971/eMarketer_FOM_Global_ Roundup.pdf

emarketer (2015), "Social Network Ad Spending to hit $23.68 Billion Worldwide in 2015", retrieved June 20, 2016 from http://www.emarketer.com/Article/Social-Network-Ad-Spending-Hit-2368-Billion-Worldwide-2015/1012357

Fama, E., French, K., (1996) "Multifactor explanations of asset pricing anomalies," *Journal of Finance,* 51(1):55–84.

Fama, E., French, F., Kenneth, R., (1993) "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics,* 33(1):3–56.

Forman, C., Ghose, A., Wiesenfeld, B., (2008) "Examining the Relationship Between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets," *Information Systems Research,* Vol. 19, 3, pp. 291-313

Fornell, C., (1992), "A National Customer Satisfaction Barometer: The Swedish Experience," *Journal of Marketing ,* 56, 1 – 21

Fornell, C., Mithas, S., Morgeson III, F. Vl., Krishnan, M. S., (2006), "Customer Satisfaction and Stock Prices: High Returns, Low Risk," *Journal of Marketing,* Vol. 70, 1, pp. 3 -14

Godes, D., Mayzlin, D., (2004) "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science* 23(4):545-560.

Godes, D., Mayzlin, D., Chen, Y., Das, S., Dellarocas, C., Pfeiffer, B., Libai, B., Sen, S., Shi, M., Verlegh. P., (2005), "The firm's management of social interactions," *Marketing Letters,* 16(3/4) 415–428.

Godes, D., and Mayzlin, D., (2009) "Firm-created word-of-mouth communication: Evidence from a field test," *Marketing Science, 28,* 4, 721–739.

Goh, K., Heng, C., Lin, Z., (2014) "Social Media Brand Community and Consumer Behavior: Quantifying the Relative Impact of User- and Marketer- Generated Content," *Information Systems Research,* 24(1):88-107

Goyal, A., Santa-Clara, P., (2003) "Idiosyncratic risk matters!," *Journal of Finance,* 58(3):975–1007.

Griffin, P. A. (1977). “The time-series behavior of quarterly earnings: Preliminary evidence.” *Journal of Accounting Research* 15 (Spring): 71-83.

Gruca, T. S., Rego, L. L., (2005) “Customer Satisfaction, Cash Flow, and Shareholder Value,” *Journal of Marketing,* Vol. 69, 115-130

Gu B, Park J, Konana P (2012) “The impact of external word-of mouth sources on Retailer sales of high-involvement products” *Information Systems Research.* 23(1):182–196.

Hanson, W., Kalyanam, K., (2007) “Internet Marketing and e-Commerce” *(Thomson South-Western, Mason, OH)*

Harvard Business Review (2010) “The new Conversation: Taking Social Media from Talk to Action,” *Harvard Business School Publishing*

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., Gremler, D. D., (2004) “Electronic Word-of-Mouth Via Consumer-Opinion Platforms: What Motivates Consumers to Articulate Themselves on the Internet,” *Journal of Interactive marketing,* 18, 1, pg. 38 – 52

Hribar, P., Folsom, D., Mergenthlaer, R, D., Peterson, K., (Forthcoming) “Principles-Based Standards and Earnings Attributes,” *Management Science*

Ho-Dac, N. N., Carson, S. J., Moore, W. L., (2013), “The Effects of Positive and Negative Online Customer Reviews: Do Brand Strength and Category Maturity Matter?” *Journal of Marketing,* Vol. 77, 37 - 53

Huffingtonpost.com, (2009) “United Breaks Guitars’: Did It Really Cost The Airline $180 Million?,” (August 24). Retrieved from http://www.huffingtonpost.com/2009/07/24/unitedbreaks-guitars-did_n_244357.html (last accessed: February 9, 2015)

Hunt, S. D., and Morgan, M. R. (1995). "The Comparative Advantage Theory of Competition," *Journal of Marketing,* 59 (April), 1-15.

Hutter, K., Hautz, J., Dennhardt, S., Fueller, J., (2013) “The impact of user interactions in social media on brand awareness and purchase intention: the case of MINI on Facebook,” *Journal of Product & Brand Management,* 342-351

Jacobson, R., Mizik, N., (2009) “The Financial Markets and Customer Satisfaction: Reexamining Possible Financial Market Mispricing of Customer Satisfaction,” *Marketing Science* 28(5) 810-819.

