Association for Information Systems
AIS Electronic Library (AISeL)

ICIS 2019 Proceedings | General Topics

Retail Firms' Use of Social Media – Insights from Analysis of Large-Scale Twitter Data

Mikhail Lysyakov
University of Maryland, mlysyakov@rhsmith.umd.edu

Kunpeng Zhang
University of Maryland, kzhang@rhsmith.umd.edu

Siva Viswanathan
University of Maryland College Park, sviswana@rhsmith.umd.edu

Lysyakov, Mikhail; Zhang, Kunpeng; and Viswanathan, Siva, "Retail Firms' Use of Social Media – Insights from Analysis of Large-Scale Twitter Data" (2019). *ICIS 2019 Proceedings*. 20.
https://aisel.aisnet.org/icis2019/general_topics/general_topics/20

Retail Firms' Use of Social Media – Insights from Analysis of Large-Scale Twitter Data

Completed Research Paper

Mikhail Lysyakov
Robert H. Smith School of Business
University of Maryland
College Park, MD, USA
mlysyakov@rhsmith.umd.edu

Kunpeng Zhang
Robert H. Smith School of Business
University of Maryland
College Park, MD, USA
kzhang@rhsmith.umd.edu

Siva Viswanathan
Robert H. Smith School of Business
University of Maryland
College Park, MD, USA
sviswana@rhsmith.umd.edu

Abstract

While social media platforms have been used by retailers for a variety of purposes, there is limited research on how traditional retailers compete on social media platforms and what the effects of such competition are on related outcomes. Our paper seeks to fill this gap by examining whether retailers that are close competitors in the traditional context adopt similar content strategies on Twitter. We find that dissimilar firms have higher online engagement and acquire new followers faster. In examining the underlying mechanism, we find that this is attributable to their ability to leverage higher-level affordances of Twitter (i.e., relationship formation, meta-voicing, interactivity, collaboration, and competition). We find that it is the use of these higher-level affordances that leads to greater online engagement compared to the use of lower-level affordances, such as self-presentation and communication. Our findings have important implications for firms' competitive strategies in online social media platforms.

Keywords: Social Media, Twitter, Firm Competition, Content Strategies, Social Media Affordances, Deep Learning

Introduction

Social media platforms (SMPs) such as Facebook, Twitter, and Instagram, that primarily focus on connecting users and facilitating sharing of content among users have increasingly become important channels for firms as well. Most businesses today have a significant presence on many SMPs (Figure 1).

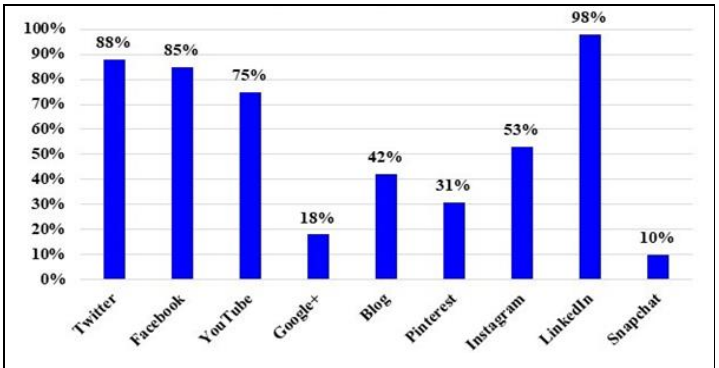

Figure 1. Fortune 500 corporate social media usage in 2017

Note. Source: umassd.edu 2017

While these social media platforms have been largely used as information dissemination channels by firms, over the last few years firms have found new and varied uses of these platforms. For example, firms use SMPs for customer service (mashable.com 2018) or use SMPs to target influential social media users and bloggers (lonelybrand.com 2017). Firms also use SMPs to build online communities around brands (marketingland.com 2016). More importantly, these social media platforms have evolved into new channels for firms to compete with their rivals. There has been a growing interest in understanding how firms use SMPs. Prior research on firms use of SMPs has focused on specific content shared by firms (Swani et al. 2013) and on the content-related differences between B2B (business-to-business) and B2C (business-to-consumer) firms (Swani et al. 2014). Firms post diverse content (news, updates about products or services, online promotions and emotional content) on social media (Swani et al. 2013, Lee et al. 2018, Bapna et al. 2019). Other research has focused on social networks and behavior of brand followers and their electronic word of mouth (Kim et al. 2014), community building strategies by large brands (Culnan et al. 2010), customers' complaints handling by firms (Einwiller and Steilen 2015), reputation management strategies on social media (Seebach et al. 2013), and offensive and defensive social media marketing strategies of firms after product-harm crises (He et al. 2017). While these streams of research examine specific use of SMPs by individual firms, there is very little research examining *how* firms compete on online social media platforms with their traditional rivals and what the effects of such competition are on related outcomes. Our paper seeks to fill this gap and focuses on studying how firms compete on Twitter, one of the more widely and more frequently used online social media platforms. In particular, we focus on firms that are identified as close competitors in the traditional context and examine whether these traditional competitors adopt *similar* content strategies on Twitter, and the impacts of such *similarity* or *dissimilarity* in competition on related outcomes. We restrict our focus to B2C (business-to-consumer) firms, which operate in the retail sector with NAICS codes 44-45, since retail firms have one of the highest levels of consumer facing social media activities (adweek.com 2015). Also, we are interested in exploring traditional competitors' strategies on social media platforms, and competition in the retail sector has one of the highest growth among all sectors, as measured by the Herfindahl-Hirschman Index (Autor et al. 2017). With the growing competition in the retail sector, and with retail firms' active presence on social media platforms, it is expected that offline competition will also be reflected in competition on SMPs.

Traditionally, competitors are identified based on two broad approaches (Clark and Montgomery 1999). The first approach is supply-based, which in general considers the following attributes of competing firms: technologies used, products offered, and strategies employed in traditional channels. This approach is specifically related to similarity in suppliers, distribution channels, resources, and size. The second approach is demand-based, which classifies competitors based on customers' attitudes and behaviors. Specifically, the demand-based approach is related to similarity in products, geo-markets, and price**,** as perceived by customers. The most common measures of competition based on supply-side or demand-side similarity are industry codes (Hoberg and Phillips 2016), and Hoover's database (Pant and Sheng 2015). These measures are extensively used by researchers in finance, strategy (Bergen and Peteraf 2002), marketing, management (Gur and Greckhamer 2018), and Information Systems, among others. Our study focuses on understanding how these traditional competitors that are characterized by a high degree of offline similarity, compete on online social media platforms.

While SMPs have rapidly emerged as the new frontiers of competition for traditional firms, SMPs also differ from traditional channels in important ways as they offer a number of capabilities that provide firms the ability to adopt strategies different from offline settings. Compared to traditional offline channels, firms have an opportunity to communicate with a large number of followers simultaneously and in a personalized manner. SMPs also enable firms to share content in real time and directly monitor users' reactions to their content. SMPs also enable firms to create communities of interest around specific products and services. While SMPs afford firms the ability to leverage this plethora of unique features, not all firms might be adept at leveraging the capabilities offered by SMPs and different firms might use SMPs in very different ways. Given the unique capabilities afforded by SMPs, our study seeks to examine if firms that compete closely in traditional channels also compete similarly on SMPs. Specifically, we focus on firms' content strategies on Twitter and examine how similar or dissimilar are firms with respect to their top traditional competitors.

In previewing our findings, we find that close competitors that have a high degree of similarity offline show greater divergence in the content strategies online. We find that the more dissimilar a firm's content strategy is to its closest rivals, the higher is its online engagement and new followers' acquisition rate.

