CSR, Big Data, and Accounting: Firms' Use of Social Media for CSR-Focused Reporting, Accountability, and Reputation Gain

GREGORY D. SAXTON

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN BUSINESS ADMINISTRATION
SCHULICH SCHOOL OF BUSINESS
YORK UNIVERSITY
TORONTO, ONTARIO

JULY 2016

Abstract

The rise of "Big Data," particularly social media, is engendering considerable disruptions in the ways in which firms and stakeholders communicate about firm-relevant issues. The effect of social media appears to be particularly strong in the domain of corporate social responsibility (CSR). This thesis presents three empirical studies on *Fortune 200* firms' use of social media to engage in CSR-related activities. All three studies rely on original 2014 data related to the 42 CSR-focused Twitter accounts maintained by the US-based *Fortune 200* companies – comprising 18,722 firm messages and 163,402 messages sent by members of the public. This thesis first examines the outcomes of firms' social media-based CSR engagement, building a theoretical argument about the reputational benefits, or *reputational capital,* acquired by firms through the messages they send on social media. It then turns to an investigation of the *public's* discussion of the companies' CSR activities; this second study relies on inductive analyses to build insights into the nature of the firm-centered CSR messages sent by members of the public, the nature of firms' reactions to these public messages, and the relationship between the two. The third and final study refines and then empirically tests the causal model developed in the second study. Collectively, these three studies shed light on the nature of the *micro-reporting* and *micro-accountability* behaviors that appear to characterize firms' CSR efforts on social media sites. The thesis concludes with a summary of the implications of these new behaviors for the accounting and CSR literatures.

Dedication

This thesis is dedicated to my brilliant wife Michelle, without whose love and support this thesis would not have been possible, and to my two amazing boys Riley and Tyler. Thanks to all three of you for all of your love, support, help, and encouragement.

Acknowledgements

I would like to thank first my committee, Dean Neu, Giri Kanagaretnam, and Jeff Everett. All were supportive and helpful. I owe special thanks to Dean, who took a chance on an unconventional PhD applicant and convinced me to switch programs to Schulich. It was a great move, and his support and guidance over the past five years has been outstanding.

Others have nudged me to learn, to think about concepts in a new light, or pushed me to do better. Here courses by Linda Thorne, Alan Richardson, Dean Neu (thanks again), Eileen Fischer, and Amin Mawani were all eye-opening in different ways. The courses I took on economics and mathematics and game theory at UB were similarly useful in further developing a different set of methodological and theoretical skills. In a different way, my colleagues at UB were also important – they provided a culture in which continual learning is seen as admirable and an environment in which everyone strives to do relevant, rigorous, and groundbreaking work. My friend and frequent co-author Chao Guo has been a great supporter of this endeavor and has always provided frank and incredibly helpful advice on my research over the years. My friend Jerome, in turn, provided helpful inspiration in talks over beer about Python and machine learning.

My family – my sisters Tracy and Karen and my Mom and Dad – each provided some combination of example, support, and food that was likely instrumental in me reaching this point. My grandparents did not go to college but without their support of my parents doing so I would likely not be at this point today.

My biggest thanks, however, are for my wife Michelle. She has been a sounding board, a supporter, a reviewer, and more. This thesis would simply not have been possible without her unconditional and continual love and support.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This thesis is largely a collection of three stand-alone manuscripts. Each is intended to be sent out for review at peer-reviewed journals and thus contains its own distinct abstract, introduction, and concluding chapters. Nevertheless, in an effort to facilitate the reading of the dissertation, I have included this introductory chapter to summarize its general questions, theoretical path, data contributions, and empirical results.

I was originally motivated to write this dissertation after noting that the rise of "Big Data," particularly social media, is engendering considerable disruptions in the ways in which firms and stakeholders communicate about firm-relevant issues. The impact of new and social media has long been of interest to me within the realm of non-profit organizations (e.g., Saxton et al., 2007; Saxton & Guo, 2011; Saxton et al., 2012). In pursuing the PhD at Schulich, I quickly saw that many of the same developments could also be found in accounting phenomena, yet had not been deeply examined. This at the same time as social media came to pick up speed and cause further disruptions in organizational and public communication in ways that were new to all academic disciplines. At the broadest level, social media have, simply put, engendered a more interactive communicative environment – a public space (Neu, 2006) where citizens, firms, and interest groups alike can debate, discuss, denigrate, deny, and dialogue about such core issues as the firm's level of corporate social responsibility.

1.1 Vignettes

The domain of corporate social responsibility (CSR) thus seemed especially ripe for an assessment of the effect and impact of social media on both stakeholders and the firms themselves. In searching for a dissertation topic, I came to see that the types of CSR activities firms engage in on Twitter presented both opportunities and challenges to the existing accounting and CSR literatures. As a brief entry point to some of the ways the accounting literature is being challenged by these activities, I present a number of brief vignettes of new forms of firm and public communication that shed light on these issues. To start, take the following tweet by Cisco's CSR-focused Twitter account *@CiscoCSR*:

> *We're proud to have supported @100khomes campaign to house 100k #homeless Americans. Congrats! http://t.co/sYrx0hegbK*

Such messages relay an account of the firm's CSR performance and thus readily fit within the existing concepts of reporting and disclosure (e.g., Cho et al., 2010; Neu et al., 1998; Patten, 2002). Yet there is something notably different about this type of reporting: it is best considered a *micro* report. It thus poses a "problem" to the existing reporting literature insofar as it represents a model in which reporting is not delivered annually but rather daily and in brief, discrete chunks.

The following tweet by Bank of America, sent on November 30th, 2014 and containing a #WorldAIDSDay video starring U2's Bono, poses a different challenge to existing CSR literature:

> *It's #WorldAIDSDay. RT this video & we'll donate to @RED. Help us get one step closer to ending #AIDS. #onestep4RED*

Namely, it is indicative of how social media is being used to actually *conduct* CSR. Corporate philanthropy conducted in this manner allows firms to engage a broader spectrum of the public in its philanthropy efforts. Whereas previously Bank of America would have simply donated $1 million and noted the donation in an annual report, directory, or filing (e.g., Lev et al., 2010), on social media the firm may donate the same amount of money but will receive a considerably larger reputational yield – the total donation is not delivered until 1 million Twitter users have come into contact with the message – read it, "liked" it, or shared it. In effect, the firm not only involves the public in its corporate philanthropy effort – thereby forming a connection with members of the public – but plausibly receives a not insubstantial boost in both of the two main dimensions of reputation – awareness and favorability (e.g., Rindova et al., 2005). The challenge to existing literature here lies in how such efforts combine corporate philanthropy, reporting on corporate philanthropy, and expanded reputational benefits.

Differently put, firms' communication on social media is often not merely reporting on its CSR activities but often *constitutes* a CSR activity in and of itself. Other types of messages, such as the following tweet from *@Microsoft_Green*, diverge from the reporting model in different ways inasmuch as it presents one-way, firm-to-public communication yet does not report on the firm's own activities:

> *From @virginia_tech: Sugar could help power #smartphones in the future via @guardian http://t.co/AbeW20Rzs8 #cleanenergy*

As with disclosure tweets, this "public education" tweet is intended to convey information; however, unlike disclosure tweets, such messages are intended not to report on the company's activities but rather to educate the public on a topic related to a CSR core area – such as technology, health, education, sustainability, diversity, or the environment. The communication and public relations are well suited to explaining such communicative forms (e.g., Taylor & Kent, 2014), yet the accounting literature has no readily available concepts to theorize about these new forms of CSR communication.

The accounting literature is similarly not conceptually prepared to account for the following two types of messages:

> *Pls join the #P4SPchat on citizen engagement. Starts in 45 minutes, Noon ET http://t.co/GobUBqy3Qe*

> *Retweet if you're taking a stand for LGBT youth by celebrating #SpiritDay #ComcastGoesPurple http://t.co/pLzKSdn13n*

The first of these, sent by *@IBMSmartCities*, relates to what is known as a "tweetchat," involving a dialogic, back-and-forth dialogue between the firm and anyone interested in the chat topic. While dialogue has been examined in the accounting literature on earlier forms of technology such as websites and discussion boards, the findings have revealed an absence of true reciprocal dialogue (Unerman & Bennett, 2004). One-way reporting predominated (e.g., Cho & Roberts, 2010). On social media this seems to not be the case.

The second of the above messages is neither informational nor dialogic. Instead, it aims to *mobilize* the public to take some CSR-related action – one that is not directly related to the firm's self-interest. Like dialogic tweets, mobilization tweets aim to go beyond one-way information. Yet instead of seeking to engage members of the public in dialogue – or, in effect, to *say* something – mobilizational messages aim to mobilize audience members to *do* something. Again, this form of communication is not accounted for in the accounting or CSR literatures but can be explained by referring to communication concepts (Saxton & Waters, 2014).

A theme in the above messages is the high level of interaction between firm and public. The flip side of the democratized, more open communication system is that firms need to be comfortable with the fact that, on social media, they cannot control the message. The public can be just as much of a player as the firm. Take the following example. When Walmart's CSR-focused account *@WalmartAction* tweeted

> *A new Walmart distribution center in Union City, #GA, will create ~400 jobs over three years: http://trib.al/uyNfixL @repdavidscott*

the following negative reply was sent:

> *@WalmartAction @repdavidscott Has anybody got back to you, yet, on how many jobs it will destroy?*

In effect, social media do appear to represent a form of interactive public space (Neu, 2006), a communication market in which ideas, information, rumors, opinions, and sentiments compete for public attention. Social media also enable the public to enter this market at little to no cost and actively engage with firms – not only praising, questioning, and chastising firms in turn for their CSR performance but entering into real two-way dialogue with the corporation.

The accounting literature is further problematized, I argue, by a lack of theorizing around how and why firms choose to respond to queries from members of the public. A final example is indicative of the types of interactions that are taking place. On February 21, 2014, a

Twitter user sent the following message targeted at Cisco:

> *@plus_socialgood @CiscoCSR @Cisco Luv #CSR's philanthropic endeavors! Has #Cisco considered teaching impoverished women how 2 make Cat5cable?*

Cisco's CSR-focused account *@CiscoCSR* responded by saying:

> *@tinacornely Good idea. We teach impoverished women how to have careers in IT through @CiscoNetAcad and help them through nonprofit partners.*

The above interaction poses an interesting problem for the accounting literature. A specific member of the public has effectively asked for – and received – an "account" (Ahrens, 1996) of the firm's actions. In an important sense, the back-and-forth conversations between firm and public represent a form of CSR-focused accountability activity that has previously been difficult to see – if indeed it existed at all.

While the extant literature has not addressed these "problems," there have been calls for greater understanding of stakeholder reactions to CSR disclosures (Moser & Martin, 2012) and of the micro-foundations of CSR efforts (Aguinis & Glavas, 2012). This study helps address such calls while building and testing theoretical explanations of the determinants and outcomes of the phenomena described above. In so doing, it constitutes a focused effort to fulfill calls to address the implications of "Big Data" for the accounting literature (Vasarhelyi et al., 2015).

1.2 Empirical Setting: Twitter

This thesis confronts the opportunities and challenges noted above in presenting results from three studies that focus on the nature, determinants, and outcomes of *Fortune 200* firms' CSR-based activities on the social media platform Twitter. I quickly recognized, that to pursue these general issues, I would need to create an extensive cross sectional data myself. Indeed, one of the contributions of this dissertation is the creation of a large n dataset (including, among other data, 163,402 public messages and 18,722 firm messages) that will be made publicly available upon future publication of this research.

In so doing, I decided to largely focus on Twitter as the main engine of communication between firms and their stakeholders. Twitter is the world's premier *message network*. Communication occurs in the form of brief, discrete messages; this is different from the traditional corporate website, for example, which is more akin to a static "brochure" than a vehicle for the provision of messages. Twitter is proving to be a powerful, networked, real-time information aggregation and dissemination platform; it is a public space (Neu, 2006) that has been found to have a robust effect in disseminating information in such contexts as disaster response efforts (Hughes & Palen, 2009), protest movements (Gaffney, 2010), and marketing campaigns (Jansen et al., 2009).

Twitter is an important setting for two reasons. First, it has become an important component of firms' information environments, as seen in the growing body of research on the impact Twitter and other forms of social media are having on the financial markets (Blankespoor et al., 2014; Saxton, 2012). Differently put, Twitter is important in its own right through the direct role it plays in the markets and the CSR-related public space.

A second reason Twitter is important is more academic: It facilitates tests of the development, flow, spread, and discussion of accounting information. To understand this we can consider the efficient markets hypothesis. The predominant capital markets literature has a relatively simple view of accounting information, focusing on the production and disclosure of information, to be sure, along with incorporation of information into price. However, within the "black box" between disclosure and incorporation into price occur variations in the degree to which information is disseminated (Blankespoor et al., 2014), variation in how it is changed, and variation in how it is discussed. I posit this pattern holds not only in the capital markets but in all facets of firms' information environments, including the CSR domain. The archiving, dissemination, modification, and discussion of accounting information can all be considered dimensions of information processing. I argue that all are important, and can result in different market, firm, and public outcomes. The discussion element is particularly omitted in extant literature, despite arguments that "Access to information is far less important, politically, than access to conversation" (Shirky, 2011).

At the same, the interactivity inherent on social media causes a change in role for the company from merely discloser of information to that of information intermediary. One of Twitter's key market roles, in fact, is as a dissemination network (Blankespoor et al., 2014). Yet this network is one in which the traditional set of information intermediaries in the capital markets – the array of sell-side analysts, mainstream media, auditors, and financial institutions – is disrupted and democratized.

Where Twitter becomes relevant for researchers is not only because of how social media users access, share, and discuss firm-relevant information. Rather, it is doubly important because it provides a venue in which (thus far unexamined) theories about reporting, disclosure, and information processing – about information seeking, archiving, sharing, and discussing – can now be tested in a non-laboratory environment.

It is in that spirit that this thesis builds a large-n and large-variable (i.e., Big Data) dataset to help deliver theoretical and empirical insights into firms' and stakeholders' CSR activities on these new communication platforms.

1.3 Individual Studies

In terms of the organization of the thesis, I next present three papers in turn on *Fortune 200* firms' use of social media to engage in CSR-related activities. Paper #1 examines the public's reactions to the CSR messages sent by firms through their Twitter accounts. Firms' efforts on these sites are posited to derive from the effort to build *reputational capital*. Empirical analyses code the *communication tactics* employed by the firm in each of the 18,722 original CSR messages sent in 2014; I then relate these communication tactics to two measures, based on public reactions to these messages, that reflect the two core dimensions of reputational capital. This paper thus contributes to the accounting literature by providing evidence of a number of new, non-reporting-based CSR communication tactics and illustrates how firms may acquire reputational capital on a micro, message-by-message, day-to-day level.

Paper #2 flips the first study on its head and examines *firms'* reactions to messages sent by *members of the public.* I argue that how firms respond and react to public comments, queries, and questions represents their commitment to public account-giving and reporting behavior. Given how novel the conceptual and empirical context is, this paper employs inductive analyses – specifically, of the 163,402 messages sent on Twitter mentioning the 42 CSR-focused accounts managed by Fortune 200 companies in 2014 – in order to help understand what drives firm reactions to CSR-related queries and comments from the public. The study innovates by incorporating not only traditional qualitative inductive methods but also "Big Data"-driven machine learning techniques to help identify the most important features of the public messages. The paper culminates in the presentation of a theoretical model of the determinants of firm reactions to public CSR messages.

Paper #3 refines the conceptual model presented in the second study, adds specific hypotheses, and then empirically tests the model using a series of multivariate logit regressions. The paper provides evidence of the message, sender, and firm characteristics that drive companies to engage in *micro-reporting* and *micro-accountability* behaviors in the CSR domain. In effect, on social media members of the public are continually "calling firms out" for their CSR actions, and firms' decisions to respond or not respond constitute a new form of public, dynamic, interactive reporting and accountability behavior that has yet to be addressed by the extant accounting literature.

The thesis concludes with a summary of the implications of these new behaviors for the accounting and CSR literatures.

Chapter 2

CSR Communication and the Micro-Accumulation of Reputational Capital

Abstract

This paper argues corporate social responsibility (CSR) performance should be understood as something that is not just reported but communicated. Beyond one-way disclosure, CSR is increasingly seen in firms' mobilizational efforts, two-way dialogue with stakeholders, public educational messages, and a variety of other discursive and conversational tactics. These tactics, I posit, can play an important role in determining dynamic, micro-level changes in reputation. In this study, I employ inductive analyses combined with machine learning algorithms to code the communication tactics and acquired reputation in the 18,722 original messages sent by Fortune 200 firms' dedicated CSR feeds in 2014. Using a series of logit regressions, I then examine the relationship between different dimensions of reputational capital and firms' CSR communication tactics. I find the awareness dimension of reputation is driven by the use of informational tactics such as disclosure, while reputational favorability is significantly influenced not by the provision of information but by the use of more interactive communicative tactics. In summary, this paper provides evidence of a number of new, non-reporting-based CSR communication tactics and illustrates how firms may acquire reputational capital on a micro-, message-by-message, day-to-day level.

Keywords: Corporate reputation, corporate social responsibility, CSR disclosure, corporate communication, stakeholder engagement, public perceptions, social media, Big Data, machine learning

2.1 Introduction

The accounting literature generally argues firms' corporate social responsibility (CSR) efforts constitute either economically valuable signaling and information disclosure (Mahoney et al., 2013; Mishra & Suar, 2010) or cynical efforts at impression management, legitimacy-seeking, and "greenwashing" (Du & Vieira, 2012; Lyon & Montgomery, 2013; Neu et al., 1998; Patten, 1991). While there may be truth in both perspectives, this dichotomy misses important aspects of firms' CSR efforts, leading to calls for more nuanced theoretical approaches (Cho et al., 2015).

Notably, there is mounting evidence many firms do not merely disclose CSR performance, they *communicate* (Castelló et al., 2015; Colleoni, 2013), going beyond one-way reporting of CSR activities to engage stakeholders in dialogue and other communicative tactics (Kent & Taylor, 2016; Unerman & Bennett, 2004). The question is, why would any rational firm do this? I argue the reason lies in how such communication influences public perceptions, which are increasingly posited as playing an important role in determining employee, regulatory, consumer, and investor outcomes (e.g., Zahller et al., 2015). Despite this evidence, the focus of CSR research, whether in the signaling (Richardson & Welker, 2001) or legitimacy (Du & Vieira, 2012) approaches, is still largely on one-way reporting. Moreover, despite research on the importance of legitimacy and reputation, we have relatively little understanding of precisely *how* public perceptions are influenced – of the underlying mechanisms involved in the acquisition of legitimacy, a strong corporate reputation, or favorable public opinion, particularly at the micro level (Aguinis & Glavas, 2012).

This study addresses both issues in a micro-level study of the relationship between CSR communication and corporate reputation. Starting with a general conceptual framework outlining the connection between CSR communication, public perceptions, and strategic outcomes, I then present a micro-level model that relates the communication tactics visible in firms' social media-based CSR messages to dynamic changes in reputation. This model guides empirical analyses aimed at addressing two research questions. First, which communication tactics are used in firms' CSR efforts on social media? Second, how do these tactics relate to the accumulation of reputational capital?

Findings are based on a combination of inductive insights and logit analyses on an original dataset comprising all 18,722 social media messages *Fortune 200* firms sent through their dedicated CSR-focused Twitter accounts in 2014. Using machine learning techniques (Go et al., 2009) that are relatively new to the field, I first code each message for visible communication tactics as well as two core dimensions of reputation: *awareness*, as reflected by how broadly the message is diffused (retweeted); and *favorability*, as reflected in the sentiment seen in public responses to firms' messages. A series of logit regressions are then used to examine the relationship between communication tactics and reputation.

This study aims to make several contributions to the CSR and accounting literatures. To start, in finding evidence of a number of non-informational communicative tactics, the study extends recent insights that CSR activities go beyond mere reporting (e.g., Schultz et al., 2013). Moreover, in illustrating how CSR-based reputational capital is accumulated on a micro-, message-by-message, day-to-day level, the study responds to calls for greater understanding of stakeholder reactions to CSR disclosures (Moser & Martin, 2012) and of the micro-foundations of CSR efforts (Aguinis & Glavas, 2012). The study opens new avenues of research by encouraging a longer-term view of CSR efforts, by expanding the number of activities that can be examined in studying CSR, and by delving into the day-to-day flows of CSR communication and reputational capital.

In the following section I lay out the theoretical perspective, which provides an overview of how the study builds on existing examinations of legitimacy and reputation, my arguments on how CSR communication goes beyond reporting, and the relationships between different communication tactics and the acquisition of reputational capital. Method and results follow, and the paper ends with a discussion of the study's practical and theoretical implications along with insights into how a Big Data-driven (Vasarhelyi et al., 2015) communication perspective provides scholars with a new set of tools for analyzing and problematizing CSR.

2.2 CSR Communication and Corporate Reputation

2.2.1 Legitimacy, Reputation, and the Importance of Public Perceptions

The legitimacy approach to CSR (e.g., Du & Vieira, 2012; Neu et al., 1998; Patten, 2002) has concentrated on legitimacy as an outcome of CSR and, to a lesser extent, reputation (the key outcome of interest here). These complementary, inter-related concepts are both "perceptions of approval of an organization's actions" (King & Whetten, 2008, p. 192). Organizations gain legitimacy via CSR disclosures or actions when those indicate they comply with the minimum, taken-for-granted societal standards in the CSR arena; they gain reputation when they become more favorably viewed with respect to the ideal standard in CSR behavior – they hold a "good reputation" when "they are viewed favorably relative to the ideal standard" (King & Whetten, 2008, p. 192). In effect, legitimacy is a more binary concept, with firms scored according to whether they are above or below the standard, while reputation is a more continuous construct, with firms scored on a range from the best possible to the worst possible behavior.

While both legitimacy and reputation reflect collective public perceptions of approval of the firm's behavior (Bebbington et al., 2008; King & Whetten, 2008), the connection to the broader notion of *public perceptions* has largely been implicit. Here the CSR literature could benefit from insights in sociology and political science, where public perceptions writ large – including public opinion, legitimacy, grassroots support, image and reputation – are explicitly seen as an important and understudied outcome of firms' nonmarket activities (McDonnell & King, 2013; Vasi et al., 2015; Walker & Rea, 2014).

Businesses themselves have come to recognize the political importance of this broad suite of public perceptions – outcomes that are mainly achievable not through corporate reporting nor through "direct" political tactics such as lobbying or political spending, but through the use of longer-term, "indirect" or "outside" communication and advocacy tactics (Guo & Saxton, 2014; Mosley, 2009) such as research, media advocacy, grassroots lobbying,

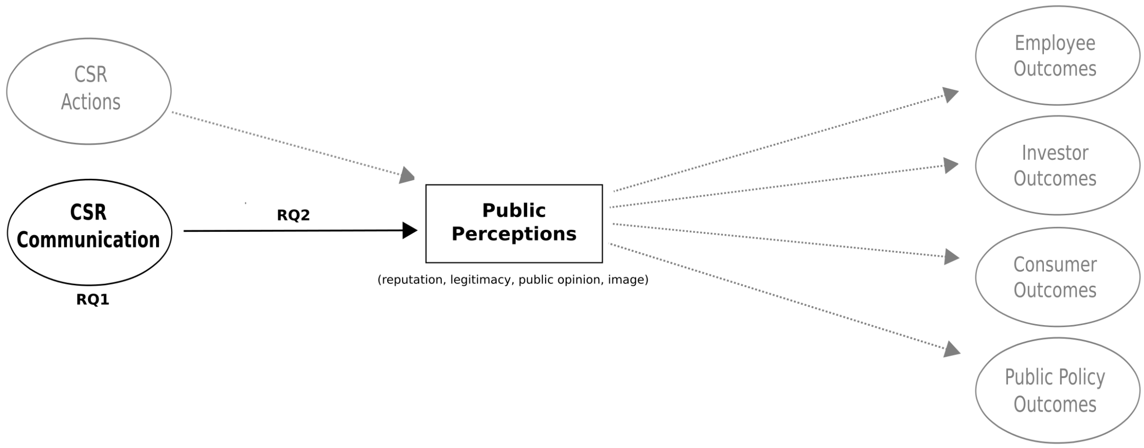


Figure 2.1: CSR Behaviors, Public Perceptions, and Key Outcomes

public education, and voter education. Not surprisingly, research finds the public affairs function of large firms has increased considerably since the 1970s (Griffin & Dunn, 2004), while Walker notes "the expanding market for grassroots lobbying services," estimating that "…nearly 40% of the Fortune 500 appears on the client lists of at least one such firm…" (Walker & Rea, 2014, pp. 292-3). Overall, social movement scholars have noted the increasing focus on reputation management" (e.g., Walker, 2009) in corporations' political engagement.

CSR Activities, Public Perceptions, and Strategic Outcomes

CSR-driven perceptions are not merely important for political goals, however. Figure 2.1, which lays out the broad conceptual starting point for the study, separates CSR into two key activities – performance[1] and communication – and presents four reasons why firms are interested in changing the perceptions that flow from these CSR behaviors. Namely, CSR-driven public perceptions directly influence *employee outcomes, investor and capital market outcomes, consumer outcomes,* and *policy outcomes.*

To start, a firm's CSR efforts – and the reputation that derives from those efforts – can

[1]By "performance," I mean the non-communicative actions such as corporate philanthropy, investment in environmental initiatives, commitment to gender equality, etc.

help attract new employees and play an important role in employee morale, motivation, and productivity (Balakrishnan et al., 2011; Waddock & Graves, 1997). As Roberts & Dowling (2002, p. 1079) argued, "ceteris paribus, employees prefer to work for high-reputation firms."

There also appears to be a strong link between CSR performance and consumer outcomes (Sen & Bhattacharya, 2001), with purchasing decisions, product evaluations, and brand loyalty all influenced by perceptions of CSR performance (Brown & Dacin, 1997; Creyer, 1997). A growing number of consumers are less willing to be associated with a brand – or be members of a brand community – related to a company with a negative corporate image or reputation.

Investor and capital market outcomes are likewise driven by CSR performance (and perceptions of that performance) in a number of ways. To begin with, the public's investing behavior is influenced by firms' CSR performance (Dawkins & Lewis, 2003), and there is evidence professional investors have important preferences for CSR information (Cohen et al., 2015). Others have noted a connection between the initiation of CSR *disclosure* and various capital market outcomes, including a lower cost of equity capital and the attraction of analyst following and dedicated institutional investors (Dhaliwal et al., 2011). While such outcomes are likely to be largely driven by the information content and value relevance of CSR disclosure, they also appear to be driven at least in part by the public's perceptions of CSR performance. For instance, Herremans et al. (1993) showed companies with better CSR reputations had higher stock market returns and lower risk compared to companies with less favorable reputations.

Lastly, a growing body of research documents the influence of public perceptions on public policy outcomes. There are several ways public perceptions influence policy outcomes. First, reputation affects how policymakers assess firms (Werner, 2015); a positive reputation can enable policymakers to avoid imposing negative regulatory and fiscal policy changes. Second, positive perceptions reflect latent support for the firm, and once mobilized, these mobilized publics can be instrumental in referenda and initiatives and also act as grassroots – or, cynically, *astroturf* – lobbyists that directly engage with and apply pressure to legislators and regulatory bodies (e.g., Lyon & Montgomery, 2013; Walker & Rea, 2014). Extending beyond the purely political, there is also evidence that higher levels of legitimacy garnered through high-quality CSR disclosure can engender social resilience to exogenous shocks such as social, political, environmental, or economic upheaval (Zahller et al., 2015).

In short, reputation and other public perceptions are important determinants of both market and nonmarket outcomes. The major task of this paper is to examine how CSR communication engenders changes in reputation (the relationship highlighted in bold in Figure 2.1). While the causal logic is similar for legitimacy and reputation, I focus on the latter. Because of reputation's theoretically more continuous nature, it is more feasible to see and assess whether reputation is gained or lost on a message-by-message basis.[2] Moreover, to emphasize the instrumental nature of the acquired reputation, I use the term *reputational capital*; the reputation is not an end in itself but rather a means to acquiring the strategic outcomes noted above.

Expanding the Conceptualization and Measurement of Reputation

Despite its importance, we do not have a solid understanding of the processes through which reputation is acquired. I argue the conceptualization and operationalization of perception-based outcome variables need to be expanded in the typical CSR study in at least three ways. First, there is little examination in the existing literature of the processes by which legitimacy or reputation is acquired; the perceptions are examined at a cross-sectional level or at a snapshot in time. Effectively, most studies look at the stock of perceptions, with the consequence that we have little conceptualization of the flow. Second, as argued by Moser and Martin in their reflection on the accounting-based CSR literature (2012, p. 801), "despite a significant amount of prior research, we do not yet fully understand ... how investors and other stakeholders react to CSR disclosures." Lastly, in a recent review of the CSR literature,

[2] An individual message could be seen as "legitimate" or "illegitimate," but it would be difficult to determine whether and how much legitimacy was gained in response to a given CSR message.

Aguinis and Glavas (2012, p. 955) argue that, given the "predominance of organizational- and institutional-level research," there is a pressing need to examine the micro-foundations of CSR. The present study addresses these three issues by leveraging data on social media-based stakeholder reactions to examine the micro-level acquisition of CSR-based reputation.

Social Media and Reputational Change

Beginning in the early 2000s with Friendster and MySpace, social media use exploded with the rise of LinkedIn in 2003, followed by Facebook in 2004, YouTube in 2005, Twitter in 2006, and Instagram and Pinterest in 2010.[3] What distinguishes these social media from older forms of new media (such as websites) is their interactivity, the focus on micro-messages, and the primacy of formal social networks (Kane et al., 2014; Scott & Orlikowski, 2012; Suddaby et al., 2015). Social media are, fundamentally, communication networks (Monge & Contractor, 2003): they are networks, with actors dynamically connecting with and dis-connecting from other actors; and what flows through these networks is communication, in the form of the stream of messages that are sent. Whether a tweet on Twitter, a status update on Facebook, a photo on Instagram, or a video on YouTube, the medium for communication is the series of visual and/or textual updates that is sent to an organization's followers. It is these continual brief, discrete dynamic updates, or "messages," that comprise the central communicative tool on all social media platforms (de Vries et al., 2012; Lovejoy & Saxton, 2012).

Because of the ability to communicate at low cost in a real-time basis to large audiences without regard for geographic distance, social media have recently become prominent tools for CSR communication (e.g., Fieseler & Fleck, 2013; Lee et al., 2013; Whelan et al., 2013). Moreover, social media represent a form of "Big Data" (Vasarhelyi et al., 2015), with one of the key implications being that it makes visible – and testable – certain phenomena that

[3] With 50 million users on Pinterest, 236 million on Twitter, 296 million on LinkedIn, 300 million on Instagram, 800 million on China's Tencent QQ, and over 1.23 billion users on Facebook, social media sites have a substantial audience (Statista, 2015).

were previously invisible or not amenable to testing (Clark & Golder, 2015). In the same fashion, it also "leads to new research questions and new ways of thinking about existing questions" (Parks, 2014, p. 356). The present study take advantage of these qualities.

Social Media and Dynamic Changes in Reputational Awareness and Favorability

Social media afford a unique opportunity for examining reputational dynamics. Not only do social media render CSR communication publicly visible but, just as critically, they make the real-time audience reactions to this communication visible as well. Scholars in marketing, communication, and public relations have hence recently developed approaches that examine the public's "liking," commenting on, replying to, and sharing of organizations' messages (e.g., Smith, 2012). Given how these audience reactions are linked to specific messages, they facilitate tests of real-time reaction to specific organizational messages (Saxton & Waters, 2014). It is this ability to measure the almost real-time public reaction to an organization's CSR messages that makes research into the micro-foundations of reputational capital possible while providing organizations with a quantitative and comparable gauge to measure the relative effectiveness of their CSR messaging strategies.

Several recent public relations studies have examined audience reactions to CSR messages (see Fraustino & Connolly-Ahern, 2015; Gómez-Vásquez, 2013). More recently, Saxton et al. (2016) extended this research by situating it in the business literature while also, more importantly, further theorizing around the nature and theoretical importance of the public reactions. Specifically, Saxton et al. (2016) argue the public's retweeting (sharing) of organizations' message constitute micro-level indicators of *reputational awareness*. Messages receive retweets with messages that contain some "pass-along value" (Lee et al., 2013), and serve to expand the number of audience members that view a given message. While the organization's number of followers on Twitter reflects the baseline number of users who potentially "see" a given firm message, each time a follower retweets that message it is rendered visible also by the follower's followers, potentially substantially expanding the reach of the message. Saxton et al. (2016) argue that, in spreading the word, retweets thus directly impact the awareness dimension of reputation. Also known as familiarity, or prominence (Barnett et al., 2006; Lange et al., 2011; Rindova et al., 2005), *awareness* is one of the core dimensions of reputation and "reflects the degree to which opinions about an organization. . . .are disseminated among its stakeholders" (Rindova et al., 2005, p. 1038). While the dissemination traditionally occurs through intermediaries and purchasing decisions, with its heavy use of retweeting, Twitter is, at heart, a message dissemination network, making it a new and potentially important influencer of corporate reputation.

A retweet is a micro-level reaction that, in the aggregate, can enhance a firm's CSR-related reputation by directly boosting its prominence. However, I argue it is beneficial to look beyond the awareness dimension of reputation. As Saxton et al. (2016) note, one of the limitations of their study was the dependent variable (number of retweets). They argued retweeting is an important reflection of the extent to which an organization's message resonates with the public. Yet as some Twitter users point out on their profiles, "retweets do not necessarily equal endorsement." In practice, the great majority of public retweeting actions in their dataset appeared to signify an implicit endorsement of the content communicated by the firm. Still, the authors argued future research could delineate retweeting actions into those that are positive and those that are neutral or negative. Research could also expand the nature of the variable to incorporate textual analysis and "Big Data" machine learning techniques (Bollen et al., 2011; Go et al., 2009) in order to tease out changes in sentiment toward the firm subsequent to an important exchange of messages.

Retweets and other public reactions also serve to highlight the increasing relevance of interactive, two-way, and stakeholder-to-firm CSR communication (Colleoni, 2013; Schultz et al., 2013). A second issue with existing approaches, then, is that it would be beneficial to look beyond how organizations are communicating CSR and analyze the public's role in this communication. In particular, we should look beyond simple "point-and-click" reactions such as liking and retweeting and begin to examine the communicative content in audience members' reactions to corporate messages – particularly comments on and replies to those messages.

I seek to address such limitations in coding the sentiment in public replies to firms' CSR messages. Two other recent studies have recently leveraged machine learning techniques to code sentiment in CSR-related messages. Castelló et al. (2015) examined sentiment in tweets mentioning a large health firm, while (Colleoni, 2013) coded sentiment of both companies and publics in CSR-related Twitter discussions. I build on this research by coding sentiment in the replies members of the public make to companies' CSR-related tweets.

I further build on earlier studies in developing greater theoretical precision about the nature of the sentiment seen in public replies. Specifically, beyond the fact that these message-level reactions can be be aggregated to get a sense of the dynamic collective affect toward a given firm, I argue this sentiment reflects the second important dimension of reputation, that related to the favorability (Lange et al., 2011), positive assessment (Barnett et al., 2006), or perceived quality (Rindova et al., 2005) of the firm's CSR efforts. The sentiment in public replies to companies' messages directly relates to the individual's feelings toward the company's message. It is a micro-level indication of the favorability with which the individual sees the organization's message.

The coding of the replies is important for another reason: unlike looking at new public messages (as in the Colleoni and Castelló et al. studies mentioned above), replies are linked to specific firm tweets, allowing the analyses to remain consistently focused on firms' CSR messages. As a result, I am able to examine both public *awareness* and *favorability* in conducting a tweet-level analysis that delivers insights into CSR message effectiveness.[4]

The above point is worth expanding on. It is only possible to develop and "see" reputational change by linking specific organizational actions to specific audience reactions. At the macro level, firm messages and audience reactions could be aggregated to generate insights

[4]In discussing the importance of reputation, Du et al. (2010, p. 8) conclude "stakeholders' low awareness of and unfavorable attributions towards companies' CSR activities remain critical impediments in companies' attempts to maximize business benefits from their CSR activities." In effect, social media afford a unique opportunity for examining the flow of both the awareness and favorability dimensions of CSR-driven reputation.

into broad communicative strategies and legitimizing effects (Castelló et al., 2015) and create firm-level measures of the outcomes of CSR communication efforts. Yet I am interested here in something new: drilling down to the message level and generating insights into how specific individual messages engender specific reputational responses from members of the public. This adds a unique, message-level perspective that yields insights into the micro-level accumulation of reputational capital.

2.2.2 Beyond Reporting: CSR Communication & Changes in Reputation

We now have a good understanding that CSR disclosures can boost legitimacy and corporate reputation (e.g., Du et al., 2010; Eberle et al., 2013). However, we have little understanding of how other, non-reporting types of CSR communication influence public perceptions.

Beyond Reporting: Non-Disclosure-Related Types of CSR Communication

Arguably, CSR communication has long included a non-reporting, more "public affairs"-type function. A sole emphasis on disclosure is unable to account, for example, for the strong role non-reporting communication and discourse plays in, for instance, corporate political activity,[5] where firms' CSR efforts play an increasingly important part (den Hond et al., 2014; Walker & Rea, 2014). Similar trends are seen in the marketing arena in firms' efforts to influence consumer sentiment and brand loyalty, etc. (Brown & Dacin, 1997).

This trend appears to have been accelerated by social media, where different platforms (e.g., Twitter, Facebook, Instagram, and LinkedIn) have engendered a more interactive communicative environment – a public space (Neu, 2006) where citizens, firms, and interest groups alike can debate, discuss, denigrate, deny, and dialogue about such core issues as the firm's level of corporate social responsibility (Castelló et al., 2015; Schultz et al., 2013). The broadened constituencies and stakeholder interactions fostered by social media have further

[5]Recent empirical research suggests discursive tactics (McCammon et al., 2007) shape regulatory outcomes and influence the mobilization and impact of policymaking and social movement efforts (Andrew & Cortese, 2011; Vasi et al., 2015).

engendered CSR-centered participative interactions (Fieseler & Fleck, 2013) and "citizenship arenas" (Whelan et al., 2013) that suggest an expanded notion of corporate citizenship (Matten & Crane, 2005). In this context, public voices can play a key role "in amplifying or negating the messages that firms craft" (Fischer & Reuber, 2014, p. 566). Stakeholders can also use social media to mobilize against poorly performing firms (Eberle et al., 2013).[6] In effect, on social media the public becomes a more important actor in the delivery of CSR.

With their ability to foster bottom-up, many-to-many communication, there is also strong evidence social media have expanded the types of communication that play a role in the CSR and nonmarket arenas. For example, social media have engendered a number of common communication roles such as that of the online "expert" (Rheingold, 2012); organizations adopting such a role employ one-way informational public education tactics (Guo & Saxton, 2014) that are quite distinct from the typical "disclosure" model. There is also strong theoretical as well as empirical evidence of two-way, stakeholder-to-firm communication; for instance, Schultz, Castelló, and Morsing (2013) argue CSR is a "communicative event" and view CSR as a venue for symbolic interaction and interactive, two-way, dialogic communication around CSR issues.[7] Others have similarly either argued for (Kent & Taylor, 2016) or found evidence of dialogue (Colleoni, 2013) in firms' CSR engagement with stakeholders. Finally, in line with the interactive, communication-focused, and networked nature of social media, there is strong *prima facie* evidence that successful nonmarket (Oliver & Holzinger, 2008) advocacy needs to incorporate not merely one-way informational or two-way dialogic approaches but a range of mobilizational (Lovejoy & Saxton, 2012) and relationship-building (Waters, 2011) tactics as well.

In brief, a growing number of scholars suggest research move beyond a uni-directional "dis-

[6] For instance, Vasi et al. (2015) found that the anti-fracking documentary *Gasland* generated considerable discourse on social media, which led to notable changes in the nature of public discourse, which in turn ultimately led to the adoption of anti-fracking policies. In short, in the policy arena, not only have discursive opportunities expanded with the diffusion of social media platforms, but these efforts appear to be having an effect.

[7] Schultz et al. (2013) regard CSR "...as communicatively constructed in dynamic interaction processes in today's networked societies" (p. 681).

closure" model and consider a broader range of CSR communication approaches (Castelló et al., 2015, 2013; Kent & Taylor, 2016; Saxton et al., 2016; Schultz et al., 2013). Yet beyond a few studies on dialogue (e.g., Colleoni, 2013; Morsing & Schultz, 2006), there exist no empirical efforts to examine what non-reporting types of communication firms are employing to communicate CSR issues and/or performance with stakeholders.

This is partially due to a lack of data. It is here that social media prove invaluable. Social media are a key venue for examining CSR communication insofar as they make such communication visible – as seen in the continual series of public messages that flow from firms and the public through their social media feeds (Saxton et al., 2016). Benefitting the researcher is the fact that we can see these messages. All are publicly visible and available for downloading for examining CSR communicative efforts. Consequently, CSR scholars (especially in public relations) are beginning to build their studies around an examination of social media messages. Thus far, at least five studies have been conducted, with scholars thus far coding firms' CSR messages for one-way disclosure vs. two-way dialogue (Castelló et al., 2015; Colleoni, 2013; Gómez-Vásquez, 2013) as well as the use of CSR topics (Fraustino & Connolly-Ahern, 2015; Saxton et al., 2016). However, we do not yet know whether and how firms use a broader range of tactics. The first task of the present study, then, as reflected in the following research question, will be to explore the range of communication tactics employed by firms in their CSR efforts:

> *RQ1: Which communicative tactics are used by Fortune 200 firms in their CSR behaviors on social media?*

To answer this question, I undertake an in-depth, inductive analysis of the types of communication tactics being employed by firms in their CSR messages on social media.

CSR Communication and the Acquisition of Reputational Capital

The second major task of this paper is to examine how communication tactics influence changes in reputation. Given the dearth of research on non-reporting forms of communication, much remains under- or un-examined with respect to our understanding of the relationship between communication and public perceptions. However, there is some evidence that dialogue and other communication tactics may be better than disclosure at enhancing reputation. To start, as alluded to above, CSR appears to be as much a communication-based public affairs or public relations function as it is a reporting-based accounting function. In line with the former, disclosure alone is typically not sufficient to convince the average non-investor; rather, opinions change through the use of communicative tactics – by informing, persuading, convincing, and linking, by developing shared meanings and identities, and by changing the predominant framing of key issues (Benford & Snow, 2000; King, 2007). At the same time, and dovetailing with this point, a sociological approach recognizes that influence may develop over time: "Rather than assuming that influence is wholly structural (as it is with the resource dependence argument), influence may develop over time as stakeholders build an infrastructure, develop resonant frames, and take advantage of the shifting opportunity structure" (King, 2007, p. 43).[8]

In effect, a more nuanced perspective encourages CSR scholars to incorporate a longer-term view of the nature and outcomes of CSR activities. Bringing the discussion back to the relationships outlined in Figure 2.1, CSR disclosures and communication are seen as not only short-term reactions to bad behavior or poor publicity in the aftermath of a public relations disaster (e.g., Patten, 1992; Warsame et al., 2002), but long-term strategies designed to, *inter alia,* boost public engagement and perceptions and reputation, to change predominant framing of core issues, and to develop broader ideological, grassroots, and customer support, thereby ultimately maximizing returns from employees as well as the political, consumer, and financial markets.

It is when seen from this longer-term, higher-level perspective that the value of certain tactics for understanding CSR activities becomes clearer. Public perceptions are not

[8]Strategies aimed at changing public perceptions can be very long term. For instance, over the 1980s and 1990s, with notions of "welfare queens" and other stories, conservatives in the US were able to slowly change predominant frames and public opinion regarding the causes of poverty and the *deservingness* ("capacities and desires") of the poor, with the frames – and the public policies – becoming increasingly anti-poor over time (e.g. Suddaby, 2010).

changed solely through disclosing information; rather, as argued in the public policy, marketing, communication, and public relations literatures, it is necessary to use a broader communicative repertoire designed to build relationships with, convince, educate, inform, and mobilize stakeholders (e.g., Kent & Taylor, 2016; Saxton & Waters, 2014). Overall, there is evidence a broader set of tactics might be better than disclosure at fostering positive public perceptions. Yet it remains an open question. The second major task of this paper – my second, theoretical research question – is therefore to examine how the broader range of communicative tactics identified through answering RQ1 influence corporate reputation. Given the two dimensions of reputation, this research question can be broken down into two parts:

> *RQ2a: How do the communication tactics employed by firms in their CSR messages relate to the acquisition of reputational awareness?*

> *RQ2b: How do the communication tactics employed by firms in their CSR messages relate to the acquisition of reputational favorability?*

In short, there are three concepts at the heart of this study – reputational awareness, reputational favorability, and communication tactics. Figure 2.2 summarizes the relationships among these core concepts and the associated research questions. Collectively, the study's research questions will deliver insights into both the communication tactics firms are using as well as the efficacy of those tactics.

In the following section I lay out the method and analysis plan for addressing the two research questions at the heart of this micro-level model.

2.3 Method

2.3.1 Sample and Data

The proxy for social media examined in this study is Twitter. Twitter is the world's premier message network and is a popular tool for disseminating and reacting to firm information

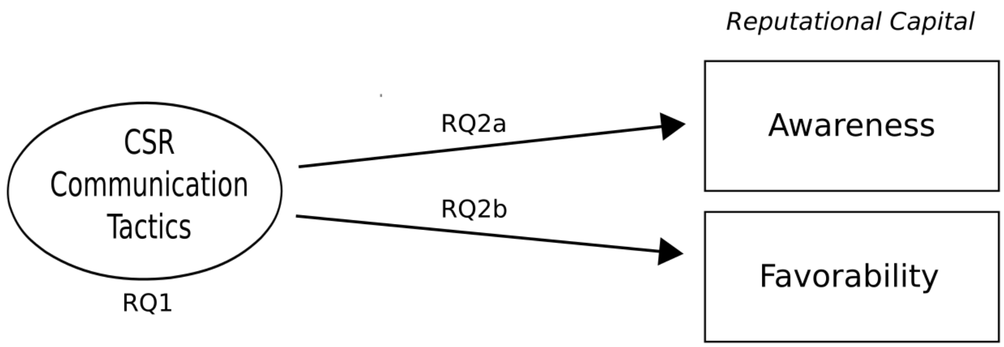


Figure 2.2: Conceptual Model of the Micro-Level Accumulation of Reputational Capital

(Blankespoor et al., 2014; Lee et al., 2015). Specifically, data for the study are derived from the 2014 CSR-focused Twitter accounts of the 200 largest firms in the 2012 *Fortune 500* index. All 200 firms maintained at least one Twitter account in 2014, typically devoted to general company news, marketing, or customer service. Some of the companies, however, manage Twitter accounts devoted entirely to CSR or sustainability issues. I gathered data from all 42 such accounts, which are listed in Table 2.1.

I accessed the Twitter application programming interface (API) using custom Python code to download all tweets sent by the 42 accounts in 2014, along with a count of the number of times each tweet was retweeted (shared) by other Twitter users. From these data I omitted any firm messages that were retweets of other users' messages; the data thus include the population of original messages (n=18,722) from all Fortune 200 firms with dedicated CSR accounts.

The Twitter API indicates whether a tweet is a reply to another Twitter message and, if so, which specific tweet is being replied to. I used this feature to also download all replies made by members of the public to any of the above 18,722 company tweets. Members of the public sent a total of 5,247 replies over the course of the year.

Table 2.1: # of Tweets sent by 42 Fortune 200 CSR Accounts, with Audience Reactions

		Firm Actions	Audience Reactions				
Twitter Account	**Company**	# Tweets Sent	# Retweets	Total # Replies	# Pos. Replies	# Neut. Replies	# Neg. Replies
3M_FoodSafety	3M	150	80	19	8	11	0
AlcoaFoundation	Alcoa	947	810	129	112	16	1
AmgenFoundation	Amgen	251	156	14	3	11	0
ATTAspire	AT&T	85	317	25	7	17	1
BofA_Community	Bank of America Corp.	1435	8445	1430	242	1012	176
CiscoCSR	Cisco Systems	2153	3234	352	128	219	5
CiscoEDU	Cisco Systems	27	38	2	0	2	0
CitizenDisney	Walt Disney	280	777	72	22	49	1
citizenIBM	International Business Machines	789	1379	84	21	62	1
ClickToEmpower	Allstate	39	32	1	1	0	0
ComcastDreamBig	Comcast	400	879	112	40	67	5
DE_Youtility	Duke Energy	1	0	0	0	0	0
Dell4Good	Dell	453	622	124	37	82	5
DellEDU	Dell	844	982	144	43	99	2
DuPont_ability	DuPont	566	647	38	7	31	0
ecomagination	General Electric	260	1487	135	19	111	5
EnviroSears	Sears Holdings	23	26	8	3	5	0
FedExCares	FedEx	11	0	2	1	1	0
FordDriveGreen	Ford Motor	86	499	101	28	64	9
FundacionPfizer	Pfizer	48	13	6	0	6	0
gehealthy	General Electric	220	337	158	32	124	2
googlestudents	Google	135	1083	102	11	88	3
GreenIBM	International Business Machines	7	0	0	0	0	0
HeartRescue	Medtronic	155	98	8	1	7	0
HoneywellBuild	Honeywell International	319	177	22	5	16	1
hpglobalcitizen	Hewlett-Packard	90	98	5	3	2	0
HumanaVitality	Humana	596	414	62	15	45	2
IBMSmartCities	International Business Machines	626	1302	127	18	108	1
Intelinvolved	Intel	621	3790	437	67	354	16
mathmovesu	Raytheon	1294	864	109	34	73	2
Microsoft_Green	Microsoft	910	752	88	29	57	2
msftcitizenship	Microsoft	1103	3155	401	133	267	1
nikebetterworld	Nike	6	33	2	0	2	0
PG_CSDW	Procter & Gamble	59	70	17	6	11	0
PPGIdeascapes	PPG Industries	198	1824	1	0	1	0
PromesaPepsiCo	Pepsi	0	0	0	0	0	0
SprintGreenNews	Sprint Nextel	39	71	19	1	10	8
TeachingMoney	Capital One Financial	145	27	7	4	3	0
TICalculators	Texas Instruments	1291	1229	260	71	180	9
VerizonGiving	Verizon Communications	1228	1407	348	81	258	9
WalmartAction	Wal-Mart Stores	457	618	129	18	99	12
WalmartGreen	Wal-Mart Stores	375	1171	147	27	110	10

Note: Table shows number of firm messages and audience reactions for each account over the course of 2014.

2.3.2 Dependent Variables: Reputational Awareness and Favorability

The analyses focus on two dependent variables designed to tap, respectively, the awareness and favorability dimensions of reputation.

Awareness

To operationalize the awareness dimension I code the number of retweets received by each company tweet. As argued above, each retweet increases the level of diffusion for that message in making the message visible not only to the company's followers but to the retweeter's followers as well. The number of retweets obtained by each tweet is returned by the Twitter *user_timeline* API that was used to download the tweets.

Favorability

To operationalize the favorability dimension of reputation I code the sentiment in the 5,247 original replies to the 18,722 company tweets. Prior studies have employed automated, or unsupervised, approaches to coding tweet sentiment (Castelló et al., 2015; Colleoni, 2013).[9] This approach is valuable insofar as it is highly reliable and easy to apply rapidly to large datasets. However, the accuracy and validity of such approaches for coding sentiment is typically lower than datasets coded with a mix of human-coded and *supervised* machine-learning techniques (González-Bailón & Paltoglou, 2015; Hopkins & King, 2010). Consequently, I chose to implement the latter, more accurate approach.

Data were coded in stages (for complete details on coding procedures see the appendix). First, a sample of 1,000 replies was selected for crowdsourced manual coding. Crowd-sourced coding – human coding by the online "crowd" on such platforms as *Mechanical*

[9] In that approach a pre-assembled dictionary of scored "positive" and "negative" words is used to analyze each tweet; each tweet receives a summary score based on the number and intensity of positive and negative words it contains.

Turk and *Crowdflower* – provides a rapid and cost-effective method for manually coding data that is becoming increasingly popular in management and accounting research (Grenier et al., 2015; Rennekamp, 2012). These 1,000 Twitter replies – along with detailed instructions and sentiment codes for 55 randomly selected replies – were uploaded to *Crowdflower* (www.crowdflower.com), where each was reply was coded by at least three coders as having either negative, neutral, or positive sentiment. Agreement between the majority crowdsourced score and my own manual codes was 89% with a Cohen's kappa score of 0.817 ($\kappa = 0.817$), indicating a high level of inter-coder agreement (Landis & Koch, 1977).[10]

I then implemented a supervised machine learning technique to code the remainder of the replies. In line with machine learning principles (e.g., Go et al., 2009), I first divided the 1,000 manually coded tweets into *training* (85% of cases) and *test* datasets (15% of cases), then trained the machine learning model by choosing between algorithms and by fine tuning model parameters. The algorithm's performance was assessed using the test data, and the trained model was then used to assign sentiment scores on all uncoded replies. Specifically, after comparing results from alternative algorithms (Naïve Bayes and Decision Tree), a support vector machine (SVM) algorithm was chosen and parameters fine-tuned until the highest level of accuracy was achieved. With the classifier trained, agreement with the test data on the 3-code sentiment variable (-1, 0, +1) was 81.3%. This fits well with expectations based on prior research that has found accuracy in sentiment classification to be around 82% (e.g., Go et al., 2009, achieved 82.2% accuracy). Accuracy is even higher with three binary variables derived from the sentiment variable: 89.0% for *positive*, 80.2% for *neutral*, and 94.5% for *negative*. In a final step, with the classification algorithm trained and tested using data on the 1,000 hand-coded replies, the algorithm was then used to generate sentiment scores for the 4,247 remaining replies.

In short, through the manual coding and supervised machine learning processes, the sentiment in each of the 5,247 replies was coded as either positive, neutral, or negative.

[10] Cohen's *kappa* values above 0.6 are considered "substantial agreement" and those above 0.8 are considered "almost perfect" (Landis & Koch, 1977).

In a final step, using these data, two binary variables were created for the analyses in the firm tweet dataset (n=18,722): *Positive Reply,* with a value of "1" indicating firm tweets that receive at least one positive reply; and *Negative Reply,* with a value of "1" indicating firm tweets that receive a negative reply. These two variables tap, respectively, the public's favorability and disfavorability to the firm's message.

2.3.3 Independent Variable: Communicative Tactics

To code the communication tactics, I first conducted a thematic content analysis of the companies' tweets. Following qualitative methodological tenets outlined by Miles and Huberman (1984) and Strauss & Corbin (1998), I analyzed the data inductively to identify communicative features unique to the CSR context on social media. Coding thus involved a multistage, iterative process of cycling back and forth among data, literature, and emergent conceptual categories (Miles & Huberman, 1984). Specifically, I coded the first 200 of these tweets in several waves, revising and refining at each stage the coding labels used. At each stage the 200 tweets were corroborated with feedback from a colleague familiar with the advocacy and communication literatures to provide feedback and reliability checks on the coding scheme.

The final coding scheme that emerged (to be discussed in depth in the results section) from the analysis of those 200 tweets contained 9 types of communicative tactics. The remainder of the 18,522 tweets were coded in a series of stages. Five of the nine tactics could be coded through a relatively automated process, whereby I developed custom algorithms for coding each of the remaining 18,522 tweets.[11] Each algorithm was refined until coding accuracy was well above 90% compared to the 200 hand-coded tweets.

The remaining four "informational" tactics needed additional hand-coding before accurate

[11]For instance, one of the tactics coded was the *politician mention,* where firms made an effort to thank, congratulate, or interact with politicians. To find mentions of politicians, a list of all Twitter users mentioned in the 18,722 tweets that had usernames starting with "@Gov," "@Sen," "@Mayor," and "@Rep," such as @RepJoeKennedy and @SenGillibrand, was compiled then verified to ensure these were Twitter accounts of politicians. Other politicians that were discovered in firm tweets, such as @NancyPelosi, were also added to the list. In the end a list of 158 verified politicians was created; at that stage automated Python code was written to identify every tweet among the 18,722 firm messages that contained one or more mentions of users from the list of politicians.

machine learning algorithms could be applied. Consequently, 1,500 tweets (the 200 hand-coded tweets plus an additional 1,300 randomly selected) were uploaded to *Crowdflower* for human coding, with each tweet being coded by 3 separate coders.[12] Using the 1,500 hand-coded tweets, a machine learning classifier was then trained and used to code the remaining 17,222 messages. The final classifier was an SVM algorithm that achieved between 81.4% and 97.2% accuracy on the four informational categories.

