



Twitter-patter: how social media drives foot traffic to retail stores

Thomas J. Weinandy¹ · Kuanchin Chen² · Susan Pozo¹ · Michael J. Ryan¹

Revised: 7 September 2022 / Accepted: 12 December 2022
© The Author(s), under exclusive licence to Springer Nature Limited 2023

Abstract

This paper answers how changes in social media activity influence customers to visit nationally known, brick-and-mortar retail stores. We consider seven measures of social media activity within a Social Impact Theory framework and test under what context does online chatter about a brand lead to higher foot traffic to those brand stores. We use hierarchical linear regression to account for the random effects of brand and store heterogeneity, which is superior to ordinary linear regression. Despite wide variation, when brand mentions increase by one standard deviation—either in likes or disagreement—then next-day foot traffic to stores of that brand will increase by 0.04 standard deviations (3–4%). This modest but meaningful effect, however, fully dissipates within 1 week. The weak cross-brand effects show that social media has distinct and larger influence on brands individually rather than universally. Our approach is novel due to (1) the large scale of data, (2) the breadth of analysis, (3) the multi-level specification, and (4) in estimating global elasticities between changes in electronic word-of-mouth (WoM) communication about brands and changes in store visits of those brands.

Keywords WoM · Multilevel analysis · Digital economics · Social impact theory · Social media · Retail

Introduction

In the summer 2019, a Twitter spat between the fast-food giants Popeyes and Chick-fil-A went viral. The so-called “Chicken Sandwich War” led to an unexpected surge in demand that caused Popeyes across the United States to run out of their newly released chicken sandwich (Sudath 2019). The heightened social media activity provided information about a product launch and increased customer engagement around its brand. This story illustrates how online word-of-mouth (WoM) communication can drive traffic to offline stores.

With sales at brick-and-mortar stores representing 85.2% of U.S. retail (*U.S. Census Bureau* 2021), it is crucial to understand the ways online activity and communication affect physical retailers. Additionally, the average internet user spends 147 min on social media a day (Statistica 2022). Current research on retail and social media is predominantly

focused on e-commerce data, and those insights do not necessarily apply to brick-and-mortar stores. The limited literature that does measure social media’s impact on physical retail only evaluates a single product, company, or product category without estimating global elasticities. Our article addresses this research gap while providing a means for modeling hierarchical data, such as products in stores or stores in countries.

Our research question is twofold: what measures of social media activity lead to changes in retail foot traffic and what are the precise magnitudes of those effects for comparable stores? The purpose of the study is to address this gap in the literature by estimating elasticities of online WoM communication (i.e., sentiment, disagreement, subjectivity, popularity, likes, followers, and recency) and offline store visits. We consider these seven metrics within the framework of the Social Impact Theory (Latané 1981) and test which can dampen or amplify the effect of a brand mention on consumer behavior. This paper is meant to inform branded retailers how to use publicly available social media data to anticipate changes to near-future demand.

Our research makes three primary contributions to measure the impact of social media activity on consumer visits to stores of nationally known retail brands. The first primary contribution is to connect brand-specific social media

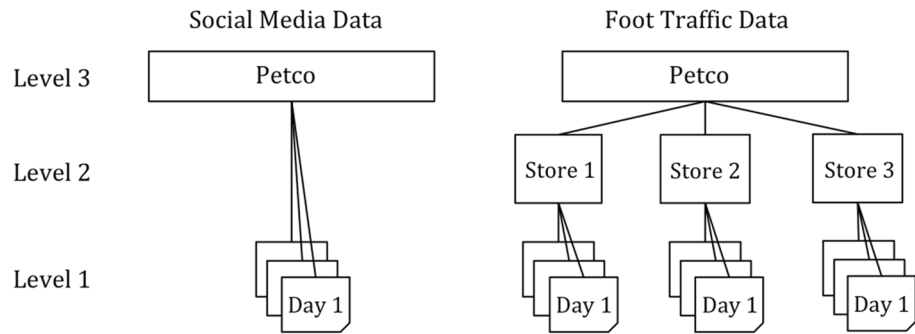
✉ Thomas J. Weinandy
thomas.j.weinandy@wmich.edu