We then examine why firms that are more dissimilar online from their closest traditional competitors experience better online outcomes. To uncover the underlying mechanism, we analyze the content of all firm-tweets and categorize them. The vast majority of firm tweets fall into a hierarchy of 10 categories that correspond to social media affordances identified in prior research (Karahanna et al. 2018). Using 20,000 manually labelled tweets, we build deep learning models that assign a majority of firm-tweets in the dataset to one of these 10 categories. The top-tier categories include tweets that emphasize cross-channel integration and tweets that involve users in co-creating value. The middle-tier categories include tweets that seek to create a community (sproutsocial.com 2018) of like-minded users around specific events, campaigns or products (themanifest.com 2018). The bottom-tier categories comprise of tweets that seek to share content with users. In examining the use of these different categories of tweets by firms, we find that firms that are dissimilar from their closest competitors are also better at leveraging Twitter's higher-level affordances, as compared to firms that are more similar to their traditional competitors. These firms that are able to better leverage Twitter for higher value-added activities relative to their closest competitors also experience better outcomes online.

Our study makes a number of important theoretical contributions and has valuable managerial implications that are discussed in the section "Implications and Conclusion".

Related Research and Theoretical Background

Prior work on SMP platforms has focused mostly on individual users, users' networks and user generated content. Users on social media start conversations, build reputation and form communities (Kietzmann et al. 2011). Other studies have examined the connections among social media users and user networks on SMPs (Susarla et al. 2011; Zeng and Wei 2012). Finally, research focusing on users behavior on SMPs has also studied user-generated content on social media (Smith et al. 2012; Luca 2015).

Some recent research (Culnan et al. 2010; Seebach et al. 2013; Swani et al. 2013; Einwiller and Steilen 2015; Lee et al. 2018) on SMPs has begun to examine firms' use of such platforms. Firms post diverse content on social media (Swani et al. 2013; Lee et al. 2018, Bapna et al. 2019), build brand communities (Culnan et al. 2010), handle complaints and manage their reputation (Seebach et al. 2013; Einwiller and Steilen 2015). While this stream of research explores specific usage of social media by firms, these studies do not examine how firms compete on social media platforms or how such competition affects outcomes. Our paper seeks to fill this gap and focuses on understanding how firms compete on SMPs by examining their content strategies on Twitter, one of the most frequently used social media platforms by firms.

A well-established body of research has examined how firms compete in traditional settings. One of the major tenets of this stream of research is "Institutional Isomorphism" (DiMaggio and Powell 1983). Isomorphism has been defined as "a constraining process that forces one unit in a population to resemble other units that face the same set of environmental conditions" (Hawley 1968). Historically, Isomorphism has been measured as the degree of adoption and/or assimilation of certain policies, standards (Bala and Venkatesh 2007), methods, norms, codes of conduct, behaviors, innovations (Hsu et al. 2012) etc. Following the seminal work of Meyer and Rowan (1977), and DiMaggio and Powell (1983), a number of studies have focused on examining the antecedents of Institutional Isomorphism, including competitive pressures, which lead to homogenization, or *similarity* of institutions. Most of the work related to Isomorphism has focused on firms in traditional settings. There is some empirical support of the effect of competitive pressures on similarity of practices adopted by firms in traditional settings (Berrett and Slack 1999; Farndale and Paauwe 2007). Our study adds to this stream of research by examining the similarity of traditional competitors' content strategies on online social media platforms.

Recently, researchers in Information Systems (Pant and Sheng 2015) have hypothesized that Isomorphism might also be observed in the Web footprints (firms' websites, online news, blogs, review platforms, co-searched firms and shared links among firms' websites) of competing firms. Pant and Sheng (2015) have found that the Web footprints of competitors are more similar than Web footprints of non-competing firms. Our study adds to this emerging stream of work by examining the similarity of firms' content strategies and its implications for firms' outcomes on Twitter. Following on the lines of prior work that examines textual content shared by firms (Pan et al. 2015; Pant and Sheng 2015), we analyze the content shared by traditional competitors on Twitter to examine the degree of similarity in content among these competitors and further test how this degree of similarity affects online engagement as well as a new followers' acquisition rate.

Our study also draws upon and builds on earlier work on social media affordances. As mentioned earlier, social media platforms are unique in several ways providing firms the ability to adopt strategies different from offline settings. Emerging research on IT affordances (Majchrzak and Markus 2012, Yoo et al. 2012) in general, and social media affordances in particular, provides a good framework to examine the capabilities provided by SMPs. An affordance is defined as an action possibility that is available to an actor in the environment (Gibson 1986), and IT affordances are possibilities for a goal-oriented action afforded to specified groups of actors by technical objects (Pozzi et al. 2014). SMPs possess unique affordances that firms can exploit to creatively engage their followers (see Karahanna et al. (2018) for a review of social media affordances). While most of existing research on social media affordances focuses on organizational use of SMPs for internal communications and knowledge sharing, our work focuses on firms leveraging social media affordances for external audiences. Of particular relevance to our context are the affordances that are specific to firms using Twitter – affordances such as *communication* and *self-presentation*, *meta-voicing*[1] *and relationship formation*, *interactivity*, *collaboration*, and *competition*.

The most common use of Twitter by firms is to communicate with users and to share *content* by leveraging SMP affordances such as *self-presentation* and *communication*. We term these lower-level Twitter affordances *content* affordances. However, social media affordances such as *relationship formation* and *meta-voicing* provide the capability for firms to not only share content online but also to create a *community* of users with similar interests around specific products or events. On microblogging platforms like Twitter in particular, social tags help to bring like-minded users around focal topics/products/events (sproutsocial.com 2018) and help users to create associations with other individuals or content (Treem and Leonardi 2012). We term these Twitter affordances *community* affordances.

Social media affordances such as *interactivity*, *collaboration*, and *competition*, enable firms to go beyond sharing content and creating a community of users to create value by engaging users in *co-creation* (Mandviwalla and Watson 2014) and combining multiple sources and channels. We term these higher-level Twitter affordances *co-creation* affordances.

While SMPs like Twitter enable a variety of affordances for all firms, these affordances are potential for action (Pozzi et al 2014) and serve as possibilities for firms to achieve different objectives. However, these affordances need to be triggered and actualized (Strong et al. 2014) by firms to achieve desired outcomes. Not all firms might be equally adept at leveraging these SMP affordances and consequently might differ in the outcomes achieved on these social media platforms. In analyzing the content strategies adopted by firms on Twitter, our study also examines the firms' ability to leverage the different social media affordances and what these imply for firms' online outcomes. The literature on social media affordances provides the framework for understanding what firms that are more similar to their closest offline competitors, do differently on Twitter and why they perform better on online outcomes relative to their competitors.

In addition to drawing upon prior work on SMP affordances, our study also contributes to this stream of research by examining how close traditional competitors differ in their ability to leverage the different affordances provided by Twitter. Our study also shows how the differences in leveraging SMP affordances impact online outcomes of interest. Our findings could be used to better understand how the effectiveness of firms' content strategies on SMPs are related to their ability to leverage specific SMP affordances.

Research Context and Data

We begin by identifying traditional close competitors, as identified by Hoover's database (hoovers.com). We also verify if these firms are also close competitors from other sources such as Mergent (mergentonline.com) and Nasdaq (nasdaq.com)[2]. In our study we focus on top three competitors, and these close competitors are also expected to have the highest offline similarity. We also confirm that these firms that are closest competitors have a high degree of similarity by analyzing their 10-K reports. Focusing on the top-3 competitors for each firm also allows us to work within the data collection restrictions imposed by Twitter API. It should be noted that Twitter API has a limit of 3,200 most recent tweets per firm account. To collect all tweets, we first open each firm's Twitter account Web page for each day, and collect each tweet ID, then using those tweet IDs we collect all tweets' text and metadata through Twitter API.

Using Twitter API, we collect data for 199 retail companies from Russel 3000 list. The list contains the largest companies in terms of market capitalization. We focus on B2C (business-to-consumer) firms and collect about 2.42 million tweets (Table 1) for these firms for the period January 2012-August 2017. The categories of firms are: fashion and apparel (including sporting goods), home supply and houseware, furniture, cosmetics, book stores, health supplement stores, supermarket chains.

Out of 2.42 million tweets (Table 1), 893,525 tweets are firm-initiated. The rest (majority) are direct responses to customers' questions and complaints (which could be identified by "@" tag at the beginning of a tweet), and retweets of tweets (minority) by a focal firm from other non-firm accounts (those tweets contain a tag "RT @"). In the analysis we focus only on firm-initiated tweets, since those tweets should reflect firms' content strategies as well as the timing of those strategies with respect to their competitors.