In brief, all 18,722 firm tweets were coded for the nine communication tactics identified through the inductive analyses. In a final step, nine binary variables were created for use in the multivariate analyses: *Disclosure, Public Education, Marketing, Dialogue, Mobilization, User Mention, Politician Mention, Stewardship Message,* and *Topic Tie.* As will be discussed in the Results section, the nine tactics were also aggregated into three broader categories: *Informational, Interactive,* and *Tie-building.*

2.3.4 Control variables

In estimating the effects of communicative tactics on reputation, I also include a series of account-level and tweet-level control variables shown to be important predictors of social media message sharing in prior literature (Bakshy et al., 2011; Saxton & Waters, 2014; Stefanone et al., 2015). First, I operationalize three variables at the tweet-level: *# of Characters,* which taps the length of the message; *URL included,* a binary variable with values of "1" if the tweet contains one or more external hyperlinks; and *Photo included,* which indicates with a value of "1" for tweets that include an image. Second, I also included three account-level controls: 1) *# of Followers* indicates the number of other Twitter users that follow the organization; 2) *Time on Twitter* indicates the number of days (as of January 1, 2014) since the firm created its Twitter account; and 3) *Broad CSR Focus* is coded a "1" for accounts that focus on a broad spectrum of CSR core areas rather than a specific CSR project or initiative or focus area such as the environment, health, or education (see

[12] The agreement between the majority Crowdflower code and my coding of the 200 tweets was 88.5%, with a Cohen's kappa value of 0.81, indicating an "almost perfect" level of agreement (Landis & Koch, 1977).

Saxton et al., 2016 for a full description of the variable). I posit that, in the social media context, these three account-level factors take the place of traditional "offline" firm-level and industry-level controls. Still to check this assumption, a series of robustness tests are also run to control for firm size and industry.

2.4 Results

2.4.1 Descriptive Analyses

To provide a broad picture of the firm actions and audience reactions, Figure 2.3 shows the daily activity in the number of original firm tweets sent over the course of 2014 along with the number of retweets of and negative and positive replies to those firm tweets; 41 of the 42 accounts are represented (one account did not tweet in 2014). As shown in the figure, firm actions and audience reactions are both dynamic, with notable spikes and dips in activity.

Reputational Awareness: Retweets

Table 2.2 includes summary statistics for all model variables. The mean number of retweets for each message was 2.08 and ranged from 0 to 1,488 (s.d. = 13.49). In total, members of the public sent 38,943 total retweets spread over 10,568 of the 18,722 firm tweets (56.4% of tweets); the other 8,154 (43.6%) of tweets received zero public retweets.

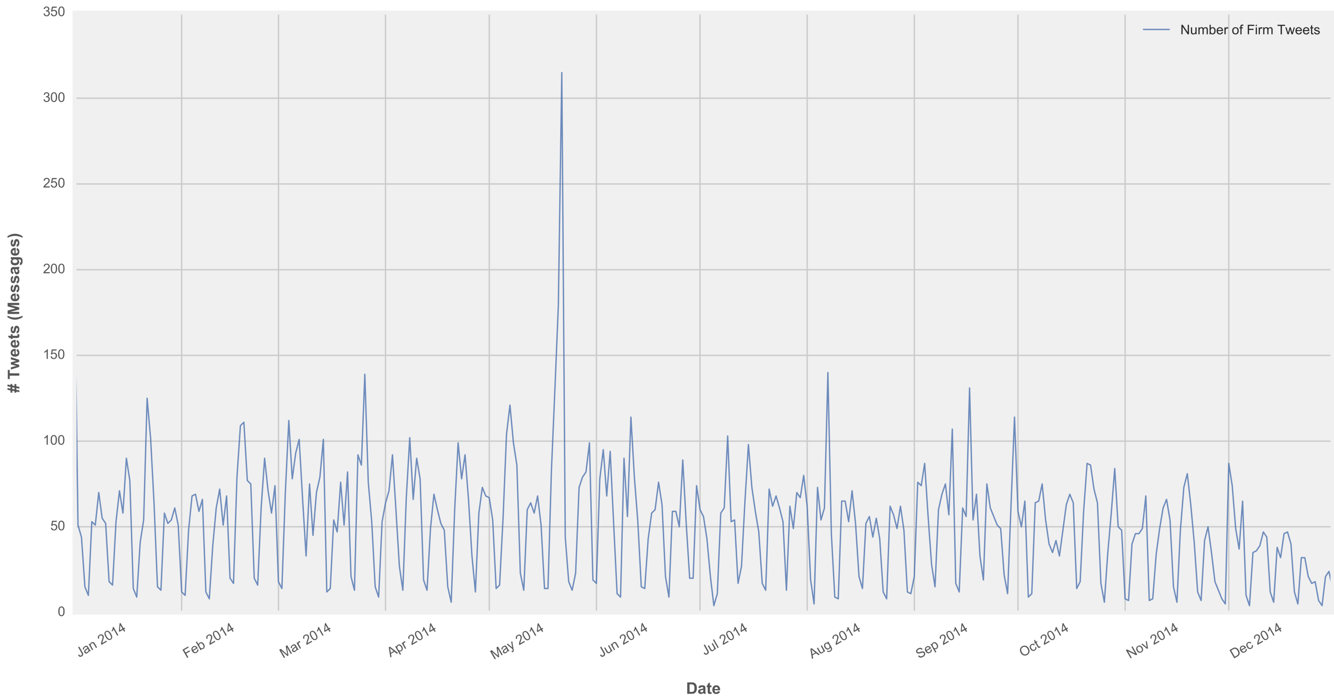
These descriptive data reflect the non-normal distribution of the variable. In fact, retweets reflect a *power law* distribution. The power law distribution (e.g., Shirky, 2003), also called the Pareto distribution, has the general shape y = 1/x.[13] In general, such distributions are noted by the small number of "winners" and the large number of "losers." Such is the case here: Only 49 tweets received 51 or more retweets; the 18,673 remaining

[13]Because of how frequently it occurs in new media consumption – the number of visitors received by websites, online book sales, etc. – Chris Anderson (2006) has referred to the power law as "the shape of our age." Power laws tend to arise under conditions of variety in content, inequality of content quality, and where network effects are present, which suppress the bad and promote the good (Anderson, 2006).

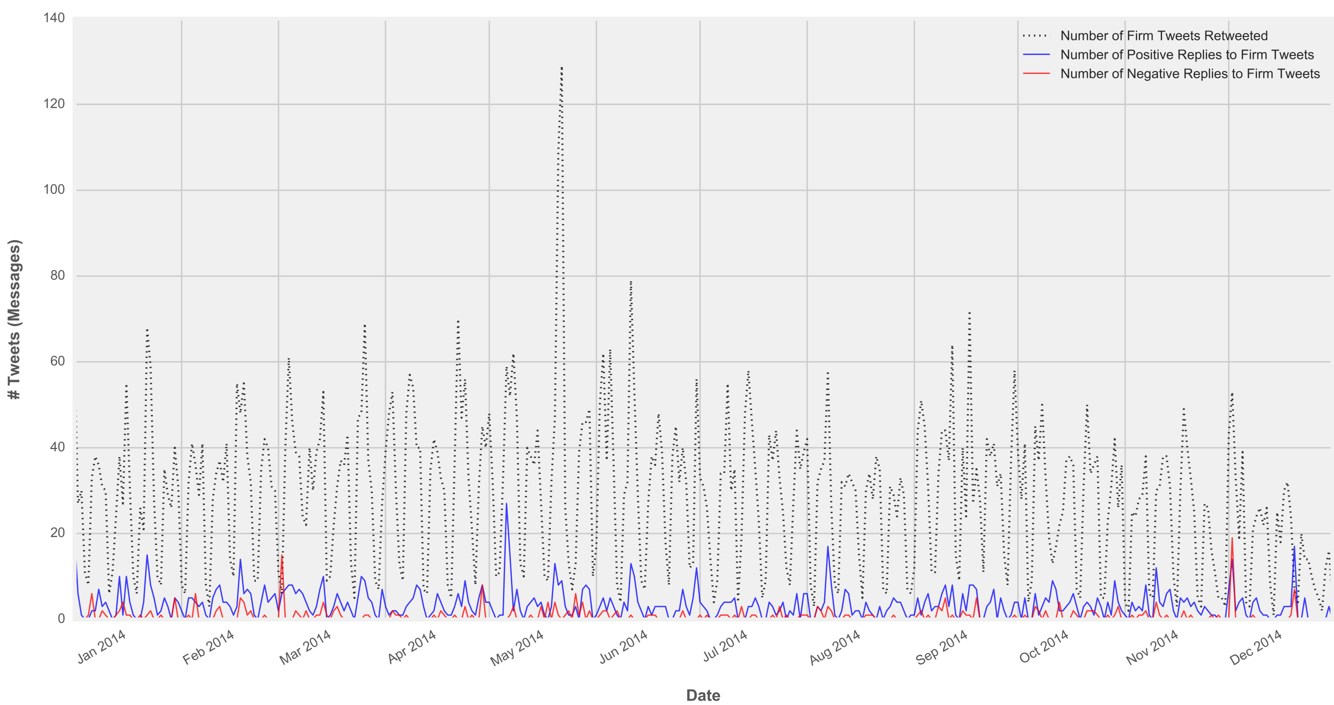
tweets received between 0 and 50 RTs. The former are "winners" in terms of social media consumption, while the latter are moderate successes as well as "losers." The most heavily retweeted message received 1,488 retweets; after this the number of retweets received drops considerably, with the second most-heavily retweeted message receiving 706 retweets (in a power law the biggest drop-off occurs, by definition, from the first- to the second-place position). The third most-heavily retweeted message received 323 retweets and the fourth 259 retweets. In effect, there is a steep drop-off in how many times a message is consumed and forwarded by audience members that is non-Normal.

Figure 2.4 shows the distribution of the number of retweets received by the 18,673 tweets that received 50 or fewer retweets (for visual clarity, the 49 most-heavily retweeted messages are omitted from the histogram). As seen in Figure 2.4, 8,154 of the tweets (43.55%) received zero retweets, while 4,266 received a single retweet (making 66.34% that receive ≤ 1 retweet). Relatively few (7.6% of tweets) receive more than 5 retweets and only 2.5% receive more than 10. The dotted line in Figure 2.4 shows the mean value of 2.08 retweets and a kernel density line indicates the basic power law shape of the distribution.

In contrast to the normal distribution, in a power law distribution a large number of messages will have low levels of reputational capital, while a few messages will generate very high levels of reputational capital. In fact, as can be seen in Figure 2.4 , the mean value of 2.08 retweets (dashed line) conveys little information in such a non-normally distributed, "rich get richer" context. Notably, 14,756 of the 18,722 tweets have fewer than the mean value; these "low awareness" tweets have an average of 0.61 retweets. By contrast, the 3,966 "high awareness" tweets (those with values higher than the mean) have an average of 7.57 retweets. While the number of retweets is but one measure of acquired reputational capital, the contrast between the low- and high-awareness messages is notable.



(a) Daily Frequency of Firm CSR Messages, 2014 (n=18,722)



(b) Daily Frequency of Various Public Reactions to 18,722 Firm Tweets

Figure 2.3: Daily Frequency of Firm CSR Tweets and Public Reactions, 2014

Note: Figure (a) shows 18,722 original tweets sent in 2014 by 42 CSR-related Fortune 200 Twitter accounts, while Figure (b) shows the public retweeting of and replies to those 18,722 firm messages.

Table 2.2: Summary Statistics

Variable	Count	Mean	Std Dev	Min.	Max
Independent Variables					
INFORMATION	18,722	0.86	0.34	0	1
Disclosure	18,722	0.59	0.49	0	1
Public Education	18,722	0.28	0.45	0	1
Marketing	18,722	.004	0.06	0	1
No information	18,722	0.14	0.34	0	1
INTERACTION	18,722	0.21	0.41	0	1
Dialogue	18,722	0.21	0.4	0	1
Mobilization	18,722	0.01	0.09	0	1
TIE-BUILDING	18,722	0.85	0.36	0	1
User mention	18,722	0.45	0.5	0	1
Stewardship message	18,722	0.11	0.31	0	1
Politician mention	18,722	0.01	0.12	0	1
Topic ties (hashtag included)	18,722	0.7	0.46	0	1
Dependent Variables					
REPUTATIONAL AWARENESS					
Message is retweeted (0,1)	18,722	0.56	0.5	0	1
REPUTATIONAL FAVORABILITY					
Message receives positive reply (0,1)	18,722	0.06	0.24	0	1
Message receives negative reply (0,1)	18,722	0.01	0.1	0	1
Controls					
ACCOUNT-LEVEL					
Time on Twitter in days (as of 1/1/2014)	18,722	1407.81	410.16	-23	2,083
Account has broad CSR focus	18,722	0.64	0.48	0	1
# of Followers	18,722	54,134.67	118,417.81	0	553,316
MESSAGE-LEVEL					
$\geq$ 1 URL included	18,722	0.69	0.46	0	1
# Characters	18,722	115.46	25.7	4	192
Photo included	18,722	0.16	0.37	0	1

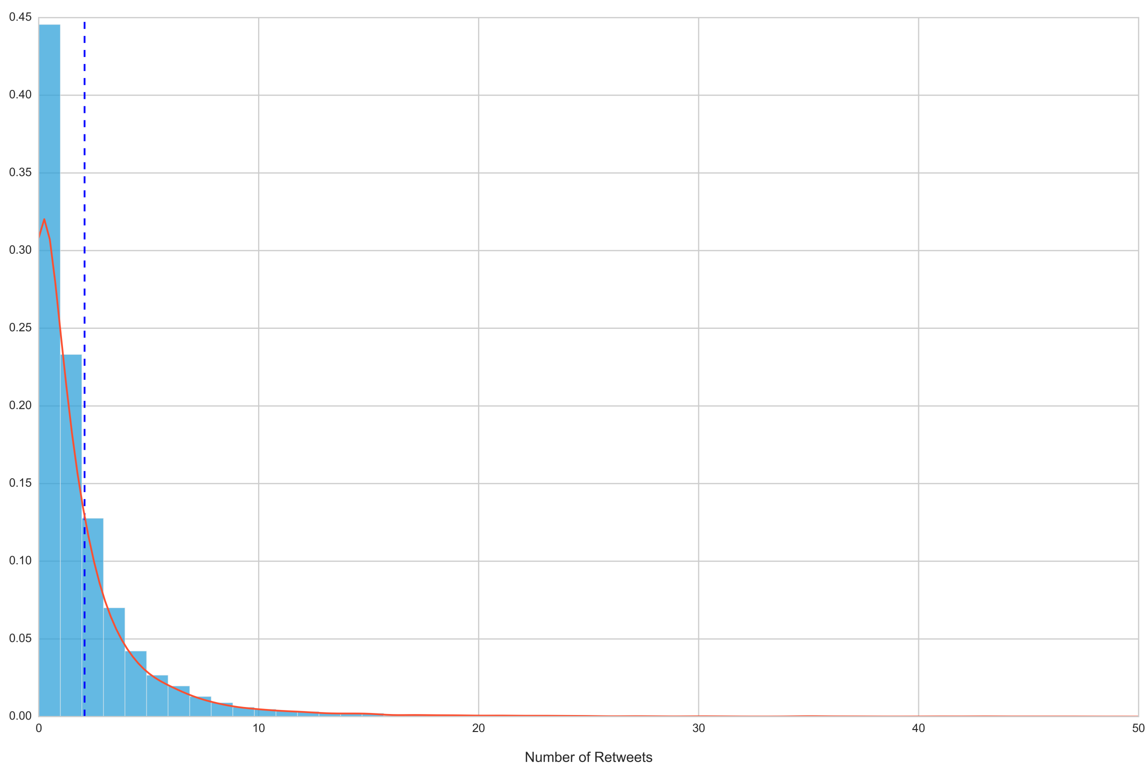


Figure 2.4: Histogram, Number of Public Retweets of Firms' Messages on Twitter

Note: Histogram of number of times each firm tweet was retweeted (shared) by other Twitter users, based on all original messages sent by the 42 CSR accounts over 2014 (n = 18,722). Dashed vertical line shows mean value of 2.08 retweets. Solid red line indicates kernel density estimation line.

Reputational Favorability: Sentiment in Replies to Firm Tweet

Members of the public sent 5,247 total replies to the 18,722 firm tweets sent in 2014. As noted above, I coded the negative, neutral, or positive sentiment conveyed in these 5,247 public replies. For instance, the following reply to Bank of America's CSR account @BofA_Community reflects strong *negative* affect toward the company:

> *@BofA_Community @CAFoodBanks @FeedingAmerica @UnitedWaysCA Trying to improve your crappy image?*

An example of a *neutral* message below occurred in a "tweetchat" held by Verizon's @VerizonGiving account; it reflects a factual, descriptive message with neither negative nor positive affect:

> *@verizongiving @CoRaft in person is critical. teachers teach children face-to-face. #disrupttheclassroom*

Finally, an example of a *positive* message is sent from a nonprofit organization thanking Cisco for its charitable funding efforts:

> *@CiscoCSR @cmtysolutions @60Minutes Thanks for funding our work to end homelessness! You guys rock!*

Figure 2.5 shows the frequency of the sentiment types in the 5,247 replies. In total there were 289 negative replies, 3,680 neutral replies, and 1,278 positive replies.[14]

As with retweets, the number of positive and negative replies approximates a power law distribution. We are interested particularly in the positive and negative replies (rather than neutral) as these positively and negatively reflect favorability to the firm's message. The mean number of negative replies per message is 0.015 with a range from 0 to 18 (s.d. = 0.24). The mean number of positive replies, meanwhile, is 0.068 with a range from 0 to 17 (s.d. = 0.33). For the binary variables used in the analyses and shown in Table 2.2, *Negative*

[14] Because some tweets received more than one reply, overall, 3,202 different firm tweets (about one in six) received a total of 5,247 replies.

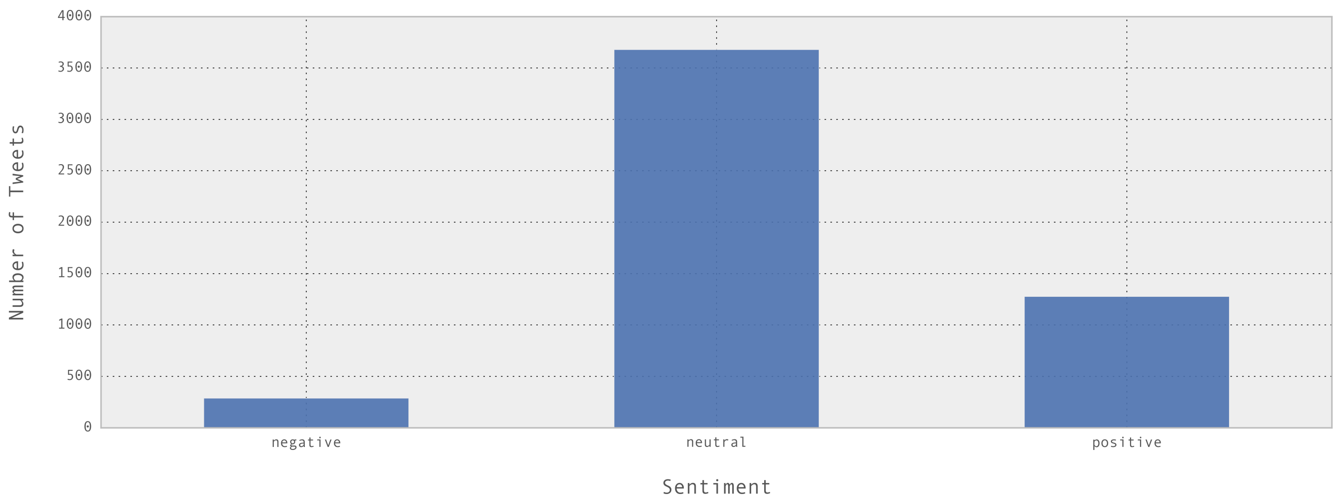


Figure 2.5: Negative, Neutral, and Positive Sentiment in Replies to Firms' CSR Messages

Note: Sentiment in the 5,247 original replies by members of the public to 18,722 tweets sent in 2014 by 42 CSR-focused Fortune 200 Twitter accounts.

Reply has a mean value of 0.01 and a standard deviation of 0.1, while *Positive Reply* has a mean of 0.06 and a standard deviation of 0.24.

Independent Variables: Communication Tactics

Figure 2.6 shows the frequency of the communicative tactics in the 18,722 firm tweets. Inductive analyses of the data led to the discovery of nine communication tactics grouped into three broader categories. First, each tweet was characterized as having one of three *informational* tactics: disclosure, public education, or marketing. Unlike the other six tactics, these three tactics were considered to be mutually exclusive and thus each tweet received a score on only one of these three variables, or else were coded as "non-informational," reflecting that the message did not convey information.[15]

First, the majority (58.6%) of tweets (n=10,967) were coded as *Disclosure*; such messages report information on the firm's CSR activities. For example, *@CiscoCSR* sent the following message noting its charitable contributions to the 100,000 Homes Campaign:

[15] In other words, each message could be coded as containing multiple tactics, but at most one informational tactic.

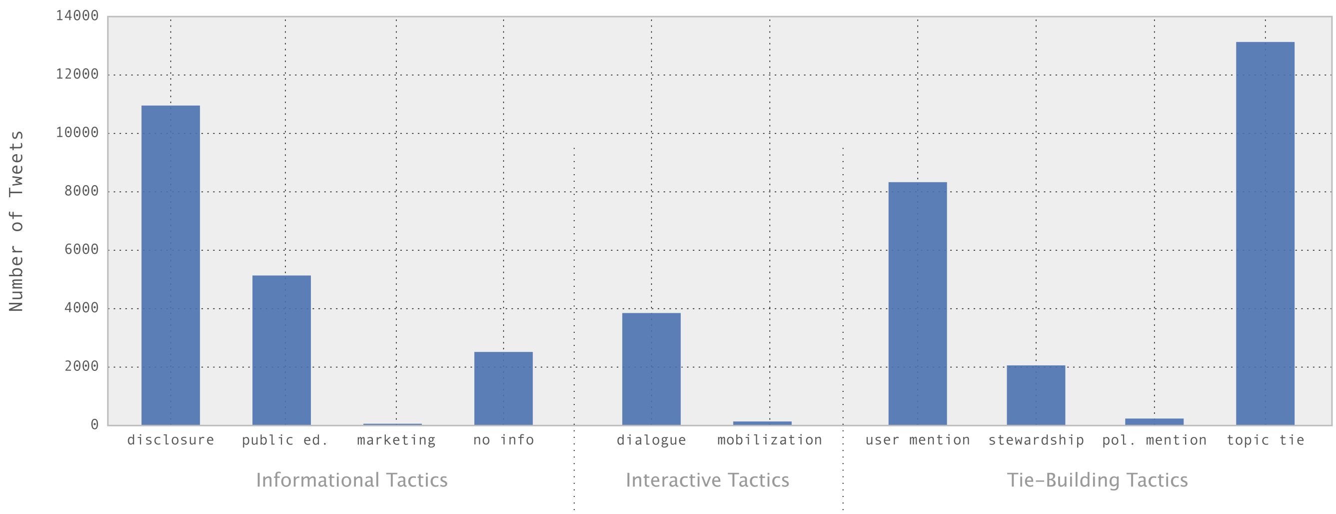


Figure 2.6: Frequency of Communicative Tactics in Firms' CSR Messages

Note: Data from coding of 18,722 original tweets sent in 2014 by 42 CSR-related Fortune 200 Twitter accounts.

> *We're proud to have supported @100khomes campaign to house 100k #homeless Americans. Congrats! http://t.co/sYrx0hegbK*

The second major category of informational tactics, covering 27.5% (n=5,149) of tweets, is what I call the *Public Education* tactic. As with disclosure tweets, public education tweets are intended to convey information; however, unlike disclosure tweets, public education messages are intended not to report on the company's activities but rather to educate the public on a topic related to a CSR core area – such as technology, health, education, sustainability, diversity, or the environment. A good example is the following tweet from *@Microsoft_ Green*:

> *From @virginia_tech: Sugar could help power #smartphones in the future via @guardian http://t.co/AbeW20Rzs8 #cleanenergy*

There is nothing in the tweet nor in the link connected to Microsoft or any of its activities. Instead, the message is intended to convey inform on environmental sustainability that Microsoft believes will be of interest to its followers.

Public education and disclosure thus constitute two distinct *one-way informational* tactics. Anything that is informational but reports on the firm's CSR activities is classified as *disclosure*; these are one-way informational messages that convey information on the company's CSR activities. *Public Education* messages, in contrast, contain information on CSR-related topics that does not relate to the firm's activities. In effect, anything that is informational but not reporting on the firm's CSR activities is classified as public education rather than disclosure. Inspirational messages and/or quotes were also considered public education.

This heavy use of the public education tactic contrasts with, for example, Colleoni (2013), who broke CSR social media communication down into "self-centered" (equivalent to disclosure) and "dialogic" categories. However, I find one-way communication need not be "self-centered;" one-way information can also serve to inform the public. Public education tactics may be a new way of building a positive reputation, and relate to studies in different contexts that audiences may be built on social media by adopting an "expertise" role in a niche area (Rheingold, 2012; Suddaby et al., 2015). In successful cases, the expertise role leads to a larger network of followers and a better network position within a strategically targeted community.

The third informational tactic is *marketing*. Such tweets do not really relate to CSR (so are somewhat out of place in the 42 CSR-focused Twitter accounts studied); instead, they are designed to sell products or services. These tweets do not convey information on the company's CSR-related activities, nor do they present information on CSR-related topics that could inform the public. Marketing messages were rare – occurring in only 74 of the 18,722 tweets. A good example is the following tweet by PPG Ideascapes marketing its "green" products:

> *For a wide array of projects, there's a PPG cool #metalcoating to protect and enhance the appearance. http://t.co/1dTl1oP8Jq*

The final informational category was "no information." Such tweets do not present information. Instead, they may ask a question, ask readers to do something, be a simple response to another tweet, or just thank, mention, or congratulate another Twitter user. Most are coded for one of the interactive or tie-building tactics outlined below. For example, *@MSFTCitizenship* tweeted this message:

> *@kanter No problem. It was great to see you at #14LCS, sorry I didn't stop by to say hi! -@susipye*

In addition to the three binary variables *Disclosure, Public Education,* and *Marketing,* for additional tests I also created a variable *Information* to indicate tweets that contain either of these three tactics. For the aggregate variable *Information*, 86.48% (n=16,190) of tweets received a score of "1," with the remaining 13.52% (n=2,532) coded as "0," or non-informational.

The second broad category of tactics are what I call *Interaction* tactics, comprising 21.4% (n=4,007) of tweets. There are two interactive tactics, encompassing *Dialogue* and *Mobilization.* Tweets coded as *Dialogue* contain a conversational tactic – typically asking a question, responding to a public query, or engaging in a "tweetchat." An example of a broadly targeted call for dialogue is the following tweet from *@IBMSmartCities*:

> *Pls join the #P4SPchat on citizen engagement. Starts in 45 minutes, Noon ET http://t.co/GobUBqy3Qe*

In contrast, the following tweet by *@GEHealthy* is a question targeted at one specific user:

> *@sarsbradley Excellent! Was it fun? :)*

In both examples, the primary feature is the tweet reflects engagement in dialogic, conversational, two-way communication (Saxton & Waters, 2014; Taylor & Kent, 2014) between firm and public. I found a relatively large proportion of tweets, 20.6% (n=3,864), to be dialogic.

Like dialogic tweets, *Mobilization* tweets aim to go beyond one-way information. Yet instead of seeking to engage members of the public in dialogue – or, in effect, to *say* something – mobilizational messages aim to mobilize audience members to *do* something. The most common mobilizational messages are those asking audience members to vote (e.g., in a contest) or to retweet a message.

> `Retweet if you're taking a stand for LGBT youth by celebrating #SpiritDay #ComcastGoesPurple http://t.co/pLzKSdn13n`

While such messages often receive a strong audience response, they are not common. Only 0.81% (n=152) of all tweets were mobilizational.

The third major category of tweets contain what I call "tie-building" tactics. Comprising 52.8% of all tweets (n=9,885), *Tie-Building* messages aim to build or reinforce a connection to another user or group of users. I found four specific tie-building tactics. First there is the *User Mention,* a common tactic found in 44.6% (n=8,344) of all messages. In these messages, the firm employs the "user mention" feature of Twitter to "talk about" another Twitter user. For example, in the following message *@ComCastDreamBig* (since renamed *@ComcastImpact*) includes user mentions of two nonprofit organizations, City Year (an AmeriCorps initiative) and Big Brothers Big Sisters:

> `An inside look at our partnerships with @BBBSA @CityYear & more http://t.co/cHCYWwNXIP #InsideCI`

In so doing, the firm (via *@ComcastImpact*) is deepening its online ties with the two users (*@BBBSA*, *@CityYear*) on Twitter. This also provides the opportunity for the firm to increase awareness by reaching users who follow *@BBBSA* and *@CityYear* but not *@ComcastImpact*; as such, it is a method for increasing its influence (Lovejoy et al., 2012).[16]

The second tie-building tactic is the "stewardship message." Stewardship is a relationship cultivation approach identified in the public relations literature (Waters, 2011). One key stewardship strategy is *reciprocity*, which involves recognizing or thanking an individual user or group of users (Kelly, 2000). The idea is that publicly recognizing, thanking, or

[16]Another important point about the above example is that it contains two tactics: *Disclosure* and *User Mention.* Not only does the tweet report on the company's charitable partnerships, which is a form of disclosure of CSR activities, but the link takes readers to the company's "Inside CI" page, which relates the company's community investment activities. As noted earlier, while the three informational tactics (disclosure, public education, and marketing) are mutually exclusive, none of the interaction or tie-building tactics are – multiple tactics could occur in any one message.

expressing gratitude to stakeholders can build loyalty and strengthen the relationship, or the tie between two users (Waters, 2011). Accordingly, I coded messages as *Stewardship* when they congratulated or expressed gratitude to a user or group of users. The following example, sent by @AlcoaFoundation, is targeting gratitude at a specific user:

> *@TheaClay That is so great to hear! Good luck today and thank you for your support w/ #makeanimpact*

Stewardship messages comprised 11.1% (n=2,052) of all tweets, and are closely related to the *User Mention.* Both seek to deepen or maintain ties with specific members of the audience.

A third tie-building tactic, the *Politician Mention,* is a specialized form of user mention. In this case, the tweet contains a mention of a specific politician.

> *The #Walmart Foundation donated 160k to Bay area #CA nonprofits to aid education & nutrition: http://t.co/VKwxOiA0G5 @NancyPelosi @RepSpeier*

This tactic was included in 1.4% (n= 252) of the 18,722 tweets. This tactic reflects the only common form of "political" message the firms made on Twitter. I did find a small number of policy advocacy messages such as the following by @WalmartAction, which supports a trade agreement among 12 Pacific Rim countries:

> *#Walmart already exports U.S. cheese, juice, chocolate & #CA wine. The Trans-Pacific Partnership would open up even more opportunities! #TPP*

Instead, these firms appear to be taking an "indirect" approach to policy advocacy; the goal is to build a constituency of supporters rather than to directly lobby on relevant political issues. The *Politician Mention* serves to help both the politician increase his or her visibility while also explicitly pointing out to the politician the good deeds the company is doing.

The final, and most common, tactic, is the *Topic Tie.* It is found in 70.2% (n=13,143) of tweets. The manifestation of links to topics are *hashtags.* Hashtags, those brief words or phrases prepended by the hash sign (#), such as *#CSR, #DoMoreEdu, #BigData, #In-*

telForChange, and *#GirlRising*, were first used in 2007 on Twitter and have since proliferated to other social media platforms. Hashtags make a message topic or theme explicit and serve as a mechanism for archiving, branding, and classifying a message while enabling the message to connect to a broader conversation or collection of messages. The use of hashtags is powerful in large part because it is participatory; the set of tags used constitutes a decentralized, user-generated organization and classification system known as a *folksonomy* (Debreceny, 2015). Not only does the hashtag classify messages and improve searchability (O'Leary, 2015), but it allows the organization to connect messages to existing networks of *ad hoc publics* (Bruns & Burgess, 2011) that form around the hashtag. Saxton et al. (2016) found that, by tapping into existing discussions and networks, tweets that employ CSR-related hashtags in particular generated higher levels of resonance with the public. In brief, including hashtags is a tactic designed to broaden the reach and resonance of the message by connecting it to existing conversations and informal groups. A good example is the following tweet from *@CiscoCSR*; it includes three hashtags, allowing the message to be linked to anyone in the Twitterverse engaged in discussions around *#Impact, #CSR,* and the *#InternetOfEvertything*:

> *Discover how @Cisco uses the #InternetOfEverything to connect and #Impact lives around the world: http://t.co/ry3Jn3x7LI #CSR*

2.4.2 Control Variables

Finally, turning to the control variables, at the account level, 64% of the tweets were from accounts that had a *Broad CSR Focus,* with 36% were from accounts focused more narrowly on a specific CSR area such as the environment. The average *# of Followers,* or the number of users following the firm, was 54,134.7 (s.d. = 118417.81), though this ranged widely from a low of 0 to a high of 553,316. And the average *Time on Twitter in Days* at the start of the year was 1407.81 (s.d. = 410.16) and ranged from a low of -23 days (for one account that was created on January 23, 2014) to a high of 2,083 days. At the tweet level, we found

69% of tweets included one or more *URL,* and 16% contained a *Photo.* The average *# of Characters* was 115.5, somewhat below the limit of 140 characters (120 characters if a photo is included).

2.4.3 Multivariate Analyses

Table 2.3 presents results from a series of six logistic regressions (Tables 2.4 and 2.5 contain variable definitions and a zero-order correlation matrix, respectively). The three binary dependent variables, *Retweeted*, *Positive Reply,* and *Negative Reply,* are each regressed first on the three-category set of independent variables (*Information, Interaction,* and *Tie-Building*) and then on the larger set of nine discrete tactics, for a total of six models. Each model contains the same set of six control variables described above. In all models the baseline category – the category against which regression parameters are compared – are the 2,532 "non-informational" messages, those coded as not containing either a disclosure, public education, or marketing tactic.

Before covering the effect size of the model variables I cover sign and significance. I start with a discussion of the regressions on *Retweeted,* the proxy for reputational awareness. Beginning with the informational tactics, *Information* receives a positive and significant coefficient in Model 1, as do *Disclosure* and *Public Education* in Model 2 (all $p < .01$). This can be interpreted as indicating that, compared to the baseline category of non-information-focused tweets, messages that include an informational tactic are more likely to be retweeted by members of the public. Only the non-CSR-related *Marketing* tactic fails to obtain significance.

The opposite occurs with the interaction variables. The aggregate variable *Interaction* is negatively associated ($p < .01$) with *Retweeted,* as is the constituent variable *Dialogue. Mobilization,* meanwhile is not significantly related to the receipt of audience retweets (p=0.12).

Tie-building tactics, in turn, have a mixed effect. Overall, the effect is positive, with the aggregate variable *Tie-Building* achieving a significant and positive relationship with

Table 2.3: Logit Regressions

	Awareness		*Favorability*			
	RT (0,1)		Positive reply (0,1)		Negative reply (0,1)	
INFORMATION	0.75**		0.03		-0.04	
	(0.06)		(0.10)		(0.26)	
Disclosure		0.37**		0.002		0.004
		(0.06)		(0.10)		(0.28)
Public education		0.40**		-0.17		-0.28
,		(0.07)		(0.12)		(0.31)
Marketing		0.12		0.34		-
		(0.25)		(0.53)		-
INTERACTION	-1.63**		0.34**		-0.95**	
	(0.05)		(0.09)		(0.26)	
Dialogue		-1.47**		0.50**		-0.87**
		(0.06)		(0.10)		(0.30)
Mobilization		0.28		0.5		-0.29
		(0.18)		(0.31)		(1.02)
TIE-BUILDING	0.32**		-0.08		-0.31	
	(0.05)		(0.09)		(0.25)	
User mention		0.15**		0.38**		-0.03
		(0.04)		(0.07)		(0.17)
Stewardship message		-0.89**		-0.16		-0.49†
		(0.07)		(0.11)		(0.29)
Politician mention		-0.17		-0.58		0.99†
		(0.14)		(0.39)		(0.53)
Topic Ties		0.47**		-0.05		-0.37*
		(0.04)		(0.08)		(0.18)
ACCOUNT-LEVEL CONTROLS						
Broad CSR Account	0.43**	0.45**	0.52**	0.47**	0.75**	0.70**
	(0.03)	(0.04)	(0.08)	(0.08)	(0.23)	(0.23)
Followers (1,000)	0.004**	0.005**	0.002**	0.002**	0.005**	0.006**
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0005)	(0.0005)
Time on Twitter (in 100 days)	-0.01*	-0.01**	-0.02*	-0.02*	-0.08**	-0.08**
	(0.004)	(0.004)	(0.008)	(0.008)	(0.02)	(0.02)
MESSAGE-LEVEL CONTROLS						
$\geq$ 1 URL included	-0.15**	-0.09*	-0.76**	-0.71**	-0.45**	-0.49**
	(0.04)	(0.04)	(0.07)	(0.07)	(0.18)	(0.18)
# Characters	0.001	0.002*	0.002	0.001	0.005	0.006
	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)
Photo included	0.62**	0.63**	0.42**	0.44**	0.75**	0.75**
	(0.05)	(0.05)	(0.08)	(0.08)	(0.18)	(0.18)
Intercept	-0.75**	-0.63**	-2.84**	-2.95**	-4.67**	-4.62**
	(0.11)	(0.11)	(0.20)	(0.21)	(0.56)	(0.56)
Model Sig. (χ^2)	3256**	3727.3**	389.12**	427.56**	262.72**	273.62**
Pseudo-R^2 (McFadden)	0.13	0.15	0.05	0.05	0.13	0.13
Log likelihood:	-11,193	-10,957	-4,015.2	-3996	-897.66	-892.21
N	18,722	18,722	18,722	18,722	18,722	18,722

†p <0.10, *p <0.05, **p <0.01; standard errors in parentheses.

Table 2.4: Variable Definitions

Variable	Description
Independent Variables	
INFORMATION	Tweet coded as containing either disclosure, public education, or marketing tactic (0,1)
Disclosure	Tweet discloses information on firm's CSR activities (0,1)
Public Education	Tweet contains information intended to educate public on CSR-related issue (0,1)
Marketing	Tweet contains information marketing the company's products or services (0,1)
No information	Tweet contains no information (0,1). Baseline (omitted) category
INTERACTION	Tweet contains dialogue and/or a mobilization tactic (0,1)
Dialogue	Tweet contains a dialogic tactic – asking a question, responding to public query, engaging in tweetchat (0,1)
Mobilization	Tweet contains a mobilizational tactic – asking audience members to, for instance, vote or retweet message (0,1)
TIE-BUILDING	Tweet coded as containing a user mention, stewardship message, political mention, and/or topic tie (0,1)
User mention	Tweet contains a mention of another user (0,1)
Stewardship message	Tweet conveys a message thanking or congratulating another Twitter user (0,1)
Politician mention	Tweet contains a user mention of a specific politician (0,1)
Topic ties	Tweet contains hashtags linking the tweet to a broader topic or conversation (0,1)
Dependent Variables	
REPUTATIONAL AWARENESS	
Message is retweeted	The firm's message is retweeted (shared) by one or more members of the public (0,1)
REPUTATIONAL FAVORABILITY	
Message receives positive reply	Tweet receives a public reply with positive sentiment (0,1)
Message receives negative reply	Tweet receives a public reply with negative sentiment (0,1)
Controls	
ACCOUNT-LEVEL	
Time on Twitter in days	# of days firm has managed Twitter account (as of 1/1/2014)
Account has broad CSR focus	Twitter account focuses on broad CSR issues rather than a narrow or specific CSR topic area (0,1)
# of Followers	Number of followers of the firm's Twitter account
MESSAGE-LEVEL	
$\geq$ 1 URL included	Tweet contains hyperlink (0,1)
# Characters	Number of characters in the message
Photo included	Tweet contains photo (0,1)

Table 2.5: Zero-order correlations

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.
1. Pos. reply	1																				
2. Retweet	.05**	1																			
3. Neg. reply	008**	.06**	1																		
4. Information	-.04**	.22**	-.001	1																	
5. Interaction	.06**	-.32**	-.02*	-.53**	1																
6. Tie-building	.02*	.01	.002	-.02*	-.05**	1															
7. Disclosure	.002	.10**	.014	.47**	-.21**	.07**	1														
8. Public ed.	-.04**	.06**	-.02+	.24**	-.17**	-.09**	-.73**	1													
9. Marketing	-.001	-.01	-.01	.02**	-.02+	-.01	-.07**	-.04**	1												
10. Dialogue	.06**	-.33**	-.02+	-.54**	.98**	-.05**	-.22**	-.17**	-.02+	1											
11. Mobilization	.008	.02+	-.003	.02*	.17**	-.02*	.03**	-.01	-.01	-.03**	1										
12. User Mention	.02*	.15**	.01	.22**	-.33**	.85**	.15**	.01	.0001	-.34**	-.01	1									
13. Stewardship	.03**	-.20**	-.002	-.42**	.43**	.33**	-.12**	-.19**	-.02+	.44**	-.02+	-.13**	1								
14. Pol. mention	-.02+	.01	.01	.03**	-.06**	-.04**	.07**	-.06**	.0001	-.05**	-.01	-.03**	-.02*	1							
15. Topic ties	-.004	.22**	.002	.28**	-.31**	.05**	.12**	.08**	-.01	-.32**	.03**	.20**	-.24**	.05**	1						
16. # Followers	.08**	.11**	.12**	-.20**	.15**	-.004	-.07**	-.07**	-.02+	.15**	-.02+	-.13**	.21**	-.04**	-.05**	1					
17. Days on Twitter	-.01	.01	-.001	.03**	.01	-.04**	.004	.016+	.02+	.01	.003	-.04**	-.001	.06**	.07**	.15**	1				
18. Broad Account	.06**	.11**	.04**	-.06**	.05**	.15**	.11**	-.16**	-.04**	.06**	-.02*	.09**	.13**	.08**	.10**	.20**	.07**	1			
19. # Characters	-.02+	.15**	.02**	.26**	-.31**	.21**	.17**	.01	.02*	-.31**	.002	.23**	-.003	.06**	.23**	.03**	-.01	.12**	1		
20. URL	-.11**	.07**	-.03**	.21**	-.30**	.03**	.09**	.06**	.02+	-.31**	-.002	.05**	-.07**	.04**	.02*	.01	.01	.06**	.29*	1	
21. Photo	.07**	.14**	.07**	.08**	-.09**	-.03**	.03**	.03**	-.02+	-.10**	-.001	.002	-.04**	-.01	.10**	.07**	.03**	-.01	.07*	-.29*	1

$^{+}$p <0.05, *p <0.01, **p <0.001

Retweeted ($p < .01$). Two of the constituent tie-building tactics, *User Mention* and *Topic Ties,* are similarly positively associated with retweeting ($p < .01$); however, the other two tie-building tactics, *Stewardship Message* and *Politician Mention,* have a negative relationship, with the latter failing to achieve significance (p=0.21).

Turning to the regressions for reputational favorability, *Positive Reply* and *Negative Reply,* we see a different pattern of relationships. None of the informational variables, neither the composite variable *Information* nor the constituent tactics *Disclosure, Public Education,* or *Marketing,* obtain a significant relationship at the p<.05 level with either of the two dependent variables. The non-significant, and small, coefficient for *Disclosure* is especially surprising given its general importance in the CSR literature (e.g., Richardson & Welker, 2001).

The interaction tactics, however, do obtain significance, yet here the signs are flipped. The composite variable *Interaction* is positively associated with *Positive Reply* and negatively associated with *Negative Reply*, as is *Dialogue* (all at $p<.01$ level). *Mobilization* fails to obtain significance at the $p<0.1$ level in either model. In short, the results are consistent: dialogic tactics are positively associated with *Positive Reply,* the proxy for reputational favorability, and negatively associated with *Negative Reply,* the proxy for reputational disfavorability.

The tie-building tactics also obtained an interesting mix of results. Overall, *Tie-Building* did not obtain significance with either *Positive Reply* or *Negative Reply* at the $p<.01$ level. As with the *Retweeted* regression, *User Mention* is positively and significantly related to *Positive Reply*; none of the other three tie-building tactics achieve a significant relationship with positive replies. However, three discrete tie-building tactics are significantly related to *Negative Reply. Stewardship Message* and (p=.093) and *Topic Ties* (p<.05) have a negative relationship, indicating such tactics are less likely to result in a disfavorable, negative reply. *Politican Mentions,* in contrast, are positively related (p=.063) to negative replies. This was the only tactic with a significant positive association with negative replies. Plausibly, audience members are reacting negatively to the more overtly political messages. For instance, @WalmartAction sent the following tweet mentioning two politicians with a link to a supportive letter:

> *LETTER: "In #Walmart, we found a neighbor eager for a productive relationship." http://walmarturl.com/KmC6uR @repgregwalden @RepSchrader*

A user then sent the following negative reply:

> *.@WalmartAction @WalmartHub @repgregwalden @RepSchrader Way 2 set a high bar! "Several neighbors were hired @ a pay scale abov the min wage"*

Similarly, when @WalmartAction sent the following tweet

> *A new Walmart distribution center in Union City, #GA, will create ~400 jobs over three years: http://trib.al/uyNfixL @repdavidscott*

the following negative reply was sent:

> *@WalmartAction @repdavidscott Has anybody got back to you, yet, on how many jobs it will destroy?*

Control Variables

The results for the controls are consistent across the five models. Save for one model in which *# of Characters* obtains a significant positive relationship with retweeting, in all other models the results are the same in terms of sign and significance: retweets and positive as well as negative replies are associated with messages that do not include a link (*URL*), that include a *Photo*, and which are sent by accounts that are newer, that have a broad (rather than narrow) CSR focus, and which a higher *# of Followers* ($p < .05$ in all cases).

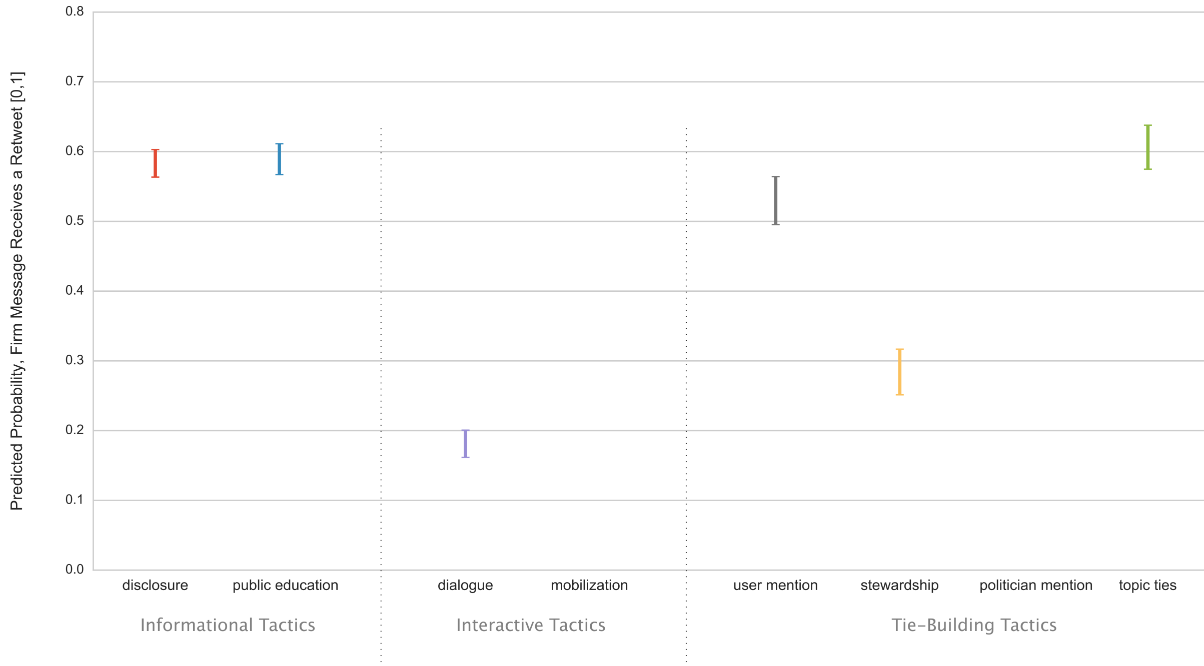
Predicted Probabilities

In sum, many of the types of communicative tactics employed in companies' CSR messages are significantly related to the measures of reputational awareness and favorability. The description above, however, focuses on a discussion of sign and significance. To summarize the above findings in a more meaningful and intuitive fashion while adding a discussion of effect size, I present predicted probabilities in Figures 2.7 and 2.8.

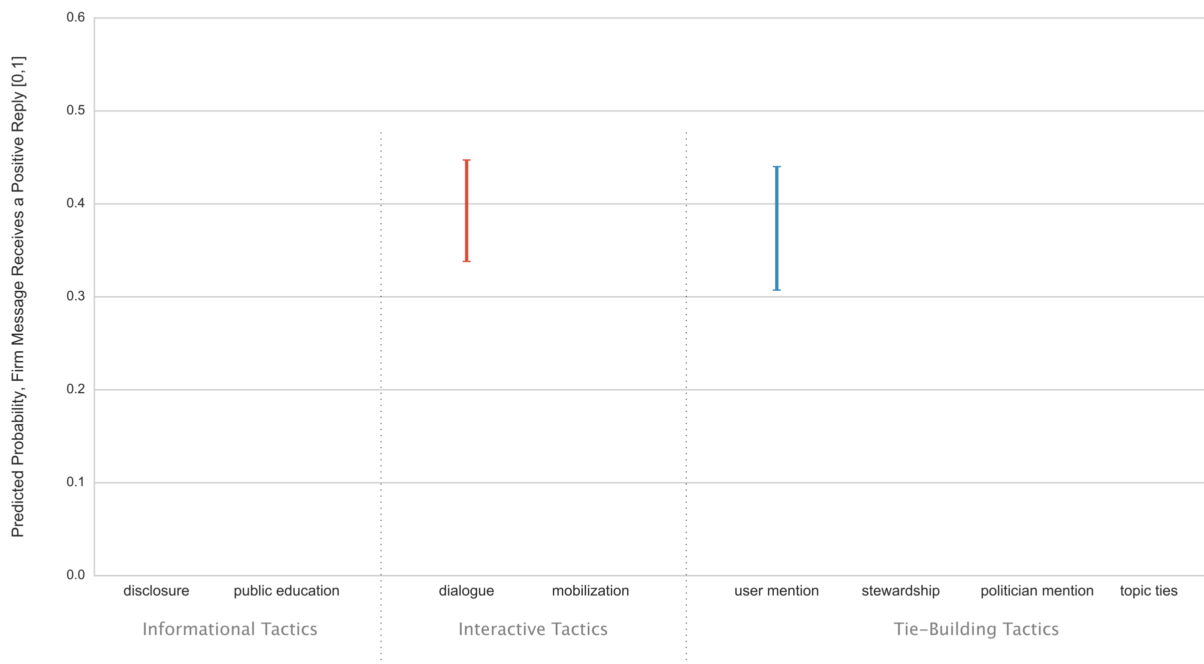
First, Figure 2.7 shows the predicted probabilities for the 8 individual CSR-focused tactics; *Marketing* is omitted because of the low number of cases (n=74) as well as, more importantly, that such messages are not related to CSR efforts. Because of the similarity of results across *Positive Reply* and *Negative Reply,* only those for the former are shown. Figure 2.7a shows probabilities for *Retweet* while Figure 2.7b shows probabilities for *Positive Reply.* The probabilities shown on the y-axis (with 95% confidence intervals) are calculated based on post-regression predictions from Model 2 (*Retweet*) and Model 4 (*Positive Reply*) in Table 2.3, holding all other independent variables constant at values of 0 and all control variables at their mean values. For example, the first bar in Figure 2.7a shows the predicted likelihood (0.583) of a message being retweeted that has a value of "1" on *Disclosure*, values of "0" on the other tactics (*Public Education, Marketing, Dialogue, Mobilization, User Mention, Stewardship Message, Politician Mention,* and *Topic Ties*), and mean values on all control variables. In brief, *ceteris paribus,* a message with a *Disclosure* tactic has a 58.3% predicted likelihood of being retweeted. *Public Education* tweets have a similar likelihood (58.9%) of receiving a public retweet. Of the two interactive tactics, in turn, *Mobilization* was not significant, while *Dialogue* messages are predicted to have only an 18.1% chance of being retweeted. Lastly, the tie-building tactics have widely ranging predicted probabilities: *User Mentions* are predicted to have a 53.0% chance of being retweeted, *Stewardship Messages* a 28.4% chance, and *Topic Ties* a 60.6% chance (*Politician Mention* was not significant).

Figure 2.7b shows the predicted probabilities of a firm message receiving a positive reply. Because of the lower number of replies compared to retweets, and to better highlight the effects of the independent variables, the probabilities shown are based on a logit regression with the same dependent variable (*Positive Reply*), but where the cases are limited to messages that receive some form of reply (n=3,202), whether negative, neutral, or positive. As in Figure 2.7a, despite having the population of 2014 tweets from the 42 accounts, I limit the examination to the significant variables. First, *Dialogue* has a predicted likelihood of 39.6% of a positive reply, while *User Mentions* have a 37.7% predicted likelihood of receiving a positive reply, both higher than the baseline category (not shown) of 26.7%.

Figure 2.8, meanwhile, shows the predicted probabilities for the three aggregate tactics *Information, Interaction,* and *Tie-building.* Probabilities for Figure 2.8a and 2.8b, respectively, are calculated based on post-regression predictions from Model 1 (*Retweet*) and Model 3 (*Positive Reply*) in Table 2.3, holding all other independent variables constant at values of 0 and all control variables at their mean values. In Figure 2.8a, showing the probabilities for *Retweeted,* the first bar shows the predicted likelihood (61.3%) of a public retweet for a firm message that has a value of "1" on *Information* (i.e., it is either *Disclosure, Public Education,* or *Marketing*), scores of "0" on *Interaction* and *Tie-building*, and mean values on all control variables. *Tie-building* tactics, meanwhile, have a predicted probability of a retweet of 50.7%, while *Interaction* tactics have only a 12.8% predicted likelihood. Turning to *Positive Reply* (Figure 2.8b), only *Interaction* is significant, with a predicted likelihood of 43%, roughly 10% higher than the baseline probability (not shown) of 32.6%.



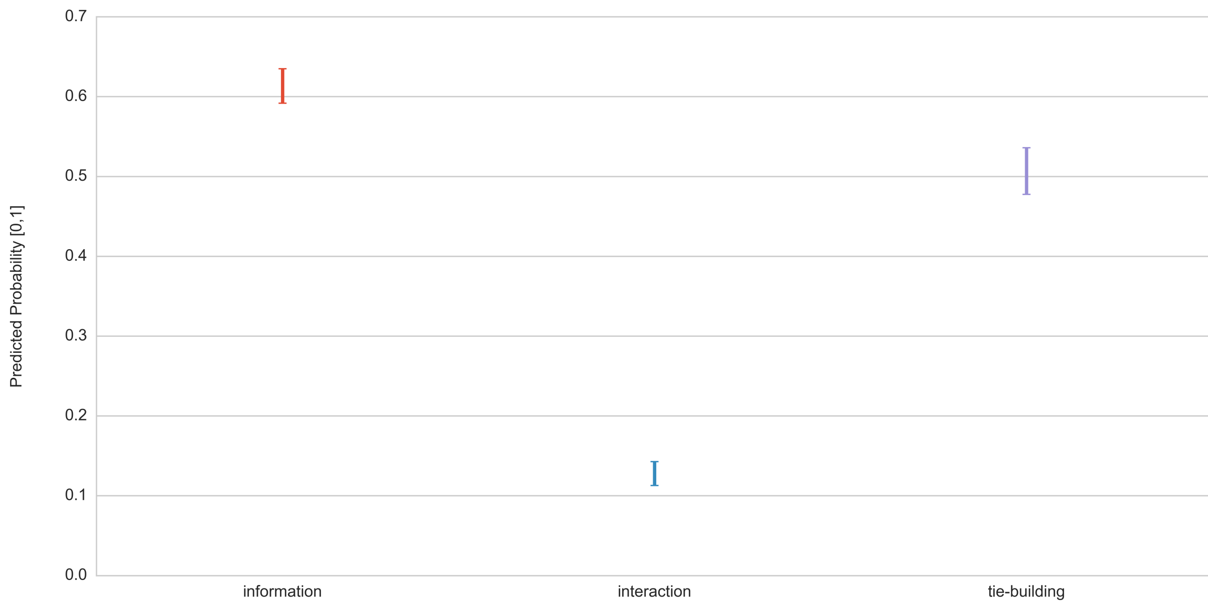
(a) Predicted Probabilities of Firm CSR Message Receiving a Retweet



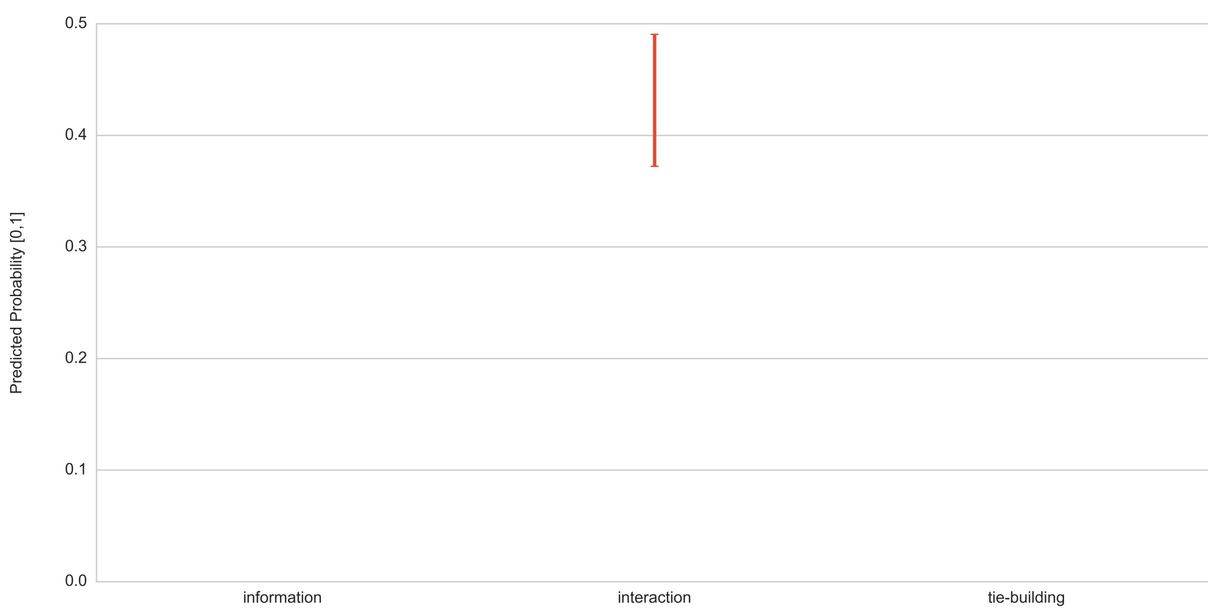
(b) Predicted Probabilities of Firm CSR Message Receiving a Positive Reply

Figure 2.7: Predicted Probabilities for 8 Communication Tactics

Note: Probabilities are calculated based on post-regression predictions from Model 2 (retweets) and Model 4 (positive reply) in Table 2.3, holding all other independent variables constant at values of 0 and all control variables at their mean values. Y-axis shows the predicted likelihood of a firm CSR message receiving a retweet (a) and a positive reply (b), for the different main categories of communication tactics (information, interaction, and tie-building). Vertical lines show the 95% confidence interval for each tactic. Non-significant tactics are not shown.