¹ Department of Economics, Western Michigan University,
1903 W Michigan Ave, Kalamazoo, MI 49008, USA

² Department of Business Information Systems, Western
Michigan University, Kalamazoo, MI, USA



Fig. 1 Hierarchical data structure for a single brand



activity to offline purchases for those brands. There is a wide literature on how social media influences e-commerce sales (Kim et al. 2019) and financial products (Antweiler and Frank 2004; Bollen et al. 2011), but these disregard where most retail sales occur.¹ Consider You et al. (2015) who performed the most extensive meta-analysis for estimating sales elasticities from electronic WoM communication. Of the fifty articles they reviewed, only two exclusively measured how online communication affected goods sold in person. Our research is meant to expand the understanding of how online chatter spills over into the physical world.

The second contribution of this paper is to assess how social media activity affects a broad array of retailers. When online activity is connected to in-person sales, it is often for a narrow type of location, such as movie theaters (Kim et al. 2019; Liu et al. 2016) or restaurants (Cheung and Thadani 2012; Luca 2016). Other research only relates electronic WoM communication to in-person sales for a solitary product type (Deloitte 2013; Sanchez et al. 2020; Zhang et al. 2012) or a single retail brand (Pauwels et al. 2016). This poses a problem in the literature by relying on results too narrow in scope and subject to the consumption patterns of an individual good, product category, or retailer. Our paper advances this research by leveraging a large dataset and considering a wider array of retail store brands. On the social media side, we collect 2.7 million tweets mentioning one of fifteen brands. Those brands span store types from general/grocery, fast food, and general merchandise that include 3870 unique locations across the United States. Our approach is less sensitive to possible variability associated with individual brands and is more generalizable to out-of-sample retailers wanting to quantify how social media activity affects store foot traffic. Additionally, we use actual foot traffic data taken from anonymized cell phone users. This is an improvement on current research that only measure purchase intention

(Mainardes and Cardoso 2019) or participants aware they are part of a study (Godes and Mayzlin 2009).

The third primary contribution is our mixed-level specification. This is critical to statistically control for the distinct relationship each store and each brand has with social media. This allows us to estimate how social media for a brand-level variable (i.e., daily measures of social media activity) can impact a store-level variable (i.e., daily customer visits), shown in the example below (Fig. 1).

This specification with hierarchical regression provides foot traffic estimates for each store and avoids the ecological fallacy of interpreting aggregated data at the individual level (Hox et al. 2010). It also avoids assumptions of homogeneity across stores (Snijders and Bosker 2011), which is especially important given the diversity of brands in our analysis.

We find that a one standard deviation increase in either the per-tweet popularity or disagreement about a brand on Twitter leads to a 0.04 standard deviation increase (3–4%) in next-day foot traffic to stores of that brand. The results are slightly stronger when measuring out the next 3 days but weaker when using less-common measures of social media activity, such as on a per-like or per-follower basis. Although this indicates social media globally plays a small role in total retail visits, the results are statistically significant and economically meaningful when extrapolated across all stores of a national brand. We qualify these results as sensitive to different measures of social media activity and only representing an average store. Our results also show that social media activity only has a statistically significant relationship with foot traffic for a few days, and any effect fully dissipates within a week. Additionally, the impacts are much stronger for individual brands and stores, indicating wide heterogeneity.

Literature

Social impact theory

Early research into social communication was pioneered by Hovland (1948) where he defined the aspects of

¹ According to the National Retail Federation (as of July 6, 2021), Amazon is only the second-largest retailer by revenue in the U.S., while the other nine of the top ten retail companies operate primarily from sales through their brick-and-mortar stores.



communication necessary to make it “social.” Hovland defined the four components of social communication (which readily maps to our research) as a communicator (sender) who transmits a stimulus (tweet) to a communicatee (user reading a tweet) who responds to that communication (considers visiting a store). Latané (1981) elaborated by defining mediators of how social connections impact our feelings, thoughts, or behaviors. His Social Impact Theory identifies three such mediators as the source *strength* (e.g., salience, power, potency), *number of sources*, and source *immediacy* (closeness in time or space). An increase in any of those dimensions will make the receiver of a message more likely to perform an action. Latané then defines an associated Social Forces Law whereby the three mediators have a multiplicative relationship. This means that the impact of a message is not only determined by the strength, number of sources, and immediacy of the message source but by the interaction of all three. Our research does not measure the direct interaction of these terms—instead, we estimate the effect of isolated changes of the variables. This is like Argo et al. (2005) but expanded on different subsets of data. The cross-tabulation approach is ideal in scenarios with insufficient data or methods too computationally demanding to estimate interaction terms.