Some firms have separate Twitter accounts for customer service (Q&A), job postings etc. In this paper we focus only on the primary official Twitter account for each firm.

Sector	**Number of firms**	**Total number of tweets**	**Average number of tweets per firm**	**Min. number of tweets for a firm**	**Max. number of tweets for a firm**	**Average (min:max) number of followers**
Retail Trade	199	2,422,968 (893,525 firm-initiated tweets)	12,176 (4,278 firm-initiated tweets)	539 (365 firm-initiated tweets)	72,137 (17,723 firm-initiated tweets)	489,026 (1,102:8,744,557)

Table 1. Descriptive statistics for the database of tweets

The Twitter data is a panel dataset with the tweets for the period from January 2012 up to August 2017. Each tweet has the text of the tweet, and such metadata as favorited tweets or "favorites" (a.k.a likes), retweets, date/time, number of followers (a static number at the time of data collection, which is August 2017). We chose 2012 as a starting year for our analysis as we find that most firms in our dataset started actively posting content on Twitter around 2012. Since our study focuses on analyzing the content shared by firms on Twitter, we ignore earlier time periods wherein the content shared by firms on Twitter is sparse.

We also collect followers' IDs for each firm separately. Next, for each follower ID we collect a follower profile. IDs and profile information are used to calculate the date when each follower starts following a firm. These dates allow us to calculate the total number of followers and new followers by quarter.

Methodology

Similarity/Dissimilarity. The methodological framework for calculating similarity/dissimilarity of content includes the following steps: first, we examine the pairwise similarity of traditional top competitors on Twitter. To this end, we first construct a term frequency-inverse document frequency (TFIDF) vector (Aizawa 2003) for tweets of each firm by quarter. Frequency of occurrence of a term from vocabulary in each firm's tweets consists of the term frequency, and the number of a firm's tweets in which a term occurs determines the inverse document frequency (as shown in equation (1) below). Such numeric value reflects how important a word is to a document in a collection/corpus. It can weigh down the effects of too frequent terms. Each firm for each quarter is thus represented by a TFIDF vector.

Next, we use TFIDF vectors to calculate pairwise cosine similarity for each pair of firms for each quarter. We chose the cosine similarity measure, because it is one of the most commonly used methods for determining text similarity (Huang 2008), and because it addresses the problem of unequal corpus lengths. We do not use Jaccard similarity because the tweet vector space is a continuous one and Jaccard similarity is specifically designed for a discrete space. Also, we do not use Euclidian distance due to its poor performance in high dimensional space. We use the "scikit-learn" library of python to calculate TFIDF vectors and cosine similarity. The cosine similarity is in the range from 0 to 1, where 1 is the most similar. Finally, we calculate average cosine similarity with top three competitors for each firm for each quarter. We also performed similar analysis by year and obtained consistent results.

To increase the quality of our data, we use pre-processing, including stop words removal and non-ASCII character deletion as well as word stemming. The text of tweets includes only firm-initiated tweets and does not include retweets by a focal firm, i.e. when a firm reposts content of some other firm or user.

The formula for computing TFIDF value for a term t_i in document d_j is provided below (Aizawa 2003):

$$TFIDF(t_i, d_j) = tf(t_i, d_j) * idf(t_i) = \frac{f_{t_i}}{|d_j|} * \log\frac{N}{N_i} \qquad (1)$$

where f_{t_i} represents the frequency of term t_i in document d_j; $|d_j|$ is the number of words in document d_j. N_i – number of documents containing term t_i, and document d_j consists of all tweets of a focal firm in a quarter; N-total number of documents.

The cosine similarity calculation for two firms is formulated by equation (2) (Pant and Sheng 2015):

$$sim(\overrightarrow{f_A}, \overrightarrow{f_B}) = \frac{\overrightarrow{f_A} \cdot \overrightarrow{f_B}}{||\overrightarrow{f_A}|| \cdot ||\overrightarrow{f_B}||} \qquad (2)$$

where $\overrightarrow{f_A}$ and $\overrightarrow{f_B}$ represent TFIDF vectors for firms A and B. $||\cdot||$ is the length of a vector. The cosine similarity is a dot product of TFIDF vectors of two firms normalized by their lengths.

Followers. Each tweet metadata provide only a static number of followers at the moment when metadata are requested from Twitter API. Bruns et al. (2014) provides a methodology to calculate the date when each follower starts following a focal firm. We obtain all followers' IDs through Twitter API "GET followers/ids" command (developer.twitter.com). These followers' IDs are provided in a very specific order - from the most recent to the earliest followers. We then collect profile information for each follower ID including username, screen name, description, location, and, most importantly, the date when a Twitter account was created by each user (i.e. follower). The ordering of followers and Twitter account date of creation are the two components used in the algorithm to calculate the "date of following" of a focal firm by each follower (due to space limitations, we do not include the method in this paper).

We then create a count of the total number of followers in each quarter and the number of new followers by quarter. To do that, we count followers who started following a focal firm by the end of each quarter.[3] The total number of followers is a dynamic number, and it could be used as an independent (control) variable in the firm fixed-effects econometric model, when all static variables are differenced out.

Next, the number of new followers by quarter is calculated as the total number of followers at the end of the quarter minus the total number of followers at the beginning of the same quarter. This is used as the dependent variable in the model to see if dissimilarity has an impact on the new followers' acquisition rate.

Hypotheses and Econometric Specification

Prior work on firm competitive strategies has found that competitive pressures lead to isomorphism in competing firms' strategies (Berrett and Slack 1999), and such isomorphism has been shown to positively impact firm performance (Brouthers et al. 2005). Particularly, when firms face competitive pressures under uncertainty, they might jump onto a bandwagon of adopting dominant strategies even if the outcomes of such strategies are ambiguous (Abrahamson and Rosenkopf 1993). For instance, when a firm enters a new domain, imitating an incumbent's strategy could improve firm's performance (Wu and Salomon 2016).

In the case of social media platforms, firms facing heightened uncertainties in decision-making with respect to social media content strategies, might jump onto a bandwagon of dominant strategies such as using similar online promotions, coupons and discounts that could drive engagement and attract more new followers. In this case, we would expect that:

Hypothesis 1 A. Isomorphism in SMP content strategies will have a positive effect on related outcomes.

A competing stream of studies has found that firms that choose divergent strategies are likely to outperform their competitors (Badir et al. 2013). According to the resource-based view of the firm (Barney 1991), firm might choose to exploit resources that are valuable, rare, imperfectly imitable, and imperfectly substitutable to differentiate themselves (Farndale and Paauwe 2007) from their competitors and such differentiation could lead to better performance (Badir et al. 2013).

Thus, in the context of social media platforms, firms might strategically choose (Farndale and Paauwe 2007) competitive divergence, and leverage unique Twitter affordances to experiment with Twitter content and use Twitter to differentiate themselves from top competitors, and such divergence will be more appealing to online users and will attract more new followers. If so, we would expect that:

Hypothesis 1 B. Divergence in SMP content strategies will have a positive effect on related outcomes.

Thus, whether isomorphism or divergence in social media content strategies leads to better online performance, remains an empirical question – one that we examine in this study.

Econometric specification. To test the effect of similarity on online engagement, we estimate the model:

$$Y_{it} = \beta_o + \beta_1 \text{Similarity}_{it} + \beta_2 \text{Tweets}_{it} + \beta_3 \text{Followers}_{it} + \alpha_i + \delta_t + u_{it} \quad (3)$$

The main independent variable (similarity) in the specification is the average cosine similarity (β_1) with top three competitors for each firm for each quarter. We chose the quarter as the period in the panel data, since we believe that quarterly data will have enough tweets to measure similarity in content strategies for each pair of firms even if posting frequencies for firms differ considerably (for example, some firms post once per day, some firms post once per week). The outcome variable (Y_{it}) is engagement on Twitter, as measured by the total number of "favorites" (first outcome) and total number of retweets (second outcome) per quarter. We report the results separately for "favorites" and retweets. The number of tweets (Tweets) and the dynamic number of followers (Followers) for a given firm for a given quarter are control variables. The variable α_i represents firm fixed effects, and the variable δ_t represents year fixed effects.