(a) Predicted Probabilities of Firm CSR Message Receiving a Retweet



(b) Predicted Probabilities of Firm CSR Message Receiving a Positive Reply

Figure 2.8: Predicted Probabilities for 3 Main Communication Tactics

Note: Predicted Probabilities are calculated based on post-regression predictions from Model 1 (retweets) and Model 3 (positive reply) in Table 2.3, holding all other independent variables constant at values of 0 and all control variables at their mean values. Y-axes in (a) and (b) show the predicted likelihood of a firm CSR message receiving, respectively, a retweet and a positive reply, for the three main categories of communication tactics (information, interaction, and tie-building). Vertical lines show the 95% confidence interval for each tactic. For example, the first bar shows the predicted likelihood of a public retweet for a firm message that has a value of '1' for the variable Information (i.e., it is either 'disclosure', 'public education', or 'marketing'), scores of '0' on Interaction and Tie-building, and mean values on all control variables. Only significant tactics shown.

2.4.4 Additional Analyses

I conducted a number of additional analyses to further explore select findings as well as to verify the robustness of the tests. These tests are summarized below; further details are available in the appendix.

Firm Size and Twitter Audience

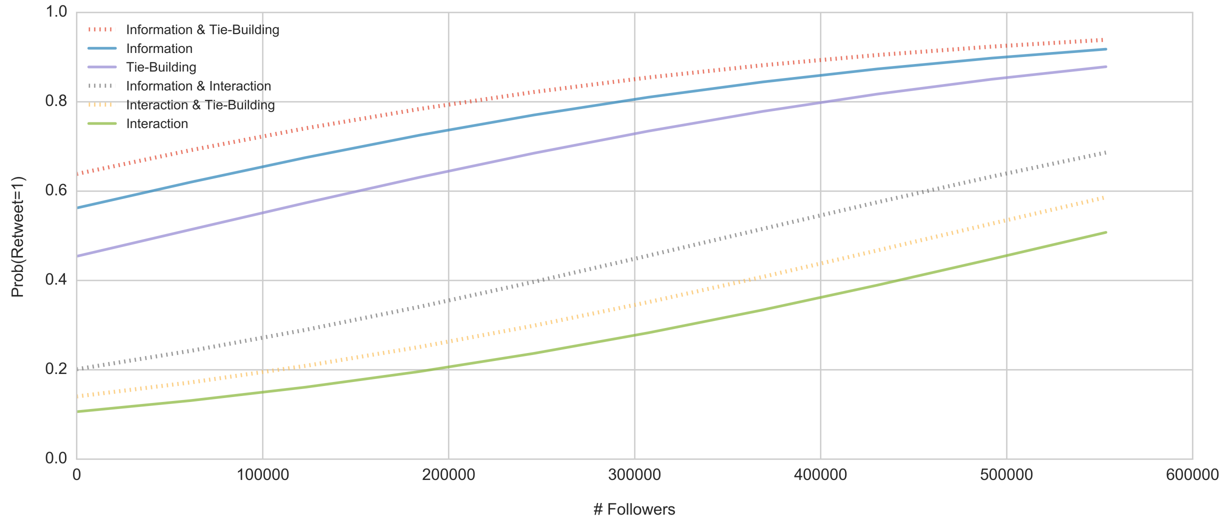
First, I incorporated alternative measures to tap the size of the firm. The argument implicitly made earlier is that it is the number of followers of the firm that is the more proximate, and hence more important, measure of company size. That is, larger companies will generally be associated with higher numbers of followers, such that these measures of firm size are more indirect proxies for the size of the audience on Twitter. As a check on this assumption I ran the above six models including number of employees and then including total assets. The inclusion of these variables resulted in only minor changes in significance levels of several independent and/or control variables in models 2, 4, and 6 (no changes in sign or significance for any model variables in models 1, 3, and 5), with the results being overall slightly more favorable to the independent variables than the results shown in Table 2.3. Both assets and employees were not significantly related to *Positive Reply* and were positively, significantly associated with *Negative Reply.* Interestingly, assets and employees have a different relationship with *Retweet:* Assets are negatively related while employees are positively related. A plausible explanation is that the public targets the companies rich in assets for negative reactions and comments, while firms with more employees are more likely to have employees among the followers, which increases the likelihood of a follower retweeting the firm's messages.

To further delve into the effect of audience size on the relationship between the independent and dependent variables, I ran predicted probabilities across the empirical range of the number of followers. These probabilities are shown in Figure 2.9. First, Figure 2.9a shows the predicted probabilities of a retweet not only for the major categories of *Information, Interaction,* and *Tie-Building* but also (given how tweets often contain multiple tactics) combinations of these tactics. This figure illustrates, first, that interaction tactics – whether alone or in conjunction with information or tie-building – have the lowest probabilities of being retweeted. The highest likelihood of retweeting comes with messages that combine information and tie-building. Second, Figure 2.9a shows the power of audience size, with the likelihood of a retweet roughly doubling on average across the empirical range of number of followers. At the same time, the figure shows the follower count is not deterministic; for instance, a company with a low number of followers is more likely to have a *Tie-Building* message retweeted than a company with 500,000 followers that sends an *Interaction* tweet. In short, tactics can trump audience size.

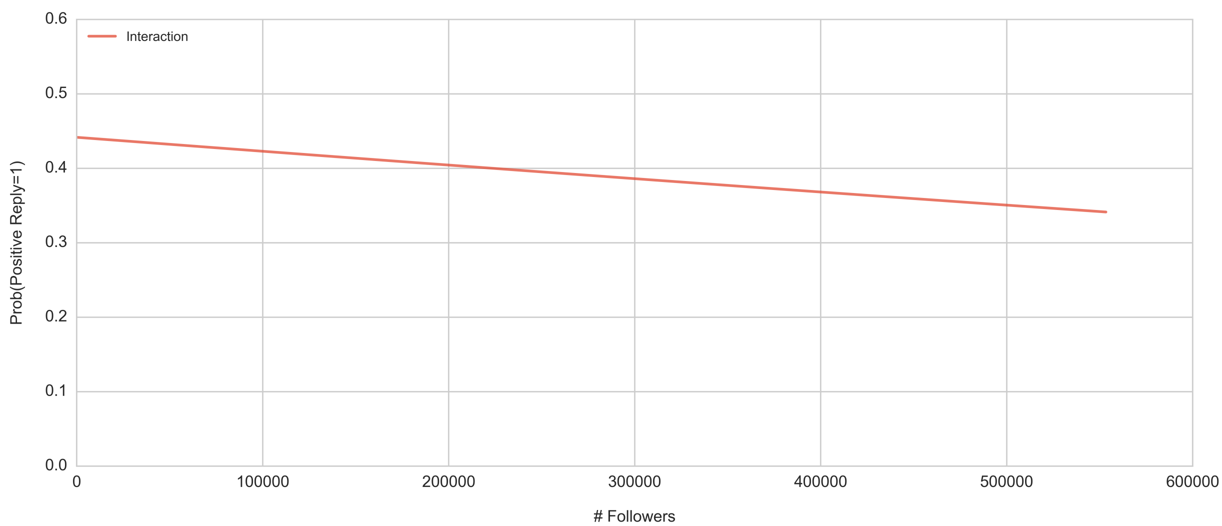
Figure 2.9b, meanwhile, is of interest for a different reason. Probabilities are derived from the same model shown in Figure 2.8b, save that the number of followers is not held at its mean value but is allowed to vary across its empirical range. The figure shows, for this sample, that the success of the only significant tactic, *Interaction,* is decreasing in follower size. This suggests a negative relationship between the size of the audience and the efficacy of the communication tactics. Given that a positive relationship between the number of Twitter followers and audience reactions is one of the most robust relationships in prior research (e.g., Bakshy et al., 2011; Saxton & Waters, 2014), this is a novel finding worth investigating further.

Alternative Dependent Variables: High Awareness

I additionally ran several logits using different versions of the binary retweet dependent variable (where n=10,568). In the original variable, *Retweeted,* 56% of tweets (n=10,568) receive a score of “1.” It would be useful to apply a higher threshold on the awareness generated and see whether the results hold. I thus ran additional logits with three different thresholds; first, with the dependent variable coded “1” on *High Awareness* if the tweet



(a) Predicted Probabilities of Firm CSR Message Receiving a Retweet



(b) Predicted Probabilities of Firm CSR Message Receiving a Positive Reply

Figure 2.9: Predicted Probabilities for 3 Main Communication Tactics (Reply Tweets)

Note: For Figure (a), predicted probabilities are calculated based on post-regression predictions from Model 1 (*Retweeted*) in Table 2.3. For Figure (b), probabilities are based off a modified version of Model 3 (*Positive Reply*) in Table 2.3, specifically, with the same independent and dependent variables but with cases limited to messages that receive some sort of reply (n=3,202). Figures (a) and (b) show the predicted likelihood of a message being retweeted and receiving a positive reply (y-axis) across the range of # Followers (x-axis) for the three main categories of communication tactics (different lines). All other control variables are held constant at their mean values. Figure (a) also shows probabilities for tweets that contain multiple tactics. For visual clarity 95% confidence intervals are omitted. Only significant tactics shown.

received 2 or more retweets, or roughly at or above the mean value of 2.08 retweets (6,302 of 18,722 tweets). The results were also estimated with the threshold at 3 or more retweets (3,966 tweets) and 5 or more retweets (1,912 tweets). In all cases the sign and significance level of the independent variables in all 6 models remains unchanged.

Alternative Regression Model: Negative Binomial Regression

I further ran the six models using a negative binomial regression, where the dependent variables are not the binary presence of a retweet or a positive or a negative reply but rather a count of the number of, respectively, retweets and positive and negative replies. The results are substantively similar to the results presented in Table 2.3.

Fixed Effects and Clustered Standard Errors

It is also possible that the relationship between communication tactics and audience outcomes differs according to the company, or according to the industry or sector in which the company operates. The above models could, in other words, potentially suffer from omitted variable bias. The remaining sensitivity analyses serve to address such concerns. First, the models were re-run with fixed effects for the 15 different industry sectors. These results were slightly more favorable to the independent variables in terms of significance and no changes in sign for any of the model variables. Second, similar results obtained when, rather than a fixed effects model, standard errors were clustered on sector. Finally, the models were re-run with sector fixed effects and robust standard errors clustered by Twitter account. Here again there are only minor changes in significance levels of several of the independent and control variables, with the results being slightly more favorable for the independent variables in model 4 and slightly less favorable in models 5 and 6. Overall, the additional tests point to the robustness of the relationships between communicative tactics and the accumulation of reputational capital shown in Table 2.3.

2.5 Discussion and Conclusions

In this paper I addressed two main research questions, one designed to shed light on the communication tactics used by Fortune 200 firms in their CSR communication and another to shed light on the relationship between these communicative tactics and the acquisition of two dimensions of reputational capital. The main propositions examined are that CSR reporting should best be seen through a communication lens, that CSR communication is conducted using a variety of tactical forms, and that one of the primary uses of these tactics is to engender positive changes in reputational capital.

2.5.1 Synthesis of Empirical Findings

Inductive analysis of the firms' CSR tweets led to the identification of nine communication tactics, one of which, marketing, was not related to CSR communication. The remaining eight tactics were grouped into three main categories. First, I found two common informational tactics, disclosure and public education. Second, there are two interactive tactics, mobilization and dialogue, with the latter being found in a fifth of all messages. Third, I identify four tie-building tactics – user mentions, stewardship messages, politician mentions, and topic ties. All, except for politician mentions, were in more than 10% of all tweets. The key insight here is that the tactics go well beyond disclosure; it is not CSR reporting so much as it CSR *communication*.

The multivariate analyses, in turn, showed how these various communication tactics relate to the accumulation of reputational capital. What do these analyses show? Four of the tactics – disclosure, public education, mobilization, and topic ties – are most highly associated with retweets, the proxy for reputational awareness. Two other tactics – dialogue and user mentions – are more highly associated with positive replies, the proxy for reputational favorability. Another tactic, the politician mention, is associated chiefly with the receipt of a negative reply, an additional (inverse) proxy for reputational favorability. Overall, some tactics are associated with increased awareness and others are associated with increased favorability. When looking at the aggregate categories, awareness is driven by the use of informational tactics and, to a lesser extent, tie-building tactics; favorability, in turn, is more closely related to the use of interaction tactics. These findings have several implications for accounting and CSR research.

2.5.2 Implication #1: CSR communication rather than reporting

The first key implication derives from the strong evidence that firms' CSR activities go well beyond disclosure. Indeed, disclosure – the most studied variable heretofore in the accounting CSR literature – appears to have no significant impact on firms' reputational favorability. Recent normative (Kent & Taylor, 2016) and empirical (Colleoni, 2013) research had documented dialogic approaches in firms' CSR activities, while conceptual work (Castelló et al., 2015; Schultz et al., 2013) had stressed the possibilities of communication and networking approaches. It was in light of this mounting evidence that led to the first research question regarding whether and how firms were using a broader range of communication tactics in their CSR behaviors. The evidence presented strongly suggests moving beyond a one-way reporting model of CSR efforts.

The current study thus helps flesh out prior conceptual work on the "communication view" of CSR (Castelló et al., 2015; Schultz et al., 2013) by documenting and further theorizing around specific manifestations CSR communication takes. At the same time, the present study departs from the approach of Schultz et al. (2013) in one important way. While the current study is sympathetic to the constructivist, symbolic interaction focus of the Schultz et al. (2013) perspective, the analyses presented here are predicated on a largely instrumentalist viewpoint. The present study argues that the newer, non-reporting forms of communication, such as dialogue, can fulfill a largely instrumental role in helping the firm reach its reputational goals.

Dovetailing with these findings are the implications for the two predominant approaches to studying CSR, the signaling and legitimacy approaches. The “blind spots” in the friction between this dichotomy have generated calls for more nuanced approaches (see Cho et al., 2015). One notable blind spot is the typical independent variable, with the signaling and legitimacy perspectives both placing the emphasis squarely on reporting or disclosure. Yet the present study strongly suggests a focus on disclosure does not do justice to firms’ CSR efforts. The typical firm studied here does not just use social media to report on its CSR activities but is assuming a CSR-focused “public education” role, engaging in dialogue around CSR issues with multiple stakeholders, and employing a number of networking and tie-building tactics to deepen and strengthen ties with constituents. In short, not only does the communication go beyond disclosure; it also goes beyond the dialogue stressed in recent public relations-focused CSR research (Colleoni, 2013; Kent & Taylor, 2016; Schultz et al., 2013). In the end, sending information is merely one type, albeit an important type, of communication. Both the signaling and legitimacy approaches are well primed to incorporate these insights with an expanded set of independent variables.

2.5.3 Implication #2: The importance of public perceptions

The second major implication of the current study lies in the role of public perceptions. Others have recognized – at least implicitly – the importance of CSR-driven public perceptions; where I extend the literature is in refining and further specifying the role of these perceptions in firms’ CSR strategies. In the process, the study helps answer a puzzle: Why would a firm choose to engage in dialogue, or adopt a public education role, or emphasize a tie-building approach? Schultz et al. (2013) offer one answer: that CSR communication is more concerned with mutual constructivism and symbolic interaction than it is with the pursuit of instrumental goals. I offer an alternative: That these alternative, non-informational, non-reporting CSR communication tactics play a role in strengthening positive public perceptions of the firm. In effect, these tactics only make instrumental sense after the importance of public perceptions is recognized, for these tactics are what help build relationships online with stakeholders (Kent & Taylor, 2016; Saxton & Waters, 2014). Of further interest is the finding that some of these additional tactics – particularly dialogue – have a significantly greater connection to reputational favorability than does disclosure.

The findings also conform with recent arguments that CSR bridges market and nonmarket strategy (Bach & Allen, 2010). If my argument is correct that CSR-driven perceptions help determine not only nonmarket policy outcomes but also market-based consumer decisions as well as investor and employee outcomes, then the "bridging" character of CSR makes sense. In fact, one key part of the answer to the puzzle of "why public education?" and other tactics is in seeing how CSR bridges market and nonmarket strategy. Namely, if consumer behavior such as purchasing decisions, product evaluations, brand loyalty, and brand communities are influenced by perceptions of CSR performance as the evidence suggests (Brown & Dacin, 1997; Creyer, 1997), then the market relevance of CSR – and, by extension, of communicative CSR tactics designed to boost consumer perceptions – becomes apparent.

Certain of the newly identified tactics also make sense from a purely nonmarket strategy perspective (Oliver & Holzinger, 2008). For instance, in the social movement literature, public education would be seen as a central, long-term advocacy tactic for reaching policy goals (Guo & Saxton, 2014). Looking at the related corporate political activity (CPA) literature, in turn, I find broad evidence to support two of Hillman and Hitt's (1999) *corporate political strategies*: information and constituency-building. Disclosure and public education conform to the information strategy, while interaction and tie-building tactics could be seen as ways of building an online constituency of followers that could eventually be mobilized to support the company's policy agenda. Social media thus helps firms utilize a *relational* approach (Hillman & Hitt, 1999) to CPA.

At the same time, my findings extend the CPA research in at least two ways. One, it suggests a broader set of tactics than that identified by Hillman and Hitt (1999) and other CPA scholars. Two, it suggests corporate political activity is not necessarily directly or overtly political; for example, if CSR were overtly "political," we would expect to see greater evidence of "direct" or "inside" lobbying tactics and even grassroots lobbying. Instead, the approach taken here is largely "indirect" or "outside" (Guo & Saxton, 2014; Mosley, 2009), with a focus on fostering favorable public perceptions and boosting ideological and public support. Firms' CSR efforts on social media are, in effect, more in line with a long-term, indirect *advocacy* agenda than they are with an overtly political agenda.

2.5.4 Practical implications

Managers and markets should first recognize that reputational capital is an important outcome of CSR activities. Analogous to brand recognition, it is an internally generated intangible asset that, despite its importance, is off the balance sheet and difficult to value. Internal valuation requires management accounts to have the cross-functional skills to be able to talk with marketing, public relations, business analytics, and data science people within the firm. There are also some apparently unique features of social media-driven reputational capital. To start, at the micro level at least, it is accumulated according to a power law distribution. There are many more "losers" than "winners." It is also likely gained and lost more quickly; the half-life of accumulated reputational capital is an unknown quantity.

In this sense accountants and marketing and public relations professionals designing CSR campaigns are more like venture capitalists choosing which companies to invest in (knowing that the success of chosen ventures is not normally distributed) rather than traditional accountants deciding what to include or exclude from a stand-alone annual CSR report. At the same time, a central insight from this study is that CSR communication is not something done annually; rather, CSR communication is something delivered on a continual basis through the day-to-day informational, interactive, and tie-building efforts made by companies. CSR communication should, in other words, be dynamic. It is also public communication; previously, some communicative and advocacy tactics involved more private discussions. Another key practical implication is for managers to understand that there are a variety of communicative tactics available and that, given that different tactics are associated with awareness and favorability, managers need to have a broader repertoire of communicative tools in order to maximize the accumulation of reputational capital. Some of the above findings also likely apply to the individual level; that is, individual managers and accountants could benefit from the above insights into how reputational capital is accumulated on a day-to-day level. Finally, the findings support the integration of market and nonmarket strategies. Managers should see CSR as a core activity that pays dividends in a number of areas. In short, CSR is not just disclosure. CSR is not just a nonmarket strategy. And CSR performance is not something reported only once a year.

2.5.5 Future research

This paper illustrates how reputational capital is accumulated on a micro, message-by-message, day-to-day level. By stressing how CSR spans market and nonmarket strategies, goes beyond one-way reporting, and is used to change public perceptions, as well as how those perceptions are influenced on a message-by-message basis, I hope to have encouraged CSR research into new and fruitful domains of research.

In particular, this study aims to serve first as an exemplar of the types of studies that can be employed to shed light on how firms change public perceptions. Reputational capital may be accumulated via the day-to-day change and/or diffusion of opinions and attitudes and sentiments that is achieved through the public's processing of firms' CSR messages – and these micro-level changes are visible by viewing the public's reaction to each individual message. There are ample opportunities for research that builds on the current study and applies Big Data and machine learning approaches to studying other micro-level phenomena.

One important unaddressed issue concerns the relationship between the micro-level accumulation of reputational capital and aggregate and more stable stocks of that capital. Differently put, the current study pertains to the flow of reputational capital; how does this flow relate to the stock? In addition to research on the lifespan and valuation of this intangible asset, research should also be conducted on the relationship between the expenditure on CSR communication and the accumulation of favorable public perceptions. While social media platforms themselves are free to adopt, the staff expenditures required to run a professional, cross-functional social media strategy are not insignificant. The most successful of the companies examined, Bank of America, integrated specific CSR campaigns with sizable donations, thus suggesting the benefits of "putting one's money where one's mouth is." Finally, research could be conducted into how reputational capital – and more broadly, public perceptions – are reported in current accounting reports. Such perceptions fit most closely with the idea of *Social and Relationship Capital* in the Integrated Reporting framework (IIRC, 2013). Future research could dig into the implications of integrating reputational capital into the framework.

The field would also now benefit from studies that extend this social media-based research with different approaches, different units of analysis, and different theories to help shed light on CSR communication on social media. While the current study is designed as a further step down this path, there are many yet to explore. One of the biggest opportunities are approaches that look further at the public side of the equation. Social media are distinguished for the heightened possibilities of interactivity and for dialogic, back-and-forth two-way communication. This study has examined the public's reactions to firms' messages, but has not looked at the broader CSR conversations the public initiates or is engaged in. We could also look at firms' reactions to *the public's* message. Overall, as others have done with activists on websites (de Bakker & Hellsten, 2013), there is a need for research that *brings the public in* and looks at firm-public interactions around CSR issues.

Appendix to Chapter 2

This appendix contains several additional analyses and figures as well as further details on the coding of the independent and dependent variables.

2.A Further Details on Coding of Variables

2.A.1 Coding of Reputational Favorability

Data were coded in three stages. First, I selected a sample of 1,000 replies for crowdsourced manual coding.[17] Crowdsourced coding – human coding by the online "crowd" on such platforms as Mechanical Turk and Crowdflower – provides a rapid and cost-effective method for manually coding data that is becoming increasingly popular in management and accounting research (Grenier et al., 2015; Rennekamp, 2012). These 1,000 tweets were uploaded to Crowdflower (www.crowdflower.com). On the site I created a set of instructions for the coders and manually coded 55 randomly selected replies as having either a negative, neutral, or positive sentiment; I also provided a brief justification for each score that was visible to the coders as both an initial practice set and for providing intermittent feedback as they proceed with coding. Figure 2.10 contains a screenshot of the instructions given to Crowdflower workers to code the sentiment in public replies, along with a sample tweet.

[Insert Figure 2.10 here]

Each of the 1,000 tweets was coded by at least three Crowdflower coders. The majority of tweets coded had 100% agreement across the three coders. Less than 1% (n=9) saw complete disagreement across the three coders, while roughly 30% saw two of three coders agree on sentiment ratings. From these values Crowdflower returns a "best score," which factors in the level of confidence in each of the three coders (based on past performance); however,

[17]Specifically, I selected the 304 replies to the 1,000 randomly selected firm tweets, plus an additional 696 randomly selected replies to firm tweets, for a total of 1,000 replies.

for the most part the "best score" is the majority vote in each case. Agreement with this crowdsourced score and my own manual codes was 89% with a Cohen's kappa score of 0.817 ($\kappa = 0.817$), indicating a high level of agreement.[18]

I then proceeded to implement a machine learning technique to code the remainder of the replies. Machine learning approaches differ from automated, unsupervised, lexicon-based approaches used in prior CSR research in that they are supervised techniques, requiring the research to make decisions and to "train" the machine learning algorithm. There are several steps involved. In these techniques the researcher takes a sample of the data, codes it, and then trains the machine learning algorithm by fine tuning select parameters. In the present study, first the 1,000 manually coded tweets were divided into training (85% of cases) and test datasets (15% of cases). Tweets were then pre-processed for machine learning by making content lowercase, by "stemming" words (e.g., "making" becomes "make") and by removing "stop words" (common words such as "the").

Three popular machine learning algorithms – Naïve Bayes, Decision Trees, and Support Vector Machines (SVM) – were trained. Each algorithm was trained separately by modifying key algorithm parameters (e.g., by selecting how many text features to consider in the algorithm)[19] and, after each round of training, assessing the classifier's accuracy by comparing the algorithm scores generated on the training dataset to those in the test dataset. In line with prior research on tweets (Go et al., 2009), the SVM algorithm achieved the highest accuracy. Naïve Bayes achieved 75.8% accuracy and the Decision Tree algorithm achieved 76.9% accuracy, both somewhat lower than initial performance with SVM. Consequently, SVM was chosen and parameters fine-tuned until the highest level of accuracy was achieved.

With the classifier fully trained, agreement on the 3-code sentiment variable (-1, 0, +1) between the manual coding and SVM coding was 81.3%. This fits well with expectations based on prior research that has found accuracy in sentiment classification to be around 82%

[18]Commonly cited guidelines by Landis & Koch (1977) refer to kappa scores of .81 or greater as "almost perfect."

[19]For instance, in training the data, the highest accuracy came with choosing 10% of the features for the algorithm.

(e.g., Go et al., 2009 achieved 82.2% accuracy). Accuracy is even higher with three binary variables derived from the sentiment variable: 89.0% for positive, 80.2% for neutral, and 94.5% for negative.

With the classification algorithm trained and tested, it was then used to generate sentiment scores for the 4,247 remaining replies. For comparison, scores were also created for the 1,000 replies that were manually coded. The agreement between the 1,000 SVM-coded variables and the 1,000 manually coded negative, neutral, and positive binary variables is 93.3%, 79.6%, and 86.1%, respectively.[20] This indicates a high level of inter-coder reliability. As a further check, I also calculated formal inter-coder reliability scores: Cohen's kappa values were from a low of 0.5827 for neutral, considered "moderate agreement" (Landis & Koch, 1977) 0.604 for negative, and 0.677 for positive, both considered "substantial agreement" (Landis & Koch, 1977).

In short, through the manual coding and supervised machine learning process each of the 5,247 replies received a value of "1" on one of the following three binary variables derived from the three-value (-1, 0, +1) sentiment variable: *negative, neutral,* or *positive*. Recall that the analysis is a message-level analysis, with the messages being the 18,722 company tweets. In a final step, the reply data were therefore merged into the company tweet dataset. Because some firm tweets could receive more than one reply, the reply-level dataset (n=5,247) was collapsed by the company tweet that was the target of the response and merged into the company tweet database (n=18,722), in the process creating the final two binary variables for analysis: *Positive Reply*, with a value of "1" indicating firm tweets that receive at least one positive reply; and *Negative Reply*, with a value of "1" indicating firm tweets that receive a negative reply.

[20] I then ran these 1,000 tweets against the unsupervised/automated lexicon-based approach ANEW (Nielsen, 2011) that was used in several prior CSR studies (Colleoni, 2013) and saw 64% agreement with the manually coded Crowdflower codes. Splitting the sentiment variable into negative, neutral, and binary variables the agreement with the hand-coded tweets is better at 89.5% (negative), 67% (neutral), and 71.5% (positive), yet still underperforms the SVM-based classifier by roughly 4%, 13%, and 14.5% on the negative, neutral, and positive variables, respectively.

2.A.2 Coding of Communication Tactics

As with reply sentiment, a machine learning classifier was trained to code the 17,222 non-hand-coded firm tweets. To improve the accuracy of the classifier only hand-coded tweets with agreement by at least 2 of 3 Crowdflower coders were used (1,257 of the 1,500 tweets). The final classifier was an SVM algorithm that achieved 80.0% overall accuracy compared to the test data set (comprising 12.5% of the 1,500 manually coded tweet); specifically, agreement with the hand-coded tweets was 81.4% with disclosure, 86.9% with public education, 97.2% with marketing, and 93.8% with no information. Given the high levels of accuracy the classifier was then used to classify the remaining tweets (i.e., assign one of the four values on these variables).

As an additional check on the reliability of the coding, scores from this round of machine learning were also compared to the 1,500 hand-coded tweets. For the four-category variable, accuracy between the 1,500 hand-coded tweets and the machine learning-coded tweets was 82.53% ($\kappa = 0.664$). For disclosure it was 82.53%, for public education it was 88.27%, for marketing it was 98.8%, and for non-informational it was 94.67%. Moreover, comparing the machine learning-coded tweets to the 200 initial hand-coded tweets, for the four-category variable agreement was 82.5%. For the four binary variables derived from this variable and used for analyses, agreement for Disclosure was 84.5% with a kappa value of 0.687, for Public Education it was 89% with a kappa value of 0.703, for Marketing it was 99% with a kappa score of 0.911, and for non-informational it was 92.5% with a kappa value of 0.702. In short, all would be considered at least "substantial agreement" by Landis and Koch (1977).

The other five (non-informational) tactics, however, could be coded through a different, more automated process. Namely, I developed custom algorithms for coding each of the remaining 18,522 tweets for the existence of the five tactics. For instance, one of the tactics coded was the *politician mention*, where firms made an effort to thank, congratulate, or interact with politicians. To find mentions of politicians, a list of all Twitter users mentioned in the 18,722 tweets that had usernames starting with "@Gov," "@Sen" "@Mayor," and "@Rep," such as @RepJoeKennedy and @SenGillibrand, was compiled then verified to ensure that these were Twitter accounts of politicians. Other politicians that were discovered to be mentioned, such as @NancyPelosi, were also added to the list. In the end a list of 158 verified politicians was created; at that stage automated Python code was written to identify every tweet among the 18,722 firm messages that contained one or more mentions of users from the list of politicians. Similar algorithms were written for the other tactics categories and were refined until coding accuracy was well above 90% compared to hand coding.

2.B Account-Level Analyses

The paper is built around message-level analyses. However, it may be of interest to see some relevant account-level analyses. Figure 2.11 contains the aggregate number of retweets and positive replies garnered by each of the accounts over the course of 2014.

[Insert Figure 2.11 here]

Figure 2.12, meanwhile, shows the number of new followers acquired over the course of 2014. While the number of followers was not a key variable in the analyses, it is an important indicator of the influence of a Twitter account.

[Insert Figure 2.12 here]

What is interesting about these figures is, first, how the distribution appears to follow a power law distribution. Second, the clear "winner" in all three figures, *BofA_ Community*, expended a lot of effort on its various CSR campaigns over the course of the year (as will be seen in the following chapter).

2.C Actor-Network Analyses

Figure 2.13 contains a depiction of a network analysis inspired by Bruno Latour's (Latour, 2005) Actor-Network Theory (A.N.T.). The figure depicts a network of ties among the actors, the tactics, and the audience reactions in the 18,722 firm tweets. Specifically, for each company tweet data were available on who sent it, which tactics were employed, and the audience reactions (retweeted, positive reply, neutral reply, negative reply, and none/ignored). Accordingly, I looped over each tweet and created *edges* based on the combination of sender, tactics, and reactions. I ran this in NetworkX in Python, exported to Gephi, and laid out the network using a "Force Atlas" layout. Edge line thickness reflects weight and node size reflects degree.

[Insert Figure 2.13 here]

A brief examination of this figure tells us the following. First, at the center of the network are three tactics – *topic ties*, *user mentions*, and *disclosure* – as well as two outcomes – *ignored* and *retweet* – as well as a handful of Twitter accounts. We can also look at individual accounts and see which tactics or outcomes they are closer to, or look at individual tactics and see which outcomes they are closer to, etc. While I ultimately decided not to analyze these data in depth, it remains a potentially fruitful avenue for future research.

2.D Robustness Tests

The following robustness tests were summarized in the main document for Chapter 1. Here I provide greater detail on these tests.

2.D.1 Additional Measures of Firm Size

Number of Employees. I conducted a number of additional analyses (not shown) to verify the robustness of the tests shown in Table 2.3. First, I incorporated alternative measures to tap the size of the firm. The argument implicitly made above is that it the number of followers of the firm is the more proximate, and hence more important, measure of company size (i.e., firms with greater assets or revenues or more employees will generally be associated with higher numbers of followers). Nevertheless, as a check on this assumption I also ran the above six models with number of employees as a control. First, when included in the three regressions with the aggregate tactics Information, Interaction, and Tie-building, the number of employees is significantly and positively associated with Retweet in model 1, and is not significant in model 3 (DV = Positive Reply) or model 5 (DV = Negative Reply). In models 1, 3, and 5 there are no changes in sign or significance for any of the other independent or control variables. In model 2, meanwhile, the number of employees does not obtain significance and there are no changes in sign for any of the other variables, but two previously non-significant variables obtain significance: Mobilization is positively associated with Retweet ($p=.082$) and Politician Mention is negatively associated ($p<.01$). Similarly, in model 4 (DV = Positive Reply) employees does not obtain significance, and Mobilization obtains a significant positive association ($p=.098$) and Politician Mention obtains a significant negative association ($p=.071$). Lastly, compared to model 6, employees is significantly, positively associated with Negative Reply, and unlike in model 6 Politician Mention no longer obtains significance ($p=0.45$); there are no other changes in sign or significance for any of the other model variables.

Assets. In place of the number of employees, total assets was also included as a size control. In model 1 assets is negatively associated with Retweet while in model 3 assets does not obtain significance; there are no other changes in sign or significance for any of the model variables. In model 5, assets is significantly, positively associated with Negative Reply, and, compared to model 5, two control variables, Broad CSR Account Focus and URL included, fail to obtain significance; all other variables retain the same size and level of significance. In the replication of model 2, assets is negative associated ($p<.01$) with Retweet and one control variable, URL included, no longer obtains significance ($p=0.140$). Lastly, in the test of model 6 with assets included, assets does not obtain significance and three tie-building variables, User Mention, Stewardship Message, and Politician Mention, no longer obtain significance, as does the control variable Time on Twitter.

2.D.2 Sector: Fixed Effects and Clustered Standard Errors.

Sector Fixed Effects. It is also possible that the relationship between communication tactics and audience outcomes differs according to the industry or sector in which the company operates. The above models could, in other words, potentially suffer from omitted variable bias. The remaining sensitivity analyses serve to address such concerns. The 42 Twitter accounts represent companies working in 15 different sectors. Three of the sectors (apparel, energy, and transportation) had a low number of tweets (n=18) and thus were combined into a miscellaneous category; binary variables were then created for the miscellaneous category plus each of the other 12 sectors, and a fixed effects model was run. In the replication of model 1 with industry controls, all variables are the same in terms of sign and significance save for the control variable # of Characters, which obtains signficance ($p<.01$). Compared to model 2, in the fixed effects regression Politician Mention gains significance ($p<.01$), as does Mobilization ($p=.061$), with no other changes in sign or significance for any model variables. In the replication of model 3, Tie-building gains significance ($p=0.067$) while Time on Twitter loses significance. In model 4, the only change is that Politician Mention gains significance ($p= 0.083$); in model 5 three controls – Time on Twitter, Broad CSR Account, and # of Characters – lose significance; while compared to model 6, Politician Mention loses significance, as do Time on Twitter, Broad CSR Account, and # of Characters.

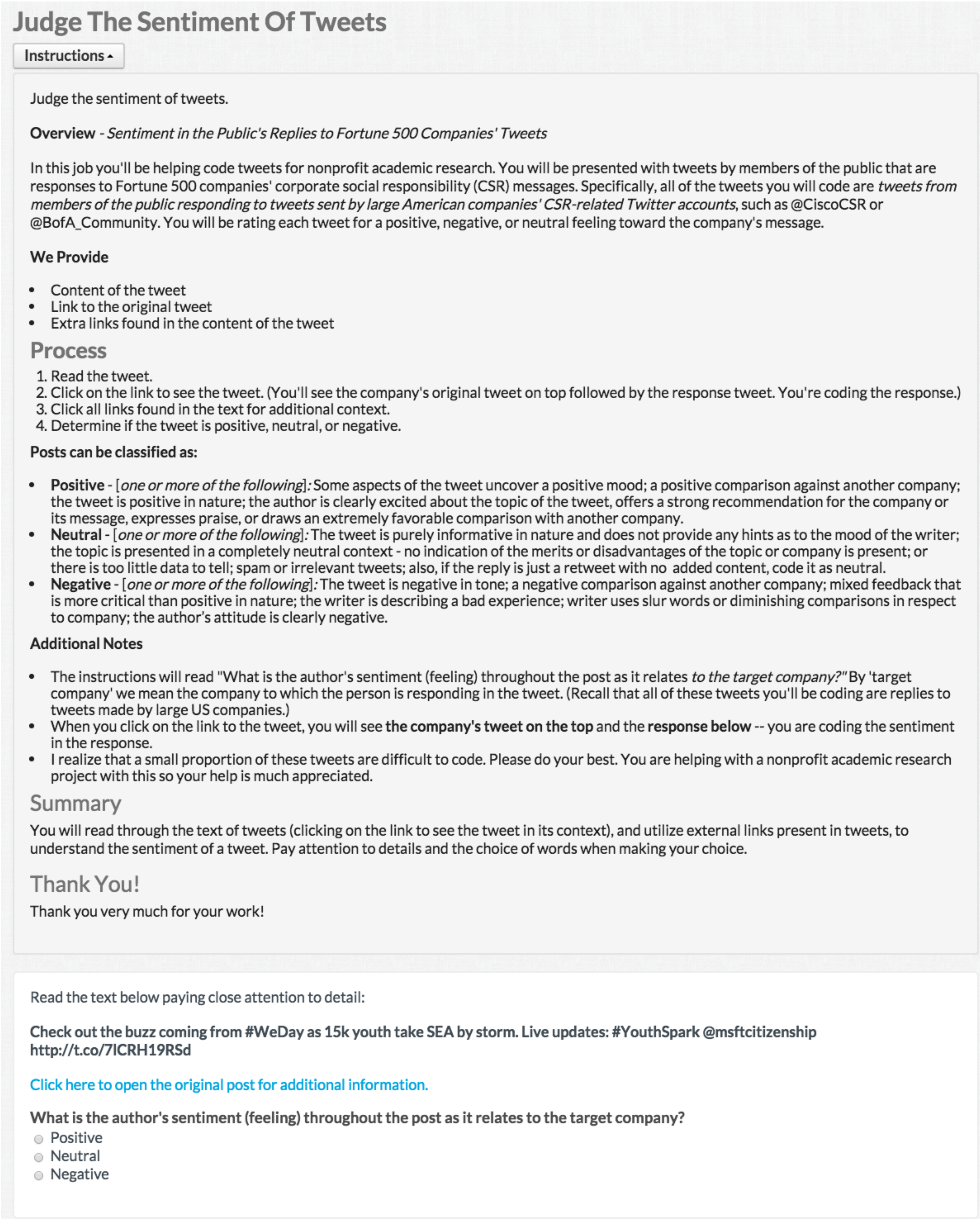
In terms of which sector dummies were significant, the results from the replication of model 2 are illustrative. With Aerospace and Defense as the omitted category, in the replication of model 2, Miscellaneous and Financials do not obtain significance, Chemicals, Media, Motor Vehicles & Parts, Retailing, Technology, and Telecommunications are positively and signficantly related to Retweet, and Health Care, Household Products, Industrials, and

Materials are negatively related.

Overall, the fixed effects results are slightly more favorable to the independent variables – with minor changes to Politician Mention and Mobilization – with the sector dummies taking away some of the explained variance from the original control variables.

Standard Errors Clustered on Sector. The six models were also re-run with robust standard errors clustered on sector. Compared to the version of model 1 shown in Table 2.3, Time on Twitter and URL included fail to obtain significance; same with model 2. Compared to model 3, Interaction and Time on Twitter do not obtain significance, while in model Interaction, Time on Twitter and Broad CSR Account do not reach significance. Compared to model 4, Mobilization obtains a positive significant relationship ($p<.01$), Politician Mention obtains a negative significant relationship ($p<.01$), while Time on Twitter loses significance. In contrast to model 5, Interaction, Broad CSR Account and Time on Twitter no longer obtain significance. Finally, compared to model 6, Dialogue loses its significant relationship, as does Broad CSR Account and Time on Twitter. These results are slightly less favorable to the independent variables, with four instances of independent variables losing significance and two instance of independent variables gaining significance.

Sector Fixed Effects with Robust Standard Errors Clustered on Account. As a final set of robustness tests, the models were re-run with sector fixed effects and robust standard errors clustered on the Twitter account. The results are the same as the above results with standard errors clustered on sector, with the only change in sign or significance being that Mobilization and Politician Mention gain significance in model 4. Overall, the additional tests point to the robustness of the relationships between communicative tactics and the accumulation of reputational capital shown in Table 2.3.



Judge The Sentiment Of Tweets

Instructions ▴

Judge the sentiment of tweets.

Overview - *Sentiment in the Public's Replies to Fortune 500 Companies' Tweets*

In this job you'll be helping code tweets for nonprofit academic research. You will be presented with tweets by members of the public that are responses to Fortune 500 companies' corporate social responsibility (CSR) messages. Specifically, all of the tweets you will code are *tweets from members of the public responding to tweets sent by large American companies' CSR-related Twitter accounts*, such as @CiscoCSR or @BofA_Community. You will be rating each tweet for a positive, negative, or neutral feeling toward the company's message.

We Provide

- Content of the tweet
- Link to the original tweet
- Extra links found in the content of the tweet

Process

1. Read the tweet.
2. Click on the link to see the tweet. (You'll see the company's original tweet on top followed by the response tweet. You're coding the response.)
3. Click all links found in the text for additional context.
4. Determine if the tweet is positive, neutral, or negative.

Posts can be classified as:

- **Positive** - [*one or more of the following*]*:* Some aspects of the tweet uncover a positive mood; a positive comparison against another company; the tweet is positive in nature; the author is clearly excited about the topic of the tweet, offers a strong recommendation for the company or its message, expresses praise, or draws an extremely favorable comparison with another company.
- **Neutral** - [*one or more of the following*]*:* The tweet is purely informative in nature and does not provide any hints as to the mood of the writer; the topic is presented in a completely neutral context - no indication of the merits or disadvantages of the topic or company is present; or there is too little data to tell; spam or irrelevant tweets; also, if the reply is just a retweet with no added content, code it as neutral.
- **Negative** - [*one or more of the following*]*:* The tweet is negative in tone; a negative comparison against another company; mixed feedback that is more critical than positive in nature; the writer is describing a bad experience; writer uses slur words or diminishing comparisons in respect to company; the author's attitude is clearly negative.

Additional Notes

- The instructions will read "What is the author's sentiment (feeling) throughout the post as it relates *to the target company?"* By 'target company' we mean the company to which the person is responding in the tweet. (Recall that all of these tweets you'll be coding are replies to tweets made by large US companies.)
- When you click on the link to the tweet, you will see **the company's tweet on the top** and the **response below** -- you are coding the sentiment in the response.
- I realize that a small proportion of these tweets are difficult to code. Please do your best. You are helping with a nonprofit academic research project with this so your help is much appreciated.

Summary

You will read through the text of tweets (clicking on the link to see the tweet in its context), and utilize external links present in tweets, to understand the sentiment of a tweet. Pay attention to details and the choice of words when making your choice.

Thank You!

Thank you very much for your work!

Read the text below paying close attention to detail:

Check out the buzz coming from #WeDay as 15k youth take SEA by storm. Live updates: #YouthSpark @msftcitizenship http://t.co/7lCRH19RSd

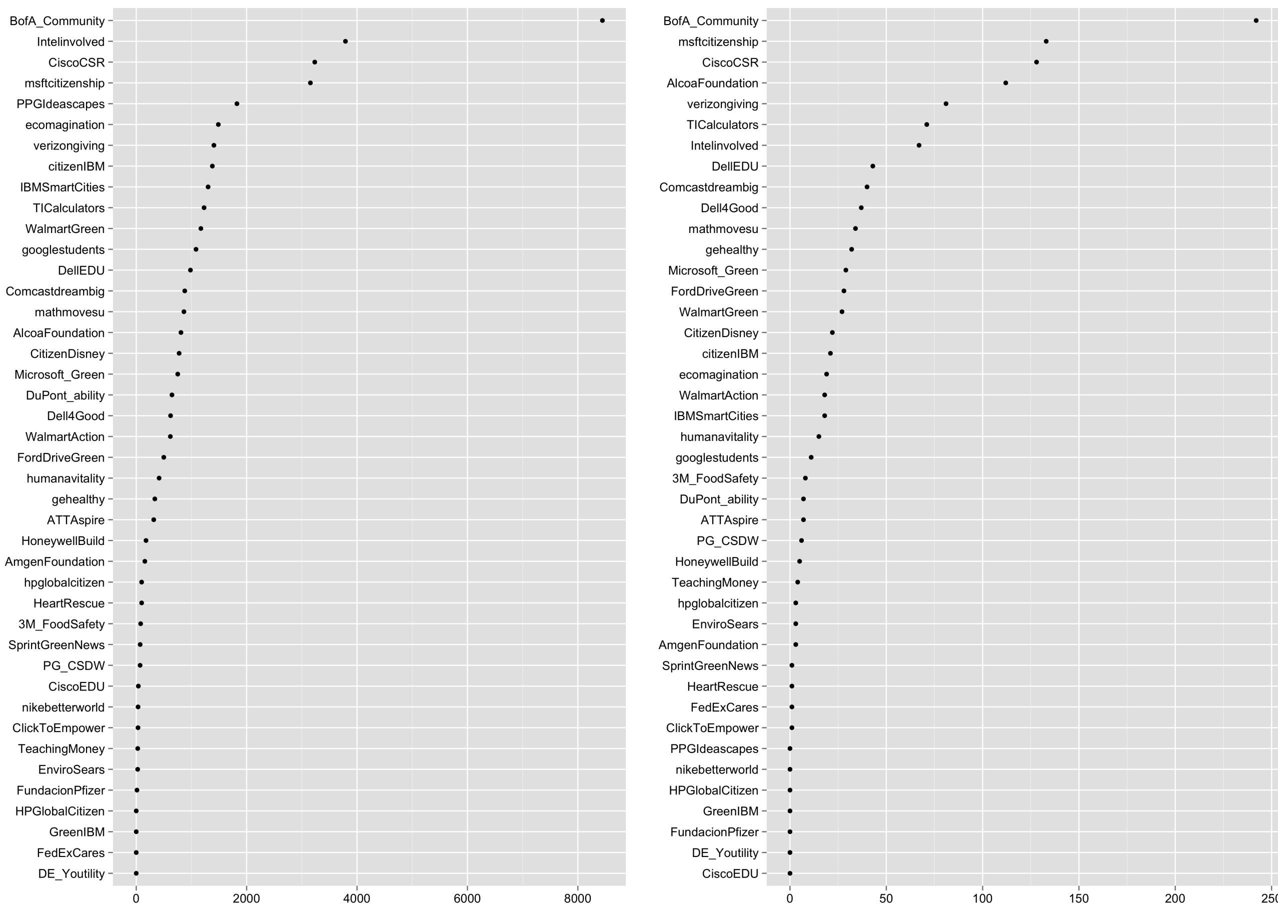
Click here to open the original post for additional information.

What is the author's sentiment (feeling) throughout the post as it relates to the target company?

- Positive
- Neutral
- Negative

Figure 2.10: Instructions for Coding Sentiment Given to Crowdflower Coders

Note: Instructions with one example.



(a) Number of Retweets   **(b)** Number of Positive Replies

Figure 2.11: Number of Retweets and Positive Replies Received per Account in 2014

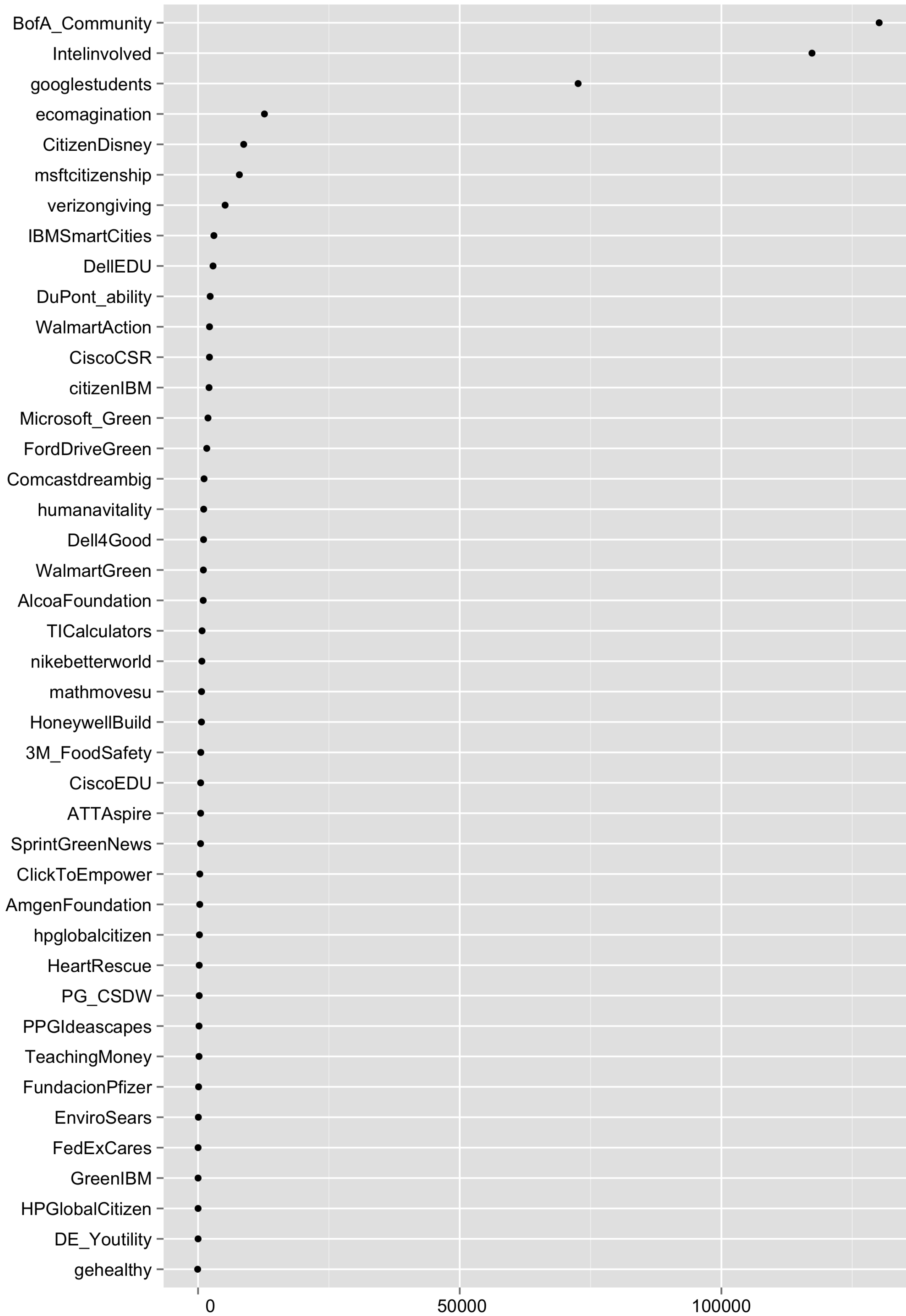


Figure 2.12: Growth in Number of Followers per Twitter Account, 2014

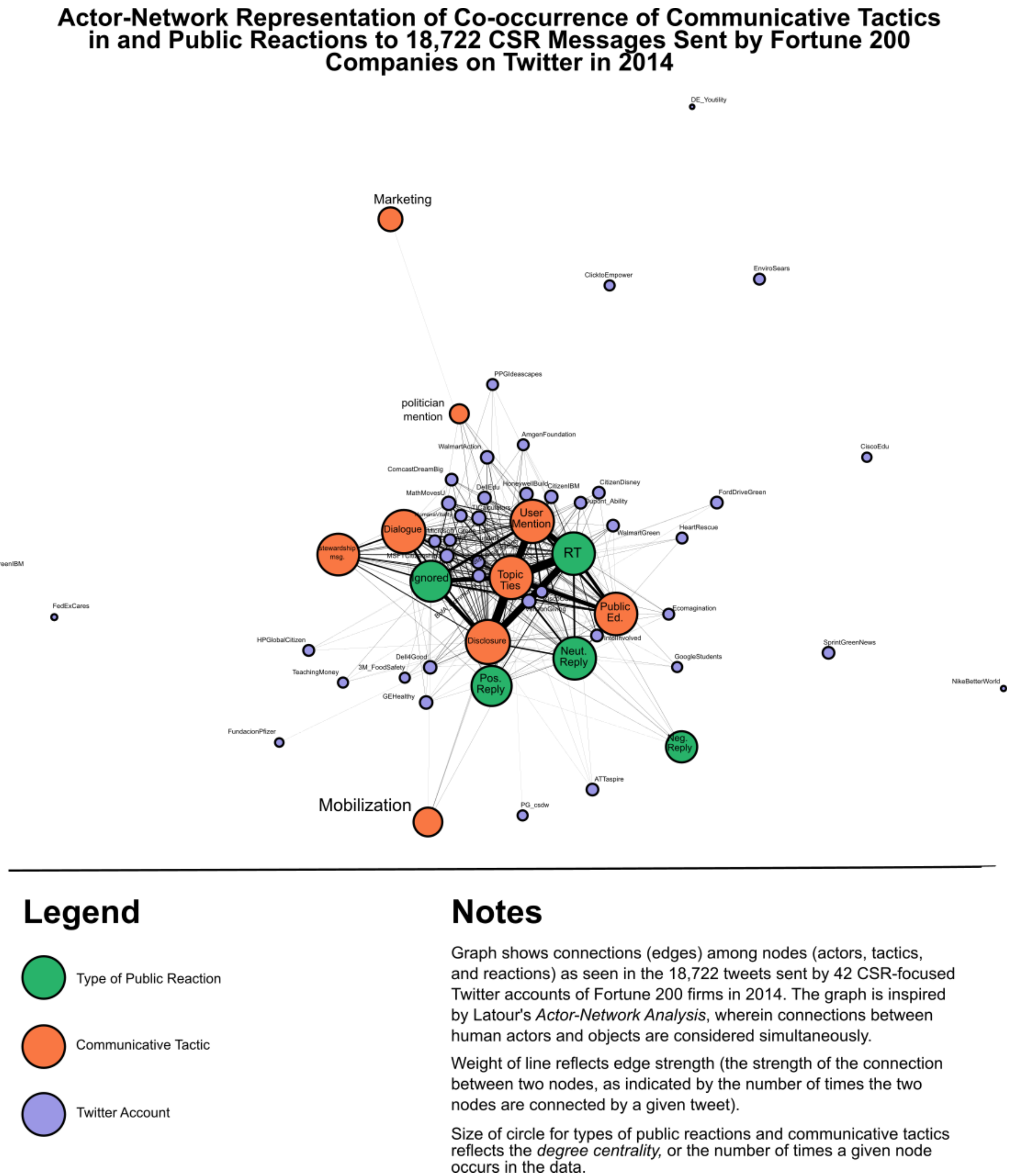


Figure 2.13: Actor-Network Theory of Tactics in and Reactions to 18,722 CSR Tweets

Chapter 3

Calling Firms Out: Exploring the Nature and Determinants of Dynamic, Interactive Micro-Reporting in *Fortune 200* Firms' CSR-related Twitter Accounts

Abstract

How firms respond and react to public comments, queries, and questions represents their commitment to public account-giving and reporting behavior. With the development of social media, such accountability efforts have been given a new and more visible platform – a micro-level reporting and accountability platform that is amenable to innovative study and theory-building by accounting scholars. Given the dearth of existing literature, this paper seeks to inductively develop a model of the determinants of firms' reactions to CSR-related messages by members of the public. To develop the model, a series of inductive analyses are undertaken of the 163,402 public messages stakeholders sent to Fortune 200 firms' CSR-focused Twitter accounts in 2014, focusing on what differentiates those messages that do from those that do not receive a reaction from Fortune 200 firms. Complementing typical qualitative inductive techniques, innovative "Big Data" machine learning techniques are employed that serve to help identify the most relevant of the examined features. Building on these qualitative and quantitative insights, the paper ends with the presentation of a proposed theoretical model of the determinants of firm reactions to public comments.