Keller, K. L., (1993) “Conceptualizing, Measuring, and Managing Customer-Based Brand Equity,” *Journal of Marketing,* 57, 1, pg. 1

Keller, K. L., and Donald R. Lehmann (2003), “How Do Brand Create Value,” *Marketing Management*, 2003 (May), (26-31).

Kerin, R. A., Sethuraman, R., (1998), "Exploring the Brand Value- Shareholder Value Nexus for Consumer Goods Companies?,” *Journal of the Academy of Marketing Science* 26 (Fall): 260-273

Kiecker, P., Cowles, D., (2002), “Interpersonal Communication and Personal Influence on the Internet: A Framework for Examining Online Word-of-Mouth,” *Journal of Euromarketing*, 11, 2, 71–88.

Kim, P., (2006) “The Forrester Wave: Brand monitoring,” Q3 *Forrester Wave (white paper),* 2006

Kim, N., Mahajan, V., and Srivastava, R. K., (1995), “Determining the Going Value of a Business in an Emerging Information Technology Industry: The Case for Cellular Communications Industry,” *Technological Forecasting and Social Change*, 49, 257–79.

King, R. A., Racherla, P., Bush, V. D., (2014) “What We Know and Don’t Know About Online Word-of-Mouth: A Review and Synthesis of the Literature,” *Journal of Interactive Marketing* 28, 167 – 183

Larkin, Y., (2013), “Brand perception, cash flow stability, and financial policy,” *Journal of Financial Economics,* 110, 232 - 253

Lev, B., and Zarowin, P., (1999), “The Boundaries of Financial Reporting and How to Extend Them,” *Journal of Accounting Research*, 37 (2), 353–85

Li, X., and Hitt, L.M. “Self selection and information role of online product reviews,” *Information Systems Research, 19,* 4 (2008), 456–474

Liu, Y. (2006), :Word of mouth for movies: Its dynamics and impact on box office revenue,” Journal of Marketing , 70 (3), 74-89.

Lorek, K.S., Willinger, G. L., (1996), “A Multivariate Time-Series Prediction Model for Cash-Flow Data,” *The Accounting Review,* Vol. 71, No. 1, pp. 81 – 102

Lorek, K.S., Willinger, G. L., (2006), “Time-Series Properties and Predictive Ability of Quarterly Cash Flows,” *Working Paper Series,* Northern Arizona University

Lorek, K.S., Willinger, G. L., (2011), “Multi-Step-Ahead Quarterly Cash-Flow Prediction Models,” *Accounting Horizons,* Vol. 25, No. 1, pp. 71 – 86

Lorek, K. S., (2014) “Trends in statistically based quarterly cash-glow prediction models,” Accounting Forum, 38, 145 - 151

Luo X (2009) "Quantifying the long-term impact of negative word of mouth on cash flows and stock prices," *Marketing Science* 28(1):148–165.

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

Luo, X., Zhang, J., Duan, W., (2013) "Social Media and Firm Equity Value," *Information Systems Research,* 24(1):146-163. http://dx.doi.org/10.1287/isre.1120.0462

Madden, T. J., Frank, F.,and Fournier, S., (2006), "Brands Matter: An Empirical Demonstration of the Creation of Shareholder Value Through Branding," *Journal of the Academy of Marketing Science*, 34 (2), 224-35.

McCarthy, N., (2014), "Social Networks Take Up the Most Online Time in the U.S." retrieved Feb 18, 2015 from http://www.statista.com/chart/2272/social-networks-take-up-the-most-online-time-in-the-us/

McKinsey and Company (2012) "The social economy: Unlocking value and productivity through social technologies"

McKinsey and Company (2012) "Turning buzz into gold – How pioneers create value from social media"

Microsoft Dynamics (2015) "Your Brand Sux – Turning Social Sentiment into Opportunity".