Regarding control variables, since top competitors are already matched by Hoover's based on offline supply-based and demand-based similarity, in our model specification we provide additional controls for firms related to their online activity such as the number of posts (tweets) and the dynamic number of followers. We also leverage the panel structure of our dataset to include the firm fixed effects in the model. Firm fixed effects control for time-invariant unobserved heterogeneity across firms. Additionally, we add year fixed effects to the model to control for potential yearly trends in firms' content strategies.

The second specification examines the effect of similarity/dissimilarity on the new followers' acquisition rate, and the model is represented by the same formula (equation 3), as described above, with the only exception that the dependent variable (Y_{it}) is new followers gained by each firm in each quarter.

It should be noted that correlations among independent variables do not go beyond 0.33 (the range is from -0.33 to 0.25) in absolute values. Additionally, Variance Inflation Factors[4] for all variables are lower than 2 (which is much lower than the "problematic" value of 10, and lower than a more conservative "problematic" value of 4). Thus, multicollinearity is not an issue in these specifications.

To address any endogeneity concerns, we also employ a system GMM model using Arellano–Bover/Blundell–Bond linear dynamic panel-data estimation with lagged dependent variables as instruments. We use a Stata xtdpdsys 2-step estimation with 2 lags of the dependent variable and with robust standard errors suggested by Windmeijer (2005). One of the prerequisites of using dependent variable lags as instruments is that there should be no second-order autocorrelation of residuals. We test that condition and confirm that the autocorrelation of the second order is not present in all model specifications (the p-value higher than 0.14). With robust standard errors, the Sargan test of overidentifying restrictions is not calculated. But this test is conducted with "gmm" errors (Stata command "estat sargan") and a 1-step estimation, and that test is passed in all model specifications (Prob > chi2 = 0.99). It is pertinent to note that adding lags of independent variables does not change the results.

Additionally, we use a 2-stage least squares estimation with heteroskedasticity-robust standard errors with the following instrumental variable – cumulative average cosine similarity with top competitors in the previous quarters up to the current quarter (excluding the current quarter). The instrumental variable (IV) should affect a firm's propensity to be dissimilar in each current quarter and should not directly affect relevant outcomes (online engagement and new followers' acquisition rate) in each current quarter. The only effect of IV on outcomes should come from dissimilarity. The first stage of the model is highly significant. Additionally, the null hypotheses of under-identification and weak identification (using Kleibergen-Paap rk LM statistic) are rejected, and Hansen J statistic cannot be used for cases of 1 IV for 1 endogenous regressor. As a robustness check, we also use another instrument – moving average similarity in the last 5 quarters before the current quarter and find the results to be consistent.

Results

We examine how the degree of similarity of a focal firm with its closest competitors is linked to the outcome variables, namely, online engagement and the acquisition rate of new followers. To recall, the independent variable is the average cosine similarity with top three competitors for each firm for each quarter, and the outcome variables are the total number of "favorites" for each quarter, the total number of retweets for each quarter and the number of new followers for each quarter.

Our results (Table 2) show that the more distant a firm is from its traditional competitors (i.e., the lower the degree of cosine similarity of content with its traditional competitors), the higher is the online engagement on Twitter. The model specifications have a negative coefficient for similarity, which indicates that higher similarity leads to lower online engagement. In other words, the more *dissimilar* a firm is from its closest traditional competitors, the higher is its online engagement.

Model	**FE**	**GMM**	**2SLS**	**FE**	**GMM**	**2SLS**	**FE**	**GMM**	**2SLS**
Variables	**Favorites**	**Favorites**	**Favorites**	**Retweets**	**Retweets**	**Retweets**	**New Followers**	**New Followers**	**New Followers**
Similarity	-31475***	-53177***	-25394***	-11553***	-8399***	-9774**	-4302^{ns}	-13626*	-16488**
St. Errors	(8132)	(10673)	(9686)	(2871)	(888.75)	(3942)	(7256)	(7331)	(7420)
Tweets	39.5***	32.6***	35.99***	18.3***	18.74***	17.51***	9.45***	10.4**	8.02***
St. Errors	(2.65)	(8.62)	(4.6)	(0.935)	(2)	(1.6)	(2.36)	(4.7)	(2.09)
Followers	0.035***	0.0084**	0.034***	0.0079***	0.0018***	0.0131***	-.0027^{ns}	0.0046^{ns}	0.052***
St. Errors	(0.002)	(0.0042)	(0.002)	(0.00073)	(0.00063)	(0.00092)	(0.0018)	(0.0032)	(0.0024)
Sample	3509	3197	3353	3509	3197	3353	3509	3197	3353
Firms	156	156	156	156	156	156	156	156	156
R-squared	0.276		0.2789	0.329		0.35	0.0054		0.4451
***p<0.01**	****p<0.05**	***p<0.1**	**ns – not significant**						

Table 2. Fixed Effects, GMM and 2SLS similarity model results

Note. Mean cosine similarity: 0.09, median: 0.08, standard deviation: 0.06, min: 0, max: 0.48; year fixed effects are not reported.

When it comes to new followers, the GMM model and the 2SLS model have a significant negative coefficient of similarity, which indicates that firms with more dissimilar content relative to their closest traditional competitors attract more followers online. Additionally, the control variables (dynamic followers and tweets) have positive coefficients in all model specifications. The exception is non-significant coefficients for the dynamic number of followers affecting the new followers' acquisition rate in GMM and FE models.

It is also pertinent to note that about 600 tweets (out of 893,525 in our dataset) generate as many as 10,000 "favorites" and/or retweets. As a robustness check, we re-estimate our models after removing top 1-5% of quarters with the highest scores for "favorites" and/or retweets and obtain consistent results.

We also perform a series of robustness checks that involve restricting the sample to firms with few/many followers, removing firms with millions of followers or very few followers, as well as splitting samples into more actively posting firms (more than 300 tweets per quarter) and less actively posting firms (less than 300 tweets per quarter). In all these cases the results are consistent.

Another robustness check is related to whether Twitter changed its timeline algorithm to feed tweets to users in a particular way (for example, using a feature "show me the best tweets first" etc.). If tweets are shown in some nonrandom manner, it could be the case that most dissimilar tweets might somehow have more impressions than similar tweets. Twitter changed its timeline algorithm in March 2016 (socialmediatoday.com 2016). Prior to that, tweets were shown in reverse chronological order (starting with the most recent tweets). As a robustness check, we restrict our data to years 2012, 2013, 2014 and 2015, and the results are consistent. Thus, we believe that the Twitter feed algorithm does not impact our results.

Why Do Dissimilar Firms Perform Better?

Our results, thus far, show that the more dissimilar a firm is to its closest traditional competitors with respect to its Twitter content, the better are its outcomes on Twitter. However, it is not clear why dissimilarity in content leads to better outcomes. To understand the underlying drivers of why dissimilarity is associated with better outcomes, we examine what dissimilar firms do differently compared to their close competitors. We manually classify 20,000 random tweets into an exhaustive set of categories[5], out of which the following 10 are the most frequent categories representing over 75% of all content.

Content: The five most common categories consist of firms sharing content relating to *products*, *product usage*, *questions, events, and coupons and promotions* through their tweets. These tweets that are primarily focused on firms sharing content with users leverage Twitter's affordances such as self-presentation and communication.

Product information. This category of tweets is used by firms to introduce a product and provide product information to online users.
Product usage tips. This category of tweets is related to specific tips on how to use a firm's products.
Questions. This category of tweets asks online users questions with a blank or without blanks.
Events. This category of tweets is related to events, where firms share information about upcoming events.
Coupons and promotions. This category of tweets is used to share information about online and offline coupons and promotions available in all stores, or in specific offline locations or exclusively online.