Keywords: Accountability, corporate social responsibility, CSR disclosure, corporate communication, stakeholder engagement, social media, Big Data, machine learning

3.1 Introduction

The accounting world is witnessing an era of heightened demands for accountability and increased pressure for CSR and sustainability reporting (e.g., Cho et al., 2015; Hopwood, 2009; Unerman & Bennett, 2004). The diffusion of new media and social media platforms has intensified the drive toward public, dynamic accountability and CSR reporting. Social media platforms such as Twitter are used not only by firms to send discrete visual and textual "micro-reports" but are also used by the public to engage with, converse, and even challenge firms for their CSR-related actions and inactions (e.g., Castelló et al., 2015; Saxton et al., 2016). However, existing literature has yet to examine the public's discussion of firm's CSR efforts and, more specifically, how firms choose to react and respond to public CSR messages. This is an important issue for the accounting literature inasmuch as it reflects a different reporting model – a model that is not performed annually but daily, a model that is driven not solely by firms but by members of the public, and a model that is not conducted behind closed doors but one that is conducted publicly.

In short, leveraging social media platforms, members of the public are continually "calling firms out" for their CSR actions, and firms' decisions to respond or not respond constitute a new form of public, dynamic, interactive reporting and accountability behavior that has yet to be addressed by the extant accounting literature. Given that it is a novel empirical and conceptual context, inductive theory-building would be especially valuable (Miles & Huberman, 1984; Strauss & Corbin, 1998). Accordingly, in this paper I undertake a series of inductive analyses of all 163,402 messages sent on Twitter in 2014 that mention the 42 CSR-focused Twitter accounts of *Fortune 200* firms, such as *@CiscoCSR* and *@Dell4Good*, and examine which features of these messages and who sends them lead firms to publicly react.

In particular, this papers aims to develop conceptual and theoretical insights into three aspects of this phenomenon: 1) the nature of the public CSR messages sent to the *Fortune 200* firms, 2) the nature of firm reactions, and 3) the relationship between the two. The analyses culminate in the presentation of a proposed theoretical model of the determinants of firm reactions to public CSR messages.

What is novel about my approach is how – in order to develop insights into the relationship between characteristics of public messages and whether and how firms react to those messages – I employ a mix of traditional qualitative inductive approaches (e.g., Miles & Huberman, 1984; Strauss & Corbin, 1998) and "Big Data"-driven, machine learning approaches (e.g. Evans, 2014). Most notably, a variety of *feature selection* algorithms are employed to reduce the 80+ identified variables into a more parsimonious theoretical model. Combined, these qualitative and quantitative inductive analyses are intended to provide fresh insights into the nature and determinants of firm reactions to public messages.

In the end, the communicative CSR interactions that occur on social media represent a new – more dynamic and interactive and public – form of accountability reporting that carries implications for accounting research beyond the domain of CSR. The current study contributes to the literature by providing insights into the nature of this new "account-demanding" and reporting venue. In so doing, it highlights a new inductive approach – notably, machine learning-driven methods for variable selection – that complement existing qualitative approaches.

The remainder of the paper is organized as follows. The second section reviews the relevant literature. The third section describes the method, including an overview of the sample, data, and plan for analysis. The fourth through ninth sections present the analyses, covering in turn aspects of *how* firms react, *who* receives a reaction, *what* types of messages receive a reaction, *when* and *where* messages receive a response, and firm characteristics that drive *why* firms react. The tenth section presents findings from a comprehensive feature selection algorithm and offers the final proposed theoretical model. The eleventh section concludes with a summary of the implications of the study.

3.2 Existing Literature

There exists a well established body on literature in accounting on the nature, determinants, and outcomes of CSR reporting (e.g., Neu et al., 1998; Patten, 1992). With the spread of the Internet in the 1990s, accounting scholars began to investigate the possibilities of early forms of new media on two-way communication between firms and the public. Yet research on websites and discussion boards revealed an absence of true reciprocal dialogue (Unerman & Bennett, 2004). Instead, websites were used for one-way reporting – and often for the purposes of enhancing legitimacy and managing impressions (e.g., Cho & Roberts, 2010).

More recent studies have begun to examine the nature of firms' CSR efforts on social media (e.g., Lee et al., 2013; Lyon & Montgomery, 2013). Social media offer something new: Not only are they tools firms use to report on their activities, but they offer a venue in which the public can react and respond to firms' CSR reporting efforts. In a way, social media represent the new "town hall," yet a town hall in which public commenting and firm reporting is done on a daily basis rather than once a year. Moreover, the account-giving is often made to *individual stakeholders* rather than to groups or to an amorphous "public." Several of the more recent of these studies have examined how the public reacts to firms' CSR reporting on social media (Castelló et al., 2015; Colleoni, 2013; Saxton et al., 2016).

Still, what remains unexamined is two things: 1) how the public talks about firms' CSR efforts and 2) how *firms* react and respond to public messages. In effect, social media have enabled a public, dynamic, interactive reporting environment, yet the existing literature has examined only one side of this relationship. The aim of this study, therefore, is to help generate theoretical insights into three things. One, the nature of the messages members of the public send on social media when they are discussing firms' CSR efforts. Two, the nature of firms' reactions to these public messages. And three, which features of the public's messages drive firms' reactions. After addressing these three issues, the paper culminates in the presentation of a proposed theoretical model of the determinants of firms' reactions to public CSR messages.

3.3 Method

3.3.1 Sample

The sample comes from the 200 firms in the 2012 *Fortune 200* list. Beginning with this list of 200 firms, a search was conducted to find all Twitter accounts of these 200 companies. All maintained a Twitter account. Some maintained more than one account, and some of these with multiple accounts maintained an account that was solely devoted to CSR or sustainability issues such as the environment, equality, diversity, human rights, education, or sustainability. The sample for this paper is derived from data on the 42 such CSR-focused accounts maintained by the 200 firms in 2014.

Table 3.1 shows the account names for the 42 Twitter accounts that comprise the sample. The table also shows the total number of public messages sent in 2014 discussing the 42 accounts along with counts of the number of these messages that were ignored by the firm; the remainder (as discussed in the following section) received some form of concrete reaction.

3.3.2 Data

Public Mentions of CSR-Focused Fortune 200 Twitter Accounts

Data are derived from several sources. The starting point is all public Twitter messages, or *tweets*, that mention any of the 42 CSR accounts over the course of 2014. Twitter has created a *search* application programming interface (API) that allows researchers to search for and download any message that "mentions" any other Twitter user over the past seven days. This search API was accessed via a Python script run twice per day over the course of 2014 to download all tweets mentioning one or more of the 42 CSR-focused Fortune 200 accounts. In total, members of the public sent 163,402 messages mentioning these accounts.

Table 3.1: 42 CSR-Focused Twitter Accounts of Fortune 200 Firms in 2014

Screen Name	Number of Public Mentions	Number Ignored by Firm
3M_FoodSafety	338	338
AlcoaFoundation	3,344	2,626
AmgenFoundation	673	663
ATTAspire	1,629	1,620
BofA_Community	48,789	48,396
CiscoCSR	12,975	12,543
CiscoEdu	244	242
CitizenDisney	3,203	3,063
CitizenIBM	4,651	4,647
ClicktoEmpower	324	324
ComcastDreamBig	3,680	3,496
DE_Youtility	7	7
Dell4Good	3,717	3,389
DellEDU	4,671	4,067
Dupont_Ability	1,372	1,350
Ecomagination	3,694	3,662
EnviroSears	210	177
FedExCares	84	82
FordDriveGreen	6,709	6,605
FundacionPfizer	176	170
GEHealthy	1,082	976
GoogleStudents	4,582	4,581
GreenIBM	3	3
HeartRescue	315	301
HoneywellBuild	910	873
HPGlobalCitizen	419	408
HumanaVitality	1,859	1696
IBMSmartCities	5,212	5167
IntelInvolved	11,164	11,034
MathMovesU	1,994	1,942
Microsoft_Green	2,588	2,434
MSFTCitizenship	12,255	11,975
NikeBetterWorld	476	476
PG_CSDW	1,062	1,016
PPGIdeascapes	1,502	1,501
PromesaPepsico	5	5
SprintGreenNews	332	328
TeachingMoney	201	189
TICalculators	4,225	3,952
VerizonGiving	5,887	5,042
WalmartAction	3,084	2,991
WalmartGreen	4,034	3,945

Note: Table shows # of public messages sent and # ignored by firm (i.e., no reaction).

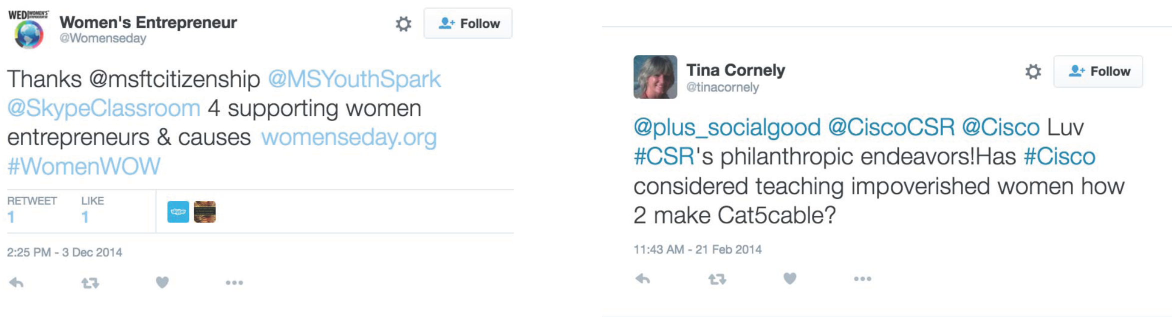


Figure 3.1: Sample Public Messages

These 163,402 messages by members of the public comprise the initial data source.

Figure 3.1 provides two examples of these tweets. What makes these messages relevant is the *user mention* convention common on social media sites such as Twitter. In the message on the left, the word *@msftcitizenship* indicates the message sender is "talking about" Microsoft's CSR-focused Twitter account *@MSFTCitizenship*. Similarly, in the tweet on the right the inclusion of *@CiscoCSR* indicates the sender is discussing Cisco's CSR-focused account *@CiscoCSR*. On Twitter this convention formally links messages to user accounts. The effect is that the targeted accounts (@MSFTCitizenship and @CiscoCSR) will receive notifications that they have been mentioned in a tweet. Anyone who follows or searches for messages mentioning the company accounts will also be able to read the message.

In effect, when a member of the public includes a user mention of one of the 42 CSR-focused accounts, it is initiating or engaging in a public conversation either with or about the Fortune 200 company. Given the focus of the 42 accounts on CSR-related issues, these public messages also relate to CSR. If we look at the collection of 163,402 public messages in their entirety, we can gain an appreciation for how members of the public are talking to and talking about firms' CSR efforts. By examining in turn how the firms react to these public messages, we can seek to understand how firms are engaging in dynamic, interactive reporting and accountability behaviors with members of the public. This paper seeks to examine both aspects.

Additional Data

Beyond the public messages, the second core set of data is related to the firms' reactions to the 163,402 public messages. Thoughout this paper (as will be discussed in a following section), the outcome variable examined is whether and what type of reaction a message from a member of the public gets.

In addition, a variety of complementary data on the characteristics of the message senders are also gathered from Twitter and supplemented by geographical and other data from additional sources. Similarly, financial and Twitter data on the message recipients – the *Fortune 200* firms with CSR-focused accounts – are also gathered. Details on data, sources, and operationalization are discussed in relevant sections throughout the paper.

3.3.3 Analytical Method and Analysis Plan

Step 1: Qualitative Inductive Analyses

Given the novelty of the data, context, and phenomenon under investigation, inductive analyses are invaluable for identifying the unique features and set of determinants involved (Miles & Huberman, 1984; Strauss & Corbin, 1998). Specifically, the inductive analyses involve a mix of qualitative and quantitative techniques undertaken in two discrete stages. In the first stage, in line with qualitative inductive methods espoused by Miles & Huberman (1984) and Strauss & Corbin (1998), the identification and coding of relevant variables involved a multistage, iterative process of cycling back and forth among data, literature, and emergent conceptual categories.

During the first stage, "Big Data"-driven, machine learning techniques are employed to *create* specific variables – namely, the identification of topics latent in the public messages using what is known as *latent Dirichet allocation*, or LDA (e.g., Arun et al., 2010). For the most part, however, the first stage is characterized by traditional – that is, qualitative – inductive techniques. The goals are to identify and conceptualize relevant variables, situate them within existing literature, and to theorize about the relationship of these concepts with the outcome variable (firm reactions).

Step 2: Quantitative *Feature Selection* Analyses

One of the opportunities and challenges with "Big Data" (e.g., Vasarhelyi et al., 2015), however, is the sheer number of possible variables that can be identified. In order to develop a more parsimonious theoretical model, we would thus benefit from complementing the above approaches with techniques designed specifically for Big Data to select the most relevant "features" (variables) of the data. In the machine learning[1] literature this process is known alternatively as *feature selection*, *variable selection*, or *model selection* (see Saeys et al., 2007).

Feature selection comprises various techniques designed to help determine which of a large number of variables should be retained for a final empirical model (Verikas et al., 2011). They are, essentially, *dimensionality reduction* methods that are ideally suited to Big Data projects characterized by a large number of potential variables. The dimensionality reduction approach effectively offers the potential for a more rigorous pruning of models than what is offered by purely qualitative approaches.

Accordingly, in the second stage, a number of primarily quantitative machine learning-based feature selection techniques are used to analyze the data. Unlike the qualitative inductive analyses – which are used here chiefly to identify and conceptualize the potentially relevant variables – the goal with feature selection is to separate the identified variables into those that are more important from those that are less important and can be removed from the model.

In line with the feature selection literature (e.g., Saeys et al., 2007), a number of different techniques are employed to help triangulate the results. In particular, I begin by employing *univariate* techniques, discarding, for instance, variables with little relevant variation. I

[1] Machine learning includes a broad array of techniques – covering prediction, classification, and feature selection, among others – and is well suited to generating fresh insights into phenomena of interest (Parks, 2014).

then incorporate two forms of *multivariate* feature selection techniques, specifically looking at multivariate correlations and bivariate visual plots, logit analyses, and Chi-square tests (see Guyon & Elisseeff, 2003). The former (multivariate correlation approaches) are useful for helping identify *redundancy* (Yu & Liu, 2004) in the nascent theoretical model, while the latter (statistical and graphical bivariate analyses) are helpful for screening variables for *relevance* (Yu & Liu, 2004). Both sets of techniques serve as *filter* methods (Guyon & Elisseeff, 2003) in the model selection process.

In terms of presentation, while all the above feature selection techniques were conducted separately from the qualitative inductive analyses, for ease of presentation the univariate and bivariate feature selection analyses are presented in each of the six main analysis sections that cover the who, what, when, where, why, and how of public messages and firm reactions (see below for details).

The final feature selection method – presented at the very end of the paper – is a comprehensive model selection algorithm known as *stability selection* (Meinshausen & Bühlmann, 2010). Once the set of potential variables has been pruned via the above qualitative and quantitative analyses, stability selection is applied simultaneously on the entire set of remaining candidate variables – more than 80 in total. The ultimate goal is to retain only those variables that are useful for predicting values of the dependent variable or, conversely, removing those variables that are not helpful. This step proves valuable in helping develop the more parsimonious theoretical model that is ultimately presented at the end of the paper in Section 10. Further details on the stability selection procedure are provided in that section.

3.3.4 Conceptual Framework

To help focus the development of inductive insights I organize my analyses into six different elements of the messages, the message senders, the reactors, and their reactions – specifically, *who* sends the messages, *where* the messages come from, *when* they are sent, *what* is included

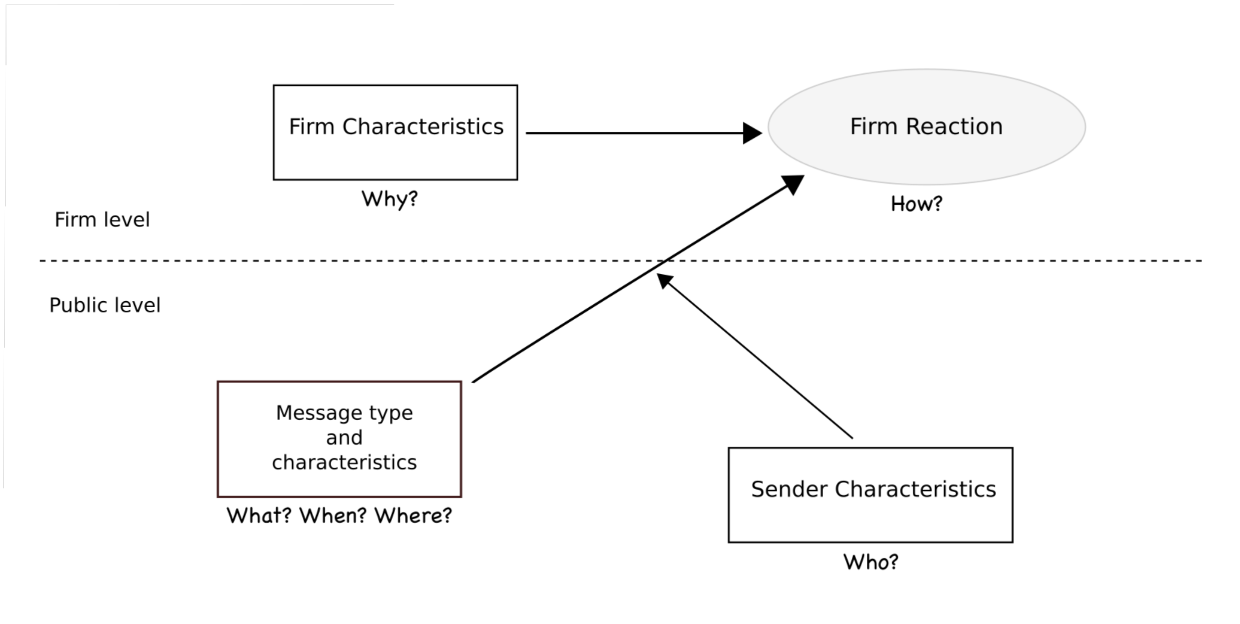


Figure 3.2: Broad Conceptual Model of Drivers of Firm Reactions to Public Messages

in the messages, firm characteristics that might explain *why* they react and, finally, *how* the firms react. In short, the inductive analyses are guided, at the highest level, by six broad conceptual categories – the "Who, What, When, Where, Why, and How" of firm reactions to public messages. This framework is summarized in Figure 3.2.

3.4 *How* Public Messages Receive a Firm Reaction

I begin the analyses with an examination of "how" – of the ways in which firms react to community-initiated messages. Based on the analyses I developed a *message reaction framework* that categorizes firm reactions into four different categories. These categories constitute the outcome variable for this study.

3.4.1 Measuring Reactions: An Information-Processing Framework

This study aims to address a key question; namely, what motivates firms to react or respond to public questions, comments, attacks, or efforts at interaction? There are a number of ways of conceptualizing firms' reactions. First, they constitute a form of *managerial attention* to stakeholders (Mitchell et al., 1997). What is perhaps unique about the reactions undertaken by firms on social media sites such as Twitter is that they are public and transparent – the message sender and other Twitter users are able to see whether and how the firm reacts to each message. As a result, firms are able to use their reactions to send signals to the message sender and the broader Twitterverse; by paying attention to some messages and not others, the firm is making certain CSR-related stakeholder claims and issues more salient.

Firms' reactions further represent a form of *information processing*. Public relations theory has posited information processing and information seeking as two key outcomes of involvement in an issue (e.g., the Situational Theory of Publics, Grunig, 1997). What separates the two outcomes is that the former is more passive while the latter is more active. In the current context (Twitter), the firms are *receiving* a stream of public messages automatically rather than conducting an active search for those messages; this makes the reactions fit within an information processing framework. In line with this research, and building on social media and information-relevant research in a variety of fields, I conceptualize message reactions as encompassing the four key information-processing behaviors shown in Figure 3.3: discussing/replying, sharing, archiving/rating, and ignoring.[2] I describe each of these in turn.

[2]Information processing research in public relations has generally conceptualized information processing as simply the amount of attention paid to a given message (Slater et al., 1992). My proposed framework extends the information processing literature in at least two notable ways. First, it expands the number of categories. Second, as will be shown, it suggests an ordering of effort in the categories of reaction. The benefit of this framework is that it facilitates greater theoretical precision around the notion of information processing – of what actors do with the messages they receive.

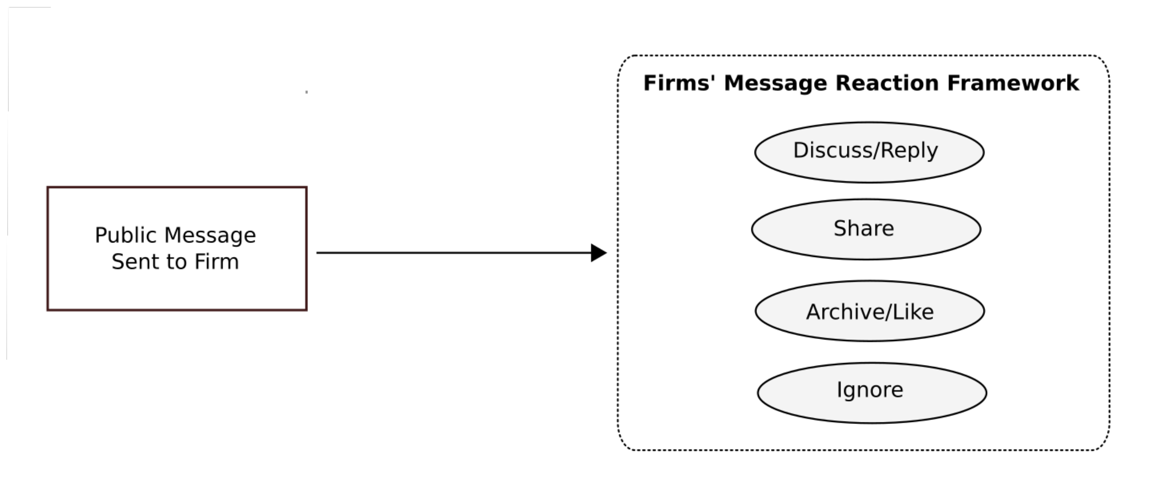


Figure 3.3: Information Processing Framework: Range of Message Recipients' Actions

3.4.2 Reaction #1: Replying

As shown in Figure 3.3, the first type of reaction is to discuss the message received. On social media, as is typical off-line, this is achieved by *replying* to the message sent by the member of the public. The Twitter API indicates whether each message is a reply to an existing message and, if so, to which tweet ID it is a response. This feature allowed me to loop over all tweets sent by the 42 CSR accounts over the course of 2014 and determine which of the 163,402 public messages received a response from the firm(s) mentioned in each of those tweets. I was thus able to determine that 1,095 of the 163,402 tweets (or 0.67% of all messages) by members of the public received a reply from one of the Fortune companies mentioned.

Figure 3.4 shows the same two tweets seen in Figure 3.1, yet this time also shows (underneath the original message) the reply the Fortune company sent to the public message. In the first, Microsoft sends a thanking, *stewardship*-type message (Kelly, 2000), while in the second message Cisco replies by sending a response to the question raised by the Twitter user. These two tweets count as two of the 1,095 tweets that received a reply from the Fortune 200 firms in 2014.

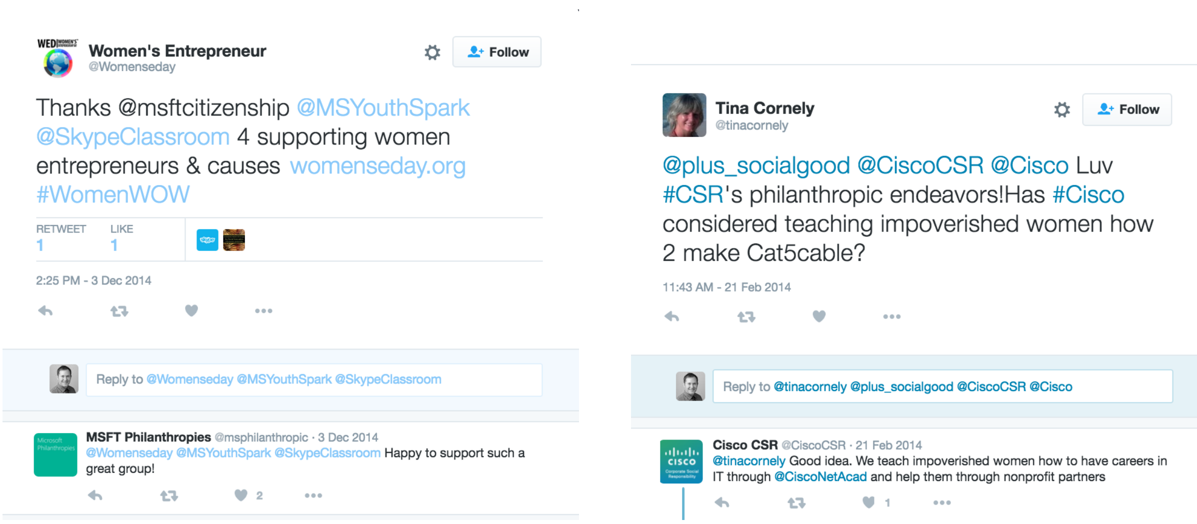


Figure 3.4: Sample Public Messages and Firm Responses

Figure 3.5 provides two additional examples of firm responses to public messages. The difference is the greater complexity of the thread of the ongoing conversation. Tweets are shown in the order in which they appear in the discussion thread. In the example on the left, the first tweet is a message by a user, @GBCEducation, as part of an online chat on CSR issues (this is indicated by the #CSRchat hashtag). The second tweet is written by @AlcoaFoundation (a firm in my sample) and is also part of the #CSRchat. The third tweet in the discussion thread – the large tweet in the middle – is treated as the "public message," given that it is the first in the thread to mention one of the 42 Fortune 200 accounts in the sample. The fourth tweet is the "firm response" to the third message, and the fifth message is @GBCEducation's response to the fourth tweet. To recap, in this paper I analyze the third, fourth, and fifth messages: the third and fifth messages constitute public messages mentioning one of the 42 CSR accounts while the fourth message constitutes a firm reaction to a public mention. The fifth message, it should be said, is considered to have been "ignored" by @AlcoaFoundation, as the tweet received neither a reply nor a retweet nor a favorite.

The messages on the right in Figure 3.5 similarly shows five tweets in the same discussion thread. The third and fourth tweets represent public messages mentioning one of the 42 CSR

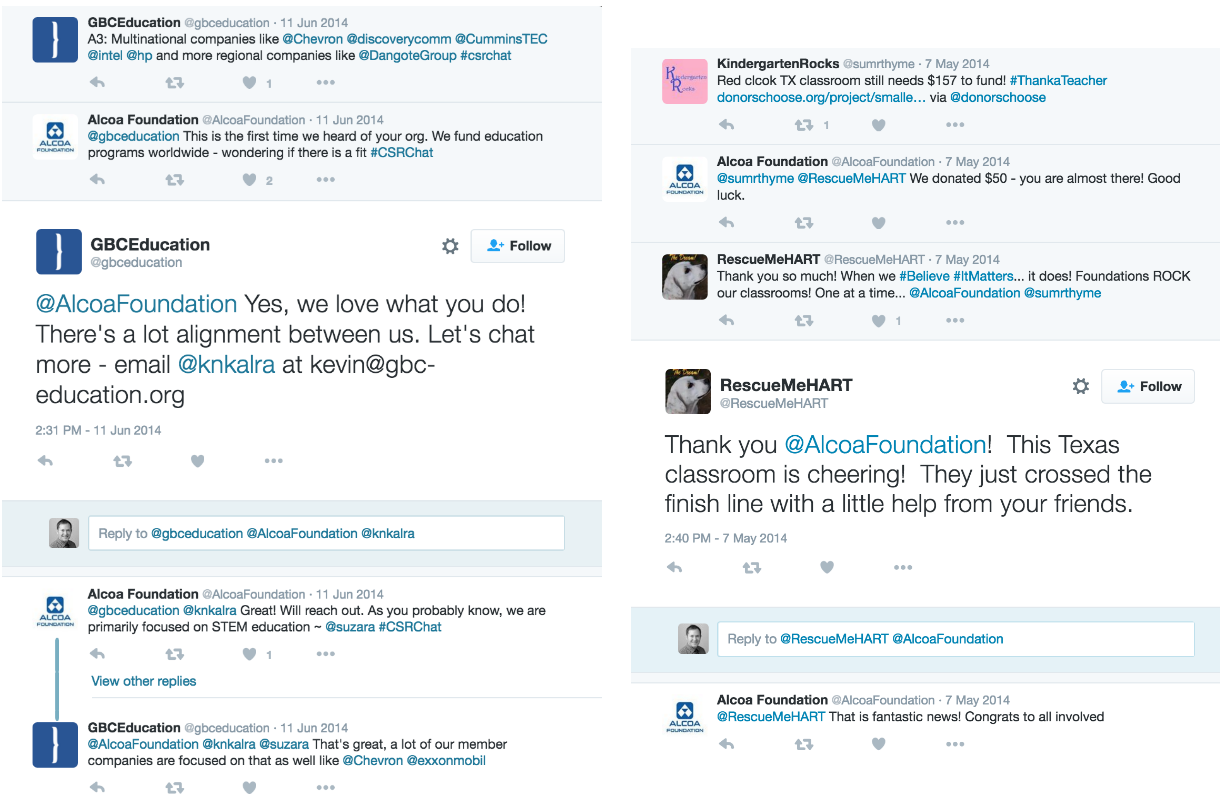


Figure 3.5: Sample Public Messages & Firm Responses – More Complex Discussion Threads

accounts (again, @AlcoaFoundation), while the fifth tweet represents @AlcoaFoundation's reply to the fourth message. Unseen here is that third tweet was favorited by @AlcoaFoundation. In effect, both public messages mentioning the @AlcoaFoundation message received a response – the first a "favorite" and the second a reply.

3.4.3 Reaction #2: Favoriting

As just noted, the third tweet on the right in Figure 3.5 was "favorited" (since 2015, this is called instead a "like" by Twitter) by @AlcoaFoundation. On the Twitter interface, who "liked" a message is visible by clicking on the "heart" icon underneath a given tweet. Figure 3.6 shows an example of what a Twitter user would see by clicking on the heart icon on that third message; indeed, it becomes visible that the message received one like, and that the like was by @AlcoaFoundation. The Twitter API allowed me to gather this information programmatically. Analogous to the process undertaken to determine which public messages

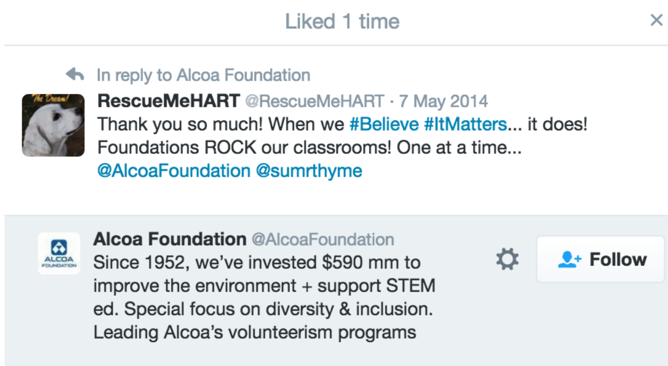


Figure 3.6: Favorite of a Message

received a reply, I was able to iterate over all of the IDs of the tweets favorited by the 42 CSR accounts over the course of 2014 and compare those to the IDs of the 163,402 public messages sent and thus determine which of those messages were favorited by one or more of the 42 CSR accounts.

3.4.4 Reaction #3: Sharing

Figure 3.7 shows a sample of one of the public messages, this one a mention of Dupont's CSR account @Dupont_ability. The bottom of the tweet shows that the message has received 2 retweets. Clicking on the hyperlink associated with the count (the number "2") will bring up a secondary screen, shown here on the right side of Figure 3.7. This shows the details of who has retweeted the message, and as shown in the figure, @Dupont_ability is one of the users that has responded.

As with replies and favorites, the Twitter API allowed me to do this programmatically for all 163,402 public mentions; in particular, I was able to gather the original tweet IDs of all the tweets retweeted by the 42 accounts over the course of 2014 and match those to the 163,402 public mentions of the 42 accounts, thereby providing a way of determining which of the 163,402 mentions were retweeted by one or more of the 42 accounts.

Figure 3.7: Sample Public Message with Retweet by Fortune 200 CSR Account

3.4.5 Reaction #4: Ignoring

The final reaction is simply to ignore the message. Perhaps not surprisingly, this was the most common reaction, occurring with 157,947 out of 163,402 tweets. Ostensibly, with "attention economies" such as social media, there may have to be something unique or pressing about a message in order for it to receive a reaction from a large Fortune 200 firm. Answering this question is a key goal for the current paper.

3.4.6 Summary of Reactions

Figure 3.8 visually summarizes the four reaction decisions possible for a firm upon seeing a message, along with counts of the number of messages that received each reaction. Going from left to right, 157,947 of the public messages are ignored, 1,961 of the messages are shared (retweeted), 2,961 of the messages are favorited (liked), and 1,095 of the messages received a reply.

For the purposes of model selection, all quantitative analyses in this paper employ a

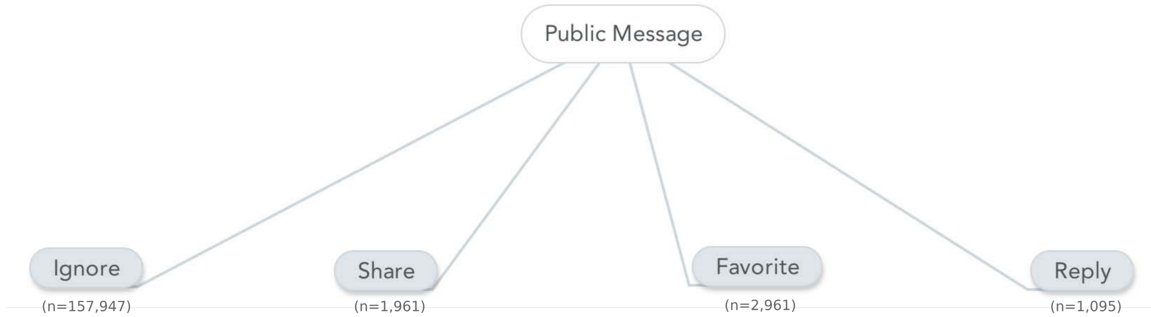


Figure 3.8: Decision Tree: Firm Reactions upon Reading Tweet from Member of the Public

binary version of the outcome variable, where values of "1" are applied to messages that receive either a like, a reply, or a retweet, and values of "0" are assigned to messages that are ignored.

3.5 *Who* Receives a Firm Reaction

The accountability literature has long raised the question of *to whom* firms and employees and nonprofit organizations are and should be accountable (Gray et al., 1987; Tower, 1993). To help understand to whom firms are paying attention in their social media-based CSR efforts, in this section I explore various features of the Twitter users sending public messages to the 42 CSR accounts.

In total there were 82,769 different Twitter users who mentioned one of the *Fortune 200* firms in their messages over the course of 2014. In this section I explore various characteristics of these message senders. There are a number of characteristics of the message sender that could be plausibly expected to influence whether a firm reacts. I explored five broad categories of sender characteristics suggested by the data: 1) features of the sender's Twitter activities, 2) characteristics of the sender's Twitter profile, 3) the gender of the message sender, 4) whether the sender was an organization or an individual, and 5) if the sender was an organization, the type of organization. I explore these categories – and their relationship to firm reactions – in turn.

3.5.1 Twitter Activity Levels

Twitter makes available five variables pertaining to the user's aggregate level of Twitter activity. First, there is the number of *followers*, or the number of other Twitter users who formally "follow" that user. Second, there is what Twitter calls the number of "friends", which reflects the number of other users followed by that user. Third, the number of *favorites* measures the number of tweets sent by other users that the user has "favorited;"[3] favoriting provides a means of both archiving a message and sending a signal to the message sender that the one favoriting the message finds the message valuable. Fourth, the number of *statuses* indicates the total number of tweets the user has sent since joining Twitter. Finally, the number of *lists* reflects the influence of the user; specifically, it reflects the number of public "lists" other Twitter users have added the user to.

Figure 3.9 illustrates these five variables in showing a screenshot of the Twitter user who sent one of the tweets shown in Figure 3.1. As shown in the middle of the figure, any visitor to the page can see five count variables that reflect the different aspects of the sender's Twitter behavior. First, *Tweets* shows the cumulative number of tweets the user @womenseday has sent. At close to 5,883 tweets over two years, the user on average sends out roughly seven messages per day, indicating an active tweeter. Second, *Following* shows the number of other Twitter users @womenseday chooses to follow. In general, the more users followed, the more the user is sending a signal that she is interested in reaching out to other users and learning from a broader community (Lovejoy et al., 2012). The relatively high number of 5,881 users followed suggests @womenseday is interested in outreach. Third, *Followers* indicates the number of other Twitter users who are following the account. At

[3] In 2015 this behavior became known instead as "liking."

Figure 3.9: Screenshot of Sender of Message Mentioning a Fortune 200 CSR Account

around 31,100 followers in early 2016, @womenseday is well above the average of the typical user on Twitter. Fourth, *Likes* indicates the number of tweets @womenseday has favorited (now called "liked"), an activity that serves both an archive function and a signal to the original message sender that @womenseday found the message useful in some regard. Liking is another means of connecting with a broader set of users. With 2,302 tweets *liked*, the implication is that @womenseday is actively reading and favoriting other users' messages. Fifth, *Lists* shows that @womenseday is subscribed to 1 public Twitter "list," called *WED regional ambassadors*, which was created by @womenseday. Twitter users who click on the Lists link can also see the number of public lists @womenseday is a member of. These are lists that other Twitter users have created and placed @womenseday on; in effect, the higher the number of lists @womenseday is a member of, the greater its level of notoriety and/or influence. In early 2016 @womenseday was on roughly 250 public lists, a relatively high number (the mean is 104).

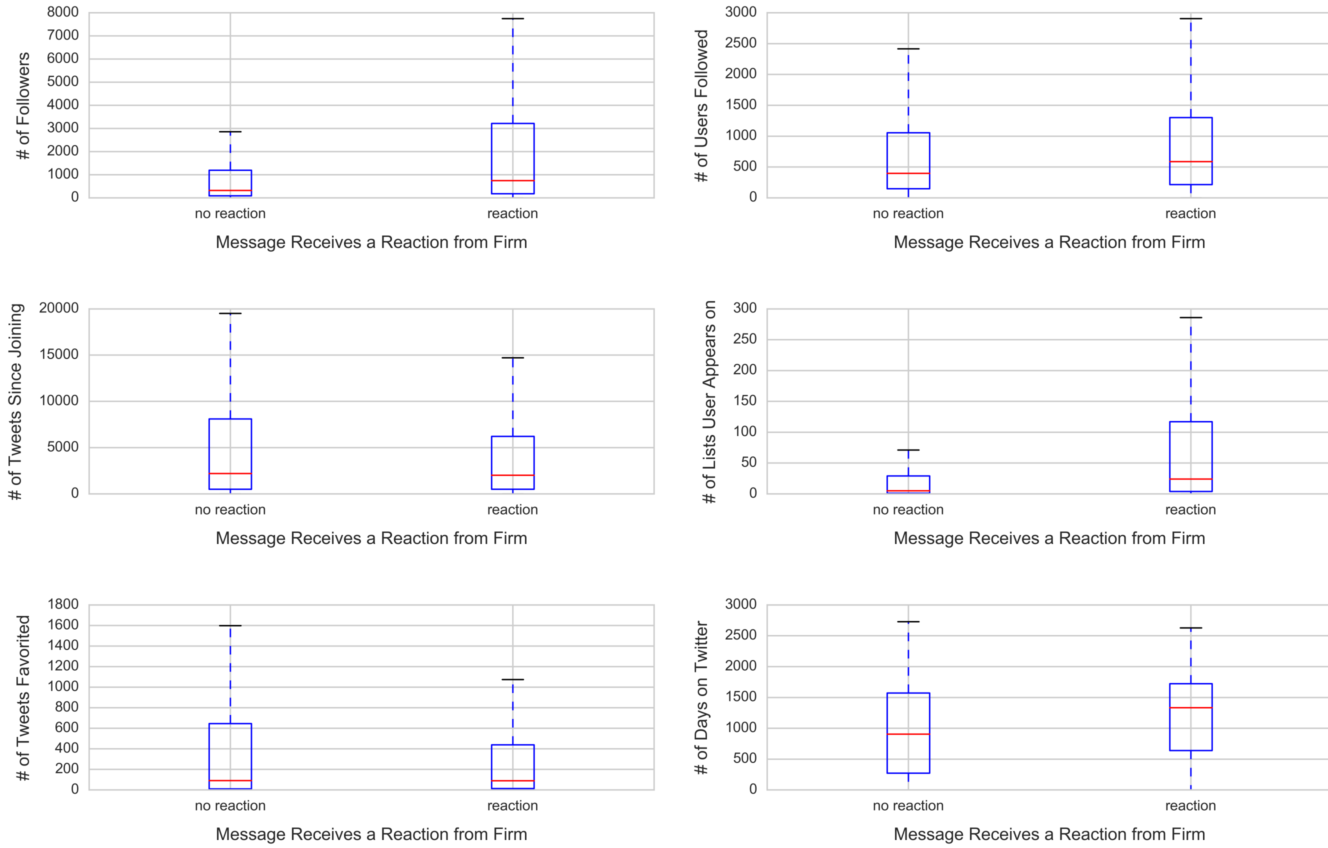


Figure 3.10: Average % of Messages Receiving a Firm Reaction based on Profile Features

Relationship between Twitter Activity and Firm Reactions

Figure 3.10 shows the proportion of messages that receive a reaction according to the Twitter activity of the message sender. Each of the above five variables are ratio-level variables; accordingly, *box plots* (aka "box and whisker plots") are shown.[4] Each plot shows the distribution of the variable first for messages that do not receive a reaction (a favorite, a retweet, or a reply) from one of the 42 Fortune-run CSR accounts and second for messages that do receive a reaction. Comparing these distributions allows us to visually inspect whether the variable appears to be related to the probability of receiving a reaction.

A visual inspection of the figure suggests positive relationships between firm reactions and the public message senders' number of followers and number of friends as well as the

[4]In a box plot, the variable is divided into quartiles. The first to third quartiles are shown in the "box," the horizontal red line in the box indicates the median value, and the vertical lines (the "whiskers") extending to the horizontal bars at the top and bottom indicate the most extreme observations not considered outliers.

number of lists the user appears on. There appears to be a slight negative relationship, in contrast, with the number of tweets the user has favorites and the cumulative number of tweets the user has sent since joining Twitter.

3.5.2 Twitter Profile Characteristics

A second category of variables reflects characteristics of the user's official Twitter profile. The first of these, *Time on Twitter,* indicates the number of days since the user created their Twitter account. Second, there is the language of the user, specifically, whether or not the tweets sent are in English. Third, I examine what Twitter refers to as an "extended" profile, a relatively new feature that allows for additional data to be included on the user's profile page. A fourth feature, unique to Twitter, is what is known as a "verified" account; these accounts are noted by the additional of a unique icon on the user's profile page, and indicate the user's status as a noteworthy individual or organization such as a celebrity, entertainer, journalist, company, author, or sport star. Finally, I examine whether the user has generated a custom profile. This variable reflects whether the user has customized their profile page or simply used the default. Using the default settings indicates a lack of sophistication and, to many users, could further indicate a "spammy" or "Twitterbot" account that is not worthy of engagement.

Relationship between Twitter Profile Characteristics and Firm Reactions

Time on Twitter is a continuous variable and thus is shown in the box-and-whisker plot in Figure 3.10. The figure suggests a positive relationship between the length of time a user has been on Twitter and the likelihood of the user's message receiving a reaction from a *Fortune 200* firm.

The other four variables analyzed are all discrete binary variables and are thus analyzed visually not through box plots but through the bar graphs shown in Figure 3.11. For each variable, the figure shows the proportion of messages that receive a reaction when the value

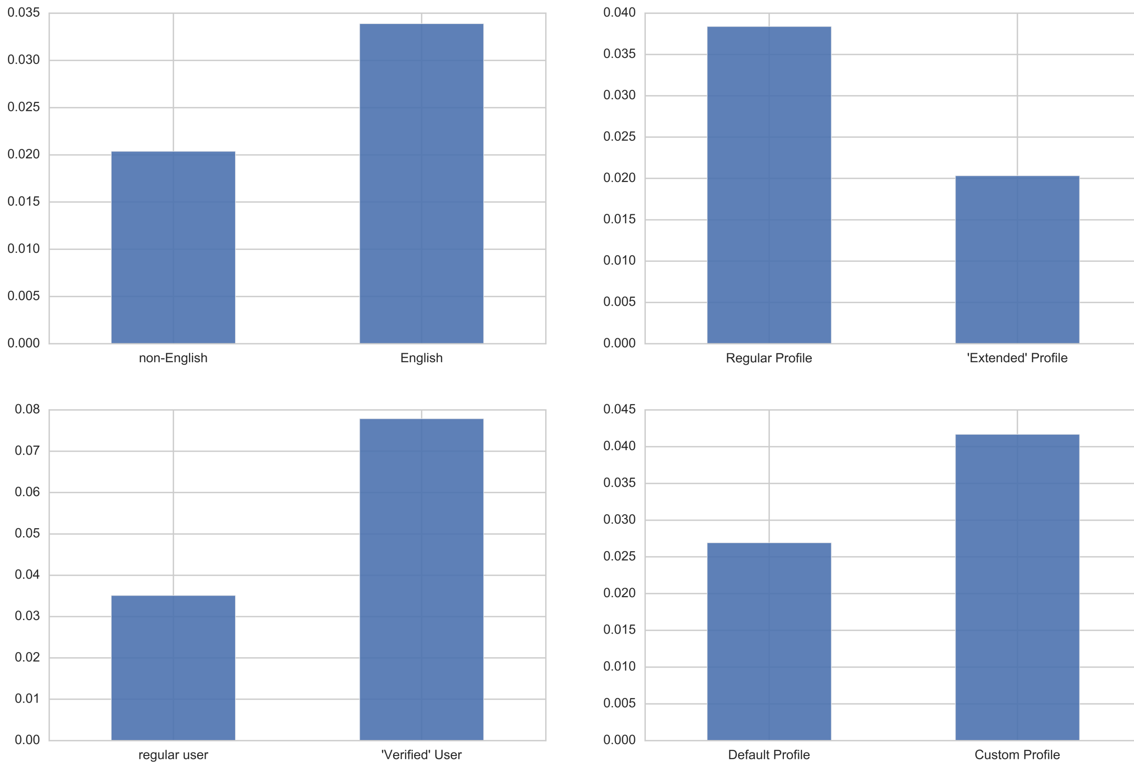


Figure 3.11: Average % of Messages Receiving a Firm Reaction based on Profile Features

of the independent variable is "0" and when it is "1." The first of the four variables is language, with the figure showing, not surprisingly given the firm's home locations, that tweets in English are more likely to receive a reply. Second, regarding "extended" profiles, the bar graph in Figure 3.11 suggests extended profiles are not related to a higher likelihood of obtaining a firm reaction (in fact, there is a negative relationship). Third, comparing messages from verified users compared to those from non-verified users shows messages from the former are much more likely to receive a reaction. The last set of bar graphs in Figure 3.11 suggests messages from users with custom profiles are more likely to receive a reaction from the *Fortune 200* firms.

3.5.3 Gender

Accounting scholars have linked the issue of gender equality with that of CSR, finding that, among other things, firms are now including information on their gender equality performance in their external reporting efforts (e.g., Grosser & Moon, 2008). It is therefore possible firms will attempt to strike a rough balance in their "micro-accountability" efforts on Twitter by responding equally to males and females. To explore this idea, the sex of each user who made one of the 163,402 mentions in 2014 was calculated by running users' first names through a gender classifier algorithm using Python code. Reliable codes were possible for 54,761 of the messages sent, with the remainder sent by organizations or by users with pseudonyms, no names, or names where sex could not be determined.[5]

Relationship between Gender and Firm Reactions

The first set of bar graphs in Figure 3.12 show the proportion of messages that receive a reaction for both male and female senders. There appears to be only a modest difference between the two groups, with females receiving a slightly higher reaction rate.

3.5.4 Individuals vs. Organizations

The final set of characteristics examined deal with whether the sender is an individual or an organization. To code this, first, 700 message senders were manually coded as being an organization or an individual, then a support vector machine (SVM) machine learning algorithm (e.g., Go et al., 2009) was trained and tested and used to code the remainder of the 82,769 users. The trained SVM algorithm achieved 88.9% accuracy compared to the manually coded users.

Organization Type

Lastly, I explored the relationship between the type of organization mentioning a *Fortune 200* firm in their tweets and the receipt of a reaction from the *Fortune 200* firm. As a

[5]The code is only sophisticated enough to identity "male" vs. "female" names. Given the self-described nature of the Twitter profiles, in most cases the name indicates the biological sex of the sender, but in other cases it will indicate self-identified gender identity. The high number of missing values is a result of a combination of missing names, pseudonyms, and names that are identified with multiples genders.

proxy for organization type I used the website domain extension, whereby organizations with a ".com" extension are assumed to be for-profits, organizations with a ".edu" extension are assumed to be learning institutions, those with ".gov" are governments, and ".org" are nonprofit organizations.

Relationship between 'Entity' Type and Firm Reactions

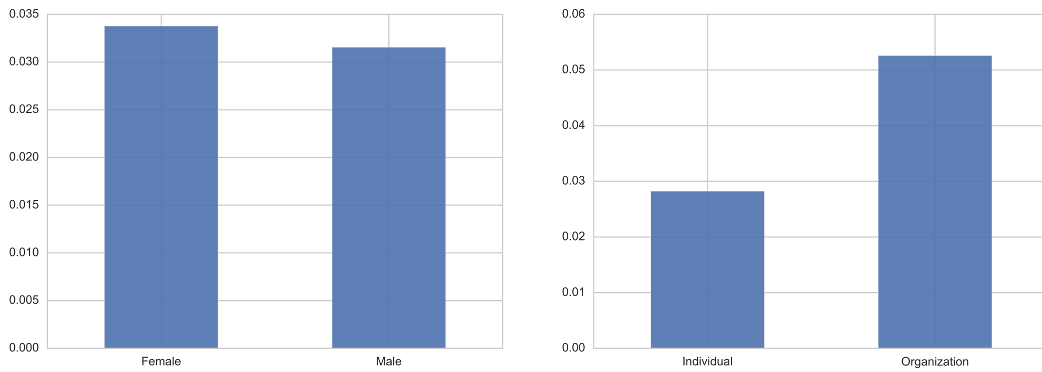
As shown in Figure 3.12, the proportion of messages that receive a reaction is higher for organizations than it is for individuals. Figure 3.12 also shows variation in the degree to which messages from the four organization types receive a reaction from the *Fortune 200* firms. Specifically, educational and nonprofit organizations have roughly double the reaction rate of for-profit and governmental organizations.

3.5.5 Bivariate Statistics and Correlations

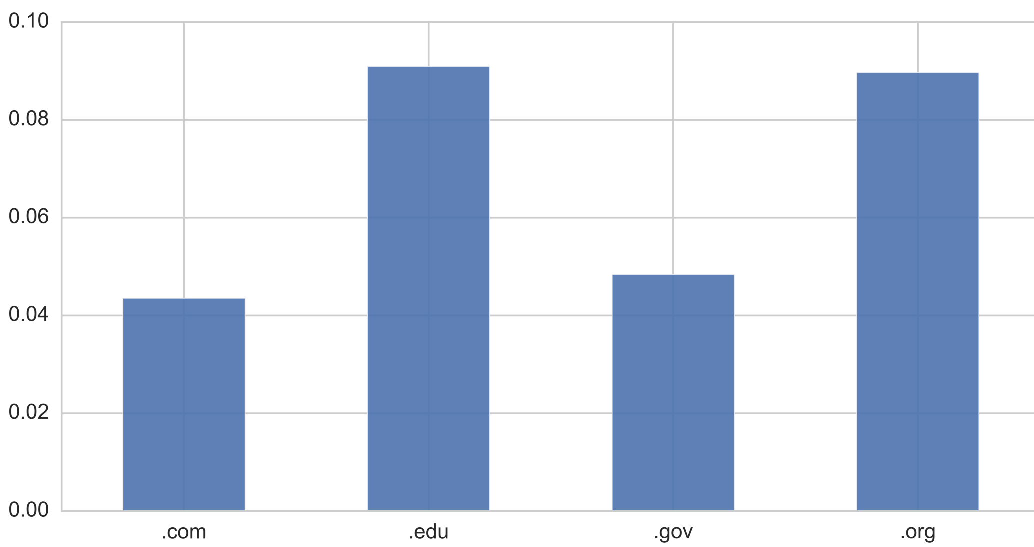
Above I have examined visually a number of potentially relevant variables related to *who* sent the message. As noted above in the Method section, in order to help identify the most relevant of the identified variables, I employ a number of multivariate *feature selection* techniques. The bivariate graphical analyses presented above are one relevant method, and in this section I summarize results from multivariable correlations and bivariate Chi-square and logit tests (see Guyon & Elisseeff, 2003). Similar analyses will be conducted in each of the remaining analysis sections.

Specifically, to further explore the relationship between Fortune reactions and sender characteristics, Tables 3.2 and 3.3 submit these variables to bivariate statistical tests. First, for the binary independent variables, Table 3.2 reports results from a series of Chi-square tests. All save for the measures of gender (*Male*) and governmental (.gov) organizations obtained a significant association with receiving a reaction.

Table 3.3, meanwhile, shows logit tests for the six continuous independent variables examined above. All are significant. Not surprisingly, the number of followers, the number



(a) *Reactions by Sex and by Individual vs. Organization*



(b) *Reactions by Organization Type*

Figure 3.12: Average % of Messages Receiving Firm Reactions for Sender Characteristics

Table 3.2: Chi-square Tests for Binary Variables - D.V. is *Fortune Reaction (0,1)*

variable	mean score, I.V. value=0	mean score, I.V. value=1	no. of obs., I.V. value=1	χ^2	sign	n
Male	0.034	0.032	29305	2.06		54761
Organization	0.028	0.053	34391	498.60**	+	163,374
.com organization	0.050	0.044	9217	6.23*	−	58,605
.edu organization	0.048	0.091	308	11.01**	+	58,605
.gov organization	0.049	0.048	124	0.04		58605
.org organization	0.044	0.090	6256	252.83**	+	58,605
English	0.020	0.034	157265	33.06**	+	163,402
Extended profile	0.038	0.020	12148	99.60**	−	104,616
'Verified' profile	0.035	0.078	2863	144.39**	+	104,616
Custom profile	0.025	0.042	71381	185.67**	+	104,616

*p<.05, **p<.01

of followees (friends), the number of days the user has been on Twitter, and the number of public lists the user is on are all positively associated with their messages receiving a reaction. On the other hand, it is somewhat surprising that two measures of Twitter activity – the cumulative number of statuses sent and the number of tweets that have been favorited – both obtained a negative coefficient with *Fortune reaction.*

Table 3.4 shows the zero-order correlations for all variables analyzed in this section. For the most part, the variables are not highly related. One exception is the number of followers and the number of lists on which the user appears. These variables have a pearson correlation coefficient of 0.78.