More recent researchers have applied Social Impact Theory (SIT) to modern retail settings. Kwahk and Ge (2012) used SIT to find how social media impacts consumers’ visit intention and purchase intention for e-commerce. Chenga and Linb (2020) expanded on SIT to find that a consumer’s perceived confidence in an online message increases purchase intention on social commerce. Finally, Naeem and Ozuem (2021) apply SIT in how social media impacts fashion retail and found that source strength and immediacy lead to greater brand engagement.

We first add to the literature by offering a newer, more economically meaningful application of Social Impact Theory to actual foot traffic across a variety of retail brands. Second, we expand on SIT by proposing and testing novel definitions of source strength and number of sources relevant to different online platforms. Our approach can be used by others looking to apply SIT to different social media contexts by leveraging data on message content and impressions.

Word-of-mouth communication

Social networks inform and influence how we live and what we buy (Jun and Park 2016). Online consumers prefer peer reviews over corporate messaging, as the former better focuses on and evaluates usage situations for a good from the perspective of the user (Kim and Kim 2018). This “word-of-mouth” communication—or simply “WoM,” the traditionally oral communication between two, non-commercial individuals about a product or service—is a key form of

social learning used in many purchase decisions. One of the earliest large-scale WoM examples comes from Antweiler and Frank (2004) who used 1.5 million posts on finance message boards to find that messages help predict stock market volatility. The subsequent research is best summarized by You et al. (2015) in a meta-analysis that considers hundreds of volume and sentiment WoM elasticities of product sales from fifty-one studies. They estimate a 0.236 volume elasticity and 0.417 sentiment elasticity (p. 19); however, they also emphasize the heterogeneity of results, and the way context matters. For example, they find both elasticities to be higher for goods that are privately consumed, have low-trialability, and are from lower-competitive industries. The researchers also find many examples where either the volume or sentiment elasticity is statistically significant but not the other.

Social media and retail

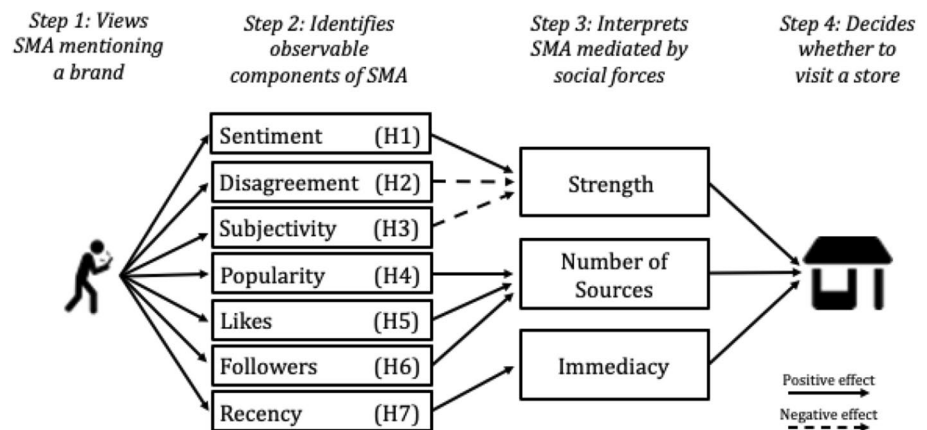
Since its launch in 2006, Twitter has become one of the most popular social media sites worldwide with 500 million tweets sent per day (Stricker 2014). In fact, 22% of U.S. adults are on Twitter and 9% of U.S. adults use the platform daily (Pew Research Center 2020). These users broadcast short messages (“tweets”) out to their subscribers (“followers”). Followers may “like” or “retweet” a post which broadcasts the message to their own followers.