Community: The next set of categories comprise of tweets where firms seek to create a community of like-minded users (sproutsocial.com 2018). All of these tweets involve the use of Twitter hashtags # (themanifest.com 2018). The categories include firm tweets relating to #expert tips, #product collections, and #special events. The use of Twitter hashtags # enable firms to create a community of like-minded followers and foster interactivity around a focal campaign/contest/event/product. In contrast to tweets without a hashtag that are primarily used to disseminate information about products/events/promotions, the use of hashtags serves as a mechanism for enabling the realization of higher-level affordances, such as relationship formation, and meta-voicing.

#Expert tips. This category of tweets is related to online collaboration with influential SMP users. These influential users provide tips related to style/look/products.
#Product collections. This category of content promotes specific product collections under a hashtag.
#Special events. This category of tweets includes sponsorship for a series of events.

Co-creation: The final set of tweet categories comprise of tweets wherein firms seek to involve users more actively to help create and share content relating to their offerings. Twitter hashtags are used to help users not only create a community around specific topics but also involve them using specific campaigns and contests. These categories include *#offline-online campaigns* and *#contests soliciting user generated content* that leverage Twitter affordances including interactivity, collaboration, and competition.

#Offline-online campaigns. This category of tweets is related to firms' cross-channel marketing efforts. Firms with offline marketing campaigns often involving celebrities, use Twitter to not only solicit creative ideas from users but also invite users to follow celebrities' example and contribute content. Typically, these campaigns originate offline, and are promoted in news sources as campaigns with a specific Twitter hashtag. Thus, firms invest resources to closely integrate their marketing and promotions across offline as well as social media platforms.[6]

#Contests soliciting user-generated content. This category consists of tweets relating to online contests asking online users to upload user-generated content in the form of advice, design suggestions, photos or videos involving a firm's products. As part of these contests, users are also encouraged to vote for other users' content, thus adding additional interactivity to the campaign.

Due to space limitation, we do not provide examples of tweets by category.

The remaining tweets (about 25% of all tweets) are categorized under "Misc.". These mostly comprise of tweets containing "thank you" messages, "birthday greetings", and public service announcements. Next, using these labeled tweets, we train two deep Recurrent Neural Network (RNN) models with Long Short Term Memory (LSTM) using the python's "keras" library – one model for 5 categories with hashtags and another model for 5 categories without hashtags. To improve the models' accuracy, we use external embeddings with the dimension 200 from GloVe (Global Vectors for Word Representation, nlp.stanford.edu), pre-trained specifically for Twitter on 2 billion tweets. The last layer uses sigmoid activation function to output a distribution of probability where each represents an input tweet being classified as the corresponding category. Further, a grid search technique is used for finding the best parameters, and the early stopping is also used to prevent overfitting. The overall first model accuracy for tweets with hashtags is 86% on the hold-out test set (20% of data, average precision is 83.7, average recall is 87.9), and 83% under the ten-fold stratified cross-validation. The overall second model accuracy for tweets without hashtags is 87% on the hold-out test set (20% of data, average precision is 84.9, average recall is 88.8), and 85% under the ten-fold stratified cross-validation. In the ten-fold cross-validation setting, the algorithm runs 10 times, and each time each training dataset uses 90% of random labelled tweets, while the test dataset includes the remaining 10% of random labelled tweets. The two models are then used to classify all remaining 873,525 firm-initiated tweets. We separate all 873,525 tweets into 2 groups – tweets with hashtags and tweets without hashtags and use two trained models to classify related categories. All later analyses are performed on all 893,525 firm-initiated tweets[7].

Hypotheses. Figure 2 illustrates the classification of the 10 tweet categories into a hierarchy of three tiers. The bottom tier consists of firm-tweet categories where the focus is on communicating information and sharing *content* with all users. All firms in our sample use Twitter affordances including *self-presentation* and *communication* to share content with their users. As mentioned earlier, we term tweets that leverage these basic set of affordances "*content*" categories.

The categories in the middle-tier leverage *relationship formation* and *meta-voicing* affordances and enable firms to not only share content but also create a *community* of users around focal themes. We term tweets that leverage these mid-tier affordances "*community*" categories.

The categories in the top tier involve *cross-channel integration* (#Offline-online campaigns) as well as engaging users in *co-creating value* (#Contests soliciting UGC). The affordances that correspond to that tier are interactivity, collaboration, and competition. Firms leverage Twitter's interactivity affordance to create offline-online interactive campaigns when users are encouraged to share their ideas related to a campaign. Firms leverage collaboration and competition affordances to design contests soliciting users to upload user-generated content. The categories in the higher tiers not only leverage the affordances in the lower tiers but also seek to create additional value and require greater investments from the firms. Given the higher cost and resources required by higher-level affordances, we expect fewer firms to leverage the higher-level "community" and "co-creation" affordances as compared to firms leveraging the lower-level "content" affordances. Thus, we hypothesize that *firms that leverage higher-level affordances are likely to be more dissimilar from firms that leverage only the affordances at lower levels in the hierarchy*. Further, since the higher-level affordances are aimed at creating a community and getting users more involved with the firms' content, we hypothesize that *higher-level affordances will have higher online engagement*.

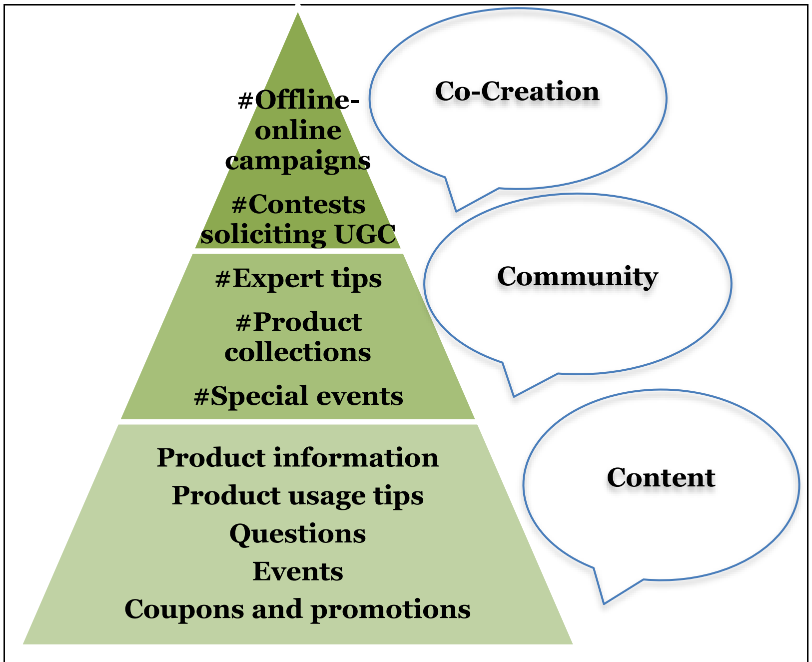

Figure 2. Hierarchy of Firm-Tweet Categories

This is also consistent with Helms et al. (2012), where the authors describe five active social media innovation strategies that are ordered according to user participation levels: *general community engagement, ideas competitions, interactive value creation, participatory design, and product design*. By creating a community, a firm stimulates online users to share experiences with like-minded people. In the context of Twitter, firms have a big community of followers, and use hashtags to create sub-communities of users with closely related interests that form around an event, a product or a campaign. Next, firms can create contests on Twitter to generate ideas and solicit user-generated content. These contests are designed to solicit novel ideas in the form of general design suggestions or desired product features. An interactive value creation occurs when a specific contest is targeted to a specific group of users, for example, when a firm runs a competition for the best post or best photo with its products. Such user-generated content receives votes from other users and might be used by a firm later in its marketing materials online and offline. Participatory design and product design involve even more focused crowdsourcing campaigns with the goal of soliciting new ideas that could be used to launch new products.

Thus, we hypothesize that "content" affordances will have the lowest level of user engagement, because they do not actively involve users in co-creation. The "community" affordances are expected to be related to a *general community engagement* strategy, and, thus, will have higher online engagement than "content" affordances". Finally, "co-creation" affordances that include *ideas contests, interactive value creation, participatory design and product design* strategies are expected to have the highest engagement.