Mizik, N., and Jacboson, R., (2008) "The Financial value Impact of Perceptual Brand Attributes," *Journal of Marketing Research,* 45 (1), 15-32

Mizik, N., and Jacboson, R., (2009) "Valuing Branded Businesses," *Journal of Marketing,* 73 137 – 153

Morgan, N.A., and Rego, L.L. (2006) "The value of different customer satisfaction and loyalty metrics in predicting business performance", *Marketing Science, 25,* 5, 426–429

Mortanges, C. P. D., Riel, A. V. (2003) "Brand Equity and Shareholder Value", *Eurpoean Management Journal,* Vol. 21, No 4, pp. 521-527

Nickell, S., (1981). "Biases in dynamic models with fixed effects," *Econometrica*, 49(6), 1417–1426

Nielsen, A.C. (2012). Global Trust in Advertising and Brand Messages: A Nielsen Report. NY: USA, ACNielsen.

O'Sullivan, D., Hutchinson. M.C., O'Connell, V., (2009) "Empirical evidence of Stock Market's (Mis)Pricing of Customer Satisfaction", *International Journal of Research in Marketin,* 26, 154-161

Ostrowski, D. A., (2010). "Sentiment mining within social media for topic identification. In Semantic Computing (ICSC)," *IEEE Fourth International Conference* on (pp. 394-401). IEEE

Pang, B., Lee, L. (2008) "Opinion Mining and Sentiment Analysis," *Foundation and Trends in Information Retrieval,* Vol. 2, Nos. 1–2, 1–135

Patton, J and Timmermann, A (2008) Portfolio Sorts and Tests of Cross-Sectional Patterns in Expected Returns . Available at http://www3.imperial.ac.uk/pls/portallive/docs/1/28779699.PDF

Pew Research Center (2014), "Social Media Use by Age Group Over Time", retrieved Feb 18th from *http://www.pewinternet.org/data-trend/social-media/social-media-use-by-age-group/*

Proulx, M., (2010) "Top Five Social Media Marketing Mistakes," *Bloomberg Business,* retrieved March 24th 2015 from http://www.businessweek.com/managing/content/jun2010/ca20100617_480316.htm

Rappaport, A., (1986), "Creating Shareholder Value," *New York: The Free Press.*

Rao, K.S. R., Bharadwaj, N., (2008) "Marketing Initiatives, Expected Cash Flows, and Shareholders' Wealth," *Journal of Marketing*, Vo. 72, 16 – 26

Ray, S., Savin, N.E., Tiwari, A., (2009), "Testing the CAPM revisited," *Journal of Empirical Finance,* 16 – 721-733

Resnick, P., Zeckhauser, R., Friedman, E., and Kuwabara, K., (2000), "Reputation Systems," *Communications of the ACM*, 43, 12, 45–8.

Richter, F., (2013) "Social Networking is the No. 1 Online Activity in the US" retrieved Feb 18th from *http://www.statista.com/chart/1238/digital-media-use-in-the-us/*

Roodman, D.M., (2009) "How to do xtabond2: an introduction to Difference and System GMM in Stata." *The Stata Journal* 9, 86–136.

Rust, R. T., Ambler, T., Carpenter, G. S., Kumar, V., Srivastava, R. K., (2004), "Measuring Marketing Productivity: Current Knowledge and Future Directions," *Journal of Marketing,* 68, 76 - 89

Scholz, M., Dorner, V., Landherr, A., Probst, Florian.(2013), "Awareness, Interest, and Purchase: The Effects of User-and Marketer-Generated Content on Purchase Decision Processes," paper presented at *34th International Conference on Information Systems*

Simon, C. J., and Sullivan, M. W. (1993) "The Measurement and Determinants of Brand Equity: A Financial Approach?" *Marketing Science* 12 (Winter): 28-52.

Singh, P. K., and Pattanayak. J. K., (2014) "Linking Customer Satisfaction with Shareholders' value: A Review," *Global Journal of Finance and Management,* 6, 5, pp. 403 - 412

Strong, E. K., (1925), "Theories of Selling," *Journal of Applied Psychology, 9, 75 - 86*

Sonnier, G. P., McAlister, L., Rutz, O. J., (2011). "A Dynamic Model of the Effect of Online Communications on Firm Sales," *Marketing Science 30 (4), 702-716*

Srivastava, R. K., Shervani, T. A., and Fahey, L., (1998), "Market-based Assets and Shareholder Value: A Framework for Analysis," *Journal of Marketing*, 62(1), 2-18.