3.5.6 Summary of *Who* Messages the Firms

In this section I have covered a large number of potentially relevant variables. Too many variables, of course, to employ in a multi-variate analysis. The first goal, therefore, is parsimony. Beyond this, the current number of variables is not optimal for several reasons. One, we are not so much interested in variables as we are in concepts. Some of the variables have

Table 3.3: Logit Tests of *Fortune reaction* on Interval-level I.V.s

variable	coeff.	sign	n
from_user_followers_count	0.000004**	+	163,402
from_user_friends_count	0.000065**	+	163,402
from_user_statuses_count	−0.000091**	−	163,402
from_user_listed_count	0.001220**	+	163,402
from_user_favourites_count	-0.000219**	−	163,402
time_on_twitter_days	0.005535**	+	163,402

*p<.05, **p<.01

Table 3.4: Zero-Order Correlations Matrix: *Who*

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11	12.	13.	14.	15.	16.
1. # of Followers	1.00	0.10	0.03	0.78	0.00	0.07	0.01	0.01	0.01	-0.00	0.00	-0.00	0.01	-0.01	0.24	0.03
2. # of users followed	0.10	1.00	0.18	0.19	0.07	0.12	0.01	0.04	0.00	-0.01	-0.01	0.01	0.01	-0.01	0.09	0.07
3. # of tweets	0.03	0.18	1.00	0.07	0.37	0.12	0.02	-0.04	-0.03	-0.01	-0.01	-0.08	0.02	0.08	0.02	0.09
4. # of lists	0.78	0.19	0.07	1.00	0.01	0.15	0.01	0.04	0.01	-0.00	0.01	0.02	0.02	-0.02	0.29	0.06
5. # of favourites	0.00	0.07	0.37	0.01	1.00	0.01	-0.03	-0.03	-0.00	-0.01	-0.01	-0.02	0.00	0.08	-0.01	0.02
6. Time on Twitter	0.07	0.12	0.12	0.15	0.01	1.00	0.01	0.10	-0.03	-0.01	-0.01	0.05	0.03	0.02	0.13	0.32
7. Male	0.01	0.01	0.02	0.01	-0.03	0.01	1.00	0.02	0.04	0.01	0.01	-0.03	-0.02	0.02	0.04	-0.03
8. SVM_org	0.01	0.04	-0.04	0.04	-0.03	0.10	0.02	1.00	0.64	0.11	0.07	0.51	0.04	-0.11	0.06	0.11
9. .com_org	0.01	0.00	-0.03	0.01	-0.00	-0.03	0.04	0.64	1.00	-0.03	-0.02	-0.15	0.03	-0.06	0.03	0.01
10. .edu_org	-0.00	-0.01	-0.01	-0.00	-0.01	-0.01	0.01	0.11	-0.03	1.00	-0.00	-0.03	0.01	-0.02	-0.02	-0.01
11. .gov_org	0.00	-0.01	-0.01	0.01	-0.01	-0.01	0.01	0.07	-0.02	-0.00	1.00	-0.02	0.00	-0.01	0.07	0.01
12. .org_org	-0.00	0.01	-0.08	0.02	-0.02	0.05	-0.03	0.51	-0.15	-0.03	-0.02	1.00	0.04	-0.10	0.01	0.04
13. English	0.01	0.01	0.02	0.02	0.00	0.03	-0.02	0.04	0.03	0.01	0.00	0.04	1.00	-0.00	0.01	0.03
14. ‘Extended’ profile	-0.01	-0.01	0.08	-0.02	0.08	0.02	0.02	-0.11	-0.06	-0.02	-0.01	-0.10	-0.00	1.00	-0.04	0.07
15. ‘verified’ account	0.24	0.09	0.02	0.29	-0.01	0.13	0.04	0.06	0.03	-0.02	0.07	0.01	0.01	-0.04	1.00	0.10
16. custom profile	0.03	0.07	0.09	0.06	0.02	0.32	-0.03	0.11	0.01	-0.01	0.01	0.04	0.03	0.07	0.10	1.00

notable conceptual overlap. Two, some of these variables may have strong inter-correlations.

Accordingly, I will first suggest some of the above variables can be organized into broader dimensions of determinants. First, *language* and *gender* are uni-dimensional and cannot be folded into meaningful higher dimensions. At the same time, the analyses above suggest the latter is not an important determinant of firm reactions and can thus be omitted from a more parsimonious theoretical model.

A second dimension is the type of *entity* represented by the user. While the analyses do suggest differences in reaction according to the type of organization sending the message, the more important first difference is whether the sender is an individual or an organizational entity. For the sake of parsimony, initial theoretical models should focus on this distinction.

The remainder of the variables are derived from Twitter data, which I divide into four dimensions. Namely, the number of users followed, the number of tweets sent, and the number of tweets favorited constitute three discrete indicators of the aggregate level of *Twitter activity*. In contrast, two other measures – the number of followers and the number of lists the user appears on – do not represent the sender's activity but rather the Twitterverse's reactions to the user's efforts. Namely, the number of followers a user acquires and the number of lists on which the user appears are both broad reflections of the *prestige* or *influence* the user has acquired on Twitter.

With respect to the Twitter profile variables, the new variable "extended profile" does not appear to be informative; in contrast, the other two profile variables – time on Twitter and custom profile – can be considered to tap the perceived *sophistication* of the user with respect to Twitter. Relatedy, the variable "verified profile" reflects some degree of *celebrity status* of the user; while verified accounts are given to those who are chiefly "online" celebrities, it is more typically indicative of an individual or organization with offline notoriety. Some of the 42 accounts in this study, for instance, are verified accounts (e.g., @DellEDU).

In sum, my analyses of the data suggest six relevant dimensions of sender characteristics: language, type of entity, Twitter activity, Twitter sophistication, Twitter prestige, and celebrity status.

3.6 *What* Receives a Firm Reaction

In this and the following two sections I turn to analyzing characteristics of the messages that are sent, and the relationships between these characteristics and the likelihood of the message receiving a firm reaction. Specifically, I look at four categories of message features: 1) the "entities" included in the messages, 2) message sentiment, 3) the originality of the message and its location in existing conversation threads, and 4) the topics covered by the message.

3.6.1 Entities included in the message

Most social media platforms facilitate the inclusion of specific types of *entities* in the social media messages, including on Twitter hashtags, users, hyperlinks, and images, with research generally finding a positive relationship between the inclusion of such entities and audience reactions (e.g., Bakshy et al., 2011; Saxton & Waters, 2014). I examine each of these four features.

To start, hashtags, or words prepended by the pound sign such as #*GirlRising* or #*IceBucketChallenge*, are used to denote topics and classify messages (O'Leary, 2015; Debreceny, 2015). Given the enhanced classification and searchability afforded by hashtags, studies have found them to be related to the extent which messages are seen and diffused on Twitter (e.g., Saxton et al., 2015). I thus created a binary measure that indicates whether a message includes one or more hashtags.

Second, I generated a binary measure to reflect the use of user mentions in the message. On Twitter, user mentions are indicated by the "@" symbol occurring before a Twitter username. User mentions are highlighted within the tweet and hyperlinked to the mentioned user's account. For instance, a tweet with "@CiscoCSR" would be considered a user mention of Cisco's CSR-focused Twitter account *@CiscoCSR*. Importantly, this user mention would formally link the tweet to the @CiscoCSR account, notifying Cisco that it has been mentioned by another Twitter user. In fact, all 163,402 messages include at least one user mention, for only those tweets that mention one or more of the 42 CSR accounts are included in the sample. Consequently, the binary variable I have created assigns a value of "1" if the message includes one or more mention of a user outside the sample of 42 accounts, otherwise "0."

Third, I generated a binary variable to indicate tweets that include one or more URLs (hyperlinks).

Fourth, given prior research showing a strong relationship between the inclusion of images and audience reactions (e.g., Saxton & Waters, 2014), I measure whether the tweet includes a photo.

Relationship between the inclusion of entities and firm reactions

Figure 3.13 shows the proportion of messages receiving a reaction that include a hashtag, a non-Fortune 200 user mention, a photo, or a URL. The graphs show, surprisingly, that each of these four entities is negatively related to a *Fortune 200* reaction. Perhaps firms view such entities as "noise."

3.6.2 Sentiment

I also posited that the sentiment of a given message might be related to the likelihood a firm reacts. Several recent CSR studies have examined sentiment in social media messages (Castelló et al., 2015; Colleoni, 2013), and (Saxton, 2016b) examined the relationship between the content of firms' CSR messages and the sentiment in public replies to those messages.

To measure sentiment I relied on a supervised machine learning technique as in existing computer science studies (Bollen et al., 2006; Go et al., 2009). Specifically, a hand-coded dataset of 1,000 tweets was used to train an SVM model. The trained model was then used

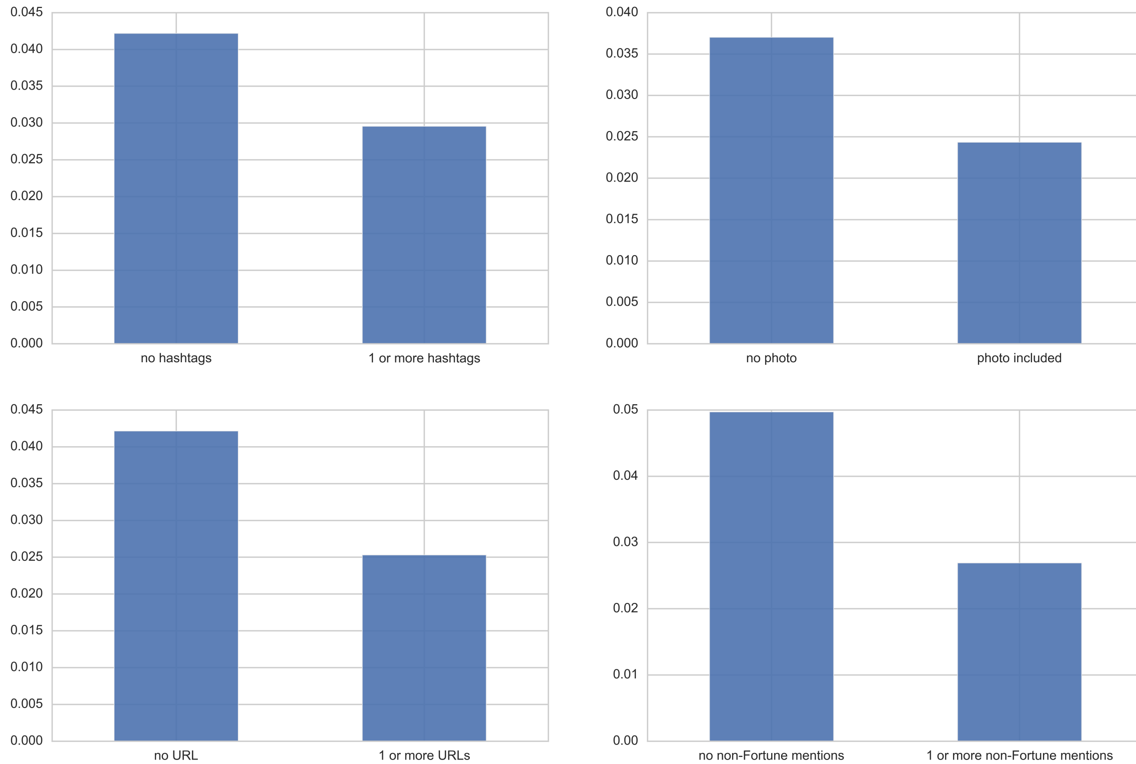


Figure 3.13: Average % of Messages with Different 'Entities' that Receive a Firm Reaction

to code each tweet as either negative, neutral, or positive sentiment. The level of accuracy compared to hand coding for these three values was 94.5%, 80.2%, and 89.0%, respectively. I found the great majority of messages (130,427, or 79.8%) were neutral in terms of sentiment; 28,948 (17.7%) had a positive sentiment, and 4,027 (2.5%) had a negative sentiment.

Relationship between sentiment and firm reactions

Figure 3.14 shows the proportion of messages that receive a reaction according to the sentiment of the message. The figure suggests a clear relationship: negative messages are least likely to receive a reaction, neutral messages are more likely to receive a reaction, and positive messages have a much stronger likelihood of receiving a reaction.

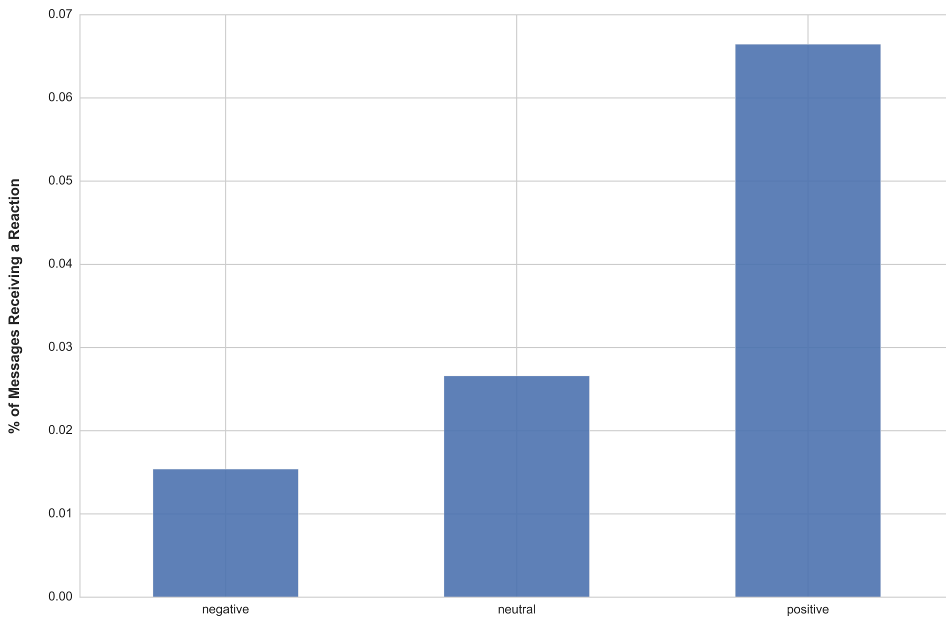


Figure 3.14: Average % of Messages Receiving a Firm Reaction based on Sentiment

3.6.3 Message Originality & Location in Existing Conversation Threads

I also found messages differ substantially in terms of where they lie in terms of existing conversations. When a user initiates a message, that message is either a *retweet* – an existing message the user wishes to share with others – or is an original message. And if it is an original message, it can either be a reply to an existing tweet (what I call a *public direct reply*) or initiate a new conversational thread. Messages beginning a new thread can either be a *public direct message* or be an "undirected" message (what I refer to here as a *mention only* message). The former are indicated by how they begin with "@USER" (e.g., *@CiscoCSR*), which indicate that the tweet is directed at the specified user and will be visible not only to the target but also all of the sender's followers and anyone following the targeted user. As such, they constitute a form of "public email."

Figure 3.15 depicts these different types of messages. In total, members of the public sent 163,402 community-initiated messages over the course of 2014, of which 58,186 are original and 105,216 are retweeted messages. Of the original messages, the three main types were

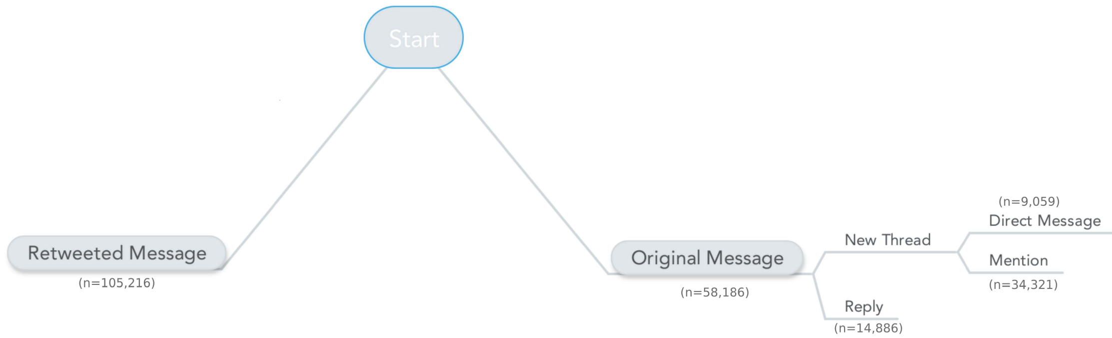


Figure 3.15: Retweeted vs. Original Messages, with Type of Message if Original

distributed as follows: 1) there were 14,886 PDRs, or public direct replies; 2) there were 9,059 PDMs, or public direct messages; and 3) and there were 32,241 "mention only" messages, or those that only mention one of the 42 CSR accounts anywhere except at the start of the message. As noted above, what differentiates the last type of message from the other two is that it represents an undirected form of message – a message that mentions or discusses one of the 42 CSR accounts but is not necessarily directed only at that user.

Relationship between originality/thread location and firm reactions

Figure 3.16 shows the proportion of messages that receive a reaction according to the originality and thread location of the message. One core finding is that firms rarely interact with retweeted messages. In fact, this is likely the strongest relationship of any examined in this study. In light of this fact, it is thus not surprising that PDRs, PDMs, and "mention only" tweets are all positively related to firm reactions.

3.6.4 Topics

Saxton et al. (2016) found a strong relationship between firms' use of CSR topics in their social media messages and audience reactions among members of the public. It thus seems likely that *firms* may likewise react differently according the topics discussed in the public's

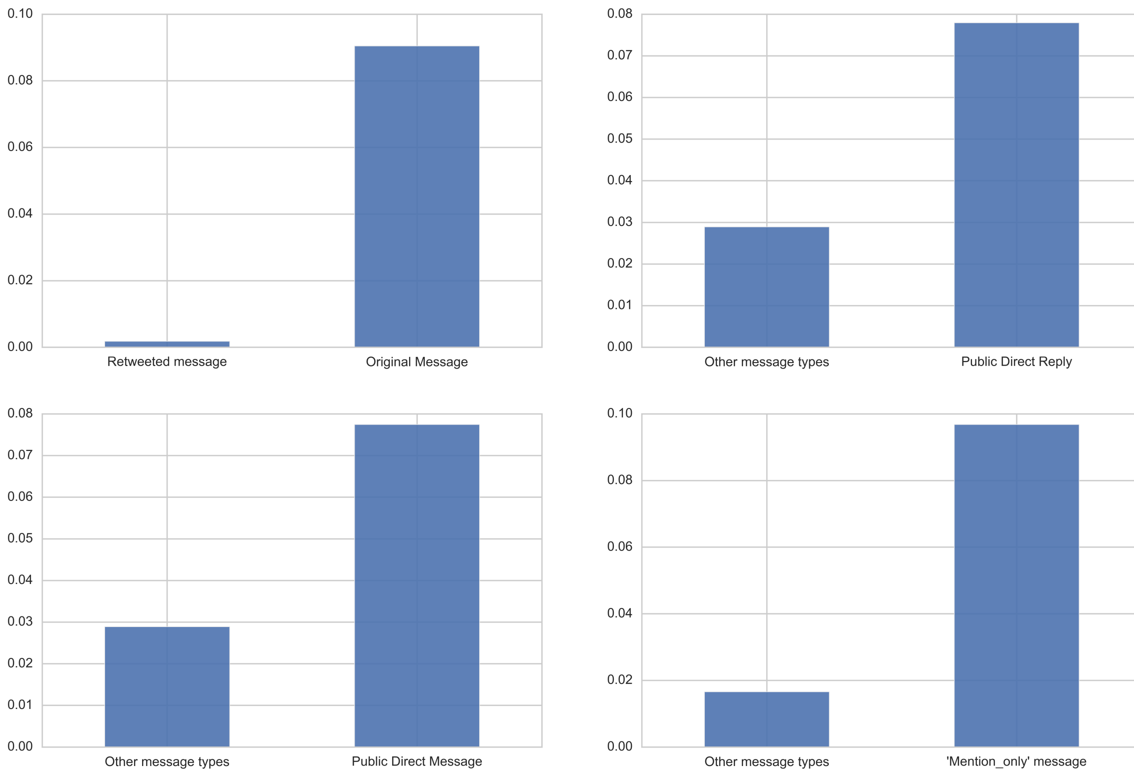


Figure 3.16: Average % of Messages Receiving a Firm Reaction based on Originality

messages. Fortunately, Big Data techniques have recently been developed and refined that facilitate the semi-automatic, or supervised, discovery of latent topics in a set of documents. In the machine learning community this is known as *topic modeling.* A popular technique is what is known as latent Dirichlet allocation (LDA) topic modeling (e.g., Evans, 2014).

All machine learning-driven topic models are algorithms for discovering latent themes in a given body of documents, or *corpus.* They can be especially helpful in discovering latent themes in the unstructured textual data that is typical of Big Data. In so doing, the algorithms can help organize the documents according to the topics and themes that are discovered. In LDA topic modeling, the algorithm "assumes that topics are latent patterns of words in the corpus, and calculates such topics as a probability distribution over words" (Evans, 2014, p. 2).[6] In effect, the LDA algorithm assumes the *topic structure* of a corpus of documents is hidden; the algorithm is employed to infer the unobserved topic structure from the observed documents. Formally, a "topic" is a "distribution over a fixed vocabulary" (Blei, 2012, p. 78).

One of the most important practical decisions made by the researcher is the number of topics to select. As a first step, I calculated several metrics to help identity the range of latent topics. In particular, two measures were used to identify the optimum number of topics. Specifically, each of these methods seek to find extremum, with Cao et al. (2009) used as a *minimization* metric and Griffiths & Steyvers (2004) used as a *maximization* metric. These metrics help find the optimum number of topics used in the LDA models. Figure 3.17 shows the output of these metrics across a range of number of topics from 2 to 100. The Cao and Griffiths metrics both suggest the ideal number of topics is somewhere between 10 and 30. Two other metrics, Arun et al. (2010) and Deveaud et al. (2014), were not informative in helping select the natural number of topics.

In a second stage, topic solutions were validated qualitatively. Specifically, I analyzed the top 50 words generated for each topic to evaluate the coherence of each topic. This was

[6] LDA is thus one type of the broader field of *probabilistic modeling* (Blei, 2012).

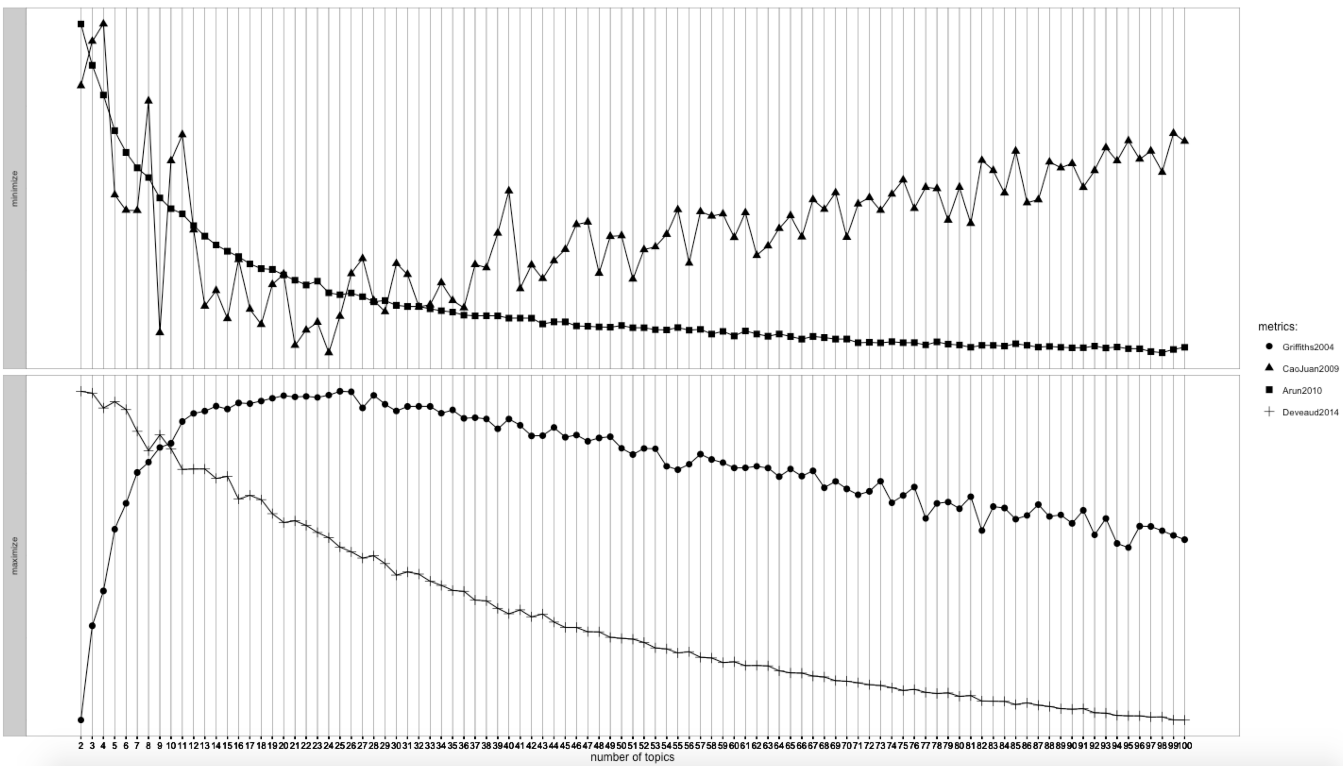


Figure 3.17: LDA metrics

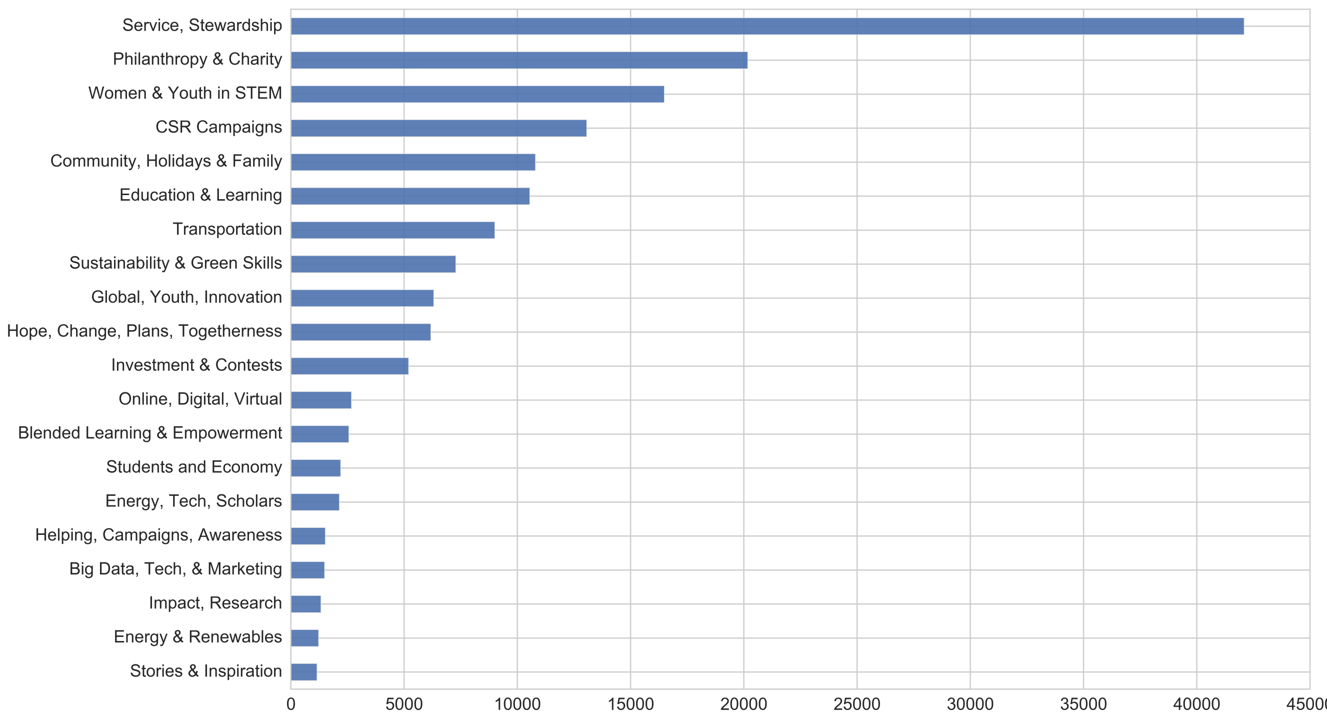


Figure 3.18: Frequency of Topics

done for topic solutions across the range from 10 to 30 topics in increments of 5. Ultimately, my inductive analyses suggested the optimal number of topics was 20. I then assigned labels to each of the 20 topics, and then used the trained LDA model to determine the primary topic for each tweet in the dataset. Figure 3.18 shows the frequencies for the 20 topics in the 163,402 tweets.

Relationship between message topic and firm reactions

Figure 3.19, meanwhile, shows the average proportion of messages that receive a reaction for each of the 20 topics. The reaction rates range from just under 2% for the topic *CSR campaigns* to just under 5% for *Big Data, Tech, & Marketing.* The broad insight, however, is firm reactions appear to vary according to the topic addressed in a public message.

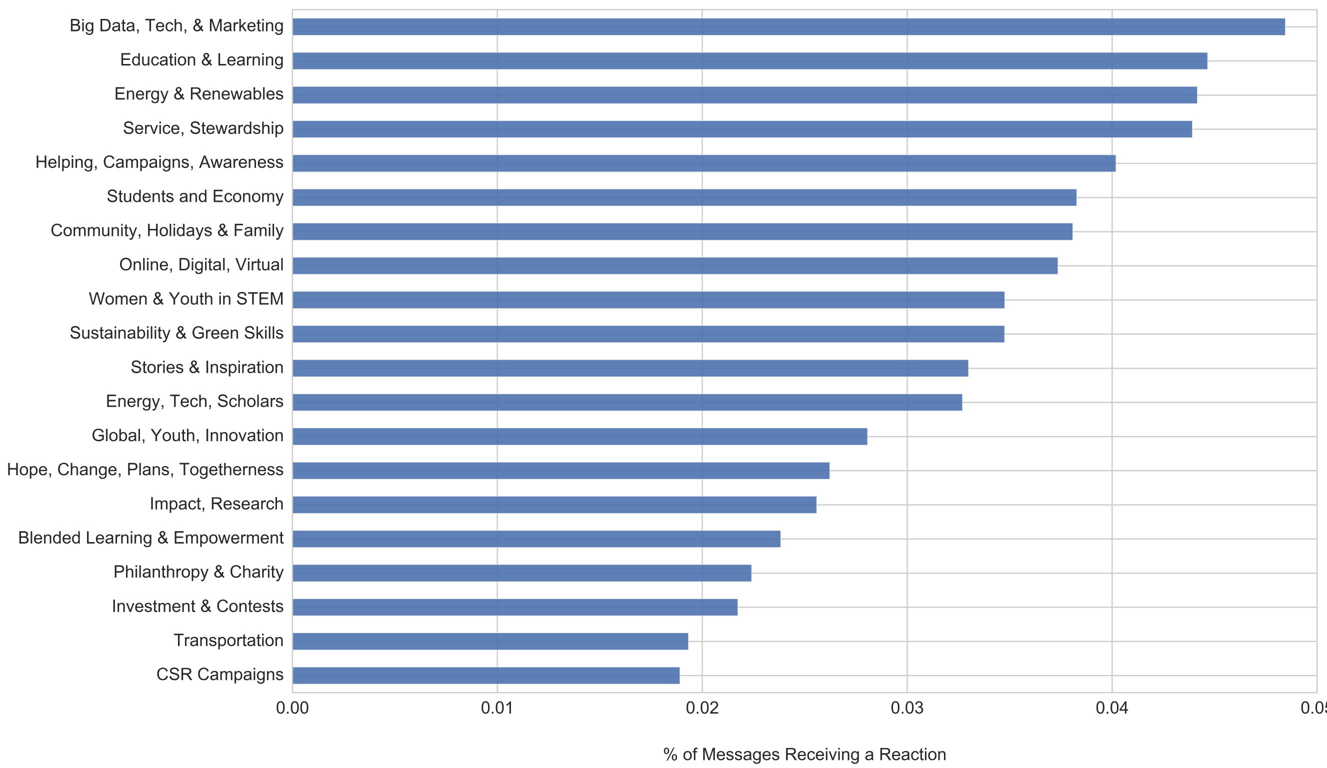


Figure 3.19: Average % of Messages Receiving Reaction, by Topic

3.6.5 Bi-variate Statistics and Correlations

All of the variables examined in this section are binary variables. Consequently, to test the bivariate statistical relationship with *Fortune reaction*, Chi-square tests were used. Table 3.5 reports the results from these Chi-square tests. All measures except for 8 of the 20 topic dummies are significantly related to a message receiving a reaction from one of the 42 *Fortune* accounts.

Table 3.6, meanwhile, shows the correlations among all the above variables save for the topic variables. All pearson correlations are less than 0.33.

3.6.6 Summary of *What* the Public Includes in their Messages

Based on the above analyses, I posit the following dimensions of *what* large firms react to. A first dimension is the entities include in the messages. I find that all four entities I examined – hashtags, hyperlinks, user mentions, and photos – were negatively related to the propensity of firms to not ignore the message. Accordingly, it is likely the four variables may be amenable to combining in a single entity index. Sentiment comprises a second dimension of public messages, one that is potentially one of the stronger determinants of firm reactions. A third dimension, topic, appears to be relevant but not so important as sentiment. A fourth dimension is the originality of the message. Lastly, a sub-dimension of originality is the location of the message in existing discussion threads.

3.7 *When* a Message Receives a Firm Reaction

In this section I analyze when firms react to public mentions. In the accounting literature the examination of timing has generally been quite restricted, with research limited to such questions as disclosure timing around specific events, such as the "timeliness" of reporting (Atiase et al., 1989). Social media represent a form of "Big Data," with one of the key impli-

Table 3.5: Chi-Square Tests for Binary Variables - D.V. is *Fortune Reaction (0,1)*

variable	mean score, I.V. value=0	mean score, I.V. value=1	no. of obs., I.V. value=1	χ^2	sign	n
Original Tweet	0.002	0.090	58,186	9127.90^{**}	+	163,402
PDR	0.029	0.078	14,886	1005.42^{**}	+	163,402
PRM	0.029	0.077	15,094	999.13^{**}	+	163,402
"Mention only" msg.	0.017	0.097	34,241	5398.44^{**}	+	163,402
Hashtag	0.042	0.030	113,806	170.38^{**}	−	163,402
Photo	0.037	0.024	46,843	166.18^{**}	−	163,402
URLs	0.042	0.025	85,086	357.61^{**}	−	163,402
non-Fortune mention	0.050	0.027	117,006	535.40^{**}	−	163,402
Positive Message	0.026	0.066	28,948	1191.76^{**}	+	163,402
Neutral Message	0.060	0.027	130,427	921.45^{**}	−	163,402
Negative Message	0.034	0.015	4,027	40.83^{**}	−	163,402
Topic 1	0.034	0.022	5,199	22.21^{**}	−	163,402
Topic 2	0.033	0.035	7,281	0.40		163,402
Topic 3	0.034	0.026	6,178	9.98^{**}	−	163,402
Topic 4	0.033	0.035	16,485	1.03		163,402
Topic 5	0.033	0.045	10,547	44.03^{**}	+	163,402
Topic 6	0.033	0.033	1,152	0.00		163,402
Topic 7	0.033	0.040	1,518	1.99		163,402
Topic 8	0.033	0.026	1,329	2.29		163,402
Topic 9	0.035	0.019	13,063	91.71^{**}	−	163,402
Topic 10	0.033	0.038	10,795	7.72^{**}	+	163,402
Topic 11	0.033	0.048	1,486	10.09^{**}	+	163,402
Topic 12	0.034	0.024	2,560	7.06^{**}	−	163,402
Topic 13	0.034	0.019	9,003	57.88^{**}	−	163,402
Topic 14	0.033	0.037	2,677	1.21		163,402
Topic 15	0.035	0.022	20,175	85.60^{**}	−	163,402
Topic 16	0.033	0.033	2,141	0.01		163,402
Topic 17	0.030	0.044	42,088	194.13^{**}	+	163,402
Topic 18	0.034	0.028	6,307	5.58^{*}	−	163,402
Topic 19	0.033	0.044	1,223	4.10^{*}	+	163,402
Topic 20	0.033	0.038	2,195	1.50		163,402

$^{*}p<.05$, $^{**}p<.01$

Table 3.6: Zero-Order Correlations Matrix: *What*

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. PDR	1.00	-0.08	-0.16	-0.27	-0.16	-0.25	0.03	0.04	-0.05	0.04
2. PDM	-0.08	1.00	-0.12	-0.11	-0.06	-0.06	-0.13	-0.02	0.01	0.04
3. Mention only	-0.16	-0.12	1.00	0.00	-0.16	0.10	-0.07	0.02	-0.01	-0.03
4. Hashtag(s)	-0.27	-0.11	0.00	1.00	0.10	0.11	0.02	-0.09	0.09	-0.02
5. Photo	-0.16	-0.06	-0.16	0.10	1.00	-0.32	0.04	0.09	-0.06	-0.07
6. URL(s)	-0.25	-0.06	0.10	0.11	-0.32	1.00	-0.09	-0.19	0.18	-0.00
7. non-Fortune mentions	0.03	-0.13	-0.07	0.02	0.04	-0.09	1.00	0.09	-0.08	0.00
8. Positive Message	0.04	-0.02	0.02	-0.09	0.09	-0.19	0.09	1.00	-0.92	-0.07
9. Neutral Message	-0.05	0.01	-0.01	0.09	-0.06	0.18	-0.08	-0.92	1.00	-0.32
10. Negative Message	0.04	0.04	-0.03	-0.02	-0.07	-0.00	0.00	-0.07	-0.32	1.00

cations being that it makes visible – and testable – certain phenomena that were previously invisible or not amenable to testing (Clark & Golder, 2015). In the same fashion, it also "leads to new research questions and new ways of thinking about existing questions" (Parks, 2014, p. 356). An example of this is the time stamps that are attached to social media artifacts, which make visible the precise times at which messages are sent. The CSR literature has thus far looked at the timing of CSR initiatives and activities at the annual level (Arya & Zhang, 2009); social media allow for insights at a much more fine-grained level. In this section I undertake analyses of the temporal characteristics of public mentions of the *Fortune 200* firms along with which temporal features are related to firms' reactions to these public mentions.

3.7.1 Daily Variation over the Calendar Year

I start by examining mentions and reactions at the daily, monthly, hourly, and day-of-the-week levels. I first examine the day-to-day variation in the number of public mentions over the course of the year. Figure 3.20 plots the number of original mentions per day over the course of 2014. There were 163,402 mentions in 2014, with 105,216 being retweets and 58,156 original messages. Retweeted messages only rarely receive any type of reaction on Twitter; instead, it is the original message that receives the like, reply, or retweet. Accordingly, Figure

3.20 shows the daily spread of the 58,146 original mentions sent by the public over the course of 2014.

There were an average of 159.3 mentions per day in 2014. However, what is immediately visible in Figure 3.20 is that the number of daily mentions is somewhat erratic, with public mentioning activity jumping up and down with some regular rhythm likely indicative of day-of-the-week effects, which are explored later on. There are also four noticeable spikes in the number of mentions, defined here as days in which there were more than 400 total original mentions of the 42 accounts. Such upsurges in activity were likely indicative of a particular event or campaign; accordingly, to provide insights into the nature of such temporal spikes I delve into the context of the four events.

August 7, 2014: Twitter Chat by @VerizonGiving

The fourth-largest spike in activity occurred on August 7th. Analyses showed that mentions of @VerizonGiving accounted for 212 of the 417 (50.8%) *Fortune 200* mentions that day. Further investigation showed that @VerizonGiving held a "Twitter chat" on August 7th. Twitter chats are discussions held at specific days and times anchored by a specific hashtag (Budak & Agrawal, 2013); they can be one-time events, such as the US Centers for Disease Control and Prevention's Ebola live Twitter chat that took place on October 8th, 2014 from 3:00 to 4:00pm EST using the hashtag #CDCchat,[7] or occur on a weekly or monthly or yearly basis.

With a Twitter chat, the mentions are indications of an active, temporally delimited conversation. Verizon announced the August 7th *#disrupttheclassroom* chat by tweeting:

> How do educators, nonprofits, & tech community work together to connect & #disrupttheclassroom? Chat w/ us & @MLPExp: http://t.co/AfN9prsvVD[8]

[7] As described on the CDC website, "CDC experts will be available to answer your questions on the response to Ebola in the first confirmed case in the United States and the epidemic in West Africa on Wednesday, October 8th, from 3:00 to 4:00pm ET. Follow @CDCgov on Twitter, and use the hashtag #CDCchat to participate in the chat." http://www.cdc.gov/features/twitterchat/

[8] @MLPExp was the Mobile Learning Partnership Initiative (MLP): "The Mobile Learning Partnerships (MLP) initiative was a project of Innovate+Educate in partnership with the Verizon Foundation to pair beta

During the chat itself, @VerizonGiving asks questions, this one about teachers' ("Ts") preferences:

> Do Ts prefer apps that provide real-time feedback to students? Are developers building apps with this functionality? #disrupttheclassroom

While members of the public provide responses:

> @verizongiving Ts need to prefer apps that give reg feedback #disrupttheclassroom @nearpod can drive entire lesson -instruction&assessment
>
> .@verizongiving teachers should not bear the brunt of that. Partner with local partners like @galvanize to teach #tech #disrupttheclassroom
>
> .@drjillbrown @verizongiving it also has to be fun & produced by educators who understand the stresses of the classroom #disrupttheclassroom
>
> @JeanMcCormick @verizongiving We had kids learn and design games to investigate the math w/college Animation Club. #disrupttheclassroom
>
> @verizongiving: A5: Not agreed, a tool if used properly should re-define the learning previously inconceivable #disrupttheclassroom

And @VerizonGiving responds to these responses:

> @m_lula Thank you! #disrupttheclassroom
>
> @mbjerede @CoRaft That's why developers need more input from educators! #disrupttheclassroom
>
> @LaToniyaAJones Thanks for joining! #disrupttheclassroom
>
> @kstoyle09 Please expand on redefining the learning process with new tools... #disrupttheclassroom

In effect, Twitter chats provide for a back-and-forth dialogue. Accordingly, the proportion of firm mentions that receive a reaction from the firm can be very high during these chat events.

---

stage mobile learning applications with schools to have the apps tested in a live classroom environment. One of the primary goals of this first year project (completed June 2015) was to empower teachers and students to provide meaningful feedback to a mobile learning tool BEFORE it goes to market thereby making the final product a 'student-tested, teacher-approved' learning tool. The education technology market is valued at nearly $8 billion with schools and educators overwhelmed by choices. How are they to know which products are high quality and support learning? The MLP project was designed to give educators and students an active voice in product development." (http://www.innovate-educate.org/mlp)

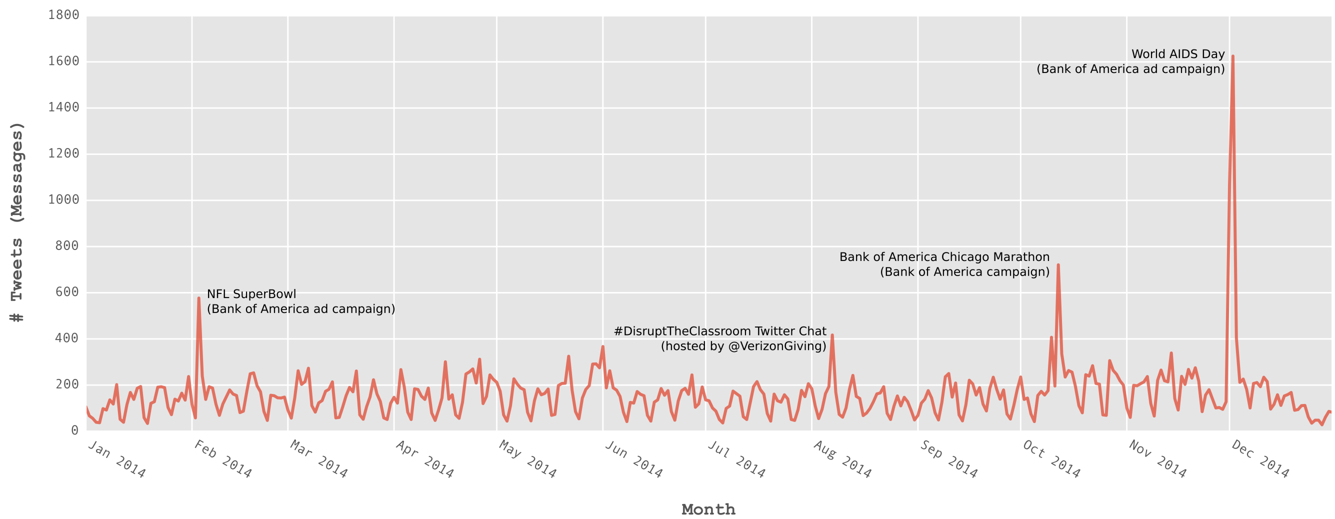


Figure 3.20: Number of Original Tweets/Day Mentioning Fortune 200 CSR Accounts, 2014

BofA_Community's(RED) campaigns

The other three spikes I will examine derive from Bank of America's partnership with (RED), an organization founded by Bobby Shriver and U2's Bono in 2006. The recipient of (RED) funds is The Global Fund, "a 21st century organization designed to accelerate the end of AIDS, tuberculosis and malaria as epidemics" (https://red.org/global-fund/). Through the partnership "Bank of America is delivering more than $10 million toward the fight against AIDS" (https://red.bankofamerica.com/our-story.html#fbid=34Qm6JvhTsF). A sample tweet from the campaign in shown in Figure 3.21.

February 2, 2014

The first is one occurring on Feb. 2, 2014: NFL SuperBowl day. @BofA_Community aired an ad with U2's Bono publicizing the free download of a new U2 song, "Invisible."[9]

[9]From the iTunes store: "All proceeds from each sale of the song featured on this page will be donated to the Global Fund to Fight AIDS, Tuberculosis and Malaria, the charitable recipient for (RED)™. The world is facing a major opportunity in the fight against AIDS–with action and urgency we can end the transmission of HIV from pregnant moms to their babies and deliver the first AIDS-free generation in over 30 years. Supporting this goal is easier than you might think. Download today."

The commercial stated that, "For every download, Bank of America donates $1 to (RED)." Mentions of @BofA_Community accounted for 494 of the 577 (85.6%) Fortune 200 mentions that day.

October 12, 2014

The second of the spikes occurred on October 12, the race day for the Bank of America Chicago Marathon 2014. Mentions of @BofA_Community accounted for 666 of the 721 (92.4%) Fortune 200 mentions that day.

A specific online campaign drove this traffic. As noted on the Bank's website, "For every mile we ran up to the weekly goal, Bank of America donated 40 cents to the global fund to fight AIDS*."[10] Moreover, as stated on the (RED) website, "(RED) and Bank of America Challenge the Global Fitness Community to Go the Distance with Turn Your Miles , Powered by Nike+ – For each Nike+ Running mile pledged to (RED) on www.nike.com/onestep4red, Bank of America will donate 40 cents – up to $1 million – toward the fight to eradicate mother-to-child HIV transmission. Nike+ Running is a free app which allows people to track, measure, share and compare their movements with a global community of athletes."[11] With this campaign as background, what specifically drove traffic was Nike's website, which had a function to "spread the word" on Facebook and Twitter. Hitting the Twitter button opened up Twitter with the following default text:

> I'm turning my Nike+ Running miles @RED with @BofA_Community to help deliver an AIDS free generation. #onestep4red http://nike.com/onestep4RED

596 of the 721 (82.7%) Fortune 200 tweets that day contained this text.

[10]"Each week, from October 12 to December 7, 2014, participants will aim to hit a combined weekly mile goal as part of the Turn Your Miles (RED) activation. For every mile recorded with Nike+ Running, Bank of America will donate $.40 to the US Fund for the Global Fund to fight AIDS, Tuberculosis, and Malaria to support and fund prevention and treatment of HIV/AIDS, up to the weekly donation goals and up to a total maximum donation of $1,000,000. Each mile each participant logs each week will be counted toward the total weekly donation goal until that weekly goal is met. No purchase necessary." (https://www.nike.com/cdp/justdoit/onestep4red/leaderboard/)

[11]http://stg.red.org/en/learn/red-and-bank-of-america-turn-your-miles-red

December 1-2, 2014

The last, and largest, spike to be examined is centered on December 1-2, 2014. Mentions of @BofA_Community accounted for 2,393 of the 2,694 (88.8%) Fortune 200 mentions that day. December 1st was World AIDS Day, and BofA_Community embarked on a tweeting campaign to spread the word regarding its related activities. The firm sent out a number of tweets mentioning celebrities who were spreading the word about the event, such as this tweet on December 1st at 3:23pm GMT:

> Thanks for helping us spread the word @JuddApatow

Meanwhile, this tweet sent on December 2nd at 12:17pm (15:17GMT) received 706 retweets from members of the public:

> Taking a #GivingTuesday #unselfie? Donate to (RED) today and we'll match your donation 2-to-1. http://go.bofa.com/28ua

What really drove the spike, however, was a tweet, shown in Figure 3.22, sent by the main corporate Twitter account, @BankofAmerica, on November 30th, which contained a #WorldAIDSDay video starring U2's Bono. The tweet was retweeted 31,000 times.

3.7.2 Month of the Year

To further explore the relationship between timing and reactions I now present a series of bar graphs with tweets and reactions aggregated at different time periods. To start, Figure 3.23 shows the total number of public mentions per month along with the total number of these messages that are ignored. This allows us to compare the month-to-month flows of public messaging and firm reactions and also to see in which months the gap between the number of messages that are sent and the number that are ignored is greater.

Mirroring the pattern shown in Figure 3.20, Figure 3.23 shows some noticeable differences in monthly tweeting and reacting, such as the October and December spikes in activity, which

Figure 3.21: Sample Tweet from @BofA_Community

Figure 3.22: Sample Tweet from @BofA_Community

are driven by the @BofA_Community campaigns described above. In terms of reactions, the great majority of messages are ignored in every month, though December appears to be when the highest proportion of messages are ignored.

3.7.3 Day of the Week

Figure 3.24 shows the number of messages and reactions aggregated by day of the week. Weekends are a slow time on Twitter on the part of members of the public, who mention *Fortune 200* firms at a considerably lower rate. In addition, the proportion of mentions that are ignored by firms is higher on Saturday and Sunday. Not surprisingly, it appears the day of the week a message is sent is an important determinant of whether it receives a reaction.

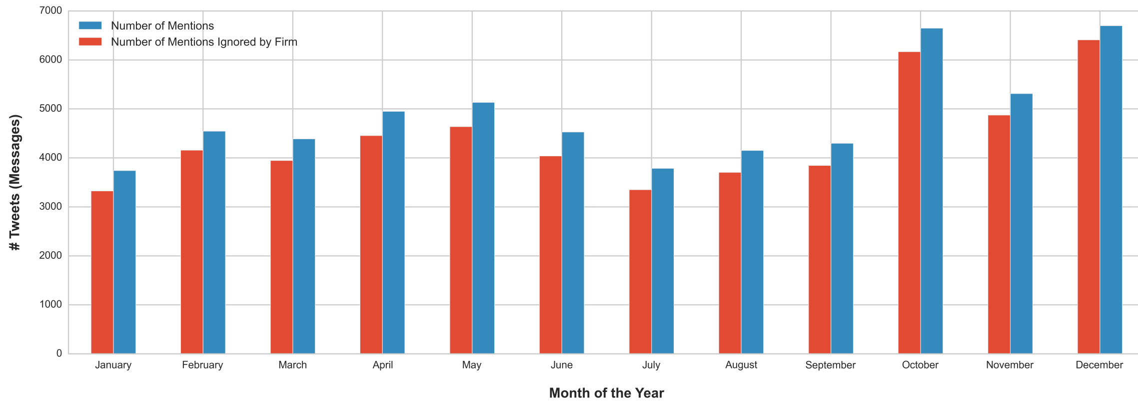


Figure 3.23: Monthly Counts of Public Tweets and Firm Reactions

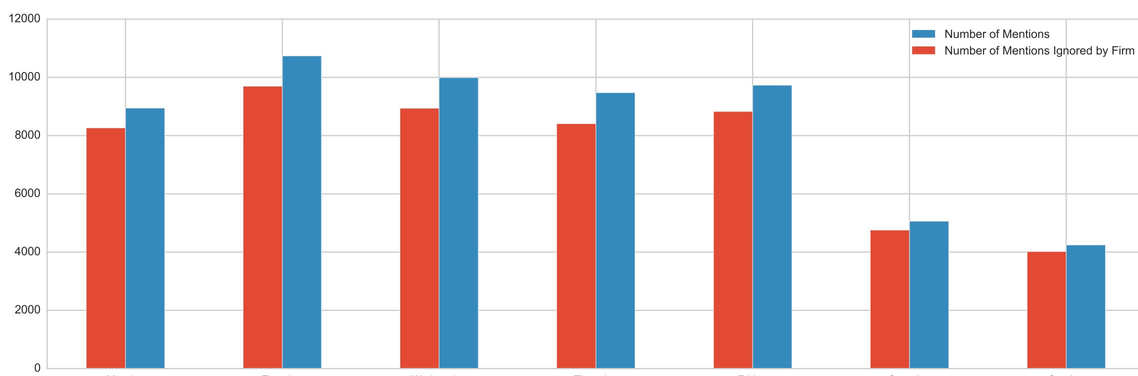


Figure 3.24: Day-of-Week Counts of Public Tweets and Firm Reactions

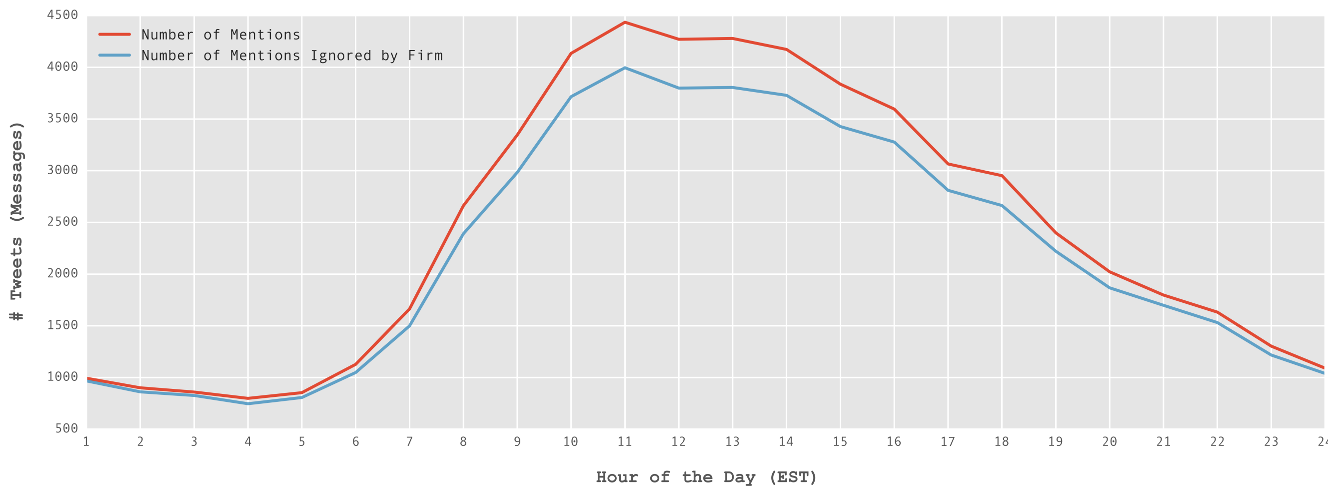


Figure 3.25: Hourly Counts of Public Tweets and Firm Reactions

3.7.4 Hour of the Day

Figure 3.25 plots the number of tweets and reactions per hour of the day in Eastern Standard Time (EST). There are more mentions in the middle of the day, roughly from 9:00am EST to 7:00pm EST. In addition, a lower proportion of mentions during this time period are ignored. By contrast, a very high proportion of mentions are ignored during the 10:00pm to 7:00am period EST. In effect, if a user wants a North American firm to react, it should avoid tweeting during the hours when most North Americans are asleep. This brings up an interesting quality of this context: On social media, reactions either come almost immediately or not at all.

3.7.5 Minute of the Hour

Figure 3.26 shows the number of messages and reactions in terms of the minute of each hour. At this level of granularity the data appear to cease being informative. Interestingly, there are spikes at the 0 minute (at the start of the new hour) as well as at the 5 minute, 10 minute, 15 minute, 20 minute, 25 minute, 30 minute, 35 minute, 40 minute, 45 minute, 50

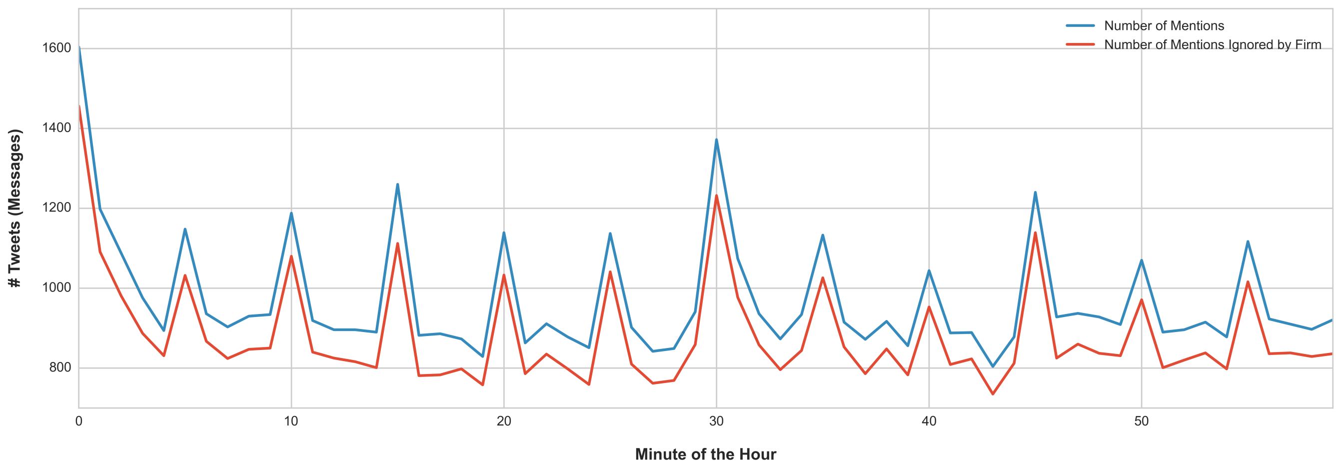


Figure 3.26: Counts of Public Tweets and Firm Reactions by Minute of the Hour

minute, and 55 minute periods. Further investigation would be needed to determine whether these spikes are due to, perhaps, automated tweet settings or, in contrast, to the degree of precision available in the Twitter timestamps.