To test these hypotheses, we use the (dis)similarity scores for all 156 firms for each quarter and calculate proportion of tweets in each category in each quarter for each firm. We estimate system GMM and 2SLS models (this model is estimated with the same instrument for dissimilarity that was used earlier), where the independent variable is the (dis)similarity score, and the dependent variable is percentage of tweets in each category (all 10 categories comprise 100%). Table 3 shows that higher dissimilarity is associated with more usage of higher-level affordances for the 2SLS model. The GMM model results are consistent.

We hypothesized that the firms leveraging higher-level affordances are likely to be more dissimilar from firms that leverage only the lower-level affordances. Our results support that hypothesis.

Percentage of tweets	#Contests soliciting UGC	#Offline-online campaigns	#Special events	#Product collections	#Expert tips
Similarity	-0.055*** (0.009)	-0.126*** (0.018)	-0.055*** (0.0107)	-0.076*** (0.0095)	-0.7*** (0.053)
P-value	0.000	0.000	0.000	0.000	0.000

Table 3. The 2SLS similarity model results for each higher-level category

Note. The results for lower-level categories are mixed (2 categories, namely "events" and "questions", have negative coefficients while other 3 categories have positive coefficients)

The second hypothesis was related to higher engagement for higher-level affordances compared to lower-level affordances. Table 4 shows the results.

Tier	Tweet Category	Favorites	Retweets	Number of tweets
I	**#Offline-online campaigns**	19.74	10.07	13017
I	**#Contests soliciting UGC**	16.67	15.09	8773
II	**#Expert tips**	15.5	8.24	78334
II	**#Product collections**	14.28	9.24	7920
II	**#Special events**	13.3	8.01	8830
III	**Product information**	12.5	6.63	228207
III	**Product usage tips**	11.44	6.57	46503
III	**Questions**	11.02	7.31	96468
III	**Events**	10.74	6.82	22439
III	**Coupons and promotions**	9.92	8.69	130872
	Misc.	12.62	8.35	234289

Table 4. Normalized engagement (number of (favorites/retweets)/tweet/100,000 followers) by category

Table 4 has all 10 categories listed in the order of corresponding tiers (1-3). We find that for "favorites", categories in tier 1 (top) have higher normalized engagement than categories in tier 2, which in turn have higher values than those in tier 3 (bottom). Regarding retweets, the general pattern is the same as for "favorites", with one exception. Tweets in the category "coupons and promotions" in tier 3 have a higher number of normalized retweets than the two categories in the 2nd tier (#expert tips and #special events).

Mediation Analysis. Our findings point to a mediation process, where dissimilarity affects online engagement and new followers' acquisition rate through the usage of higher-level affordances. To test for full or partial mediation of the effect of dissimilarity on online engagement and new followers' acquisition rate, we use the structural equation modeling method (Stata "SEM" package). We combine all 5 higher-level categories into a new variable by summing up proportions of each of the individual higher-level categories for each firm for each quarter (using the average value of those proportions would give the same result). If higher-level categories fully mediate the effect of dissimilarity on online engagement and new followers' acquisition rate, then the direct effect of dissimilarity on those dependent variables should become non-significant when "higher-level categories" variable is included in the model as a mediator.

The conceptual mediation equation is shown below:

sem (MV <- IV CV1 CV2) (DV <- MV IV CV1 CV2) (4)

where MV refers to the mediator variable (higher-level categories); DV refers to the dependent variables (favorites, retweets or new followers); IV refers to the independent variable (similarity); CVs are covariates (the number of tweets and the number of followers).

Table 5 below illustrates the results.

Dependent Variables	Direct effect of similarity	Indirect effect of similarity	Total effect of similarity	Direct effect of higher-level categories	Indirect effect of higher-level categories	Total effect of higher-level categories
Favorites	-13,455 (p = 0.126)	-5,802 (p=0.011)	-19,257 (p = 0.028)	57.61 (p = 0.001)	No Path	57.61 (p = 0.001)
Retweets	-6,325.5 (p = 0.07)	-1,610.72 (p=0.021)	-7,936.33 (p=0.016)	15.99 (p = 0.021)	No Path	15.99 (p = 0.021)
New followers	-13,787.6 (p=0,078)	-12,870.5 (p=0.000)	-26,658.1 (p=0.000)	127.8 (p=0.000)	No Path	127.8 (p=0.000)

Table 5. Structural Equation Modeling mediation results

Table 5 shows that higher-level categories fully mediate the effect of dissimilarity on online engagement and new followers' acquisition rate. The direct effect of similarity becomes non-significant in the presence of the mediator (i.e. "higher-level categories", which has a statistically significant direct effect) in the model. The direct effect of similarity on the use of higher-level categories (not reported in the table 7) is negative "–100.7" with the p-value of less than 0.001, which means that dissimilarity is associated with the usage of higher-level categories. As a robustness check, we operationalize higher-level categories as counts of tweets in specific categories for each firm for each quarter (not as proportions), and the results are consistent.

Implications and Conclusion

Our study seeks to understand how close traditional competitors compete with each other in online social media platforms and consequences of such competition for outcomes in these platforms. In doing so, our study introduces a new measure of online social media competition based on similarity of content with top competitors. We find that divergence in content strategies from a firm's closest competitors leads to higher online engagement and attracts more new followers for the focal firm. While earlier research (Pant and Sheng 2015) shows that close traditional competitors are more likely to adopt similar content strategies compared to other firms, our study focuses on the differences in content strategies among close traditional competitors. We find that although close traditional competitors have a high degree of similarity offline, there is greater dissimilarity in their online content strategies on Twitter, and that these differences have important consequences for firms' online outcomes on Twitter.

To understand the underlying mechanism behind the positive effect of divergence on related outcomes, we classify content on social media into 10 categories that map to 3 tiers of social media affordances. We find that some firms not only adopt different content strategies on Twitter as compared to their closest rivals but are also adept at leveraging higher-level social media affordances and that higher-level social media affordances lead to higher online engagement.

Our study makes several important contributions. First, there is a growing body of research literature related to how firms use social media platforms. However, there is very little research related to competitors' strategies on SMPs. Our paper seeks to fill the gap by exploring traditional competitors' strategies on Twitter. Next, the results of our study make contribution to research literature on isomorphism of traditional competitors, and specifically to research related to dynamic modern methods of competitors analysis using Web footprints of rivals. The imitator (Wu and Salomon 2016) or innovator (Zheng Zhou 2006) dilemma has been described in the research literature related to traditional firms in offline channels. We find that it is divergence (i.e. innovator strategy) that leads to better online outcomes. Thus, traditional rivals can overcome the pressure to be isomorphic on social media platforms and use the strategy of divergence that leads to higher engagement with their followers and attracts more new followers.

Our findings related to how divergence leads to related outcomes make contribution to research literature on the use of SMP affordances by competing firms. While previous research identifies SMP affordances, it does not explore the differential impact of affordances on outcomes. We propose a 3-tier strategic framework of tweet categories and affordances and show that higher-level categories with higher-level affordances not only include lower-level categories with lower-level affordances, but also create additional value for the dissimilar firms that use higher-level categories in higher proportion compared to their more similar rivals. Hierarchically, the value of content categories increases when firms shift their strategies from only sharing content to creating community of like-minded users to co-creating value with online users.

Our study also contributes to the research literature on early mover advantages. While the close traditional competitors in our sample are well-established firms that compete closely with their traditional rivals, we find that not all of them are equally adept at leveraging the different affordances of Twitter. We find that early movers, who have been able to leverage Twitter's higher-level affordances, better than their closest competitors, experience better online outcomes. Whether these early mover advantages are sustainable would be an interesting topic for future research.

Our findings have some important managerial implications. Managers can exploit the ranking of competitors by dissimilarity of content on social media platforms to determine which of their top traditional competitors have a potential social media competitive advantage. The dissimilar rivals are more likely to leverage higher-level social media affordances and experience better online outcomes. Additionally, managers can use our hierarchy of content categories to better design their social media strategies. While bottom-tier affordances are relatively less costly for firms, their impact on engagement is limited. Higher-level affordances require higher investments from firms, but those investments pay off in the form of higher online engagement and higher number of followers. Another implication is that firms not only need to leverage the interactive affordances of SMPs, but also need to provide mechanisms for users to keep track of, and engage with, the firms' interactive online campaigns over time.