Surowiecki, J. (2005). "The wisdom of crowds: why the many are smarter than the few," *Abacus*.

Tetlock, P., (2007) "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *The Journal of Finance* , 62, 1139-1168.

Tellis, G., and Johnson, J., (2007) "The value of quality," *Marketing Science, 26,* 6,758–773

Thecmosurvey., (2014) "Tweet this: Social media important to company performance but difficult to prove" retrieved Feb 18th, 2015 from *http://www.cmosurvey.org/blog/tweet-this-social-media-important-to-company-performance-but-difficult-to-prove/*

Thecmosurvey., (2016), "CMO Survey Report: Highlights and Insights", *www.cmosurvey.org*

Thelwall, M., Wouters, P., and Fry, J., (2008) "Information centered research for large scale analyses of new information sources," *Journal of the American Society for Information Science and Technology,* 59(9), 1523-1527.

Thelwall, M., Buckley, K., and Paltoglou, G., (2011) "Sentiment in Twitter events," *Journal of the American Society for Information Science and Technology,* 62(2), 406-418

Tim, A., (1997) "How much brand equity is explained by trust?," *Management Decisions,* 35.4, 283 - 292

Tirunillai S, Tellis G., (2012) “Does chatter matter? The impact of online consumer generated content on a firm’s financial performance,” *Marketing Science,* 31(2):198–215

Triest, S. V., Bun, M. j. G., Raaij, E. M. V., Vernooij, M. J. A., (2009), “The impact of customer –specific marketing expenses on customer retention and customer profitability,” *Marketing Letters,* 20, 125 - 138

Trusov, M., Bucklin, R., and Pauwels, K. (2009). “Effects of word-of-mouth versus traditional marketing: findings from an internet social networking site.” *Journal of Marketing*, 73 (5), 90- 102

Tuli, K. R., Bharadwaj, S. G., (2009). “Customer satisfaction and stock returns risk,” *Journal of Marketing,* 73 (6) 184–197

Villarejo-Ramos, A. F., Sanchez-Franco, M. J., “Impact of marketing communication and price promotion on brand equity,” *Journal of Brand Management*, 12 (6) (2005), pp. 431–444

Wansbeek, T., (2012), “On the remarkable success of the Areallano-Bond estimator,” *AENORM,* vol. 20 (77) pp. 15 - 19

Wilde, A., (2014). “Social Media and Brand Management”, Retrieved Feb 18, 2015, from http://www.ttuhub.net/2014/11/social-media-brand-management/

Wilson, G.P. (1986). “The relative information content of accruals and cash flows: Combined evidence at the earnings announcement and the annual report release date.” *Journal of Accounting Research* 24 (Supplement): 165-200.

Wilson, G.P. (1987). “The incremental information content of the accrual and funds components of earnings after controlling for earnings.” *The Accounting Review* 62 (April): 293-322.

Wirthman, L., (2013) “Why Businesses Should Listen to Customers on Social Media,” *ForbesBrandVoice,* Retrieved March 24th From http://www.forbes.com/sites/ups/2013/08/08/why-businesses-should-listen-to-customers-on-social-media/

Wooldridge J (2002). “Econometric Analysis of Cross-Section and Panel Data.” *MIT press.*

Wooldridge, J.M., (2009) “Introductory Econometrics: A Modern Approach”. *South Western College Publications*, fourth ed.

Wright, A. (2009) "Mining the Web for feelings, not facts," *New York Times* retrieved on Feb 18, 2015 from http://www.nytimes.com/2009/08/24/technology/internet/24emotion.html?pagewanted=all

Ye, Q., Law, R., Gu, B., Chen, W., (2011) "The influence of user-generated content on traveler behavior: An empirical investigation on the effect of e-word-of-mouth to hotel online bookings" *Computers in Human Behavior,* 27, 634 - 639

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

Zendesk.com, (2013) "The impact of customer service on customer lifetime value" retrieved on Feb 18 2015 from https://www.zendesk.com/resources/customer-service-and-lifetime-customer-value