3.7.6 Second of the Minute

As with the previous figure, Figure 3.27 is not terribly useful. It is included primarily to show the granularity with which the data can be examined. The spike in mentions and reactions at the beginning of the minute suggests either auto-scheduled tweets or, alternatively, an idiosyncracy of the Twitter recording mechanism, with perhaps the "0 second" recording being some kind of default.

3.7.7 Relationship between timing and firm reactions

Based on the above analyses I generated dummy variables to indicate tweets that include a "chat" hashtag, tweets that are sent during business days (Monday - Friday), tweets that are sent during typical business hours (9:00am - 5:00pm), and those that are sent during

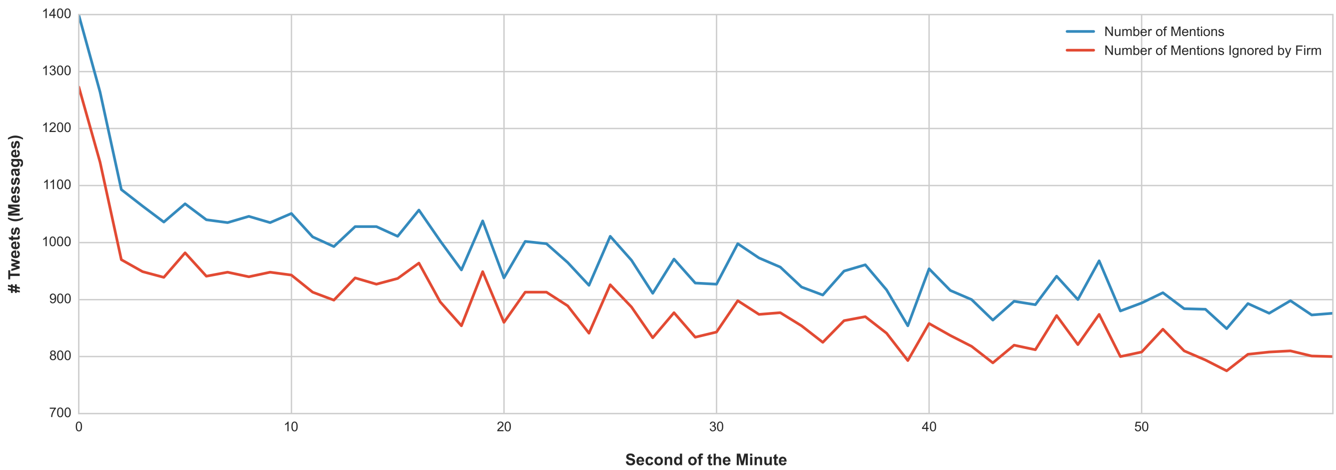


Figure 3.27: Counts of Public Tweets and Firm Reactions by Second of the Minute

Table 3.7: Zero-Order Correlations Matrix: *When*

	Chat tag	Weekday	Bus. hours (7 days)	Bus. hours (M-F)
Chat tag	1	0.02	0.01	0.02
Weekday	0.02	1	0.01	0.44
Bus. hours (7 days)	0.01	0.01	1	0.83
Bus. hours (M-F)	0.02	0.44	0.83	1

business hours on business days.

Table 3.7 shows correlations between the four variables. Figure 3.28 shows bar graphs plotting the proportion of public mentions that receive a reaction for each of the four variables both when the value of the variable is "0" and when the value of the variable is "1." Table 3.8, meanwhile, presents the results from Chi-square tests relating each of these four variables with *Fortune Reaction.* As suggested by the figure and shown in the table, all four achieve a significant coefficient and are positively related to the firms' reactions.

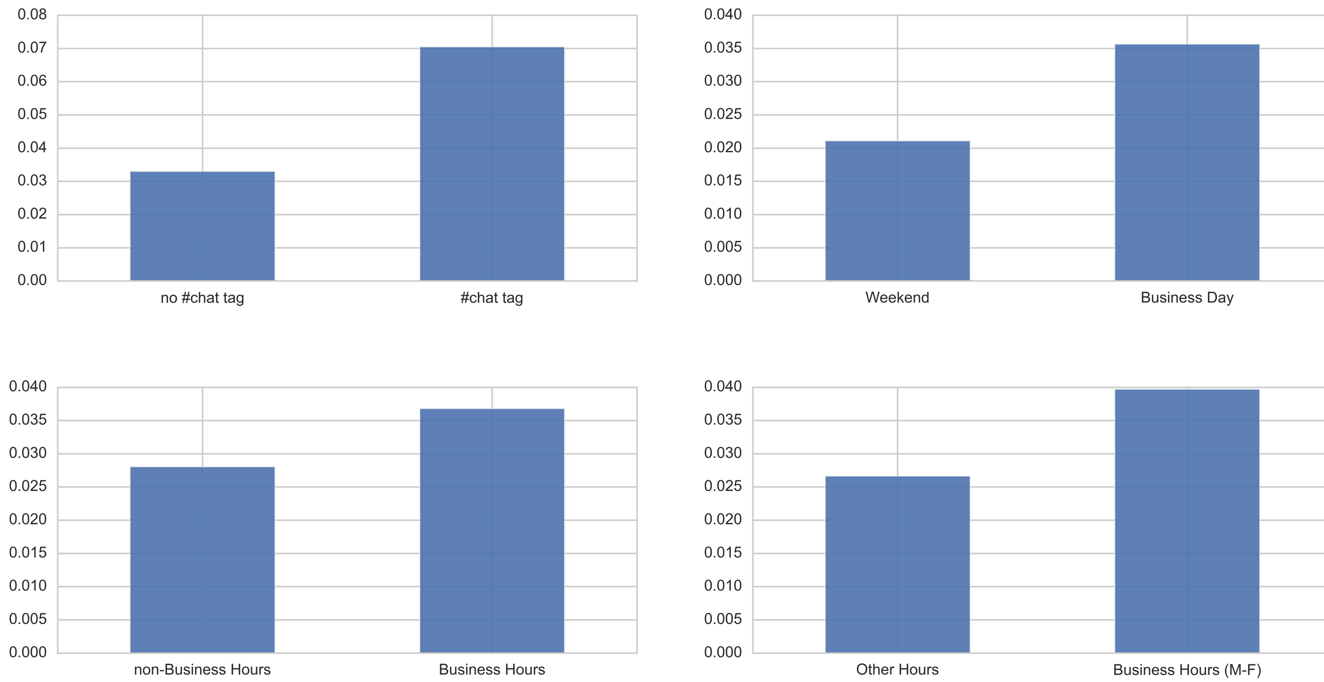


Figure 3.28: Average Firm Reaction by Timing

Table 3.8: Chi-Square Tests for Binary Variables - D.V. is *Fortune Reaction (0,1)*

variable	mean score, I.V. value=0	mean score, I.V. value=1	no. of obs., I.V. value=1	χ^2	sign	n
#chat hashtag	0.033	0.070	1731	73.45^{**}	+	163,402
Business Day	0.021	0.036	137930	141.07^{**}	+	163,402
Business Hours	0.028	0.037	99572	92.32^{**}	+	163,402
Business Hours (M-F)	0.027	0.040	84409	215.42^{**}	+	163,402

*p<.05, **p<.01

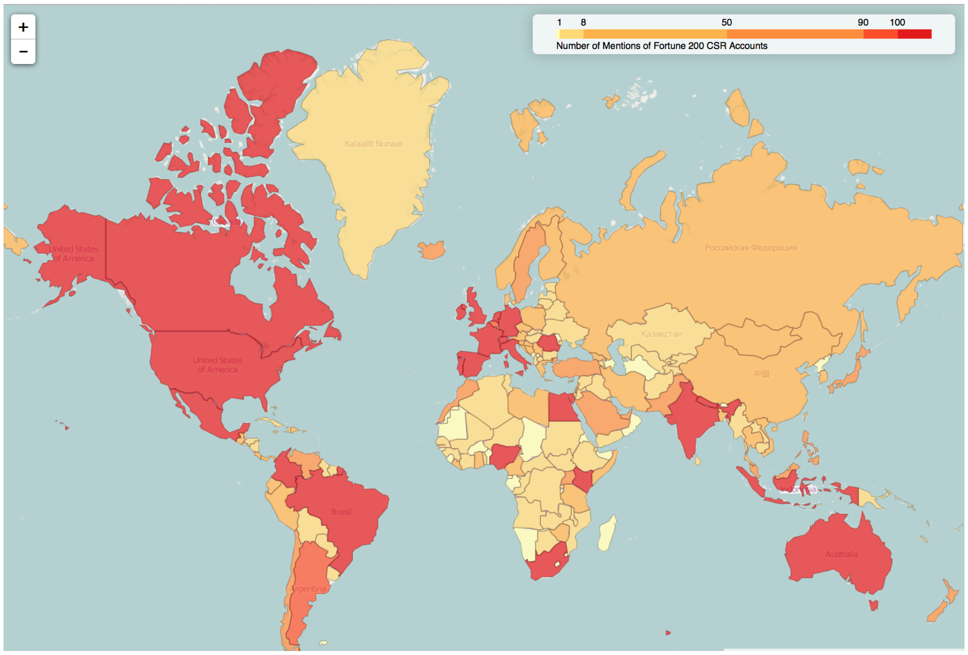
3.7.8 Summary of *When* Messages Receive a Reaction

Through the analyses above I have identified three dimensions of timing. First, there is *business hours*. The three variables I generated (weekday, mid-day hours, and business hours) are all related; the most parsimonious model would choose the strongest of these variables, which here appears to be the variable reflecting business hours during the weekday. A second dimension is *chats*, with mentions made in the context of Twitter chats being expected to receive higher levels of reactions. The third dimension is *campaigns*; as seen above, the most successful campaigns caused notable spikes in activity.

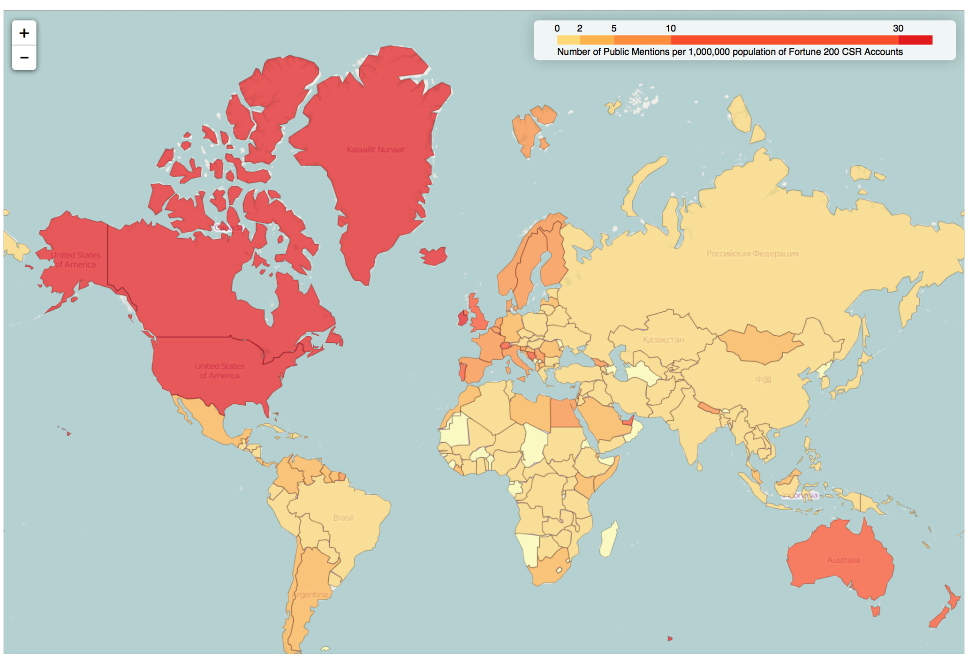
3.8 *Where* Messages Receive a Firm Reaction

As with time stamps, geo-coded messages are likewise enabled by many social media sites. At the same time, geographic data has recently garnered interest in the broader accounting literature (e.g., O'Brien & Tan, 2015). To explore the relationship between the geographic location of the public message senders and firm reactions, I first ran users' self-described location through geographical databases (*GeoNames* and *ArcGIS*) in order to find the city and country in which the user is located as well as, where possible, the geographical coordinates.

I then created a number of what are known as *choropleth* maps. Choropleth maps aggregate numerical data to a specified geographical unit and assign a different color code to each unit according to the intensity of the variable. Figures 3.29, 3.31, and 3.33 show the total number of public mentions of the 42 CSR accounts for, respectively, countries of the world, U.S. states, and U.S. counties; while Figures 3.30, 3.32, and 3.34 show the aggregate number of firm reactions to public messages at the three levels. In each of the six figures, subfigure (a) shows the aggregate number while subfigure (b) shows the per capita number of mentions or reactions.

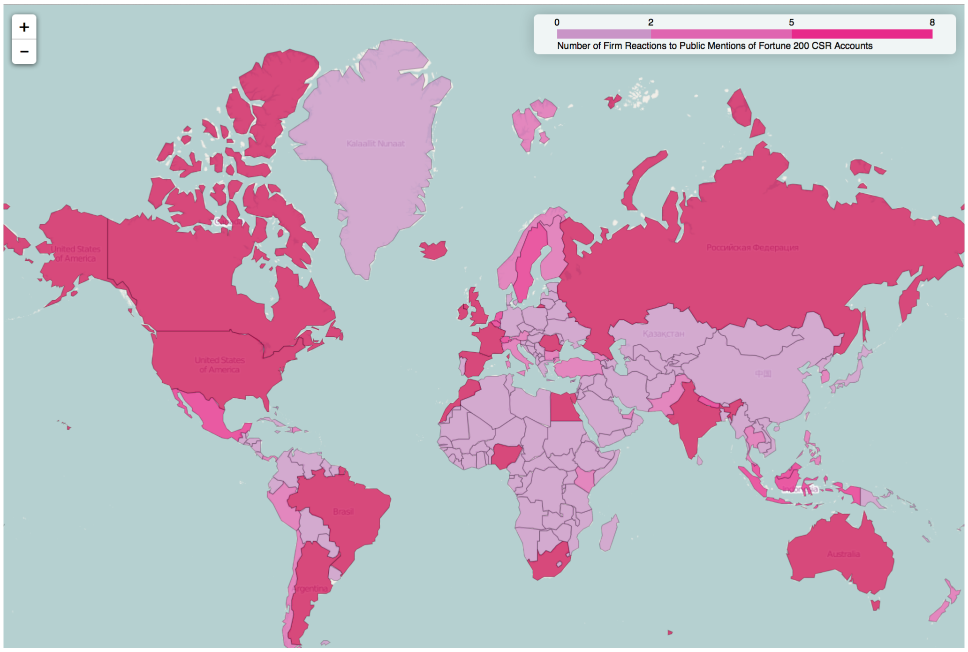


(a) Number of Public Mentions

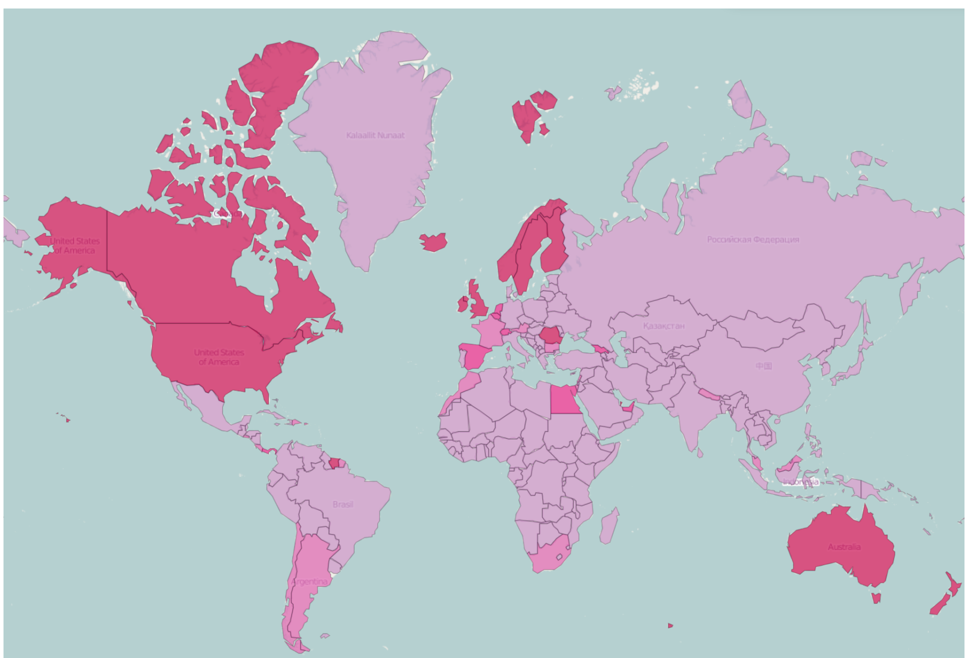


(b) Number of Public Mentions per capita

Figure 3.29: # of Public Mentions of CSR Accounts per Country

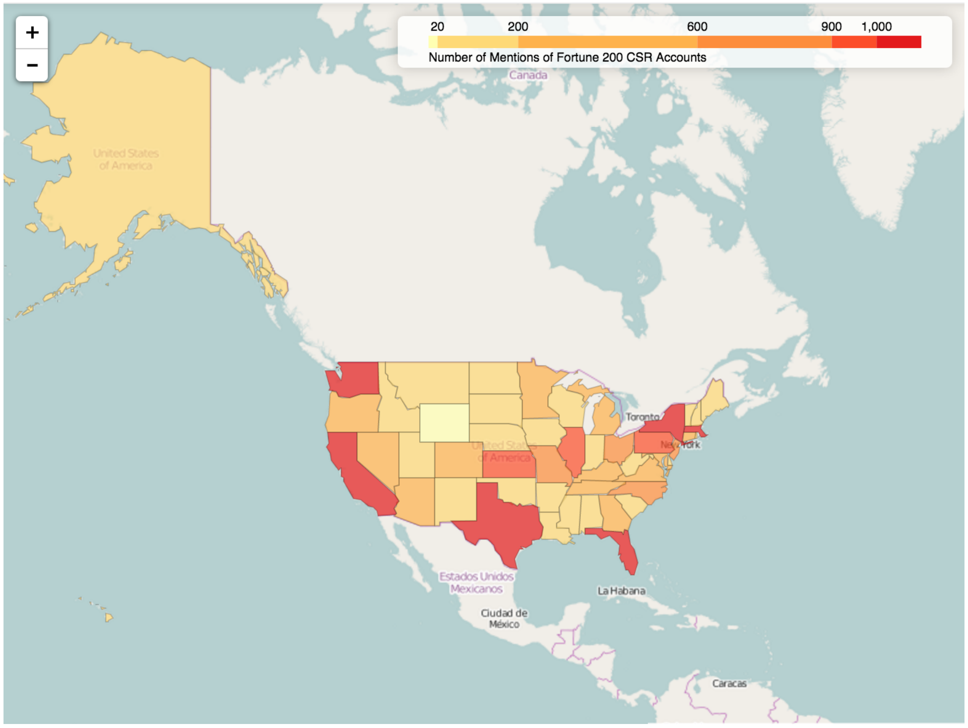


(a) Number of Firm Reactions to Public Messages

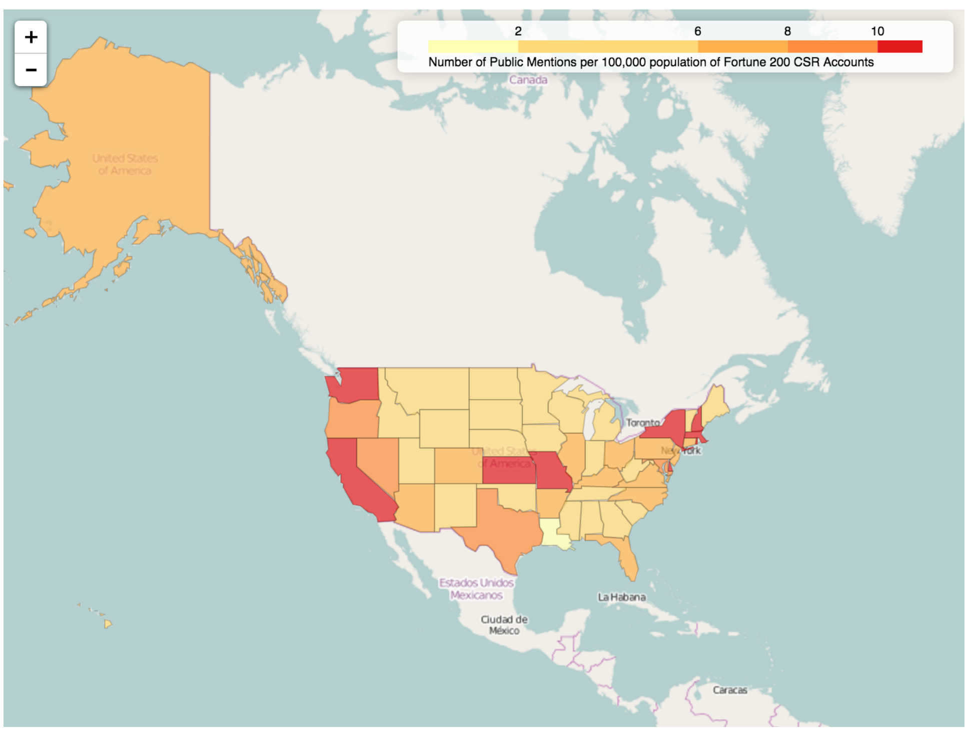


(b) Number of Firm Reactions per capita

Figure 3.30: # of Firm Reactions to Public Messages per Country

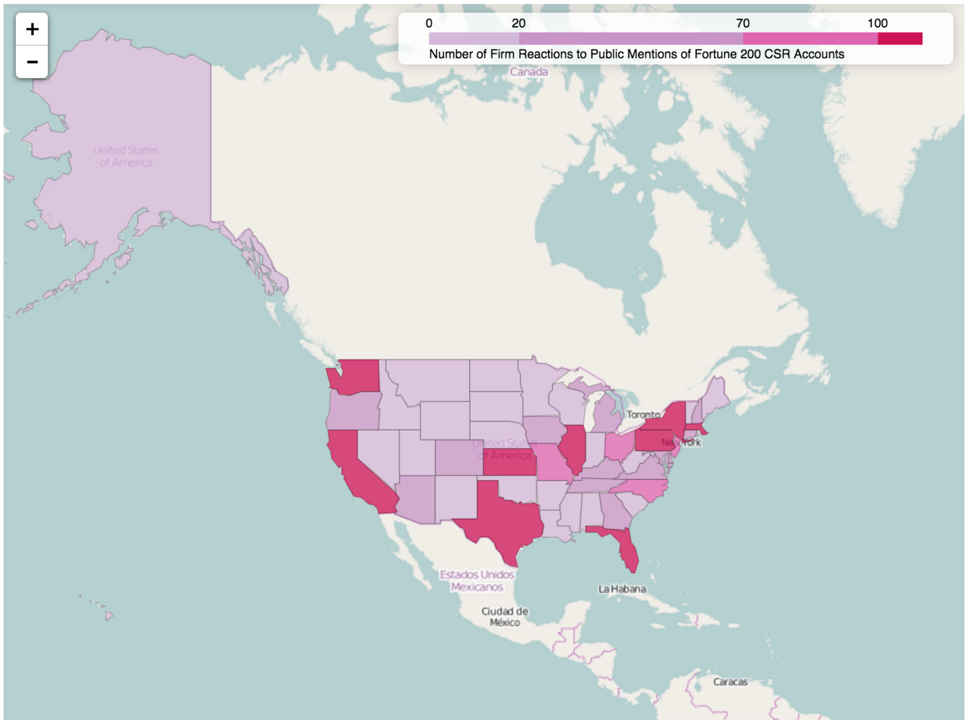


(a) Number of Public Mentions

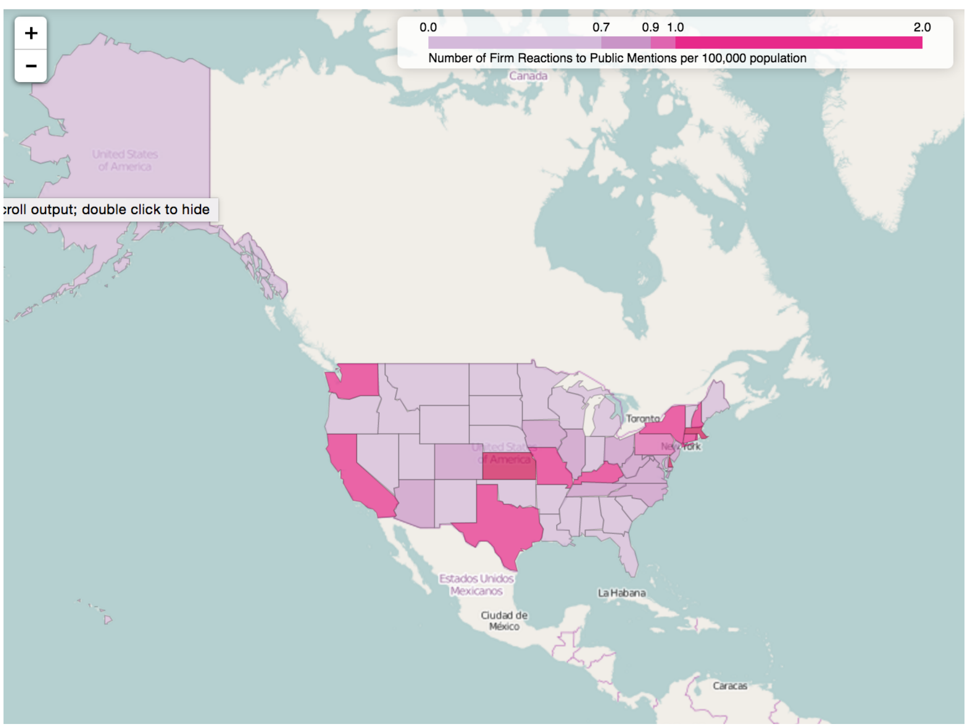


(b) Number of Public Mentions per capita

Figure 3.31: # of Public Mentions of CSR Accounts per State

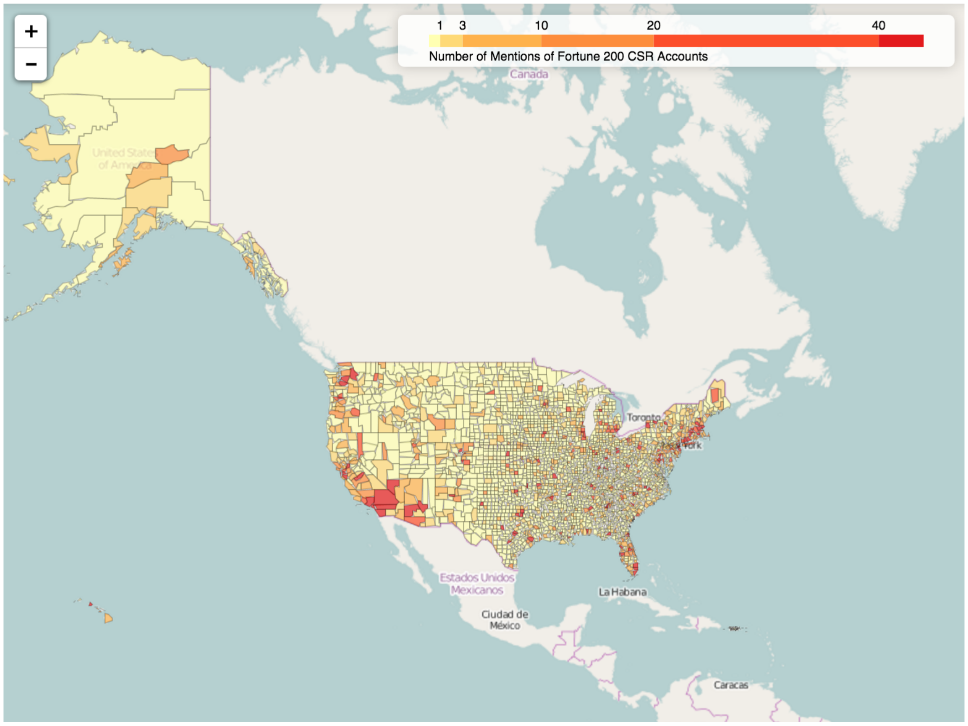


(a) Number of Firm Reactions to Public Messages

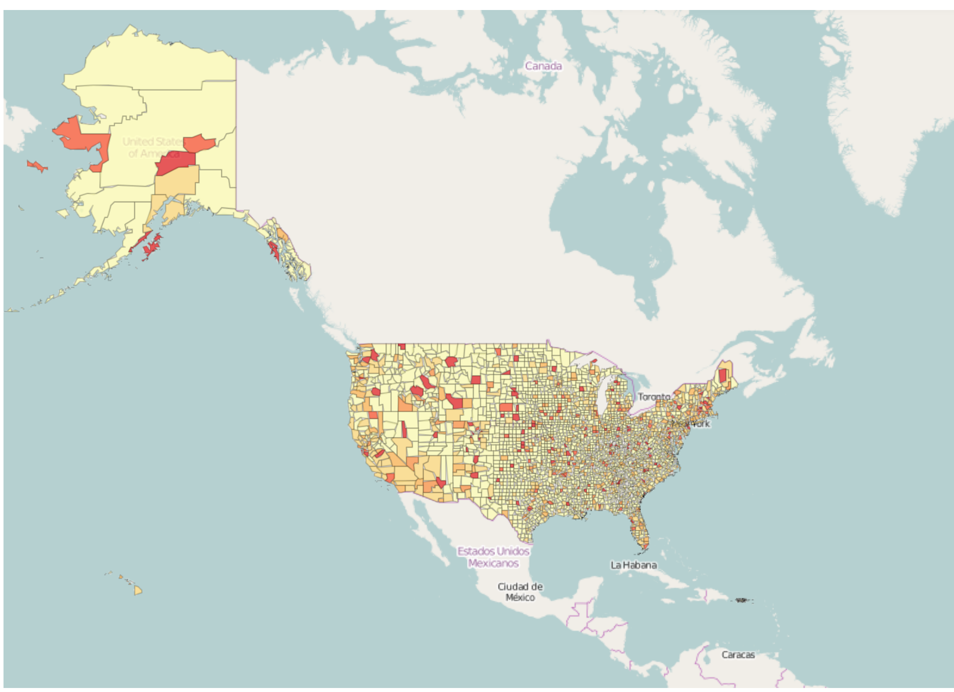


(b) Number of Firm Reactions per capita

Figure 3.32: # of Firm Reactions to Public Messages per State

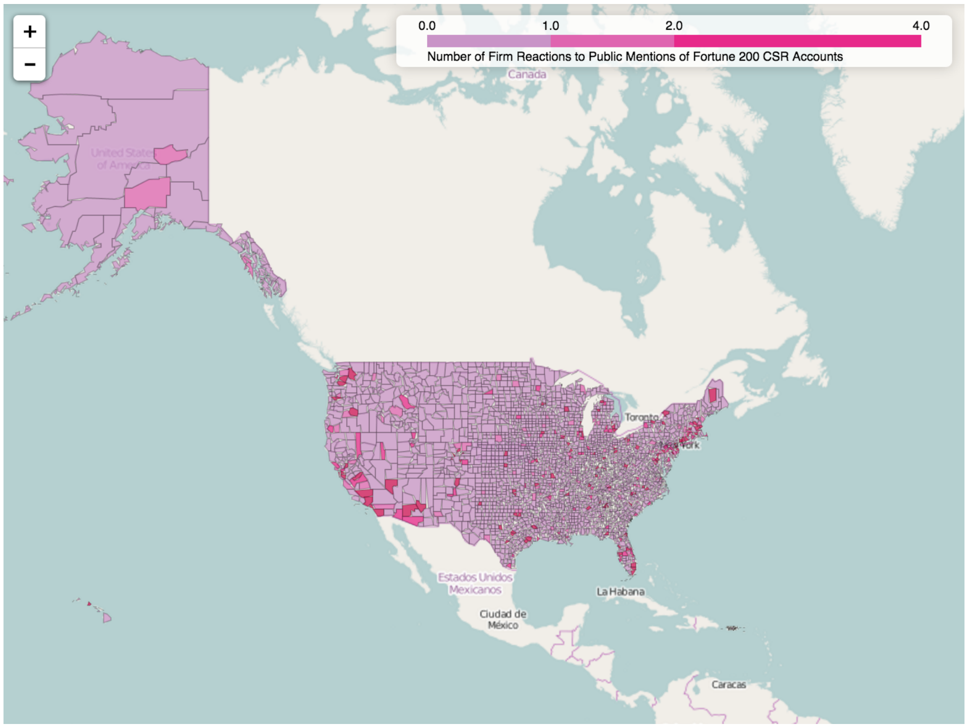


(a) Number of Public Mentions

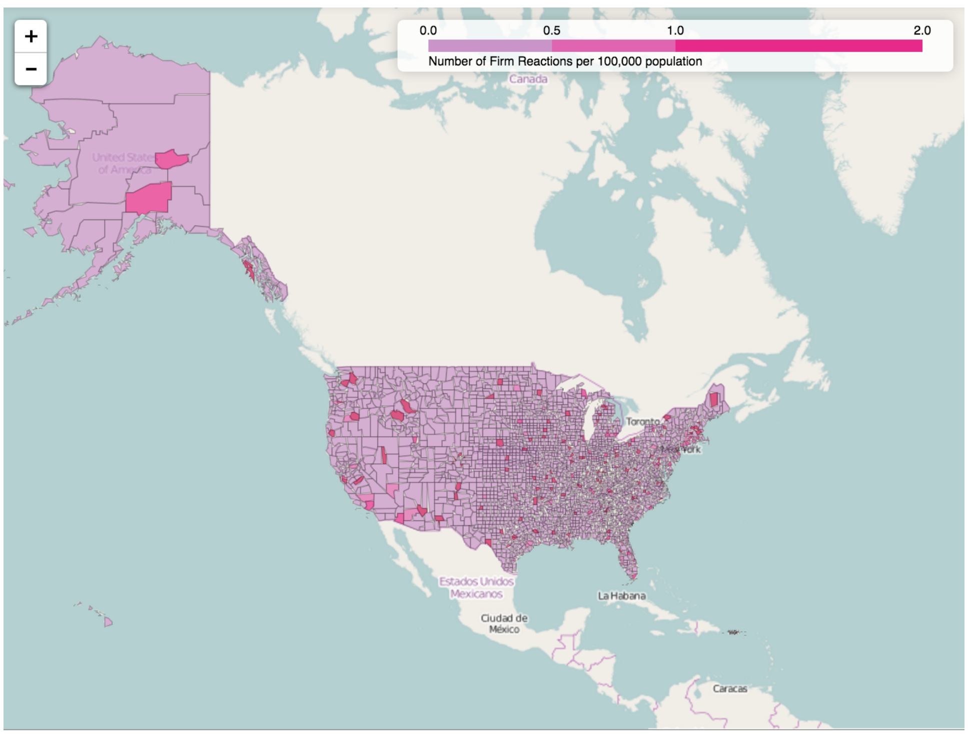


(b) Number of Public Mentions per capita

Figure 3.33: # of Public Mentions of CSR Accounts per County



(a) Number of Firm Reactions to Public Messages



(b) Number of Firm Reactions per capita

Figure 3.34: # of Firm Reactions to Public Messages per County

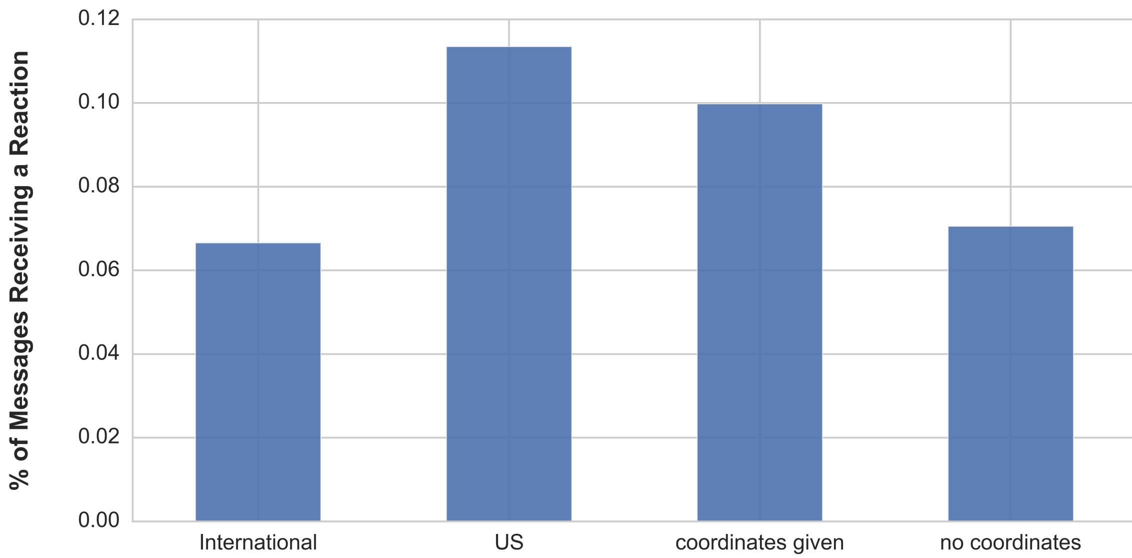


Figure 3.35: Average % of Messages Receiving Reaction by Location Indicators

3.8.1 Bivariate Statistics and Correlations

To more rigorously explore the relationship between location and reaction I created two indicator variables that tap, first, whether a tweet includes an identifiable location and, second, whether the user is located in the United States. Table 3.9 shows the correlation matrix of these two variables as well as the dependent variable. Figure 3.35, in turn, shows the proportions of messages receiving a reaction for each of these binary variables. The figure suggests having a recognizable geographical location in the user's profile (as reflected by "coordinates given" in the figure) is positively associated with firm reactions, as is being location in the United States. The latter is not surprising given the location of the *Fortune 200* firms studied here. Table 3.10 further corroborates the two relationships, where the Chi-square tests show significant positive relationships with *Fortune reaction.*

Table 3.9: Zero-Order Correlations Matrix: *Where*

	Fortune reaction (0,1)	coordinates available (0,1)	USA (0,1)
Fortune reaction (0,1)	1.00	0.05	0.07
coordinates available (0,1)	0.05	1.00	0.65
USA (0,1)	0.07	0.65	1.00

Table 3.10: Chi-Square Tests for Binary Variables - D.V. is *Fortune Reaction (0,1)*

variable	mean score, I.V. value=0	mean score, I.V. value=1	# obs, I.V. value=1	χ^2	sign	n
Coordinates available	0.0227	0.0396	103,359	337.32^{**}	+	163,402
United States	0.0228	0.0480	68,606	785.09^{**}	+	163,402

$^{*}p<.05$, $^{**}p<.01$

3.8.2 Summary of *Where* Firms React to Public Messages

Above I have identified two potentially relevant dimensions of location: 1) whether an identifiable location for the sender is made available and 2) whether the sender is located in the United States. The two variables I used to tap these dimensions correlated highly; in a more parsimonious test using only one might be preferable. I posit the second of the two variables is more relevant, which could be generalized to indicate whether the message sender is in the firm's home country.

3.9 *Why* Firms React to Public Messages

I posit that *why* a message receives a response is reflective of the particular characteristics of the given firm. In this section I thus explore a variety of firm traits, such as industry and size.

Both variables are seen in the accounting literature as key determinants of reporting (e.g., Atiase et al., 1989; Tsang, 1998). Similar to the analysis of message senders in an earlier section, I also examine levels of Twitter activity such as the number of tweets, friends, and followers.

3.9.1 Financial Characteristics

To start, I examine four indicators of the size and profitability of the firm provided by *Fortune*: the number of employees, market value, total assets, and profits.[12] Figure 3.36 shows box plots for each of these four variables, showing the distributions for the variables both when a message receives a reaction and when it does not. A visual inspection of the figures suggests very weak or no relationship for assets, market value, and profits. The number of employees, however, does appear to be lower on average when messages receive a reaction, suggesting smaller firms are more likely to be engaged with members of the public.

3.9.2 Industry

I next examine industry. In total 27 distinct industries are represented by the firms maintaining the 42 Twitter accounts. Figure 3.37 shows the average proportion of messages that receive a reaction for the 27 industries represented in the dataset. What sticks out is the high reaction rate for four industries: telecommunications, metals, health care, and computer hardware.[13] In contrast, IT services, commercial banks, and chemicals have lower reaction rates than other industries.

3.9.3 Twitter Characteristics

As noted above, I also examine a suite of measures that reflect the firm's level of Twitter activity. These are the same measures as those explored earlier under the category of "who,"

[12]Data come from the firms' balance sheet, income statement, and 10-K filings.

[13]There is only one firm in the metals industry, Alcoa, such that the high proportion of reactions seen in that industry is entirely due to @AlcoaFoundation's level of responsiveness to public messages.

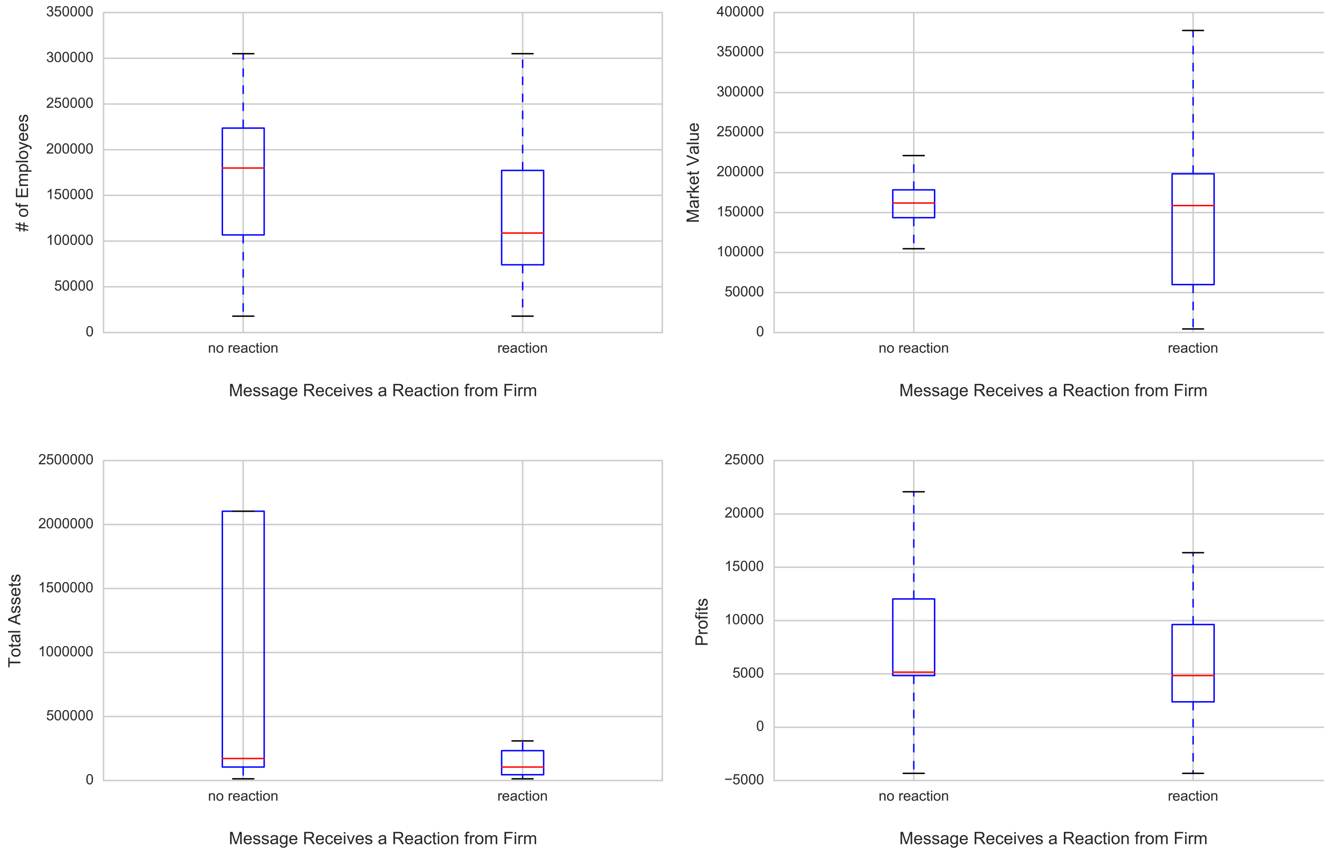


Figure 3.36: Average % of Messages Receiving Firm Reaction based on Financial Features

except that instead of indicating the public message sender's Twitter activity, these indicate the message *reactor's* – that is, the firm's – Twitter indicators.

To make these variables more concrete, Figure 3.38 is a screenshot of the Twitter profile page for Cisco's CSR-related handle @CiscoCSR. As shown in the middle of the figure, any visitor to the page can see five count variables that reflect different aspects of Cisco's Twitter behavior. First, *Tweets* shows the cumulative number of tweets @CiscoCSR has sent. At close to 10,000 tweets over six years, the firm on average sends out roughly 4 messages per day, a somewhat modest number for such a large firm. Second, *Following* shows the number of other Twitter users @CiscoCSR chooses to follow. In general, the more users followed, the more Cisco is sending a signal that it is interested in reaching out to other users and and learning from a broader community (Lovejoy et al., 2012). The relatively low number of 653 users followed suggests Cisco is not as interested in outreach as other accounts. Third,

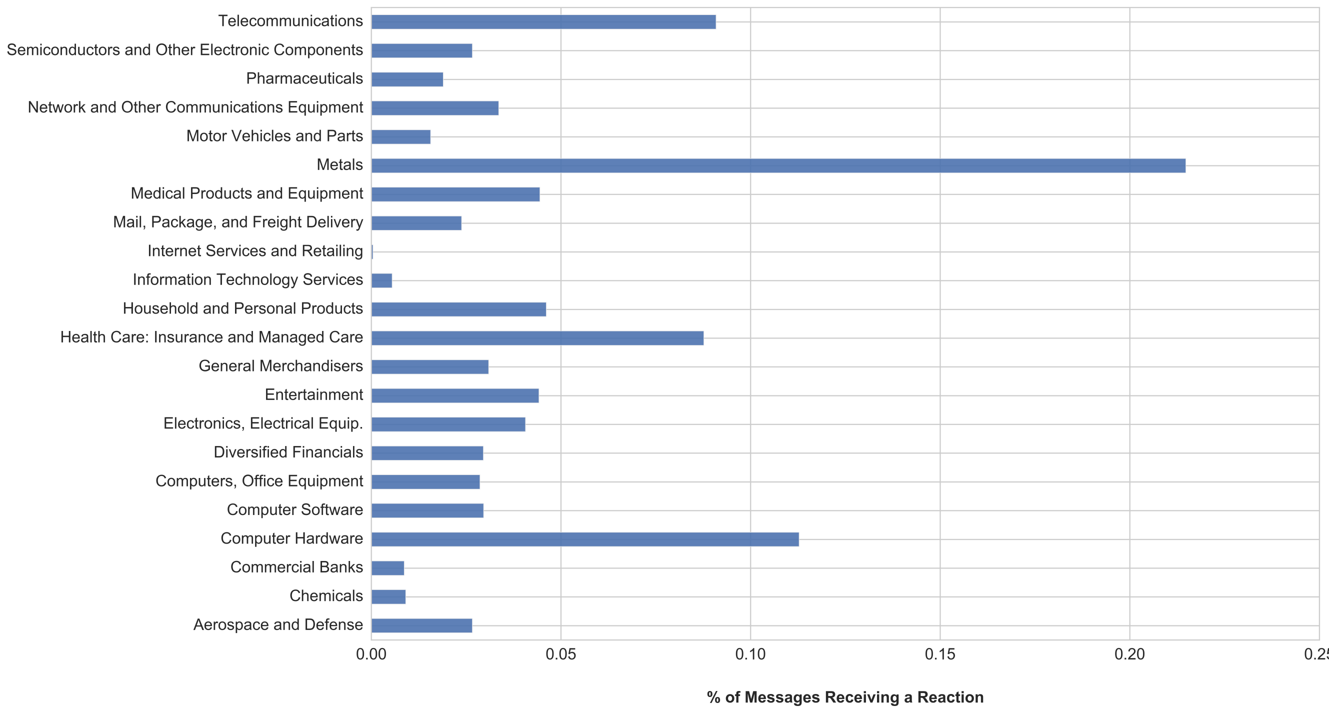


Figure 3.37: Mean Reaction by Industry

Followers indicates the number of other Twitter users who are following the account. At only 9,218 followers in early 2016, Cisco is well below the average of the large firms examined in this study. Fourth, *Likes* indicates the number of tweets @CiscoCSR has favorited (now called "liked"), an activity that serves both an archive function and a signal to the original message sender that Cisco found the message useful in some regard. Liking is another means of connecting with a broader set of users. With only 744 tweets *liked*, the implication is that Cisco is not as active as many other firms in the sample. Fifth, *Lists* shows that @CiscoCSR is subscribed to 10 public Twitter "lists," including *Veterans* and *Critical Human Needs*, all created by either @CiscoCSR or @Cisco. Twitter users who click on the Lists link can also see the number of public lists @CiscoCSR is a member of. These are lists that other Twitter users have created and placed @CiscoCSR on. The higher the number of lists @CiscoCSR is a member of, the greater its level of prestige and influence. In early 2016 @CiscoCSR was on 384 public lists, a moderate number that is below the mean of 928 for all 42 CSR accounts

Figure 3.38: Screenshot of @CiscoCSR's Twitter Profile Page

in this sample.

As with the sender characteristics examined in an earlier section, I also use the profile information reflected in the "Joined" date shown on the profile page (for @CiscoCSR, this is the "Joined February 2010" statement, which is automatically created by Twitter) to generate a measure of the number of days since each account joined Twitter. The intuition is that companies that have been on Twitter longer may be more sophisticated in their use of the platform.

Figure 3.39 shows the distribution of these six Twitter activity variables when messages do and do not receive a reaction. A visual inspection of the box plots suggests a slight negative relationship between the number of followers a firm has and whether it reacts to messages from the public, while there is a slight positive relationship between reactions and the number of users the firm chooses to follow. There does not appear to be any relationship between reactions and the cumulative number of tweets the firm has sent since it joined Twitter. The number of public lists a firm appears on, in turn, does seem to be somewhat negatively related to the likelihood of reacting. The number of tweets favorited by the firm appears to have a positive relationship with firm reactions while, finally, there appears to

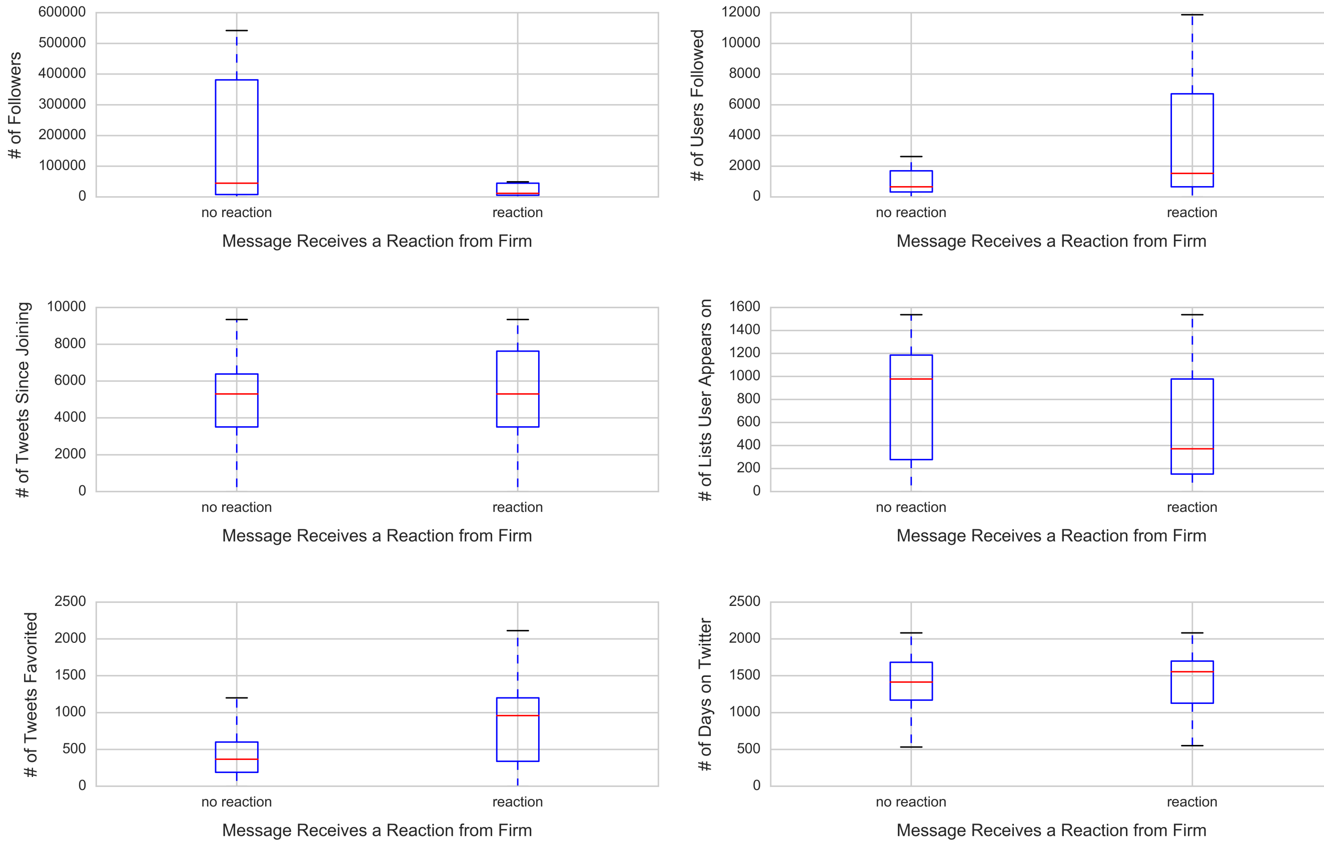


Figure 3.39: Average % of Messages Receiving a Firm Reaction based on Profile Features

be at best a modest positive relationship between reacting and the length of time since the firm has joined Twitter.

3.9.4 Correlations and Bivariate Statistics

Table 3.11 reports results from Chi-square tests of the relationship between *Fortune reaction* and all the binary variables covered in this section. Namely, of the 27 industry dummies, 17 obtain a significant coefficient, suggesting variation in the reaction rates that are worth exploring further.