Our study is not without limitations. First, we explore firms' social media content strategies in the retail sector. Future research could investigate firms' strategies in other industry sectors such as Arts, Entertainment, and Recreation or Finance and Insurance. Firms' social media strategies in these other sectors might differ from strategies in the retail sector, and it would be interesting to examine how firms in these sectors compete on SMPs. Next, in our study we do not investigate details of followers' behavior for each firm's official Twitter account. Future research could explore each firm's followers' activity to see if dissimilar firms not only better engage their current followers but also attract more engaged distinct loyal followers (brand fans) that follow only a focal dissimilar firm and do not follow competitors. Those loyal brand followers might be a source of strong online and offline word of mouth and might act like brand ambassadors. Future research can further explore the detailed role of social media affordances in creating and sustaining higher online engagement, as well as the value to firms of combining online strategies with offline marketing campaigns. Future extensions to this work can examine whether online strategies impact offline metrics, for example, firm sales or stock prices. Finally, our study focuses on firms' use of one social media platform, Twitter. Future work can analyze firms' strategies on other social media platforms, such as Facebook, and examine if firms' competitive behaviors are similar across these platforms.

Endnotes

[1]Meta-voicing is defined as "engaging in the ongoing online knowledge conversation by reacting online to others' presence, profiles, content and activities" (Majchrzak et al 2013)

[2]While the last two sources provide an average of 15 to 20 competitors for each focal firm, these sources do not identify the top competitors. However, Hoover's database specifically highlights top three competitors for each firm, and we confirm that these competitors are also listed as competitors in the other sources

[3]Alternatively, and more conservatively, one could calculate the total number of followers for each quarter using the beginning of each quarter as a cut-off. We use both approaches and obtain consistent results.

[4]Those VIFs were calculated in a simple linear regression

[5]We identified about 100 categories of tweets. The 90 categories are less frequent, while the 10 categories are the most frequent. The first author classified all 20,000 random tweets, while the second author classified a random subsample of 1,000 tweets out of those 20,000 tweets. The percent agreement between two raters varies between 86% and 88%.

[6]We search offline press reports to confirm that the tweets in this category involve firms' offline campaigns and that these offline campaigns include the Twitter hashtags as well.

[7]For the category "#offline-online campaigns" we confirm that predicted tweets' hashtags appear in the news.

References

Abrahamson, E., and Rosenkopf, L. 1993. "Institutional and Competitive Bandwagons: Using Mathematical Modeling as a Tool to Explore Innovation Diffusion," *The Academy of Management Review* (18:3), pp. 487–517. (https://doi.org/10.2307/258906).

Adweek.com. 2015. 91% of Retail Brands Use Two or More Social Media Channels. Retrieved September 14, 2017, https://www.adweek.com/digital/yesmail-retail-brands-social-media-channels/

Aizawa, A. 2003. "An Information-Theoretic Perspective of Tf–Idf Measures," *Information Processing & Management* (39:1), pp. 45–65. (https://doi.org/10.1016/S0306-4573(02)00021-3).

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. 2017. "The Fall of the Labor Share and the Rise of Superstar Firms," Working Paper No. 23396, Working Paper, National Bureau of Economic Research, May. (https://doi.org/10.3386/w23396).

Badir, Y., Charoenngam, C., Siengthai, S., and Teeratansirikool, L. 2013. "Competitive Strategies and Firm Performance: The Mediating Role of Performance Measurement," *International Journal of Productivity and Performance Management* (62:2), pp. 168–184. (https://doi.org/10.1108/17410401311295722).

Bala, H., and Venkatesh, V. 2007. "Assimilation of Interorganizational Business Process Standards," *Information Systems Research* (18:3), pp. 340–362. (https://doi.org/10.1287/isre.1070.0134).

Bapna, S., Benner, M., and Qiu, L. 2019. "Nurturing Online Communities: An Empirical Investigation," *Management Information Systems Quarterly* (43:2), pp. 425–452.

Barney, J. 1991. "Firm Resources and Sustained Competitive Advantage," *Journal of Management* (17:1), pp. 99–120. (https://doi.org/10.1177/014920639101700108).

Bergen, M., and Peteraf, M. A. 2002. "Competitor Identification and Competitor Analysis: A Broad-Based Managerial Approach," *Managerial and Decision Economics* (23:4–5), pp. 157–169. (https://doi.org/10.1002/mde.1059).

Berrett, Ti., and Slack, T. 1999. "An Analysis of the Influence of Competitive and Institutional Pressures on Corporate Sponsorship Decisions," *Journal of Sport Management* (13:2), p. 114. (https://doi.org/10.1123/jsm.13.2.114).

Brouthers, L. E., O'Donnell, E., and Hadjimarcou, J. 2005. "Generic Product Strategies for Emerging Market Exports into Triad Nation Markets: A Mimetic Isomorphism Approach," *Journal of Management Studies* (42:1), pp. 225–245. (https://doi.org/10.1111/j.1467-6486.2005.00495.x).

Bruns, A., Woodford, D., and Sadkowsky, T. 2014. "Towards a Methodology for Examining Twitter Follower Accession," *First Monday* (19:4). (http://firstmonday.org/ojs/index.php/fm/article/view/5211).

Clark, B. H., and Montgomery, D. B. 1999. "Managerial Identification of Competitors," *Journal of Marketing* (63:3), pp. 67–83. (https://doi.org/10.2307/1251776).

Culnan, M. J., McHugh, P. J., and Zubillaga, J. I. 2010. "How Large U.s. Companies Can Use Twitter and Other Social Media to Gain Business Value," *MIS Quarterly Executive* (9:4), pp. 243–259.

DiMaggio, P. J., and Powell, W. W. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields," *American Sociological Review* (48:2), pp. 147–160. (https://doi.org/10.2307/2095101).

Einwiller, S. A., and Steilen, S. 2015. "Handling Complaints on Social Network Sites – An Analysis of Complaints and Complaint Responses on Facebook and Twitter Pages of Large US Companies," *Public Relations Review* (41:2), Digital Publics, pp. 195–204. (https://doi.org/10.1016/j.pubrev.2014.11.012).

Farndale, E., and Paauwe, J. 2007. "Uncovering Competitive and Institutional Drivers of HRM Practices in Multinational Corporations," *Human Resource Management Journal* (17:4), pp. 355–375. (https://doi.org/10.1111/j.1748-8583.2007.00050.x).

Gibson, J. J. 1986. *The Ecological Approach to Visual Perception*, Lawrence Erlbaum Associates.

Gur, F. A., and Greckhamer, T. 2018. "Know Thy Enemy: A Review and Agenda for Research on Competitor Identification," *Journal of Management*, p. 0149206317744250. (https://doi.org/10.1177/0149206317744250).

Hawley, A. H. 1968. "Human Ecology//International Encyclopedia of Social Sciences. L," *NY: Mancmillan*, pp. 328–337.

He, S., Rui, H., and Whinston, A. B. 2017. "Social Media Strategies in Product-Harm Crises," *Information Systems Research*. (https://doi.org/10.1287/isre.2017.0707).

Helms, R. W., Booij, E., and Spruit, M. R. 2012. "Reaching Out: Involving Users In Innovation Tasks Through Social Media," *ECIS 2012 Proceedings*. (https://aisel.aisnet.org/ecis2012/193).

Hoberg, G., and Phillips, G. 2016. "Text-Based Network Industries and Endogenous Product Differentiation," *Journal of Political Economy* (124:5), pp. 1423–1465. (https://doi.org/10.1086/688176).

Hsu, C., Lee, J.-N., and Straub, D. W. 2012. "Institutional Influences on Information Systems Security Innovations," *Information Systems Research* (23:3-part-2), pp. 918–939. (https://doi.org/10.1287/isre.1110.0393).

Huang, A. 2008. "Similarity Measures for Text Document Clustering," in *Proceedings of the Sixth New Zealand Computer Science Research Student Conference (NZCSRSC2008), Christchurch, New Zealand*, pp. 49–56.

Karahanna, E., Xin Xu, S., Xu, Y., and Zhang, N. (Andy). 2018. "The Needs–Affordances–Features Perspective for the Use of Social Media," *MIS Quarterly* (42), pp. 737–756. (https://doi.org/10.25300/misq/2018/11492).

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., and Silvestre, B. S. 2011. "Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media," *Business Horizons* (54:3), SPECIAL ISSUE: SOCIAL MEDIA, pp. 241–251. (https://doi.org/10.1016/j.bushor.2011.01.005).

Kim, E., Sung, Y., and Kang, H. 2014. "Brand Followers' Retweeting Behavior on Twitter: How Brand Relationships Influence Brand Electronic Word-of-Mouth," *Computers in Human Behavior* (37), pp. 18–25. (https://doi.org/10.1016/j.chb.2014.04.020).

Lee, D., Hosanagar, K., and Nair, H. S. 2018. "Advertising Content and Consumer Engagement on Social Media: Evidence from Facebook," *Management Science* (64:11), pp. 5105–5131. (https://doi.org/10.1287/mnsc.2017.2902).

Lonelybrand.com. 2017. The Gap's Styld.by Campaign Uses Bloggers to Build Connections. Retrieved September 14, 2017, https://lonelybrand.com/blog/the-gaps-styld-by-campaign-uses-bloggers-to-build-connections/

Luca, M. 2015. "Chapter 12 - User-Generated Content and Social Media," in *Handbook of Media Economics* (Vol. 1), Handbook of Media Economics, S. P. Anderson, J. Waldfogel, and D. Strömberg (eds.), North-Holland, pp. 563–592. (https://doi.org/10.1016/B978-0-444-63685-0.00012-7).

Majchrzak, A., Faraj, S., Kane, G. C., and Azad, B. 2013. "The Contradictory Influence of Social Media Affordances on Online Communal Knowledge Sharing," *Journal of Computer-Mediated Communication* (19:1), pp. 38–55. (https://doi.org/10.1111/jcc4.12030).

Majchrzak, A., and Markus, M. L. 2012. "Technology Affordances and Constraints in Management Information Systems (MIS)," SSRN Scholarly Paper No. ID 2192196, SSRN Scholarly Paper, Rochester, NY: Social Science Research Network. (https://papers.ssrn.com/abstract=2192196).

Mandviwalla, M., and Watson, R. 2014. "Generating Capital from Social Media," *MIS Quarterly Executive* (13:2), pp. 97–113.

Marketingland.com. 2016. 8 companies doing social media right and what marketers can learn from them. Retrieved September 14, 2017, https://marketingland.com/8-companies-social-media-right-marketers-can-learn-198228

Mashable.com. 2018. Airlines try but most can't respond to all your frustrated tweets. Retrieved February 10, 2018, https://mashable.com/2018/01/10/airlines-social-media-customer-service/#kLZRM57hHZq3

Meyer, J. W., and Rowan, B. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony," *American Journal of Sociology* (83:2), pp. 340–363.

Pan, Y., Huang, P., and Gopal, A. 2015. "Enemies (Almost) at the Gate: New Entry Threats and R&D Investments in the U.S. IT Industry," SSRN Scholarly Paper No. ID 2607674, SSRN Scholarly Paper, Rochester, NY: Social Science Research Network, May 18. (https://papers.ssrn.com/abstract=2607674).

Pant, G., and Sheng, O. R. L. 2015. "Web Footprints of Firms: Using Online Isomorphism for Competitor Identification," *Information Systems Research* (26:1), pp. 188–209. (https://doi.org/10.1287/isre.2014.0563).

Pozzi, G., Pigni, F., and Vitari, C. 2014. "Affordance Theory in the IS Discipline: A Review and Synthesis of the Literature," in *AMCIS*.

Seebach, C., Beck, R., and Denisova, O. 2013. "Analyzing Social Media for Corporate Reputation Management: How Firms Can Improve Business Agility," *International Journal of Business Intelligence Research (IJBIR)* (4:3), pp. 50–66. (https://doi.org/10.4018/ijbir.2013070104).

Smith, A. N., Fischer, E., and Yongjian, C. 2012. "How Does Brand-Related User-Generated Content Differ across YouTube, Facebook, and Twitter?," *Journal of Interactive Marketing* (26:2), pp. 102–113. (https://doi.org/10.1016/j.intmar.2012.01.002).

Socialmediatoday.com. 2016. Twitter's Algorithm Timeline Switches to Opt-Out, Coming to All Users. Retrieved May 30, 2019, https://www.socialmediatoday.com/social-networks/twitters-algorithm-timeline-switches-opt-out-coming-all-users

Sproutsocial.com. 2018. How to Use Hashtag Marketing to Dramatically Boost Brand Awareness. Retrieved September 14, 2018, https://sproutsocial.com/insights/hashtag-marketing

Strong, D., Volkoff, O., Johnson, S., Pelletier, L., Tulu, B., Bar-On, I., Trudel, J., and Garber, L. 2014. "A Theory of Organization-EHR Affordance Actualization," *Journal of the Association for Information Systems* (15:2). (https://aisel.aisnet.org/jais/vol15/iss2/2).

Susarla, A., Oh, J.-H., and Tan, Y. 2011. "Social Networks and the Diffusion of User-Generated Content: Evidence from YouTube," *Information Systems Research* (23:1), pp. 23–41. (https://doi.org/10.1287/isre.1100.0339).

Swani, K., Brown, B. P., and Milne, G. R. 2014. "Should Tweets Differ for B2B and B2C? An Analysis of Fortune 500 Companies' Twitter Communications," *Industrial Marketing Management* (43:5), pp. 873–881. (https://doi.org/10.1016/j.indmarman.2014.04.012).

Swani, K., Milne, G. R., Cromer, C., and Brown, B. P. 2013. "Fortune 500 Companies' Use of Twitter Communications: A Comparison between Product and Service Tweets," *International Journal of Integrated Marketing Communications* (5:2), pp. 47–56.

Themanifest.com. 2018. How to Create a Hashtag That Gets Your Brand Noticed. Retrieved November 14, 2018, https://themanifest.com/social-media/how-create-hashtag-gets-your-brand-noticed

Treem, J. W., and Leonardi, P. M. 2012. "Social Media Use in Organizations: Exploring the Affordances of Visibility, Editability, Persistence, and Association," SSRN Scholarly Paper No. ID 2129853, SSRN Scholarly Paper, Rochester, NY: Social Science Research Network. (https://papers.ssrn.com/abstract=2129853).

Umassd.edu. 2017. The 2017 Fortune 500 Go Visual and Increase Use of Instagram, Snapchat, and YouTube. Retrieved September 14, 2017, https://www.umassd.edu/cmr/socialmediaresearch/2017fortune500/#d.en.963986

Windmeijer, F. 2005. "A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators," *Journal of Econometrics* (126:1), pp. 25–51.

Wu, Z., and Salomon, R. 2016. "Does Imitation Reduce the Liability of Foreignness? Linking Distance, Isomorphism, and Performance," *Strategic Management Journal* (37:12), pp. 2441–2462. (https://doi.org/10.1002/smj.2462).

Yoo, Y., Boland, R. J., Lyytinen, K., and Majchrzak, A. 2012. "Organizing for Innovation in the Digitized World," *Organization Science* (23:5), pp. 1398–1408. (https://doi.org/10.1287/orsc.1120.0771).

Zeng, X., and Wei, L. 2012. "Social Ties and User Content Generation: Evidence from Flickr," *Information Systems Research* (24:1), pp. 71–87. (https://doi.org/10.1287/isre.1120.0464).

Zheng Zhou, K. 2006. "Innovation, Imitation, and New Product Performance: The Case of China," *Industrial Marketing Management* (35:3), New Product Development in Asia, pp. 394–402. (https://doi.org/10.1016/j.indmarman.2005.10.006).