Table 3.12 shows a correlation matrix of all the interval-level variables examined in this section, while Table 3.13 shows results from bivariate logit tests for each of these variables. As shown in the table, each of the four financial and size variables obtains a significant negative coefficient with *Firm reaction.* Of the Twitter activity variables, all except the

Table 3.11: Chi-Square Tests for Industry Variables - D.V. is *Fortune Reaction (0,1)*

variable	mean score, I.V. = 0	mean score, I.V. = 1	no. of obs., I.V. = 1	χ^2	sign	n
Aerospace and Defense	0.033	0.027	1,989	2.62		163,402
Apparel	0.033	0.000	473	15.4^{**}	–	163,402
Chemicals	0.034	0.009	2870	52.8^{**}	–	163,402
Commercial Banks	0.044	0.009	48,988	1322.5^{**}	–	163,402
Computer Hardware	0.029	0.113	8,299	1705.5^{**}	–	163,402
Computer Software	0.034	0.030	14,824	7.1^{**}	–	163,402
Computers, Office Equip.	0.033	0.029	419	0.16		163,402
Diversified Financials	0.034	0.030	4,776	2.15		163,402
Electronics, Elect. Equip.	0.033	0.041	910	1.28		163,402
Entertainment	0.033	0.044	3,191	11.4^{**}	+	163,402
Food Consumer Products	0.033	0.000	5	0.69		163,402
General Merchandisers	0.034	0.031	7,302	1.32		163,402
Health Care	0.033	0.088	1,859	170.1^{**}	+	163,402
Household & Personal Prod.	0.033	0.046	1,062	5.0^{*}	+	163,402
Information Technology Svcs	0.035	0.006	9,864	252.5^{**}	–	163,402
Insurance: Ppty & Casualty	0.034	0.000	324	10.2^{**}	–	163,402
Internet Svcs & Retailing	0.034	0.0004	4,582	157.5^{**}	–	163,402
Mail, Package, & Frt Delivery	0.033	0.024	84	0.03		163,402
Medical Products & Equip	0.033	0.044	315	0.88		163,402
Metals	0.030	0.215	3,334	3463.8^{**}	+	163,402
Motor Vehicles and Parts	0.034	0.016	6,709	67.6^{**}	–	163,402
Network & Comm. Equip.	0.033	0.034	13,159	0.01		163,402
Pharmaceuticals	0.033	0.019	844	5.0^{*}	–	163,402
Semiconductors & Other Elec	0.034	0.027	15,356	23.7^{**}	–	163,402
Telecommunications	0.029	0.091	11,519	1268.2^{**}	+	163,402
Utilities: Gas & Electric	0.033	0.000	7	0.31		163,402
Miscellaneous	0.033	0.000	338	10.7^{**}	–	163,402

$^{*}p<.05$, $^{**}p<.01$

Table 3.12: Zero-Order Correlations Matrix: *Why*

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Assets	1.00	-0.03	-0.36	-0.05	0.68	-0.21	0.08	0.22	-0.23	-0.34
2. Employees	-0.03	1.00	0.29	0.25	-0.14	-0.05	0.01	-0.13	-0.24	0.05
3. Profits	-0.36	0.29	1.00	0.86	-0.15	0.37	0.27	0.31	0.07	0.28
4. Market Value	-0.05	0.25	0.86	1.00	0.05	0.30	0.32	0.60	-0.01	0.16
5. # of Followers	0.68	-0.14	-0.15	0.05	1.00	-0.25	-0.03	0.46	-0.03	0.08
6. # of Users Followed	-0.21	-0.05	0.37	0.30	-0.25	1.00	0.60	0.41	0.12	0.17
7. # of Tweets	0.08	0.01	0.27	0.32	-0.03	0.60	1.00	0.35	0.27	0.12
8. # of Times 'Listed'	0.22	-0.13	0.31	0.60	0.46	0.41	0.35	1.00	-0.12	0.15
9. # of Favorites	-0.23	-0.24	0.07	-0.01	-0.03	0.12	0.27	-0.12	1.00	0.42
10. Time on Twitter	-0.34	0.05	0.28	0.16	0.08	0.17	0.12	0.15	0.42	1.00

number of days on Twitter are significant, with the number of followers and number of lists obtaining negative coefficients and the number of friends, number of tweets, and number of favorites obtaining positive coefficients.

3.9.5 Summary of Firm Characteristics and Firm Reactions

With the goal of presenting a more parsimonious model, I propose the following relevant dimensions. First, there is *firm size*, represented here by several measures that could be combined into a single index. Two, as with sender characters discussed in an earlier section, the number of users followed, the number of tweets sent, and the number of tweets favorited constitute three discrete indicators of the aggregate level of firms' *Twitter activity.* Likewise, a third dimension, represented by the number of followers and the number of lists the account is placed on, relates to the *prestige or influence* of the firm's CSR-focused Twitter account. All three of the above dimensions appear to be related to the likelihood of the firm reacting to the public messages it receives. In contrast, measures of account "sophistication" do not appear relevant, unlike what was seen with public senders. In large part this is due to the relative lack of variation in sophistication – all 42 accounts are relatively professional with customized profiles, etc. Time on Twitter, moreover, was not relevant and can likely be

Table 3.13: Logit Tests of *Fortune reaction* on Interval-level I.V.s

variable	coeff.	sign	n
Assets	-0.000009**	–	163,402
Employees	-0.000005**	–	163,402
Profits	-0.000489**	–	163,402
Market Value	-0.000039**	–	154457
# of Followers	-0.000046**	–	163,402
# of Users Followed	0.000136**	+	163,402
# of Tweets Sent	0.000278**	+	163,402
# of Times "Listed"	-0.008816**	–	163,402
# of Tweets Favorited	0.010558**	+	163,402
Time on Twitter in Days	0.000223		163,402

*p<.05, **p<.01

dropped from future quantitative analyses.

A fourth dimension is *industry.* From the analyses above it appears as if industry is related to firm reactions and should therefore be included in multivariate analyses. In line with the typical decision in accounting research (e.g., Dhaliwal et al., 2011), I recommend controlling for potential omitted variable bias by including the industry dummies as control variables in a fixed effects regression model or by clustering standard errors on industry.

3.10 Model Selection

The above statistical and inductive analyses led to the identification of a large number of potentially relevant variables. At more than 80 in total, it is important to develop a more parsimonious model. A first step, summarized in the individual sections above, was to omit the variables that analysis of the bivariate relationships and inter-variable correlations suggested were not relevant or were not priority variables. In this category are variables such as gender, weekday, organization type, "extended" profile, identifiable location, market value, the number of Twitter lists the sender appears on, and the length of time the firm has been on Twitter. Most of these variables were dropped from further consideration because of no meaningful association with the dependent variable, while others (notably, market value and the number of Twitter lists message senders appear on) were dropped because of conceptual and/or empirical overlap (high correlation) with other, more meaningful variables, as described above.

In some ways, this problem is a normal part of any theory-building endeavor. In any typical qualitative inductive analysis designed to develop theory on a new phenomenon, the initial set of variables is larger than the final proposed set of explanatory variables (e.g., Eisenhardt, 1989), with the slate of variables pared down upon each iteration. Potentially, the same process could be continued here to eventually reach a more focused model. On the other hand, even after dropping the above variables, 75 remain for consideration. This is a larger number of variables than is typical in qualitative inductive analyses. Big Data often engender such *high-dimensional* datasets, which first begs the question of whether traditional inductive methods are well suited to reducing the dimensionality of the data (i.e., developing a more parsimonious explanatory model). Second, unlike many typical theory-building situations, I was able to obtain fairly reliable quantitative measures of all of the above concepts, which opens different possibilities in the search for ways to prune the model. For both reasons, to help further reduce the set of candidate variables, in this study I have employed – as shown in the previous six analysis sections – a number of quantitative feature selection techniques that are better suited to the size of the data utilized here. With the initial set of variable selection conducted in the preceding sections, I now employ one additional, comprehensive machine learning-based feature selection method called *stability selection*.

3.10.1 Feature Selection Algorithm

A feature selection (variable selection) algorithm is a machine learning technique designed to help *select*, but not *fit* a model. In other words, feature selection is valuable for deciding which variables should be included in an explanatory model, but is not a replacement for the statistical analysis of that model. Feature selection cannot therefore replace the fitting of a model that is provided for by, say, the ordinary least squares or logistic regression models that are typically employed in accounting studies. Feature selection is uniquely suited for *building* rather than *testing* theory.

The specific technique chosen was *stability selection* (Meinshausen & Bühlmann, 2010), a relatively recent feature selection technique in the machine learning community. Stability selection is a "wrapper" method – or a technique built upon an existing method – designed to enhance the performance of a given feature selection algorithm.[14] The core idea involved is *subsampling*; the stability selection algorithm applies a particular feature selection algorithm – in this case logistic regression – on subsamples of both *data* and *variables*. More specifically, with this technique random subsets of observations and variables are taken, and on each random subset a "penalized" logistic regression (e.g., the increasingly popular "lasso" regression shrinkage technique, Tibshirani, 1996) is fitted to each random subset; the penalty applied in each penalized regression moves the coefficients of the least informative variables toward zero. Variables with a coefficient of 0 constitute the "unselected" variables in a given random subset. After repeated iterations, feature selection scores can be aggregated and an overall "stability" score given to each feature (variable). Stronger, more frequently selected variables will have scores closer to 1, while weaker variables will have scores closer to 0, and the weakest will have scores of 0.[15] In effect, variables that occur in a large portion of the subsamples are chosen (given a non-zero stability score), while the zero-score features

[14]The term "stability" refers to the search for a stable set of variables – those that are selected with a high degree of likelihood. "Stability Selection is a very general technique designed to improve the performance of a variable selection algorithm." (Samworth & Shah, 2012). Stability selection improves on the "lasso" technique (Tibshirani, 1996), for instance, by decreasing the likelihood that a group of correlated variables will be analyzed appropriately; "The limitation of L1-based sparse models [e.g., lasso] is that faced with a group of very correlated features, they will select only one. To mitigate this problem, it is possible to use randomization techniques, reestimating the sparse model many times perturbing the design matrix or sub-sampling data and counting how many times a given regressor is selected." This is what a stability selection model is designed to do.

[15]Results are then aggregated and variables that are selected frequently are kept and given a "stability" score closer to 1, whereas unselected variables and less relevant variables are given scores closer to 0.

are almost never among the selected variables and can, as a result, be considered largely irrelevant.

In sum, in contrast to typical statistics (such as a usual regression), which have a goal of *fitting* a model, randomized logistic regression has as a goal helping decide which variables should be included in the model in the first place.[16] The goal, in effect, is "feature selection" – retaining only those variables that are useful for predicting values of the dependent variable or, conversely, removing those variables that are not helpful.

Stability Selection Results

With a binary dependent variable such as the one used here, stability selection is achieved via a *randomized logistic regression* model.[17] I thus ran a randomized logistic regression model with the binary reaction variable as the dependent variable and the remaining 75 candidate independent variables. Table 3.14 shows the 30 variables that were selected from the algorithm – that is, variables with a non-zero stability score – while Table 3.15 shows the 45 variables that were not selected (stability scores of 0).

Discussing first Table 3.14, several interesting points are revealed. To start, as expected, not all variables are selected; 45 of 75 variables are not among the variables with a non-zero stability score. What would be more informative would be efforts to link the 30 selected variables to latent concepts or sub-dimensions. Accordingly, in the second column of Table 3.14 I include both the broad category for the variable (e.g., "who") along with a provisional conceptual label (e.g., "sentiment") based on the analyses conducted in earlier sections. Starting at the broadest "category" level, we see that aspects of all five independent variable categories – *who, what, when, where, why* – are among the selected features.

In terms of concepts, the following 14 are represented, organized by category:

[16] Succinctly, the algorithm may be described as follows: "Take random subsets [of observations and variables], fit a penalized [lasso] model to each and collate the results. Variables that come up frequently are selected." http://stats.stackexchange.com/questions/130970/the-differences-between-randomized-logistic-regression-and-plain-vanilla-logisti

[17] For these analyses I used the *RandomizedLogisticRegression* method in Python's *scikit-learn* library.

Table 3.14: Variables Selected by Stability Selection Test

Variable	Variable Category & Dimension	Stability Score
Positive message	*what* (sentiment)	1.000
"Mention only"	*what* (originality/thread location)	1.000
Industry: Telecommunications	*why* (industry)	1.000
Industry: Metals	*why* (industry)	1.000
Industry: Computer Hardware	*why* (industry)	1.000
Public Direct Reply	*what* (originality/thread location)	1.000
Public Direct Message	*what* (originality/thread location)	1.000
Non-Fortune User Mention	*what* (tweet entities)	0.980
# of Followers (Firm Account)	*why* (Twitter prestige)	0.695
URL included	*what* (tweet entities)	0.640
Assets	*why* (size)	0.580
Time on Twitter in days	*who* (Twitter sophistication)	0.565
USA	*where* (location)	0.535
"Verified" Twitter Account	*who* (celebrity)	0.510
# of Lists Firm Account Appears on	*why* (Twitter prestige)	0.490
Industry: IT Services	*why* (industry)	0.480
Industry: Commercial Banks	*why* (industry)	0.460
Business Hours	*when* (timing)	0.445
Organization	*who* (entity)	0.435
English	*who* (language)	0.350
Profits	*why* (size)	0.250
Industry: Health Care – Insurance	*why* (industry)	0.240
# of Employees	*why* (size)	0.135
Industry: Internet Services & Retailing	*why*(industry)	0.060
# of Tweets Sent (Firm Account)	*why* (Twitter activity)	0.055
Industry: Entertainment	*why* (industry)	0.040
Custom Twitter Profile	*who* (Twitter sophistication)	0.025
Topic 17	*what* (topic)	0.015
Topic 9	*what* (topic)	0.010
Industry: General Merchandisers	*why* (industry)	0.005

Scores derived from stability selection procedure using randomized logistic regression.

I. *Who*

1. Language
2. Twitter sophistication
3. Entity type (organization vs. individual)
4. Celebrity

II. *What*

5. Originality/thread location
6. Tweet entities (e.g., hashtags)
7. Sentiment
8. Topic

III. *When*

9. Timing

IV. *Where*

10. Location

V. *Why*

11. Industry
12. Size
13. Twitter prestige
14. Twitter activity

There are several practical conceptual and operational insights from this analysis. One is that I propose indices be used to tap several of the concepts. For instance, there are two selected measures of the sender's level of Twitter sophistication (Time on Twitter and custom profile); these could be combined into a single index in order to further refine the model. Similarly, there are three indicators of firm size (employees, assets, and profits) that could similarly be combined into a single variable. I also propose composite indices of tweet entities (URL, non-Fortune user mention) and the firm's level of Twitter prestige (# of followers and # of lists it appears on). These steps would reduce the number of variables to be tested in an empirical test from 30 to 25.

The findings further show that, while two tweet topics are selected, the stability scores are quite weak; at 0.010 and 0.005, the two topic dummies have the second and third weakest stability scores of the 30 selected variables. With only two of 20 topic dummies among the selected variables – and with weak stability scores at that – topic can likely safely be omitted from multi-variate analyses without losing much explanatory power.

The findings from the stability selection test also reinforce the ideas noted earlier concerning industry. Table 3.14 shows industry to be relevant, with 9 of 30 variables being industry dummies. In line with earlier recommendation, I propose including all the industry dummies as controls in order to run a fixed effects model that is more typical of accounting research (e.g., Dhaliwal et al., 2011). With these 9 binary variables removed from the main analysis and after creating the four composite indices just described, 16 independent variables would remain. The end result is a much more manageable and more parsimonious model.

With respect to the non-selected variables, Table 3.15 also contains some important insights. To start, of the 45 variables (all with stability scores of 0), 18 are topic dummies. Overall, the broad topic covered in a tweet does not seem to carry much explanatory value in this model. Similarly, 18 of the unselected variables are industry dummies. While 9 of the 27 industry variables were selected (Table 3.14), the majority were not. Industry appears to be only moderately important for determining how firms react to public messages.

Of the nine remaining variables, we see that two represent tweet entities (photo, hashtag), one represents timing (tweet-chat hashtag), one represents sentiment (negative message),[18] three represent the sender's level of Twitter activity (# of tweets sent, # of tweets favorited, # of users followed), one represents the sender's Twitter prestige (# of followers), and one represents the firm's level of Twitter activity (# of users followed). Most surprising of all these is the measure of Twitter prestige, for the number of followers is one of the strongest variables in the social media literature (Bakshy et al., 2011, e.g.,). The most plausible explanation is that, among the selected variables, the related concepts of Twitter sophistication (represented by Time on Twitter and custom profile) and celebrity (represented by "verified"

[18] As a brief example, we can interpret the non-selection of *negative message* as follows: while the bivariate analyses suggested negative messages were less likely to receive a reaction, the stability selection tests suggest the difference is due to the fact that the difference between negative and neutral messages is not sufficiently strong to warrant inclusion in the model; seen from a different angle, the variable is not selected because of a relative lack of variation – negative tweets rarely receive a response from the Fortune firms.

Table 3.15: Variables *Not* Selected by Stability Selection Test

Variable	Variable Category & Dimension	Stability Score
Topic 10	*what* (topic)	0
Topic 8	*what* (topic)	0
Topic 7	*what* (topic)	0
Topic 6	*what* (topic)	0
Topic 5	*what* (topic)	0
Topic 4	*what* (topic)	0
Topic 3	*what* (topic)	0
Topic 20	*what* (topic)	0
Topic 19	*what* (topic)	0
Topic 18	*what* (topic)	0
Topic 16	*what* (topic)	0
Topic 15	*what* (topic)	0
Topic 14	*what* (topic)	0
Topic 13	*what* (topic)	0
Topic 12	*what* (topic)	0
Topic 11	*what* (topic)	0
Topic 2	*what* (topic)	0
Topic 1	*what* (topic)	0
Photo	*what* (tweet entities)	0
Negative Message	*what* (sentiment)	0
Industry: Miscellaneous	*why* (industry)	0
Industry: Utilities: Gas & Electric	*why* (industry)	0
Industry: Semiconductors & Other Elec. Comp.	*why* (industry)	0
Industry: Pharmaceuticals	*why* (industry)	0
Industry: Network & Other Comm. Equipment	*why* (industry)	0
Industry: Motor Vehicles & Parts	*why* (industry)	0
Industry: Medical Products & Equipment	*why* (industry)	0
Industry: Mail, Package, & Freight Delivery	*why* (industry)	0
Industry: Insurance: Property & Casualty	*why* (industry)	0
Industry: Household & Personal Products	*why* (industry)	0
Industry: Food Consumer Products	*why* (industry)	0
Industry: Electronics, Electrical Equip.	*why* (industry)	0
Industry: Diversified Financials	*why* (industry)	0
Industry: Computers, Office Equipment	*why* (industry)	0
Industry: Computer Software	*why* (industry)	0
Industry: Chemicals	*why* (industry)	0
Industry: Apparel	*why* (industry)	0
Industry: Aerospace & Defense	*why* (industry)	0
Hashtag included	*what* (tweet entities)	0
# of Tweets Sent	*who* (Twitter activity)	0
# of Users Followed	*who* (Twitter activity)	0
# of Followers	*who* (Twitter prestige)	0
# of Tweets Favorited	*who* (Twitter activity)	0
# of Users Followed (Firm Account)	*why* (Twitter activity)	0
Tweet-Chat Hashtag	*when* (timing)	0

Scores derived from stability selection procedure using randomized logistic regression.

account) are more important indicators – at least for firms – of *who* is worthy of a reaction.

3.10.2 Proposed Theoretical Model

Bringing together all of the above findings, I propose the explanatory model presented in Figure 3.40. The figure shows 14 explanatory concepts, organized by the five broader categories, which in turn are organized into three even larger categories of variables: 1) firm characteristics, 2) message characteristics, and 3) sender characteristics. I propose the concepts shown under these three categories represent important determinants of whether and how a CSR-related message from a member of the public will receive a reaction from a large *Fortune* firm.

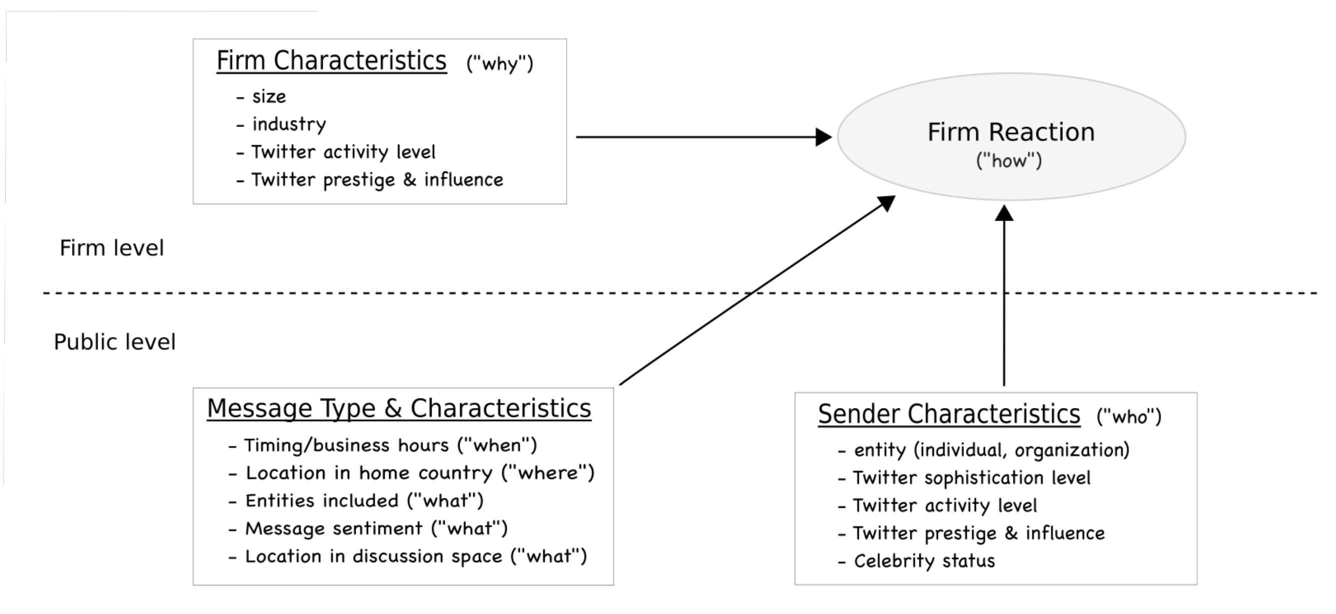


Figure 3.40: Theoretical Model: Determinants of Firm Reactions to Community Messages

3.11 Conclusions

In this paper I have sought to address three questions. One, what is the nature of the public's messages sent to *Fortune 200* firms' CSR-related social media accounts? Two, how do firms react to these public messages? And three, what are the determinants of whether a public message receives a firm reaction? I will summarize the findings and implications in each of these three areas.

3.11.1 The Nature of Public CSR Messages

In the course of conducting the inductive analyses, the study has delivered insights on more than 80 different features of the messages the public sends on Twitter, in which they are either talking about firms' CSR performance with others or talking directly to firms about the company's CSR performance. I have examined a number of relevant areas, including characteristics of the messages, the message senders, and the reacting firms. In so doing, I have provided considerable data on the nature of firm-targeted CSR messages the public is sending on Twitter, including where they come from, when they are sent, the topics and types of content included in the messages, as well as who is sending them and who is reacting.

Many of these variables are new to the CSR literature and, even more so, to the accounting literature. In one sense this is due to the nature of social media-based Big Data, which allow for the quantitative investigation of new variables such as time stamps and geo-location. In another sense, these variables are new to the literature because the phenomenon itself has been changed by social media. Simply put, the conduct of "CSR" is different on social media than it is on prior platforms. For instance, the public messages I have examined here represent a form of what marketing scholars would refer to as *user-generated content* (Smith et al., 2012) – a type of public-to-firm communication for which there was no need to theorize in the accounting and CSR literatures until recently. While my inductive analyses have provided important first steps, considerable future work is needed to flesh out the range of empirical, methodological, theoretical, and practical implications of this new public forum for the practice of CSR.

3.11.2 The Nature of Firm Reactions

The data showed four general types of reactions the firms can make to each message. They can favorite the message, they can retweet the message, they can reply to it, or they can simply ignore it. The analyses showed that firms are most likely to take the latter approach. Much more likely. In the end, firms, similar to investors in the capital markets, have *limited attention* (Hirshleifer & Teoh, 2003) and there may be considerable noise in the stream of data flowing to firms on Twitter.

At the same time, 3.3% of messages *do* receive a reaction. Not all the messages are ignored by these firms, and preliminary analyses suggest an important degree of back-and-forth dialogue between firms and stakeholders on Twitter that was found to be missing from earlier technology platforms (Unerman & Bennett, 2004).

At the same time, it is important to note that two-way dialogue manifests itself in only one of the three active reactions: replies. Of the three reactions, a reply is the most intensive way of devoting managerial attention (Mitchell et al., 1997) to public stakeholders. Moreover, it is the reaction with the most readily available conceptual analogue (namely, dialogue) in the CSR (e.g., Kent & Taylor, 2016) and accounting (Unerman & Bennett, 2004) literatures. Retweets, in turn, generalize to the notion of *dissemination* (Blankespoor et al., 2014) and represent a powerful tool for diffusing relevant messages. There exists no ready analogue in the accounting literature for the "like" or the "favorite." Yet favoriting on Twitter provides two functions. One, it allows the firm to archive the message. Two, and more importantly, it allows the firm to send a signal to the message sender that it values the message that has been sent. It is therefore likely a key determinant of the composition of the firm-centered network that ultimately takes shape around the company's Twitter account. Future research is needed to deliver insights regarding such ideas as well as, more broadly, to further weave the three key reactions into the accounting and CSR literatures.

3.11.3 The Determinants of Firm Reactions

The bulk of the analyses were devoted to the third task – developing insights into what drives firms to react to public messages. To help understand this behavior, I have provisionally framed it as related to two accounting phenomena: reporting and accountability. The highly interactive nature of the social media technology has engendered a system in which there is little to no barrier to the typical user logging on to Twitter and demanding an "account" (Ahrens, 1996) from the firm for its CSR actions. Same for a member of the public who simply wishes to engage in dialogue with the firm about the best way to reach some CSR-related societal goal. In each case, the firm's reaction determines the type and level of accountability or reporting quality that takes place.

Yet the existing literature was not well situated to bringing theoretical insights to bear on such an interactive, message-based system of reporting and accountability. For this very reason, inductive theory-building was the most appropriate approach. The current study has innovated in this task by incorporating a number of quantitative machine learning methods into the usual set of qualitative techniques. Such an approach was well suited to this study given the nature of the Big Data examined. The techniques employed could meaningfully be used as a template in future studies that rely on Big Data, which is becoming increasingly relevant in a number of areas of accounting research (Vasarhelyi et al., 2015).

Together, these qualitative and quantitative theory-building analyses culminated in the presentation of a proposed explanatory model that contains 14 distinct concepts spread across 3 broad categories. The next step will be to refine and test the proposed model. That is the task to which I turn in the following paper.

Chapter 4

Why Micro-Reporting Happens: The Determinants of ***Fortune 200*** Firms' Reactions to CSR-related Public Comments on Social Media

Abstract

Social media represents a form of public space, a communication market in which ideas, information, rumors, opinions, and sentiments compete for public attention. Social media also enable the public to enter this market at little to no cost and actively engage with firms – not only praising, questioning, and chastising firms in turn for their CSR performance but entering into real two-way dialogue with the corporation. I posit how firms respond to these messages from the public represents a new form of *micro-accountability* and *micro-reporting* behavior that has yet to be examined by the accounting literature. To examine such behavior, this paper uses data on 163,402 CSR-focused public tweets to refine and then test a model designed to explain which types of messages from public audiences garner some form of "reaction" from large firms. Specifically, the study employs data on the Twitter actions taken by 42 *Fortune 200* firms' CSR-focused Twitter accounts, which are then linked to the 163,402 messages sent in 2014 by members of the public mentioning, discussing, or addressing these 42 accounts. Reactions are conceptualized using a four-category information processing framework in which a public message can be shared, replied to, "liked," or ignored. These reactions become the dependent variables in a series of logistic regressions used to test the theoretical model. The implications for the accounting literature of these new forms of public, dynamic, and interactive reporting and accountability behaviors are discussed.

Keywords: Corporate accountability, corporate social responsibility, CSR disclosure, micro-reporting, stakeholder engagement, social media, Big Data

4.1 Introduction

Over the past several decades, the financial accounting world has undergone a significant change in terms of the amount of off-cycle and discretionary reporting (e.g., Verrecchia, 1983) that takes place in areas ranging from earnings surprises (Kasznik & Lev, 1995) to sustainability reporting (Dhaliwal et al., 2011). Over the same period, the field has been disrupted by a proliferation in reporting media such as XBRL (Debreceny, 2001; Hodge et al., 2004), open conference calls (Bushee et al., 2003), and websites (Trabelsi et al., 2004). Yet despite the considerable developments, such changes in management reporting decisions and styles and technological developments merely represent different ways of reporting or disclosing information – that is, different vehicles for the one-way transmission of information from firms to stakeholders. Until very recently, there have been few venues in which two-way communication or interactivity could occur – fora in which the public and other stakeholders might publicly react to and comment on firms' performance. As a result, seen as a field of communication, the accounting literature has conceptualized financial reporting as exclusively *one-way*, and largely regularly scheduled, communication.

I posit social media have disrupted this reporting field by offering a more dynamic, more interactive *public space* (Neu, 2006) in which discussion of firms' performance can take place. In this space citizens, firms, and interest groups alike can debate, discuss, denigrate, deny, and dialogue about such core issues as a firm's level of corporate social responsibility (CSR). This communicative space effectively offers a number of challenges as well as opportunities for scholars interested in corporate reporting and accountability. To help advance this discussion, the present study examines the relationship between the CSR-related public messages audience members send on social media and firms' reactions to these messages.

Specifically, it refines and then tests a model of the determinants of firm reactions to public CSR messages developed by Saxton (2016a). To test the model, I gather data on all 163,402 Twitter messages sent in 2014 that mention the 42 CSR-focused Twitter accounts managed by *Fortune 200* firms. I then relate characteristics of these public messages to whether the message receives a reaction from the firm in the form of a *retweet* (share), a *like*, or a reply. A series of logistic regressions are then used to test the model.

The results largely confirm the theoretical model. In finding a series of relationships between firm reactions and various characteristics of the messages sent, the users who sent them, and the firms targeted, the study contributes to the literature on to whom and for what firms are accountable. In the end, firms' decisions to respond or not respond constitute a new form of public, dynamic, interactive reporting and accountability behavior that has yet to be addressed by the extant accounting literature. The current study contributes to and seeks to help propel this nascent body of research.

The remainder of the paper is structured as follows. Section 2 examines the existing CSR literature, particularly as it pertains to reporting decisions on new media. Section 3 presents the theoretical model and hypotheses. Section 4 presents the method, including the sample, data, and variable measurement. Section 5 presents the results from the logistic regressions. Section 6 discusses the implications of the findings for the accounting and CSR literatures.

4.2 Social Media and CSR Micro-Reporting

4.2.1 Existing CSR Literature

There now exists an extensive body of research on CSR. One influential stream of research deals with whether CSR disclosure can influence firm performance (e.g., Cochran & Wood, 1984; Mishra & Suar, 2010). An equally large (and more critical) stream of research has questioned firms' motivations for engaging in CSR disclosure or communication, including impression management, legitimacy, and "greenwashing" (Du & Vieira, 2012; Lyon & Montgomery, 2013; Neu et al., 1998; Patten, 2002; Warsame et al., 2002). Public relations and communications approaches, in turn, have generally examined CSR in terms of the different stakeholder communication strategies employed by the firms (Morsing & Schultz, 2006).

Recently, scholars have begun to delve into the role new and social media may have on

CSR communication. In terms of older forms of new media, Unerman & Bennett (2004) conducted an early study (1999 data) of stakeholder engagement on CSR (Fieseler & Fleck, 2013; Fieseler et al., 2010). More recently, studies have begun to emerge on social media such as Twitter, Facebook, and LinkedIn, with studies looking at the role of social media in creating CSR-centered participative interactions and "citizenship arenas" (Fieseler & Fleck, 2013; Whelan et al., 2013). Research has also explored the potential for social media to be used for "greenwashing" CSR efforts (Lyon & Montgomery, 2013). Initial research has also been done on the determinants of firms' use of Twitter for CSR communications (Lee et al., 2013).

Perhaps one of the most significant implications of this early research concerns the directionality of the reporting that takes places. At the broadest level, social media seem to have engendered a more interactive communicative environment (e.g., Kane et al., 2014; Scott & Orlikowski, 2012). While research on earlier platforms such as websites and discussion boards showed heavy use of one-way reporting (Cho et al., 2010) and an absence of true dialogue between firms and the public (Unerman & Bennett, 2004), the potential for dialogic and interactive engagement has expanded with the diffusion of social media platforms sites as LinkedIn (2003), Facebook (2004), YouTube (2005), Twitter (2006), Instagram (2010), and Pinterest (2010). With 50 million users on Pinterest, 236 million on Twitter, 296 million on LinkedIn, 300 million on Instagram, 800 million on China's Tencent QQ, and over 1.23 billion users on Facebook (Statista, 2015), social media platforms have a vast audience.

Yet the term "audience" is misleading; what distinguishes social media is that the public is not only the passive recipient of content but an active participant in the creation of that content (e.g., Smith et al., 2012). Social media sites can thus effectively be distinguished from the first generation Internet technologies (such as the traditional website) in terms of the substantially heightened opportunities for direct interactivity, two-way exchange of information, network connectivity, and the creation and exchange of user-generated content (Scott & Orlikowski, 2012; Suddaby et al., 2015).

Within such a heavily interactive arena, it is perhaps not surprising that Saxton (2016b) found firms were using Twitter not strictly as a reporting vehicle but as a *communication* vehicle – a place for not only one-way reporting but also public education, mobilization, and interactive two-way dialogue. It is within this context that the idea of firm/public dialogue has garnered greater interest from CSR scholars (Castelló et al., 2015; Colleoni, 2013; Gómez-Vásquez, 2013; Kent & Taylor, 2016). Several of these studies examine the public reaction to firms' CSR messages on social media (Castelló et al., 2015; Colleoni, 2013; Gómez-Vásquez, 2013), yet except for several working papers (Saxton et al., 2016; Saxton, 2016b), none are in the field of accounting. There is thus a need to more explicitly link the communicative phenomena to existing accounting issues and concepts.

As others have done with activists on websites to engender change in CSR-related issues (de Bakker & Hellsten, 2013), there is also a need for research that *brings the public in* and looks at more intense firm/public interactions around CSR issues. In the present study I aim to do this by looking at something new: how firms respond to messages from the public. Social media essentially constitute a "virtual town hall" in which the public is able to easily comment on and ask questions of firms' CSR performance. What's more, the social media platforms allow anyone with access to the site to view and gain access to these public messages and also see whether and how firms respond. This will be the first study in the CSR literature in any field, accounting or otherwise, to examine firms' reactions to public CSR messages.

Such analyses are facilitated by the core characteristics of social media. Especially relevant is the fact that the key affordance of social media are *messages* – whether a video on YouTube, an image on Instagram, a "pin" on Pinterest, a tweet on Twitter, or a status update on Facebook – the heart of social media are the brief dynamic updates that users send to their audiences on a regular basis (e.g., de Vries et al., 2012; Jansen et al., 2009; Waters & Jamal, 2011). This characteristic means the interactivity that occurs on these sites is message-based interactivity (Rafaeli & Sudweeks, 1997); in order to understand the

CSR communication, it is thus necessary to focus on the dynamic sending and receiving of these micro-messages. Social media are also notable for a second characteristic: the high degree of connectivity (e.g., Kane et al., 2014). Any interactions that occur are thus mediated by the series of formal connections among users. Any attempt to connect firms' CSR activities on social media to the accounting literature needs to incorporate these two features.

4.2.2 Micro-Reporting and Micro-Accountability on Social Media

In regard to the relationship to the accounting literature, I posit that firms' reactions to public questions, comments, and messages represents a form of behavior that spans two phenomena important to accounting scholars: accountability and reporting. Given the centrality of tweets and other forms of micro-messages noted above, each reaction is a form of *micro-reporting* or *micro-accountability* effort. In fact, given the argument that accountability occurs around and through specific accounting *artifacts* such as formal reports and documents (Roberts, 1991), tweets can be seen to represent a new, micro form of accounting artifact.

At the same time, given the importance of the formal network connections among users, what we see on social media is *connected* accounting, reporting, and accountability behaviors. Accountability fundamentally involves the giving and receiving of accounts (Ahrens, 1996; Roberts & Scapens, 1985), yet in prior research it has been difficult to see who is "demanding" the accounts made of firms. Social media makes these demands – and the accounting artifacts through which the demands occur – public visible and measurable. In effect, because of how the messages flow from and are received by specific, publicly identifiable users, we are able to see not only *for what* but also *to whom* firms are engaged in reporting and "account-giving" behaviors.

In short, I argue how firms respond and react to public comments, queries, and questions represents their commitment to public *account-giving* and *reporting* behavior. With the development of social media, accounting scholars can now study such micro-level forms of

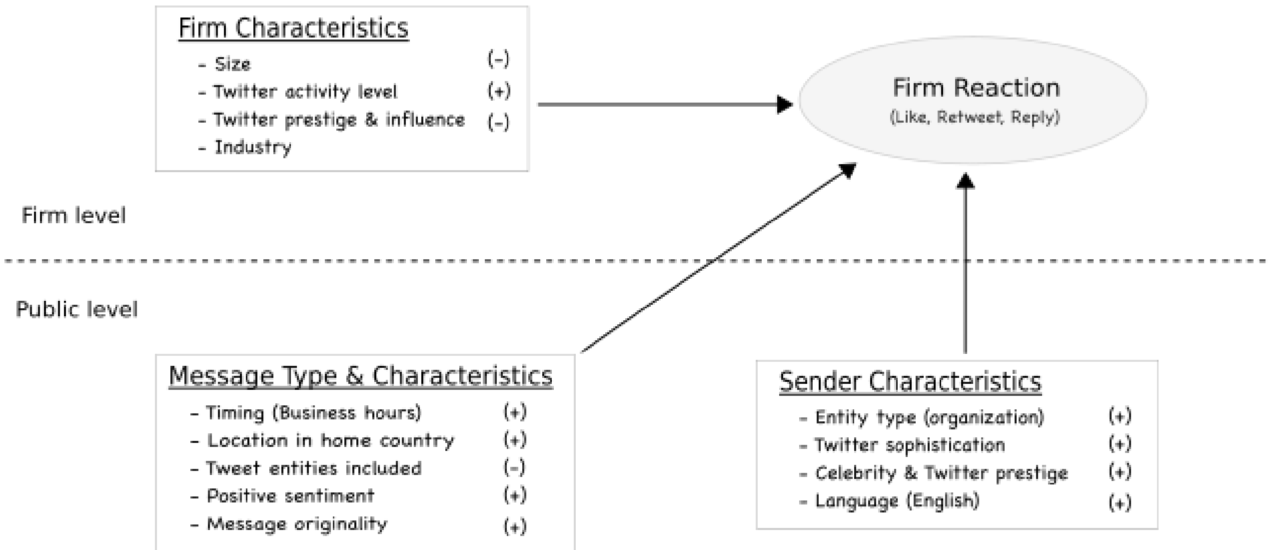


Figure 4.1: Hypothesized Determinants of Firm Reactions to Public Messages

accountability and reporting efforts.

Because it is a novel phenomenon the literature is undeveloped. To develop conceptual and theoretical insights Saxton (2016a) thus conducted detailed inductive analyses into the nature of public messages and firm responses as well as the relationship between the two. That paper culminated in the presentation of a general model of the determinants of firms' reactions to public CSR messages. In this paper I refine and then test the model.

4.3 Theoretical Model and Hypotheses

In this section I refine the Saxton (2016a) model in developing a set of testable hypotheses. The refined model, which adds the proposed directionality of the relationships, is shown in Figure 4.1.

4.3.1 Firm Reactions

The end point of the model – the dependent variable – is the *firm reaction*. Saxton (2016a) found there were four types of reactions on Twitter: each message could be "liked" (favorited), retweeted (shared), replied to, or ignored. Saxton (2016a) contains a detailed discussion of each of the four types of reactions. The great majority of messages were found to have been ignored, while roughly equal amounts of messages were liked, retweeted, or given a reply.

Building on Saxton (2016a), I argue that, at the broadest level, the three active tactics should be seen as reflections of firms' managerial attention (Mitchell et al., 1997) to public stakeholders. In line with stakeholder theory, who and what firms are paying attention to is a reflection of priorities. Given that the managerial effort is in the domain of CSR reporting and accountability, the managerial attention that takes place via these reactions constitutes CSR-focused micro-accountability and micro-reporting efforts. In line with the ideas noted earlier, however, what occurs is not solely reporting but, more generally, *communication*, covering not merely one-way reporting or account-giving but public education, two-way dialogue, and other forms of communication (see Saxton, 2016b).

These ideas are posited on the notion that firms' goals from engaging in these CSR-related efforts are multi-faceted. Not only do CSR activities – given their effects on employee, consumer, investor, and public policy outcomes – generally span market and non market strategy (Bach & Allen, 2010), but these efforts may improve public perceptions of the firms' CSR performance (Saxton, 2016b) and help enhance accountability.

In brief, the firms' CSR-focused Twitter accounts provide a venue in which firms – through their replies, retweets, and liking actions – can give accounts of their behavior, send micro-reports, and engage in dialogue with concerned stakeholders, while at the same time actually *conducting* CSR by helping provide fora for public education about CSR issues. Of course, there is the very real possibility low-performing firms will engage in social media-based CSR to "greenwash" their performance (Lyon & Montgomery, 2013) or engage in impression management (Neu et al., 1998), yet that is beyond the scope of the current study. The focus of this paper is not on the broad underlying motivation for engaging in online CSR but is instead on what leads firms to react in different ways to the messages they receive on these CSR platforms.

As shown in Figure 4.1, the model proposes three main sets of features as determinants of firm reactions: firm characteristics, message characteristics, and sender characteristics. Using this model as a base, I develop and present a series of 12 hypotheses to be tested in this paper.

4.3.2 Sender Characteristics

The first main element focuses on characteristics of who sends the message. The accountability literature has long raised the question of to whom firms and governments and nonprofit organizations are and should be accountable (Gray et al., 1987; Tower, 1993). Leaving aside the normative question of to whom organizations *should* be accountable, stakeholder theory (Mitchell et al., 1997) has argued firms are more likely to devote managerial attention to more powerful stakeholders. This idea is captured in various ways through the first three hypotheses.

To start, Saxton (2016a) proposes the entity type of the message sender helps determine whether the message receives a reaction from the firm. Specifically, the argument is that messages from organizations will be more likely to receive a response than those from individuals. In line with Mitchell et al. (1997), I argue organizations will generally be seen as more important and more powerful stakeholders than individuals, making firms more likely to react to messages sent by organizations. My first hypothesis (H1) is thus:

Hypothesis 1. *Firms are more likely to react to messages sent by organizations.*

A second relevant characteristic of the message senders is their level of sophistication on Twitter. Accounts that fail to customize their profile page, for example, are more likely to be considered a "spam" or "Twitter-bot" account. Even if they are "real" users, the perceived lack of sophistication may render firms less likely to believe they need to engage with the given user, viewing the user as a less relevant, more low-quality stakeholder or information intermediary. Saxton & Guo (2015) argue the key immediate outcomes of firms' efforts on social media is the acquisition of social media-based social capital. Given that users with less sophisticated Twitter accounts are less likely to be able to increase the firm's social capital, firms may be more prone to ignore such users. This argument is captured in H2:

Hypothesis 2. *Firms are more likely to react to messages sent by users with more sophisticated Twitter accounts.*

A related sender-focused hypothesis concerns the user's relative level of prestige or influence. Certain users on Twitter will simply be more valuable: a "celebrity" (e.g., a user with a "verified" Twitter account, Saxton, 2016a), or a user with a large number of followers (e.g., Bakshy et al., 2011), will bring considerably more attention to the firm's activities than a user with only a handful of followers. H3 therefore argues for a link between prestige and reactions:

Hypothesis 3. *Firms are more likely to react to messages sent by more prestigious Twitter users.*

Lastly, Saxton (2016a) argued that messages sent in English will be more likely to receive a reaction. H4, the most intuitive of all the hypotheses, relates reactions to the language employed by the message sender:

Hypothesis 4. *Firms are more likely to react to messages sent by users who communicate in English.*

4.3.3 Message Characteristics

The second main element focuses on characteristics of the messages that are sent. Saxton (2016a) found three sets of characteristics: 1) timing, 2) location, and 3) content.

Timing

With the prevalence of event studies in accounting research, the issue of timing is in some ways common in the accounting literature. It has yet to gain much traction in the CSR literature, however. Arya & Zhang (2009) have examined the timing of CSR initiatives and activities at the annual level. However, the time stamps that are included with social media artifacts make visible the precise times at which the messages are sent. This enables a more fine-grained analysis of the relationship between timing and reactions. I posit one hypothesis to capture this relationship. Specifically, I look at the relationship between reactions and whether the message is sent during regular business hours. The reason why business hours are important is in one way obvious: that is when the employees of the firm are most likely to be working. Yet there is another reason unique to social media: On Twitter and other social media sites, the empirical reality is that messages either get a reaction almost immediately or they do not get one at all (see Saxton, 2016b). My fifth hypothesis (H5) expresses this idea:

> **Hypothesis 5.** *Firms are more likely to react to messages sent during regular business hours.*

Location

Geographic data has recently gained interest in the accounting literature (e.g., O'Brien & Tan, 2015). Geolocation data is another data feature exposed by social media, facilitating more granular analyses of location data. As a first test of this idea in the CSR context, I propose that tweets sent by users in the United States will be more likely to garner a reaction. Given that this is the home base for the firms in the sample, this idea makes intuitive sense. Moreover, in line with Mitchell et al. (1997), firms may plausibly view members of the home country as more "legitimate," which would make them more likely to receive managerial attention. H6 is thus:

Hypothesis 6. *Firms are more likely to react to messages sent from the USA.*

Content

The next three hypotheses relate to characteristics of the content of the public messages that are sent. This is an area that is largely unexamined in the accounting literature; given the lack of dialogic and interactive possibilities in reporting venues before the spread of social media, there was no feasible way of measuring what types of "account-demanding" messages firms responded to. The three hypotheses presented here thus help encourage this nascent line of research.

The first concerns the originality of the message. Twitter is in large part a dissemination network (Blankespoor et al., 2014), with shared messages, or *retweets*, comprising a large proportion of the messages that are sent. With H7, I posit that firms are more likely to ignore retweeted messages and instead devote their energy to original messages:

Hypothesis 7. *Firms are more likely to react to original messages.*

Next, most social media platforms facilitate the inclusion of specific types of *entities* in the social media messages, including on Twitter hashtags, user mentions, hyperlinks, and images (Bakshy et al., 2011; Debreceny, 2015). The social media literature has generally found a positive relationship between the inclusion of such entities and audience reactions (e.g., Bakshy et al., 2011; Saxton & Waters, 2014). Saxton (2016a), however, found evidence the relationship may be negative. I posit the difference lies in how reactions have been measured in the extant literature; namely, by looking at the *number of different users* who have liked, shared, or commented on a given organizational message. Tweet entities broaden the potential reach of a given message by, for instance, connecting the tweet to additional users and topics (e.g., Bakshy et al., 2011). While this makes sense when a message sender is trying to reach the maximum number of audience members, it may not make as much sense when a user is targeting a specific user – in this case, one of the 42 CSR accounts. In this context, the inclusion of more tweet entities may simply detract from the core message and make the firm less likely to see itself as the true focus and target of the message. Accordingly, H8 argues the following:

Hypothesis 8. *Firms are less likely to react to messages with tweet entities.*

Sentiment is a third relevant feature of message content. The negative, neutral, or positive sentiment of social media messages has recently gained attention from CSR scholars (Castelló et al., 2015; Colleoni, 2013; Saxton, 2016b). Saxton (2016a) found preliminary evidence that positive messages in particular were more likely to receive a reaction from *Fortune 200* firms' CSR accounts. H9 formally expresses this idea:

Hypothesis 9. *Firms are more likely to react to positive messages.*

Firm Characteristics

The third and final main element of the model focuses on characteristics of the firms that are making the reactions. I generate hypotheses related to three concepts examined by Saxton (2016a).

The first concept is firm size, which is generally seen in the accounting literature as a key determinant of reporting (e.g., Atiase et al., 1989; Tsang, 1998). In line with preliminary findings by Saxton (2016a), H10 argues for a negative relationship between size and the likelihood of reacting:

Hypothesis 10. *Larger firms are less likely to react to public messages.*

I further posit that firms with more prestigious Twitter accounts – those that are on more public "lists" (see Saxton, 2016a) and have more followers – are less likely to be responsive to public messages. The logic is that such firms have already developed substantial *social media capital* (Saxton & Guo, 2015), resulting in a less pressing need to engage with a high proportion of stakeholder messages. Accordingly, H11 is the following:

Hypothesis 11. *Firms with more prestigious Twitter accounts are less likely to react to public messages.*

Lastly, I posit that more active firms – particularly those that more regularly send out tweets – are more likely to respond to public messages. Prior research has pointed to wide variation in how frequently firms use social media, with some accounts being largely inactive (e.g., Gómez-Vásquez, 2013; Saxton et al., 2016; Suddaby et al., 2015). Given that inactive accounts are less likely to respond, my final hypothesis (H12) is therefore:

Hypothesis 12. *Firms with more active Twitter accounts are more likely to react to public messages.*

4.4 Method

4.4.1 Sample

The starting point for the data are all firms on the 2012 US *Fortune 200* list. I then searched for all Twitter accounts of these 200 companies. All were found to maintain a Twitter account, but I was specifically interested in accounts devoted exclusively to CSR-related issues, such as Cisco's *@CiscoCSR* or Verizon's *@VerizonGiving*. As summarized in Table 4.1, I found 35 of the 200 firms maintained such an account. Because several firms maintained more than one CSR-focused account (e.g., Microsoft's *@Microsoft_ Green* and *@MSFTCitizenship*), there were 42 accounts in total.

Table 4.1: Sample

	Firms	Accounts	Messages*
US *Fortune 200* Firms	200		
Firms without CSR-Focused Twitter Account in 2014	(165)		
Firms with CSR-Focused Twitter Account in 2014	35	42	163,402

*All Twitter messages sent by members of the public that mention the 42 CSR accounts in 2014.

4.4.2 Data and Measurement

The heart of the analysis, however, is not the accounts per se but rather the messages the public sends to these 42 CSR accounts. As noted in Table 4.1, there were 163,402 public Twitter messages, or tweets, that mentioned one or more of the 42 CSR accounts over the course of 2014. The Twitter application programming interface (API) allowed me to download all 163,402 messages through a Python script I ran twice per day over the course of 2014.

Saxton (2016a) provides examples and further details on these messages. The relevant point to recall is that, when a member of the public includes a user mention of one of the 42 CSR- focused accounts, it is initiating or engaging in a public conversation either with or about the *Fortune 200* company's CSR efforts. Examining this collection of 163,402 public messages allows me to see how members of the public are talking to and talking about firms' CSR efforts. That is the first component of the data.

Dependent Variables

The second component comprises firms' reactions to these 163,402 public messages. By examining how the firms react to these messages, we can understand how firms are engaging in dynamic, interactive reporting and accountability behaviors with members of the public. The Twitter API allowed me to download all 2014 tweets, retweets, and "likes" by the 42 accounts; matching the tweet IDs of these actions to those in the messages sent by members of the public allowed me to determine which of the 163,402 public messages the firm had "liked," retweeted, replied to, or ignored. Table 4.2 summarizes the number of public messages and firm reactions in 2014 for each of the 42 accounts.

The study is not conducted at the firm level but rather the message level. Accordingly, I created four binary (1,0) dependent variables, each measured at the level of the individual message: *Liked* is assigned a value of "1" if the mentioned firm liked the message (mean

Table 4.2: Public Mentions and Firm Reactions for Fortune 200 CSR-Focused Accounts

		Firm Reactions			
Screen Name	Public Mentions	# Replied to	# Favorited	# Retweeted	# Ignored
3M_FoodSafety	338	0	0	0	338
AlcoaFoundation	3,344	123	289	383	2,626
AmgenFoundation	673	0	3	7	663
ATTAspire	1,629	2	0	7	1,620
BofA_Community	48,789	82	0	313	48,396
CiscoCSR	12,975	102	305	67	12,543
CiscoEdu	244	1	1	1	242
CitizenDisney	3,203	0	139	1	3,063
CitizenIBM	4,651	0	0	4	4,647
ClicktoEmpower	324	0	0	0	324
ComcastDreamBig	3,680	39	0	151	3,496
DE_Youtility	7	0	0	0	7
Dell4Good	3,717	84	192	179	3,389
DellEDU	4,671	57	469	204	4,067
Dupont_Ability	1,372	7	2	18	1,350
Ecomagination	3,694	21	8	6	3,662
EnviroSears	210	1	29	3	177
FedExCares	84	1	0	1	82
FordDriveGreen	6,709	8	94	13	6,605
FundacionPfizer	176	6	0	0	170
GEHealthy	1,082	104	1	1	976
GoogleStudents	4,582	0	0	1	4,581
GreenIBM	3	0	0	0	3
HeartRescue	315	1	5	10	301
HoneywellBuild	910	6	7	25	873
HPGlobalCitizen	419	7	0	4	408
HumanaVitality	1,859	78	90	3	1696
IBMSmartCities	5,212	10	9	29	5167
IntelInvolved	11,164	9	115	13	11,034
MathMovesU	1,994	17	38	4	1,942
Microsoft_Green	2,588	18	120	22	2,434
MSFTCitizenship	12,255	90	135	76	11,975
NikeBetterWorld	476	0	0	0	476
PG_CSDW	1,062	8	20	21	1,016
PPGIdeascapes	1,502	0	0	1	1,501
PromesaPepsico	5	0	0	0	5
SprintGreenNews	332	0	2	2	328
TeachingMoney	201	3	0	9	189
TICalculators	4,225	89	54	150	3,952
VerizonGiving	5,887	85	744	80	5,042
WalmartAction	3,084	14	49	32	2,991
WalmartGreen	4,034	11	0	80	3,945

Note: Table shows number of public mentions and firm reactions for each account over the course of 2014.

= 0.018); *Retweeted* receives a value of "1" if the message was retweeted (mean = 0.012); *Replied* has a value of "1" if the firm replied to the message (mean = 0.007); and *Reaction* is given a value of "1" if the message received either a like, a retweet, or a reply (mean = 0.033). In effect, only 3.3% of all tweets (n=5,455) received a reaction; the other 96.7% (n=157,947) were ignored. Table 4.3 shows descriptive statistics for these and all other all model variables.

Independent Variables: Message Sender Characteristics

A total of 15 independent variables were also created using a variety of sources. First, I examine features of the senders of the public messages. In total there were 82,769 different Twitter users who mentioned one of the *Fortune 200* firms in their messages over the course of 2014.

To code data on sender characteristics, I gathered additional data from the Twitter *user* API. To start, I use the language meta-data provided by the Twitter API to create the binary variable *English*.

To code for whether the message sender is an organization or an individual, I first used data in the users' self-described description field to hand code 700 message senders as being an organization or an individual, and then trained and tested a support vector machine (SVM) machine learning algorithm (e.g., Go et al., 2009) that was then used to code the remainder of the 82,769 users.[1] The binary variable *Organization* is assigned a value of "1" for all messages sent by users from an organization, with those from individuals having a value of "0." As shown in Table 4.3, the mean is 0.211 (s.d. = 0.408); 34,391 of the messages were from organizations rather than individuals.

Third, I created a composite index to measure the Twitter sophistication of the message sender. Two variables were used to create the index. One, the number of days a user has been on Twitter as of January 1, 2014 was calculated using the account creation date information

[1] The trained SVM algorithm achieved 88.9% accuracy compared to the manually coded users. Cases where the Twitter user did not provide a description were not codeable.

Table 4.3: Summary Statistics

	count	mean	std	min	max
Dependent Variables					
Any Reaction	163402	0.03	0.18	0	1
Liked	163,402	0.02	0.13	0	1
Retweeted	163,402	0.01	0.11	0	1
Replied to	163,402	0.01	0.08	0	1
Sender Characteristics					
Language					
English	163,402	0.96	0.19	0	1
Entity					
Organization	163,374	0.21	0.41	0	1
Twitter Sophistication					
# of Days on Twitter	163,402	908.41	704.48	-360	2,727
Custom Profile	104,616	0.68	0.47	0	1
Twitter Sophistication Index	104,616	0.09	1.60	-3.26	3.26
Prestige/Celebrity					
'Verified' Twitter Account	104,616	0.03	0.16	0	1
# of Followers	163,402	7,672.28	113,581.57	0	13,478,517
Message Characteristics					
Sentiment					
Positive Sentiment	163,402	0.18	0.38	0	1
Originality & Thread Location					
Original Tweet	163,402	0.36	0.48	0	1
'Mention Only' Tweet	163,402	0.21	0.41	0	1
Public Direct Reply	163,402	0.09	0.29	0	1
Public Direct Message	163,402	0.06	0.23	0	1
Tweet Entities					
# Entities	163,402	2.22	0.89	0	4
Timing					
Weekday Business Hours	163,402	0.52	0.50	0	1
Location					
Located in USA	163,402	0.42	0.49	0	1
Firm Characteristics					
Twitter Prestige					
# of Followers	163,402	173,744.87	195,229.50	4	542,203
# of Lists Account Appears On	163,402	927.97	922.50	0	4,832
Twitter Prestige Index	163,402	0.00	1.71	-1.90	5.36
Twitter Activity					
# of Tweets Sent	163,402	5,208.78	2,951.65	1	16,871
Firm Size					
Assets	163,402	730,156.06	902,617.13	13, 209.00	2,104,534.00
Profits	163,402	8,494.43	5,938.32	-4326.00	22,074.00
Employees	163,402	258,989.75	42,3495.10	17,900	2,200,000
Size Index	163,402	0.00	1.67	-3.40	5.33

provided by the Twitter API. Two, I measured whether the user had created a custom Twitter profile (again, using information provided by the Twitter API). Both variables were standardized (z-scores) and values added together to create the *Twitter Sophistication Index* variable (mean = 0.088, s.d. = 1.597).

Lastly, two different variables are used to tap the prestige or celebrity of the message sender. One relies on a feature that is known in Twitter as a "verified" account; these accounts are noted by the addition of a unique icon added by Twitter to the user's profile page, which indicates the user's status as a noteworthy individual or organization such as a celebrity, entertainer, journalist, company, author, or sport star. Messages sent from a user with a verified account received a score of "1" on *Verified Account*, otherwise "0." The mean value of 0.027 (s.d. = 0.163) reflects the 2.7% of messages (2,863 of 104,616 coded tweets) that were sent from verified accounts. Because values are missing from almost 60,000 of the tweets, an alternative measure, the *Number of Followers*, is used for sensitivity analysis. This variable is supplied by the Twitter API and indicates the number of other Twitter users who choose to follow the message sender on Twitter. It ranges from a low of 0 to a high of 13,478,517 followers (for *@ESPN*), with a mean of 7,672.3 (s.d. = 113,581.6).

Independent Variables: Message Characteristics

I measure eight variables reflecting different features of the messages that are sent. First, *Positive Sentiment* is a binary measure with values of "1" indicating messages conveying positive affect. To measure sentiment I relied on a supervised machine learning technique popular in computer science studies (e.g., Go et al., 2009). Specifically, a hand-coded dataset of 1,000 tweets was used to train an SVM model. The trained model was then used to code each tweet as either negative, neutral, or positive sentiment.[2] I found 17.7% (n = 28,948) had a positive sentiment and thus coded as "1" on the variable *Positive Sentiment*, with the remainder of the messages being either neutral (n = 130,427, or 79.8%) or negative (n =

[2] The level of accuracy compared to hand coding for the 3 values was 94.5%, 80.2%, and 89.0%, respectively.

4,027, or 2.5%) and thus coded as "0."

Next, to tap originality, *Original Tweet* is coded "1" if the tweet is not a retweet; 35.6% (n = 58,186) of the 163,402 messages were original, with the remainder being retweeted messages. For additional tests, I also operationalize three variables that reflect the three main types of original tweets in terms of where they lie in existing conversational threads. Each original tweet was assigned a value of "1" on one of the following three variables: *Public Direct Reply*, or a tweet that is a reply to an existing tweet (mean = 0.09); *Public Direct Message*, or a tweet that begins a new conversation thread but targets the message at a specific Twitter user (mean = 0.06); and *"Mention Only" Tweet*, or a tweet that initiates a new conversation thread but is "undirected," or not targeted at any one user (mean = 0.21).[3]

I then operationalized a composite index to capture the number of different tweet entities included in each public message. Specifically, I combined four binary variables that indicated whether a tweet contained, respectively, a hashtag, a URL, a non-Fortune user mention, and a photo. The range for the index variable *# of Entities* is thus from 0 to 4; the mean value is 2.22 (s.d. = 0.89).

Weekday Business Hours, in turn, is a binary variable that indicates whether the tweet was sent from 9am - 5pm on Monday to Friday. Measurement was facilitated by the timestamp provided by the Twitter API for each tweet. As shown in Table 4.3, 52% of tweets have a value of "1" on this variable (s.d. = 0.50).

The final message characteristic operationalized is the binary variable *Located in USA*. Values of "1" are assigned to messages sent from users whose self-described location is within the United States. To create the variable I ran users' self-described locations through two geographical databases (*GeoNames* and *ArcGIS*) to extract the country in which the user is located. Interestingly, only 42% (s.d. = 0.49) of messages were sent from users located within the USA.

[3] For a more in-depth discussion of these three variables, see Saxton (2016a).

Independent Variables: Firm Characteristics

The final three variables reflect firm characteristics. The first is a *Twitter Prestige Index* that combines values on two variables, the number of followers of the firm's account and the number of public "lists" the account appears on. The latter is an idiosyncratic feature of Twitter that shows the number of lists other Twitter users have created and placed the firm on. The higher the number of lists on which the firm is placed, the higher its level of prestige and/or influence. The number of followers taps the same concepts. Consequently, they are combined into an additive index after standardizing scores. Table 4.3 shows summary statistics for the *Twitter Prestige Index* as well as its two constituent elements (*# of Followers, # of Lists Account Appears on*).

# of Tweets Sent, meanwhile, captures the number of tweets the firm has sent through the account since joining Twitter. It is an indicator of the general level of activity of the account. The variable ranges from a low of 1 tweet sent to a high of 16,871 and has a mean value of 5,208.8 (s.d. = 2,951.6).

The final variable is an index of firm size built by adding composite standardized scores on total assets, profits, and number of employees. Table 4.3 shows summary statistics for the *Size Index* as well as its three constituent elements.

In addition to the descriptive statistics for all model variables shown in Table 4.3, I also compute zero-order correlations. The correlation matrix for all model variables is shown in Table 4.4.

Analysis Plan

Given the binary nature of the dependent variables, the hypotheses will be analyzed through a series of logistic regressions. In the following section I present results from a total of six regressions – three with *Reaction* as the dependent variable and one each with *Liked*, *Retweeted*, and *Replied*, respectively, as the outcome variable.

Table 4.4: Zero-Order Correlations Matrix

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11	12.	13.	14.	15.	16.	17.	18.
Any Reaction	1	0.73	0.59	0.44	0.09	0.18	0.08	0.07	0.07	0.04	0.04	0.06	0.01	0.02	-0.09	-0.10	0.06	-0.11
Liked	0.73	1	0.12	0.09	0.06	0.13	0.07	0.04	0.05	0.01	0.02	0.04	0.01	0.01	-0.06	-0.09	0.05	-0.09
Retweeted	0.59	0.12	1	0.06	0.06	0.16	-0.01	0.04	0.05	0.05	0.03	0.05	0.01	-0.01	-0.02	-0.05	0.04	-0.06
Replied	0.44	0.09	0.06	1	0.04	0.03	0.09	0.06	0.01	-0.00	0.02	0.01	0.00	0.03	-0.09	-0.04	0.02	-0.05
Positive Sentiment	0.09	0.06	0.06	0.04	1	0.02	0.04	-0.02	0.06	0.01	0.06	0.02	0.07	-0.02	-0.07	0.02	0.01	-0.01
'Mention Only' Tweet	0.18	0.13	0.16	0.03	0.02	1	-0.16	-0.12	0.09	0.05	0.00	0.09	-0.00	0.06	-0.06	0.04	0.14	-0.01
Public Direct Reply	0.08	0.07	-0.01	0.09	0.04	-0.16	1	-0.08	0.00	-0.01	-0.01	-0.05	-0.16	0.01	-0.34	0.02	-0.02	0.04
Public Direct Message	0.07	0.04	0.04	0.06	-0.02	-0.12	-0.08	1	0.03	-0.01	-0.01	0.01	-0.06	0.01	-0.19	0.02	-0.01	0.02
Sender in USA	0.07	0.05	0.05	0.01	0.06	0.09	0.00	0.03	1	0.05	0.09	0.11	0.07	0.03	-0.02	0.03	0.19	-0.06
'Verified' Account	0.04	0.01	0.05	-0.00	0.01	0.05	-0.01	-0.01	0.05	1	0.02	0.06	0.01	-0.02	0.01	0.01	0.14	-0.04
Weekday Business Hours	0.04	0.02	0.03	0.02	0.06	0.00	-0.01	-0.01	0.09	0.02	1	0.05	0.03	0.06	-0.03	-0.00	0.05	-0.03
Organization	0.06	0.04	0.05	0.01	0.02	0.09	-0.05	0.01	0.11	0.06	0.05	1	0.04	0.05	0.01	-0.06	0.11	-0.12
English	0.01	0.01	0.01	0.00	0.07	-0.00	-0.16	-0.06	0.07	0.01	0.03	0.04	1	-0.00	0.10	-0.01	0.04	-0.02
Firm Twitter Activity	0.02	0.01	-0.01	0.03	-0.02	0.06	0.01	0.01	0.03	-0.02	0.06	0.05	-0.00	1	-0.02	0.21	0.04	0.19
# Tweet Entities	-0.09	-0.06	-0.02	-0.09	-0.07	-0.06	-0.34	-0.19	-0.02	0.01	-0.03	0.01	0.10	-0.02	1	-0.02	0.00	-0.06
Firm Size	-0.10	-0.09	-0.05	-0.04	0.02	0.04	0.02	0.02	0.03	0.01	-0.00	-0.06	-0.01	0.21	-0.02	1	-0.01	0.28
Twitter Sophistication	0.06	0.05	0.04	0.02	0.01	0.14	-0.02	-0.01	0.19	0.14	0.05	0.11	0.04	0.04	0.00	-0.01	1	-0.05
Firm Twitter Prestige	-0.11	-0.09	-0.06	-0.05	-0.01	-0.01	0.04	0.02	-0.06	-0.04	-0.03	-0.12	-0.02	0.19	-0.06	0.28	-0.05	1

4.5 Results

Table 4.5 shows a series of six logistic regressions. Each regression includes industry fixed effects with standard errors clustered by firm CSR account. The first three regressions (Models 1-3) have *Any Reaction* as the dependent variable and present variations in the set of independent variables employed. Models 4 through 6, in turn, have *Liked*, *Retweeted*, and *Replied*, respectively, as the outcome variable.

The results are quite robust and consistent across all six models. Accordingly, for a more parsimonious presentation I will present the results on a variable-by-variable basis for the six models as a whole, pointing out important differences along the way.

The first hypothesis, asserting a positive relationship with reactions if the message sender is an organization, receives only partial support. The coefficient on *Organization* is positive and significant ($p<0.5$) in Models 1 and 2 (with dependent variable *Any Reaction*) and Model 5 (with *Retweeted* as dependent variable) yet does not obtain significance in Model 3 nor in Models 4 and 6, which have *Liked* and *Replied*, respectively, as the dependent variable. In effect, it appears as if firms are more likely to retweet messages sent by organizations, but when it comes to deciding whether to like or reply to messages, firms make no distinction

Table 4.5: Logistic Regressions, Dependent Variables are Firm Reactions

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	(Any Reaction: Like, Retweet, or Reply)			(Liked)	(Retweeted)	(Replied)
Sender Characteristics						
Entity						
Organization	0.21**	0.20**	0.12	0.06	0.51***	-0.04
	(0.09)	(0.08)	(0.08)	(0.09)	(0.12)	(0.16)
Twitter Sophistication						
# of Days on Twitter			0.0002***			
			(0.00004)			
Sophistication Index	0.08***	0.08***		0.08**	0.07*	0.05*
	(0.02)	(0.02)		(0.03)	(0.04)	(0.03)
Prestige/Celebrity						
'Verified' Twitter account	0.50**	0.50**		0.04	0.97***	-0.19
	(0.22)	(0.22)		(0.18)	(0.25)	(0.24)
# of Followers (1,000s)			0.0005**			
			(0.0002)			
Language						
English	0.60***	0.57**	0.68***	0.84***	0.23	0.62*
	(0.23)	(0.22)	(0.22)	(0.23)	(0.38)	(0.34)
Message Characteristics						
Timing						
Weekday Business Hours	0.25***	0.25***	0.27***	0.09	0.29***	0.36**
	(0.08)	(0.08)	(0.08)	(0.11)	(0.09)	(0.15)
Location						
Located in USA	0.26***	0.25***	0.30***	0.25**	0.32***	-0.04
	(0.09)	(0.09)	(0.08)	(0.11)	(0.11)	(0.11)
Originality & Thread Location						
Original Tweet	3.92***					
	(0.23)					
'Mention Only' Tweet		3.99***	4.05***	3.62***	4.95***	3.66***
		(0.24)	(0.28)	(0.26)	(0.18)	(0.43)
Public Direct Reply		3.68***	3.75***	3.53***	3.20***	4.09***
		(0.21)	(0.22)	(0.25)	(0.20)	(0.35)
Public Direct Message		3.92***	3.97***	3.34***	4.71***	4.41***
		(0.25)	(0.26)	(0.31)	(0.20)	(0.42)
Tweet Entities						
# of Entities	-0.18	-0.22**	-0.18**	-0.09	0.09	-0.77***
	(0.11)	(0.09)	(0.08)	(0.10)	(0.11)	(0.13)
Sentiment						
Positive Sentiment	0.72***	0.74***	0.76***	0.66***	0.81***	0.45***
	(0.17)	(0.17)	(0.16)	(0.11)	(0.25)	(0.16)
Firm Characteristics						
Firm Size						
Size index	-0.41***	-0.42***	-0.43***	-0.44***	-0.56***	-0.18*
	(0.13)	(0.14)	(0.13)	(0.15)	(0.16)	(0.11)
Twitter Prestige						
Twitter Prestige Index	-0.19	-0.19	-0.14	-0.02	-0.18	-0.64***
	(0.12)	(0.12)	(0.12)	(0.14)	(0.17)	(0.16)
Twitter Activity						
# of Tweets Sent	0.00008*	0.00008*	0.00007	0.00006	0.00003	0.0002***
	(0.00005)	(0.00005)	(0.00005)	(0.00006)	(0.00007)	(0.00005)
Industry						
Industry Fixed Effects	yes	yes	yes	yes	yes	yes
Intercept	-9.32***	-9.22***	-9.27***	-10.67***	-10.68***	-10.49***
	(0.83)	(0.80)	(0.78)	(0.89)	(0.79)	(0.85)
χ^2	10,442.0***	10,482.0***	15,589.0***	6,298.7***	4,150.4***	2,922.7***
Pseudo R^2	0.32	0.32	0.33	0.32	0.29	0.30
Log-Likelihood:	-11,097.0	-11,077.0	-16,103.0	-6,844.0	-5,082.4	-3,456.9
N	104,616	104,616	163,374	104,616	104,616	104,616

*p<.10, **p<.05, ***p<.01. Coefficients are in log-odds format, with standard errors clustered on CSR Twitter account in parentheses between organizations and individuals.

H2 posited firms would be more likely to react to messages sent by users with more sophisticated Twitter accounts. This hypothesis is supported by the results in all six models, which all see significant, positive coefficients for the sophistication measure. One potential limitation in Models 1-2 and 4-6 is that the number of cases (n = 104,616) is much lower than the total possible number of observations (n = 163,402), which is due to the sophistication index as well as the prestige index. To check the robustness of the model to alternative specification of these variables and to a greater number of observations (n = 163,374), Model 3 includes alternative measures of sophistication and prestige. The alternative measure of sophistication is *Number of Days on Twitter*; as with the sophistication index in the other five models, the coefficient for *Number of Days on Twitter* is significant and positive ($p<.01$). The evidence suggests firms are more likely to reply to, like, and retweet messages sent by members of the public with more sophisticated Twitter accounts.

Relatedly, H3 presented the idea that more prestigious and influential Twitter users would be more likely to receive a firm reaction. In Models 1-2 and 4-6 this concept was measured by a dichotomous variable indicating messages sent from users with "verified" accounts. The coefficient on *Verified* is significant and positive in Models 1, 2 and 5 yet fails to obtain significance in Models 4 (with *Liked* as dependent variable) and 6 (with *Replied* as dependent variable). As alluded to above, in Model 3 *Verified* was replaced with the variable *Number of Followers*, which has a greater number of available observations. *Number of Followers* has a significant, positive coefficient in the model ($p<.05$). Overall, the results support the idea that sender prestige is related to the likelihood of a message being retweeted by the firm, but not related to the probability of receiving a like or a reply.

The final message sender characteristic hypothesized concerned language, with H4 positing firms will be more likely to react to messages from senders corresponding in English. This hypothesis is largely supported, with positive, significant coefficients for the variable *English* in all models save for Model 5.

The next five hypotheses made arguments concerning the effects of characteristics of the messages that are sent. To start, H5 argued for a positive relationship between firm reactions and messages being sent during regular business hours. The findings largely support this hypothesis, with the variable *Weekday Business Hours* receiving a positive and significant coefficient in all regressions ($p<.05$) except for Model 4, which has *Liked* as the dependent variable.

H6, meanwhile, posited tweets sent from users that could be identified as being in the United States would be more likely to receive a firm reaction. *Located in USA* obtains a significant, positive coefficient in all models ($p<.05$) save for Model 6. The hypothesis is thus largely supported, with the caveat that being located in the USA does not appear to be related to the likelihood of the message receiving a reply.

I then hypothesized (H7) firms will be more likely to react to original messages rather than retweets. The significant, positive coefficient on *Original Tweet* in Model 1 supports H7. For the remaining models I replace *Original Tweet* with three separate types of original messages that vary according to whether the tweet initiates a new conversation thread and whether the message is directed at a specific Twitter user (as indicated by the tweet beginning with *@USER*) or is undirected. *Mention Only Tweets* are undirected messages, *Public Direct Messages* are tweets that initiate a new discussion thread and are directed at the *Fortune* CSR account, and *Public Direct Replies* are directed messages that are replies to an existing tweet. The omitted (baseline) category for the regressions are the 105,216 non-original (retweeted) messages. All three variables obtain significant and positive coefficients ($p< .01$) in all five remaining models. The results suggest firms are more likely to react to original messages regardless of whether the message is a reply or initiates a new thread and regardless of whether the message is directed (targeted specifically at the firm) or is undirected (only mentioning or discussing the firm).

H8 posited a negative relationship between the number of tweet entities included in a message – hashtags, user mentions, URLs, and photos – and the likelihood of the message receiving a reaction. This hypothesis is only partially supported, obtaining a negative, significant coefficient ($p<.05$) only in Models 2, 3, and 6. Including entities appears to reduce the likelihood of receiving a reply but has no relationship with the likelihood of receiving a like or a retweet.

The final message feature hypothesis (H9) argued firms will be more likely to react to messages with positive sentiment. This hypothesis is strongly supported, obtaining a significant positive coefficient in all six models.

The final three hypotheses tap the relationship between characteristics of the firm and the likelihood the firm will react to messages sent from members of the public. H10 posited a negative relationship between firm size and firm reactions. With a significantly negative coefficient in all six models, H10 is supported.

Results for the *Twitter Prestige Index*, in contrast, are largely insignificant. The variable obtains a significant negative coefficient only in Model 6. H11, which posited a negative relationship between the prestige of the firm's Twitter account and the likelihood it will react to public messages, thus only receives support with respect to the likelihood of a reply.

The final variable, *Number of Tweets Sent*, similarly appears to be chiefly associated with the likelihood of replying to public messages. *Number of Tweets Sent* is significantly, positively associated with reactions in Models 1, 2, and 6 and insignificant in the other three models. In brief, H12, which posited firm reactions would be positively associated with the level of Twitter activity, receives partial support.

4.6 Discussion and Conclusions

4.6.1 Summary and Discussion of Empirical Findings

The findings suggest firms are more likely to react to messages that are positive, original, contain fewer tweet entities, are sent from the United States during regular business hours, and are sent by organizations, those writing in English, and those with more sophisticated and prestigious Twitter accounts. Moreover, smaller firms, as well as those that are more active on Twitter, are more likely to react to public messages.

The findings also suggest some differences in the drivers of the three main reactions – likes, retweets, and replies. For instance, whether the message comes from an organization and the prestige of the message sender only appear to be relevant to whether the message receives a retweet. In contrast, language is only relevant for likes and replies but not retweets. Meanwhile, tweet entities, the prestige of the firm's Twitter account, and the level of firm Twitter activity only appear to be associated with replies. Timing, in turn, does not appear to be relevant to likes, and location does not seem to be associated with the probability of receiving a reply.

4.6.2 Theoretical Implications and Future Research

The Meaning of the Different Reactions for Accounting Research

In effect, I found that what drives the three main reactions differs according to whether it is a like, a retweet, or a reply. This begs the question of what these three different reactions mean for the accounting and CSR literatures. At the broadest level, the three reactions can be seen as three distinct reflections of firms' managerial attention (Mitchell et al., 1997) to public stakeholders.

As argued by Saxton (2016a), liking, first of all, sends a signal to the message sender that the firm views the message as valuable or informative. It is, in some ways, analogous to a "buy" signal in the capital markets (e.g., Drake et al., 2011); with Twitter representing a market of information and communication (Saxton, 2012), liking can thus be seen as a *buy signal* for an accountability, CSR-focused communication market. More specifically, it is telling its audience members that it values the particular type of communication that has been sent. As a result, liking plausibly plays a strong role in determining the ultimate make-up of the communication network the firm fosters on Twitter – both in terms of type of content and type of users.

A retweet, in turn, represents a type of buy signal that differs from a like in several respects. One, it represents a somewhat stronger signal; with a retweet causing the message to be sent out to all of the firm's followers, the threshold of value needs to be higher in order that the firm not be seen as "polluting" its Twitter feed with irrelevant messages. A retweet signal is also more likely to be employed when the public message conveys *information* as opposed to the other forms of communication tactics (e.g., dialogue) that Saxton (2016b) found to be common in firms' CSR tweets. At the same, retweeting causes a change in role for the company from discloser of information to *information intermediary*. One of Twitter's key market roles, in fact, is as a dissemination network (Blankespoor et al., 2014). Yet this network is one in which the traditional set of information intermediaries in the capital markets – the array of sell-side analysts, mainstream media, auditors, and financial institutions – is disrupted and democratized.

A reply, finally, represents something different altogether. Of the three reactions, it represents the highest level of effort. It also sends a signal to the message sender, but is closer to two other analogues in the accounting literature – reporting and accountability – which I cover next.

Micro-Reporting

The replies a firm sends effectively constitute a micro form of reporting or disclosure. The reporting is orders of magnitude smaller than the typical stand-alone report (e.g., Dhaliwal et al., 2011) used to convey CSR performance. Yet it is also orders of magnitude more frequent, being conducted on a daily rather than annual basis. Furthermore, given that it is a reply to an existing message, the "micro-report" is often directed at a specific, publicly identifiable user; this distinguishes these micro-reports from the typical annual report, which is disclosed to an amorphous public rather than communicated to specific stakeholders. Social media-based micro-reporting effectively offers a new reporting and disclosure paradigm that presents challenges as well as opportunities for the accounting literature along a number of dimensions.

Micro-Accountability

The replies firms send on Twitter also represent a form of micro-accountability. In fact, evidence from Saxton (2016a) and Saxton (2016b), looking at public and firm messages, respectively, suggests the micro-messages are not merely disclosing – that is, sending information – but just as often *communicating* with their audiences. What occurs is often two-way CSR conversations rather than one-way CSR reporting.

In effect, not only do social media platforms offer tools firms can use to report on their activities, but they offer a venue in which the public can react and respond to firms' CSR reporting efforts. In a way, social media represent the new "town hall," yet a town hall in which public commenting and firm reporting is done on a daily basis rather than once a year. Furthermore, the account-giving is often made to individual stakeholders rather than groups or to a generalized "public." Moreover, given the importance of the formal network connections among users, what we see on social media is *connected* accounting, reporting, and accountability behaviors. Given how the messages flow from and are received by specific, publicly identifiable users, we are able to see to whom firms are engaged in reporting and "account-giving" behaviors. It is these qualities that have enabled me here to address the question of *to whom* organizations are accountable, an issue that has long been of interest to the accountability literature (Gray et al., 1987; Tower, 1993).

Firm Reactions and the Conduct of CSR

Not only do the firms' reactions constitute companies' efforts to report on and give accounts of their CSR activities, but the firms' reactions to public messages also *constitute* CSR. That is, if one of the goals of a given firm's CSR efforts is to produce better societal-level environmental outcomes, or to produce a stronger civil society, or to expand educational opportunities for underserved communities, then the firm's Twitter-based efforts to educate the public on these issues, to build an active and concerned issue network, and to mobilize this network to action could be seen as integral to the attainment of the firm's CSR objectives. The current study supports arguments that, in some important respects, CSR *is* inherently communicative (Schultz et al., 2013). While an in-depth examination of this issue was beyond the scope of the present study, this idea deserves further investigation in the CSR literature.

4.6.3 Future Research

This paper has aimed to present a novel approach to understanding the types of CSR activities firms are conducting on social media platforms. The next step would be to further develop some of the issues that have been left under-examined. For instance, future research could extend the analyses conducted here and undertake more detailed examinations of the type of stakeholders to which firms provide micro-accounts as well as the types of content that are worthy of firm responses. Furthermore, this study has looked at whether the firm replies to public messages but not the nature of those replies. Future research could meaningfully build on some of the above ideas and code those replies for whether they are disclosing information, engaging in "account-giving" activities, or instead related to broader community-building or public education-related communicative efforts.

Another aspect deserving of inquiry is the outcomes of firms' micro-reporting and micro-accountability efforts. Saxton (2016b) framed audience reactions to firm messages in terms of corporate reputation. It remains to be seen whether firm reactions to public messages are similarly related to legitimacy (Colleoni, 2013) or reputation (Cho et al., 2012) or are instead more closely aligned with perceptions of accountability (Gray et al., 1987) or responsiveness (Saxton et al., 2007) or CSR performance (Mahoney et al., 2013).

4.6.4 Conclusions

This paper has provided evidence of the message, sender, and firm characteristics that drive companies to engage in *micro-reporting* and *micro-accountability* behaviors in the CSR domain. In effect, on social media members of the public are continually "calling firms out" for their CSR actions, and firms' decisions to respond or not respond constitute a new form of public, dynamic, interactive reporting and accountability behavior that has yet to be addressed by the extant accounting literature. This paper has sought to help propel this nascent body of research.

In the end, the firms' CSR-focused Twitter accounts provide a venue in which firms – through their replies, retweets, and liking actions – can give accounts of their behavior, send micro-reports, and engage in dialogue with concerned stakeholders, while at the same time actually *conducting* CSR by helping provide fora for public education about CSR issues. The tweet represents a new accounting artifact.

Chapter 5

Conclusions

I began this thesis with examples of different types of firm and public messages that suggested how, on social media, the nature of communication – and therefore the practice of CSR – is different. Through the course of three separate studies, I have sought to make several contributions to the accounting and CSR literatures.

First, the thesis has provided evidence of a number of new, non-reporting-based CSR communication tactics encompassing not only the one-way disclosure of information but also public educational messages, two-way dialogue, and messages meant to mobilize the public through calls to action. The key insight gained is that the tactics go well beyond disclosure; in essence, firms often move beyond CSR reporting towards more nuanced CSR communication. The findings thus strongly suggest accounting and CSR scholars move beyond a one-way reporting model of CSR efforts.

Second, this thesis has built and tested theoretical explanations of the reputational outcomes of engaging in various CSR communication tactics. In effect, both communication and reputation are multi-dimensional. Just as there are multiple forms of communication, there are two key dimensions of reputation – awareness and favorability (Rindova et al., 2005). The newly identified non-informational, non-reporting forms of CSR communication seem to play a particularly strong role in strengthening positive public perceptions of the firm.

Moreover, in illustrating how CSR-based reputational capital is accumulated on a micro-, message-by-message, day-to-day level, the study responds to calls for greater understanding of stakeholder reactions to CSR disclosures (Moser & Martin, 2012) and of the micro-foundations of CSR efforts (Aguinis & Glavas, 2012). The study opens new avenues of research by encouraging a longer-term view of CSR efforts, by expanding the number of activities that can be examined in studying CSR, and by delving into the day-to-day flows of CSR communication and reputational capital.

Third, this thesis has provided the first study of which I am aware of firms' reactions to public CSR-focused messages. I argued, at the broadest level, that firms' reactions should be seen as reflections of firms' managerial attention (Mitchell et al., 1997) to public stakeholders. I have also theorized more specifically about how the various reactions might be conceptualized and more fully incorporated into the accounting literature. Notably, I have posited that the replies a firm sends to public messages can constitute a micro form of reporting or disclosure, and that social media-based micro-reporting offers a new reporting and disclosure paradigm that presents challenges as well as opportunities for the accounting literature along a number of dimensions.

I also argue that, in other circumstances, the replies a firm sends may constitute a form of micro-accountability. In effect, not only does social media constitute a ready-made reporting tool, but it offers a venue in which members of the public can react and respond to firms' CSR reporting efforts. In a way, social media represent the new *Agora*, a public forum in which commenting and firm reporting is done on a daily basis rather than once a year.

At the same time, the account-giving is often made to individual stakeholders rather than to groups or to a generalized public. Moreover, given the importance of the formal network connections among users, with each micro report sent or micro account given the firms are making a message and/or network connection to these individual stakeholders. It is in this sense that what we see on social media is relational or *connected* accounting, reporting, and accountability behaviors. And given how the messages flow from and are received by specific, publicly identifiable users, we are able to see to whom firms are engaged in reporting and "account-giving" behaviors, an issue that has long been of interest to the accountability literature (Gray et al., 1987; Tower, 1993).

In finding a series of relationships between firm reactions and various characteristics of the messages sent, the users who sent them, and the firms targeted, the study has contributed to the existing literature on to whom and for what firms are accountable. Moreover, this thesis has provided evidence of the message, sender, and firm characteristics that drive companies to engage in *micro-reporting* and *micro-accountability* behaviors in the CSR domain. Ultimately, firms' decisions to respond or not respond constitute a new form of public, dynamic, interactive reporting and accountability behavior that had yet to be addressed by the extant accounting literature. The current study has sought to contribute to and help propel this nascent body of research.

Fourth, throughout the dissertation, I have sought to incorporate empirical, conceptual, and methodological insights from other disciplines, especially those that are better situated to analyze Big Data such as that offered by social media. All three studies employed various machine learning techniques in order to either develop measures or identify variables of interest. Study #2, notably, employed machine learning to help build inductive insights into the nature of public messages and the determinants of firms' reactions to comments and queries from members of the public. There are ample opportunities for research that builds on the current study and applies these techniques to studying other micro-level phenomena. Indeed, in many ways, a Big Data-driven (Vasarhelyi et al., 2015) communication perspective provides scholars with a valuable new set of tools for analyzing and problematizing CSR.

Similarly, throughout this thesis I have incorporated conceptual and operational insights from a broad range of non-accounting disciplines – including computer science, communication, marketing, public relations, sociology, and political science – in order understand the nature and dynamics of CSR communication. In so doing, many of the variables identified through the inductive analyses and/or tested in the empirical examinations are new to the CSR literature and, even more so, to the accounting literature. In one sense this is due to the nature of social media-based Big Data, which allow for the quantitative investigation of new variables such as time stamps and geo-location. In another sense, these variables are new to the literature because the phenomenon itself has been changed by social media. Simply put, the conduct of "CSR" is different on social media than it is on prior platforms.

In the end, the firms' CSR-focused Twitter accounts provide a setting in which firms – through their replies, retweets, and liking actions – can give accounts of their behavior, send micro-reports, and engage in dialogue with concerned stakeholders, while at the same time building reputational capital and actually *conducting* CSR by helping provide fora for public education about CSR issues. The communicative CSR interactions that occur on social media represent a new, more dynamic and interactive and public form of accountability reporting that carries implications for accounting research beyond the domain of CSR. The current study contributes to the literature by providing insights into the nature of this new reputation-building, "account-demanding," and reporting venue.

By the same token, a central insight from this study is that CSR communication is not something done annually; rather, CSR communication is something delivered on a continual basis through the day-to-day informational, interactive, and tie-building efforts made by companies. Social media ultimately offer a more public, relational, interactive, and dynamic information and accountability and reporting environment. It will continue to "disrupt" and produce new sets of winners and losers not only for the companies studied but also the accounting firm, the report, and the accountant – not to mention the public at large.

It is in this way that the thesis highlights the ways in which social media are *dynamizing* accounting – engendering a reporting and reacting system that is substantially less static and temporally delimited than previously seen. In sum, this thesis posits and empirically supports the proposition that the conceptual and methodological possibilities deriving from a more dynamic accounting information environment are not inconsequential.

References

Aguinis, H. & Glavas, A. (2012). What we know and don't know about corporate social responsibility: A review and research agenda. *Journal of Management*, *38*, 932–968.

Ahrens, T. (1996). Styles of accountability. *Accounting, Organizations and Society*, *21*, 139–173.

Anderson, C. (2006). *The long tail: Why the future of business is selling less of more*. New York, NY: Hachette Digital, Inc.

Andrew, J. & Cortese, C. (2011). Accounting for climate change and the self-regulation of carbon disclosures. *Accounting Forum*, *35*, 130–138.

Arun, R., Suresh, V., Veni Madhavan, C. E., & Narasimha Murthy, M. N. (2010). On finding the natural number of topics with Latent Dirichlet Allocation: Some observations. In Zaki, M. J., Yu, J. X., Ravindran, B., & Pudi, V., eds., *Advances in Knowledge Discovery and Data Mining: 14th Pacific-Asia Conference, PAKDD 2010, Hyderabad, India, June 21-24, 2010. Proceedings. Part I*, 391–402. Berlin, Heidelberg: Springer Berlin Heidelberg.

Arya, B. & Zhang, G. (2009). Institutional reforms and investor reactions to CSR announcements: Evidence from an emerging economy. *Journal of Management Studies*, *46*, 1089–1112.

Atiase, R. K., Bamber, L. S., & Tse, S. (1989). Timeliness of financial reporting, the firm size effect, and stock price reactions to annual earnings announcements. *Contemporary Accounting Research*, *5*, 526–552.

Bach, D. & Allen, D. (2010). What every CEO needs to know about nonmarket strategy.

MIT Sloan Management Review, *51*, 41–48.

Bakshy, E., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Everyone's an influencer: Quantifying influence on Twitter. In *Proceedings of the fourth ACM International Conference on Web Search and Data Mining*, 65–74. New York, NY, 65-74.

Balakrishnan, R., Sprinkle, G. B., & Williamson, M. G. (2011). Contracting benefits of corporate giving: An experimental investigation. *The Accounting Review*, *86*, 1887–1907.

Barnett, M. L., Jermier, J. M., & Lafferty, B. A. (2006). Corporate reputation: The definitional landscape. *Corporate Reputation Review*, *9*, 26–38.

Bebbington, J., Larrinaga, C., & Moneva, J. M. (2008). Corporate social reporting and reputation risk management. *Accounting, Auditing & Accountability Journal*, *21*, 337–361.

Benford, R. D. & Snow, D. A. (2000). Framing processes and social movements: An overview and assessment. *Annual Review of Sociology*, *26*, 611–639.

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

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*, 77–84.

Bollen, J., Mao, H., & Pepe, A. (2011). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *ICWSM*.

Bollen, L., Hassink, H., & Bozic, G. (2006). Measuring and explaining the quality of Internet investor relations activities: A multinational empirical analysis. *International Journal of Accounting Information Systems*, *7*, 273–298.

Brown, T. J. & Dacin, P. A. (1997). The company and the product: Corporate associations and consumer product responses. *Journal of Marketing*, *61*, 68–84.

Bruns, A. & Burgess, J. E. (2011). The use of Twitter hashtags in the formation of ad hoc publics. *European Consortium for Political Research, Reykjavik, 25-27 Aug. 2011.*

Budak, C. & Agrawal, R. (2013). On participation in group chats on Twitter. In *Proceedings of the 22nd International World Wide Web Conference Committee (IW3C2)*, 165–176.

International World Wide Web Conferences Steering Committee.

Bushee, B. J., Matsumoto, D. A., & Miller, G. S. (2003). Open versus closed conference calls: The determinants and effects of broadening access to disclosure. *Journal of Accounting & Economics*, *34*, 149–180.

Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. *Neurocomputing*, *72*, 1775–1781.

Castelló, I., Etter, M., & Årup Nielsen, F. (2015). Strategies of legitimacy through social media: The networked strategy. *Journal of Management Studies*, *53*, 402–432.

Castelló, I., Morsing, M., & Schultz, F. (2013). Communicative dynamics and the polyphony of corporate social responsibility in the network society. *Journal of Business Ethics*, *118*, 683–694.

Cho, C. H., Guidry, R. P., Hageman, A. M., & Patten, D. M. (2012). Do actions speak louder than words? An empirical investigation of corporate environmental reputation. *Accounting, Organizations and Society*, *37*, 14–25.

Cho, C. H., Laine, M., Roberts, R. W., & Rodrigue, M. (2015). Organized hypocrisy, organizational façades, and sustainability reporting. *Accounting, Organizations and Society*, *40*, 78–94.

Cho, C. H. & Roberts, R. W. (2010). Environmental reporting on the internet by America's Toxic 100: Legitimacy and self-presentation. *International Journal of Accounting Information Systems*, *11*, 1–16.

Cho, C. H., Roberts, R. W., & Patten, D. M. (2010). The language of US corporate environmental disclosure. *Accounting, Organizations and Society*, *35*, 431–443.

Clark, W. R. & Golder, M. (2015). Big Data, causal inference, and formal theory: Contradictory trends in political science? *PS: Political Science & Politics*, *48*, 65–70.

Cochran, P. L. & Wood, R. A. (1984). Corporate social responsibility and financial performance. *Academy of Management Journal*, *27*, 42–56.

Cohen, J. R., Holder-Webb, L., & Zamora, V. L. (2015). Nonfinancial information preferences of professional investors. *Behavioral Research in Accounting*, *27*, 127–153.

Colleoni, E. (2013). CSR communication strategies for organizational legitimacy in social media. *Corporate Communications: An International Journal*, *18*, 228–248.

Creyer, E. H. (1997). The influence of firm behavior on purchase intention: Do consumers really care about business ethics? *Journal of Consumer Marketing*, *14*, 421–432.

Dawkins, J. & Lewis, S. (2003). CSR in stakeholder expectations: And their implication for company strategy. *Journal of Business Ethics*, *44*, 185–193.

de Bakker, F. G. A. & Hellsten, I. (2013). Capturing online presence: Hyperlinks and semantic networks in activist group websites on corporate social responsibility. *Journal of Business Ethics*, *118*, 807–823.

de Vries, L., Gensler, S., & Leeflang, P. S. H. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of Interactive Marketing*, *26*, 83–91.

Debreceny, R. S. (2001). The production and use of semantically rich accounting reports on the Internet: XML and XBRL. *International Journal of Accounting Information Systems*, *2*, 47–74.

Debreceny, R. S. (2015). Social media, social networks, and accounting. *Journal of Information Systems*, *29*, 1–4.

den Hond, F., Rehbein, K. A., de Bakker, F. G. A., & Lankveld, H. K.-v. (2014). Playing on two chessboards: Reputation effects between corporate social responsibility (CSR) and corporate political activity (CPA). *Journal of Management Studies*, *51*, 790–813.

Deveaud, R., SanJuan, E., & Bellot, P. (2014). Accurate and effective latent concept modeling for ad hoc information retrieval. *Document numérique*, *17*, 61–84.

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review*, *86*, 59–100.

Drake, M. S., Rees, L., & Swanson, E. P. (2011). Should investors follow the prophets or the bears? Evidence on the use of public information by analysts and short sellers. *The Accounting Review*, *86*, 101–130.

Du, S., Bhattacharya, C. B., & Sen, S. (2010). Maximizing business returns to corporate social responsibility (CSR): The role of CSR communication. *International Journal of Management Reviews*, *12*, 8–19.

Du, S. & Vieira, E. (2012). Striving for legitimacy through corporate social responsibility: Insights from oil companies. *Journal of Business Ethics*, *110*, 413–427.

Eberle, D., Berens, G., & Li, T. (2013). The impact of interactive corporate social responsibility communication on corporate reputation. *Journal of Business Ethics*, *118*, 731–746.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, *14*, 532–550.

Evans, M. S. (2014). A computational approach to qualitative analysis in large textual datasets. *PloS one*, *9*, e87908.

Fieseler, C. & Fleck, M. (2013). The pursuit of empowerment through social media: Structural social capital dynamics in CSR-blogging. *Journal of Business Ethics*, *118*, 759–775.

Fieseler, C., Fleck, M., & Meckel, M. (2010). Corporate social responsibility in the blogosphere. *Journal of Business Ethics*, *91*, 599–614.

Fischer, E. & Reuber, A. R. (2014). Online entrepreneurial communication: Mitigating uncertainty and increasing differentiation via Twitter. *Journal of Business Venturing*, *29*, 565–583.

Fraustino, J. & Connolly-Ahern, C. (2015). Corporate associations written on the wall: Publics' responses to Fortune 500 ability and social responsibility Facebook posts. *Journal of Public Relations Research*, *27*, 452–474.

Gaffney, D. (2010). #iranElection: Quantifying online activism. In *Web Science Conference 2010, April 26-27, 2010, Raleigh, NC, USA.*

Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford*, 1–12.

Gómez-Vásquez, L. (2013). Me gusta o te sigo : Análisis de la comunicación de prácticas de Responsabilidad Social Corporativa a través de los medios sociales. *Correspondencias & Análisis*, *3*, 89–109.

González-Bailón, S. & Paltoglou, G. (2015). Signals of public opinion in online communication: A comparison of methods and data sources. *The ANNALS of the American Academy of Political and Social Science*, *659*, 95–107.

Gray, R., Owen, D., & Maunders, K. (1987). *Corporate social reporting: Accounting and accountability*. Englewood Hills, NJ: Prentice-Hall International.

Grenier, J. H., Lowe, D. J., Reffett, A., & Warne, R. C. (2015). The effects of independent expert recommendations on juror judgments of auditor negligence. *Auditing: A Journal of Practice & Theory*, *34*, 157–170.

Griffin, J. J. & Dunn, P. (2004). Corporate public affairs: Commitment, resources, and structure. *Business and Society*, *43*, 196–220.

Griffiths, T. L. & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, *101*, 5228–5235.

Grosser, K. & Moon, J. (2008). Developments in company reporting on workplace gender equality? *Accounting Forum*, *32*, 179–198.

Grunig, J. E. (1997). A situational theory of publics: Conceptual history, recent challenges and new research. In Moss, D., MacManus, T., & Vercic, D., eds., *Public relations research: An international perspective*, volume 3, 3–38. London, England: International Thomson Business Press.

Guo, C. & Saxton, G. D. (2014). Tweeting social change: How social media are changing nonprofit advocacy. *Nonprofit and Voluntary Sector Quarterly*, *43*, 57–79.

Guyon, I. & Elisseeff, A. (2003). An introduction to variable and feature selection. *The Journal of Machine Learning Research*, *3*, 1157–1182.

Herremans, I. M., Akathaporn, P., & McInnes, M. (1993). An investigation of corporate social responsibility reputation and economic performance. *Accounting, Organizations* *and Society*, *18*, 587–604.

Hillman, A. J. & Hitt, M. A. (1999). Corporate political strategy formulation: A model of approach, participation, and strategy decisions. *Academy of Management Review*, *24*, 825–842.

Hirshleifer, D. & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting & Economics*, *36*, 337–386.

Hodge, F. D., Kennedy, J. J., & Maines, L. A. (2004). Does search-facilitating technology improve the transparency of financial reporting? *The Accounting Review*, *79*, 687–703.

Hopkins, D. & King, G. (2010). A method of automated nonparametric content analysis for social science. *American Journal of Political Science*, *54*, 229–247.

Hopwood, A. G. (2009). Accounting and the environment. *Accounting, Organizations and Society*, *34*, 433–439.

Hughes, A. L. & Palen, L. (2009). Twitter adoption and use in mass convergence and emergency events. *International Journal of Emergency Management*, *6*, 248–260.

IIRC (2013). The International <IR> Framework. Technical report, IIRC, International Integrated Reporting Council, IIRC, International Integrated Reporting Council.

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology*, *60*, 2169–2188.

Kane, G. C., Alavi, M., Labianca, G. J., & Borgatti, S. (2014). What's different about social media networks? A framework and research agenda. *MIS Quarterly*, *38*, 274–304.

Kasznik, R. & Lev, B. (1995). To warn or not to warn: Management disclosures in the face of an earnings surprise. *The Accounting Review*, *70*, 113–134.

Kelly, K. S. (2000). Stewardship: The missing step in the public relations process. In Heath, R. L., ed., *Handbook of Public Relations*, 279–289. Thousand Oaks, CA: Sage.

Kent, M. L. & Taylor, M. (2016). From Homo Economicus to Homo Dialogicus: Rethinking social media use in CSR communication. *Public Relations Review*, *42*, 60–67.

King, B. G. (2007). A social movement perspective of stakeholder collective action and influence. *Business & Society*, *47*, 21–49.

King, B. G. & Whetten, D. A. (2008). Rethinking the relationship between reputation and legitimacy: A social actor conceptualization. *Corporate Reputation Review*, *11*, 192–207.

Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159–174.

Lange, D., Lee, P. M., & Dai, Y. (2011). Organizational reputation: A review. *Journal of Management*, *37*, 153–184.

Latour, B. (2005). *Reassembling the Social - An Introduction to Actor-Network-Theory*. New York, NY: Oxford University Press.

Lee, K., Oh, W.-Y., & Kim, N. (2013). Social media for socially responsible firms: Analysis of Fortune 500's Twitter profiles and their CSR/CSIR ratings. *Journal of Business Ethics*, *118*, 791–806.

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

Lev, B., Petrovits, C., & Radhakrishnan, S. (2010). Is doing good good for you? How corporate charitable contributions enhance revenue growth. *Strategic Management Journal*, *31*, 182–200.

Lovejoy, K. & Saxton, G. D. (2012). Information, community, and action: How nonprofit organizations use social media. *Journal of Computer-Mediated Communication*, *17*, 337–353.

Lovejoy, K., Waters, R. D., & Saxton, G. D. (2012). Engaging stakeholders through Twitter: How nonprofit organizations are getting more out of 140 characters or less. *Public Relations Review*, *38*, 313–318.

Lyon, T. P. & Montgomery, A. W. (2013). Tweetjacked: The impact of social media on corporate greenwash. *Journal of Business Ethics*, *118*, 747–757.

Mahoney, L. S., Thorne, L., Cecil, L., & LaGore, W. (2013). A research note on standalone corporate social responsibility reports: Signaling or greenwashing? *Critical Perspectives on Accounting*, *24*, 350–359.

Matten, D. & Crane, A. (2005). Corporate citizenship: Toward an extended theoretical conceptualization. *Academy of Management Review*, *30*, 166–179.

McCammon, H. J., Newman, H. D., Muse, C. S., & Terrell, T. M. (2007). Movement framing and discursive opportunity structures: The political successes of the U.S. women's jury movements. *American Sociological Review*, *72*, 725–749.

McDonnell, M.-H. & King, B. G. (2013). Keeping up appearances: Reputational threat and impression management after social movement boycotts. *Administrative Science Quarterly*, *59*, 387.

Meinshausen, N. & Bühlmann, P. (2010). Stability selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *72*, 417–473.

Miles, M. B. & Huberman, A. M. (1984). *Qualitative data analysis: A sourcebook of new methods*. Beverly Hills, CA: Sage.

Mishra, S. & Suar, D. (2010). Does corporate social responsibility influence firm performance of Indian companies? *Journal of Business Ethics*, *95*, 571–601.

Mitchell, R. K., Agle, B. R., & Wood, D. J. (1997). Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of Management Review*, *22*, 853–886.

Monge, P. R. & Contractor, N. S. (2003). *Theories of communication networks*. New York: Oxford University Press.

Morsing, M. & Schultz, M. (2006). Corporate social responsibility communication: Stakeholder information, response and involvement strategies. *Business Ethics: A European Review*, *15*, 323–338.

Moser, D. V. & Martin, P. R. (2012). A broader perspective on corporate social responsibility research in accounting. *The Accounting Review*, *87*, 797–806.

Mosley, J. E. (2009). Institutionalization, privatization, and political opportunity: What tactical choices reveal about the policy advocacy of human service nonprofits. *Nonprofit and Voluntary Sector Quarterly*, *40*, 435–457.

Neu, D. (2006). Accounting for public space. *Accounting, Organizations and Society*, *31*, 391–414.

Neu, D., Warsame, H., & Pedwell, K. (1998). Managing public impressions: Environmental disclosures in annual reports. *Accounting, Organizations and Society*, *23*, 265–282.

Nielsen, F. Å. (2011). A new ANEW: evaluation of a word list for sentiment analysis in microblogs. In Rowe, M., Stankovic, M., Dadzie, A.-S., & Hardey, M., eds., *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages. Volume 718 in CEUR Workshop Proceedings: 93-98. 2011 May.*

O'Brien, P. C. & Tan, H. (2015). Geographic proximity and analyst coverage decisions: Evidence from IPOs. *Journal of Accounting and Economics*, *59*, 41–59.

O'Leary, D. E. (2015). Crowdsourcing tags in accounting and finance: Review, analysis and emerging issues. *Journal of Emerging Technologies in Accounting*, *12*, 93–115.

Oliver, C. & Holzinger, I. (2008). The effectiveness of strategic political management: A dynamic capabilities framework. *Academy of Management Review*, *33*, 496–520.

Parks, M. R. (2014). Big Data in communication research: Its contents and discontents. *Journal of Communication*, *64*, 355–360.

Patten, D. M. (1991). Exposure, legitimacy, and social disclosure. *Journal of Accounting and Public Policy*, *10*, 297–308.

Patten, D. M. (1992). Intra-industry environmental disclosures in response to the Alaskan oil spill: A note on legitimacy theory. *Accounting, Organizations and Society*, *17*, 471–475.

Patten, D. M. (2002). The relation between environmental performance and environmental disclosure: A research note. *Accounting, Organizations and Society*, *27*, 763–773.

Rafaeli, S. & Sudweeks, F. (1997). Networked interactivity. *Journal of Computer-Mediated Communication*, *2*, 0. doi: 10.1111/j.1083–6101.1997.tb00201.x.edu/vol.

Rennekamp, K. (2012). Processing fluency and investors' reactions to disclosure readability.

Journal of Accounting Research, *50*, 1319–1354.

Rheingold, H. (2012). *Net smart: How to thrive online.* Cambridge, MA: MIT Press.

Richardson, A. J. & Welker, M. (2001). Social disclosure, financial disclosure and the cost of equity capital. *Accounting, Organizations and Society*, *26*, 597–616.

Rindova, V. P., Williamson, I. O., Petkova, A. P., & Sever, J. M. (2005). Being good or being known: An empirical examination of the dimensions, antecedents, and consequences of organizational reputation. *Academy of Management Journal*, *48*, 1033–1049.

Roberts, J. (1991). The possibilities of accountability. *Accounting, Organizations and Society*, *16*, 355–368.

Roberts, J. & Scapens, R. (1985). Accounting systems and systems of accountability–understanding accounting practices in their organisational contexts. *Accounting, Organizations and Society*, *10*, 443–456.

Roberts, P. W. & Dowling, G. R. (2002). Corporate reputation and sustained superior financial performance. *Strategic Management Journal*, *23*, 1077–1093.

Saeys, Y., Inza, I., & Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics*, *23*, 2507–2517.

Saxton, G. D. (2012). New media and external accounting information: A critical review. *Australian Accounting Review*, *22*, 286–302.

Saxton, G. D. (2016a). Calling firms out: Exploring the nature and determinants of dynamic, interactive micro-reporting in Fortune 200 firms' CSR-related Twitter accounts. Working paper.

Saxton, G. D. (2016b). CSR communication and the micro-accumulation of reputational capital. Working paper.

Saxton, G. D., Gomez, L. M., Ngoh, Z., Lin, C., & Dietrich, S. (2016). Do CSR messages resonate with the public? Examining stakeholder reactions to firms' CSR efforts on social media. Working paper.

Saxton, G. D. & Guo, C. (2011). Accountability online: Understanding the Web-based accountability practices of nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, *40*, 270–295.

Saxton, G. D. & Guo, C. (2015). Social Media Capital: Conceptualizing the nature, determinants, and outcomes of social media-based organizational resources. Working paper.

Saxton, G. D., Guo, C., & Brown, W. A. (2007). New dimensions of nonprofit responsiveness: The application and promise of Internet-based technologies. *Public Performance & Management Review*, *31*, 144–173.

Saxton, G. D., Kuo, J.-S., & Ho, Y.-C. (2012). The determinants of voluntary financial disclosure by nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, *41*, 1051–1071.

Saxton, G. D., Niyirora, J., Guo, C., & Waters, R. D. (2015). #AdvocatingForChange: The strategic use of hashtags in social media. *Advances in Social Work*, *16*, 154–169.

Saxton, G. D. & Waters, R. D. (2014). What do stakeholders 'like' on Facebook? Examining public reactions to nonprofit organizations' informational, promotional, and community-building messages. *Journal of Public Relations Research*, *26*, 280–299.

Schultz, F., Castelló, I., & Morsing, M. (2013). The construction of corporate social responsibility in network societies: A communication view. *Journal of Business Ethics*, *115*, 681–692.

Scott, S. V. & Orlikowski, W. J. (2012). Reconfiguring relations of accountability: Materialization of social media in the travel sector. *Accounting, Organizations and Society*, *37*, 26–40.

Sen, S. & Bhattacharya, C. B. (2001). Does doing good always lead to doing better? Consumer reactions to corporate social responsibility. *Journal of Marketing Research*, *38*, 225–243.

Shirky, C. (2003). Power laws, weblogs, and inequality. *Retrieved from http://www.shirky.com/writings/herecomeseverybody/powerlaw_weblog.html.*

Shirky, C. (2011). The political power of social media: Technology, the public sphere, and political change. *Foreign Affairs*, 28–41.

Slater, M. D., Chipman, H., Auld, G., Keefe, T., & Kendall, P. (1992). Information processing and situational theory: A cognitive response analysis. *Journal of Public Relations Research*, *4*, 189–203.

Smith, A. N., Fischer, E., & Yongjian, C. (2012). How does brand-related user-generated content differ across YouTube, Facebook, and Twitter? *Journal of Interactive Marketing*, *26*, 102–113.

Smith, B. G. (2012). Public relations identity and the stakeholder-organization relationship: A revised theoretical position for public relations scholarship. *Public Relations Review*, *38*, 838–845.

Statista (2015). Leading social networks worldwide as of March 2015, ranked by number of active users (in millions).

Stefanone, M. A., Saxton, G. D., Egnoto, M. J., Wei, W. X., & Fu, Y. (2015). Image attributes and diffusion via Twitter: The case of #guncontrol. In *Proceedings of the 48th Annual Hawaii International Conference on Systems Science, Kauai, HI, Jan. 5-8, 2015.*

Strauss, A. & Corbin, J. (1998). *Basics of qualitative research*. Thousands Oaks, CA: Sage, 2 edition.

Suddaby, R. (2010). Challenges for institutional theory. *Journal of Management Inquiry*, *19*, 14–20.

Suddaby, R., Saxton, G. D., & Gunz, S. (2015). Twittering change: The institutional work of domain change in accounting expertise. *Accounting, Organizations and Society*, *45*, 52–68.

Taylor, M. & Kent, M. L. (2014). Dialogic engagement: Clarifying foundational concepts. *Journal of Public Relations Research*.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*, 267–288.

Tower, G. (1993). A public accountability model of accounting regulation. *British Accounting Review*, *25*, 61–85.

Trabelsi, S., Labelle, R., & Laurin, C. (2004). The management of financial disclosure on corporate websites: A conceptual model. *Canadian Accounting Perspectives*, *3*, 235–259.

Tsang, E. W. K. (1998). A longitudinal study of corporate social reporting in Singapore: The case of the banking, food and beverages and hotel industries. *Accounting, Auditing & Accountability Journal*, *11*, 624–635.

Unerman, J. & Bennett, M. (2004). Increased stakeholder dialogue and the Internet: Towards greater corporate accountability or reinforcing capitalist hegemony? *Accounting, Organizations and Society*, *29*, 685–707.

Vasarhelyi, M. A., Kogan, A., & Tuttle, B. M. (2015). Big Data in accounting: An overview. *Accounting Horizons*, *29*, 381–396.

Vasi, I. B., Walker, E. T., Johnson, J. S., & Tan, H. F. (2015). "No fracking way!" Documentary film, discursive opportunity, and local opposition against hydraulic fracturing in the United States, 2010 to 2013. *American Sociological Review*, *80*, 934–959.

Verikas, A., Gelzinis, A., & Bacauskiene, M. (2011). Mining data with random forests: A survey and results of new tests. *Pattern Recognition*, *44*, 330–349.

Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting & Economics*, *5*, 179–194.

Waddock, S. A. & Graves, S. B. (1997). The corporate social performance-financial performance link. *Strategic Management Journal*, *18*, 303–319.

Walker, E. T. (2009). Privatizing participation: Civic change and the organizational dynamics of grassroots lobbying firms. *American Sociological Review*, *74*, 83–105.

Walker, E. T. & Rea, C. M. (2014). The political mobilization of firms and industries. *Annual Review of Sociology*, *40*, 281–304.

Warsame, H., Simmons, C. V., & Neu, D. (2002). Responding to "discrediting" events: Annual report disclosure responses to environmental fines. *Accounting and the Public Interest*, *2*, 22–40.

Waters, R. D. (2011). Redefining stewardship: Examining how Fortune 100 organizations use stewardship with virtual stakeholders. *Public Relations Review*, *37*, 129–136.

Waters, R. D. & Jamal, J. Y. (2011). Tweet, tweet, tweet: A content analysis of nonprofit organizations' Twitter updates. *Public Relations Review*, *37*, 321–324.

Werner, T. (2015). Gaining access by doing good: The effect of sociopolitical reputation on firm participation in public policy making. *Management Science*, *61*, 1989–2011.

Whelan, G., Moon, J., & Grant, B. (2013). Corporations and citizenship arenas in the age of social media. *Journal of Business Ethics*, *118*, 777–790.

Yu, L. & Liu, H. (2004). Efficient feature selection via analysis of relevance and redundancy. *The Journal of Machine Learning Research*, *5*, 1205–1224.

Zahller, K. A., Arnold, V., & Roberts, R. W. (2015). Using CSR disclosure quality to develop social resilience to exogenous shocks: A test of investor perceptions. *Behavioral Research in Accounting*, *27*, 155–177.