Twitter is an ideal source of WoM data for its written and public communication format. Other top U.S. social media sites are less accessible due to their privacy walls (Facebook, Snapchat, WhatsApp), visual content emphasis (Instagram, Pinterest, YouTube), lower popularity (Nextdoor, Reddit), or professionalism (LinkedIn). Alternatively, Twitter is an “information intermediary” allowing users to create new information, compile existing information, and disseminate it to their followers. This includes daily chatter, conversations, sharing information, and reporting news (Java et al. 2009). The platform is also important for retail since, relative to other social networks, users are more likely to keep the brand central to a message and less likely to include self-promotion (Smith et al. 2012). Due to its culture of public sharing, Twitter makes our own research more replicable to academics and practitioners.

Microblogging sites like Twitter also reflect the non-virtual world, predicting changes in stocks (Bollen et al. 2011) and can even be an early detector of earthquakes (Sakaki et al. 2010). Current research, however, is limited to social media impacting e-commerce sales (Kim et al. 2019) or even just a consumer’s purchase intention (Mainardes and Cardoso 2019). If an article does address in-person sales, it is often an individual product, such as soft drinks (Sanchez et al. 2020), video games (Deloitte 2013), or movies (Kim et al. 2019; Liu et al. 2016). There is no research on whether social media drives foot traffic to multiple retail brands;



Fig. 2 User's path from social media activity (SMA) to retail visit



however, adjacent work by Kalyanam et al. (2018) found higher Google Ads spending leads to higher sales volumes in various brick-and-mortar stores.

Theory

Models of changed behavior

This paper considers whether online behavior on Twitter impacts offline foot traffic to brick-and-mortar retailers. We apply the Social Impact Theory, discussed in “[Social impact theory](#)” section, to a four-step process that describes the journey a customer goes from viewing social media activity mentioning a brand to the decision of whether to visit a brand store. The diagram, shown below, applies to physical or virtual stores as well as user-generated or sponsored content. It is most like the work of Cheung and Thadani (2012, p. 464) who construct an integrative framework of the impact of electronic WoM communication on purchases. We provide added value by offering a more parsimonious representation based on the Social Impact Theory.

The framework considers a single user on social media seeing WoM communication around a specific brand over a defined time interval. They observe seven components of that social media activity where the first three relate to the message content. *Sentiment* is the average tone of posts and represents how positive, negative, or neutral the text in the messages is. *Disagreement* is the distribution of sentiment, which also represents the diversity of opinions around a brand. *Subjectivity* is the average degree to which text in posts mentioning a brand use opinion-based language instead of fact-based. Together these three form the source *strength* of discussing a brand.

The next components of social media activity reflect the *number of sources* speaking about a given brand. This is visible to the user in three ways: *popularity* as the volume

of messages about that brand, *likes* as the number of likes a message receives, and *followers* as the number of followers the user has who sent a branded message.

Finally, there is *recency*, which is how soon social media activity is from the decision to visit a retail store. Recency leads singularly to the *immediacy* of social communication. These seven components of a social media activity then form the three dimensions of the Social Impact Theory to predict the directional impact on retail visits. Latané (1981) conjectures that an increase in either source strength, number of sources, or immediacy will increase the likelihood communication leads to an action.

Hypotheses

The arrows from Step 2 to Step 3 in Fig. 2 show how social media activity (SMA) mediates an individual’s perception of a brand, making them more or less likely to visit a particular retailer. We use these channels of influence to hypothesize the expected directional relationship between a component of SMA and a use’s visit decision (Fig. 2).

H1 Sentiment of brand tweets positively impacts store visits.

When brand opinions on social media improve, *sentiment* increases, which is expected to increase store traffic. Positive changes in online sentiment increases video game sales (Deloitte 2013), box office sales, (Liu et al. 2016), and restaurant patronage (Luca 2016). We expect a more positive online discussion of a brand increases its source *strength*, making a consumer more likely to visit a store as peers speak highly of that brand.

H2 Disagreement of brand tweets negatively impacts foot traffic.

Recall that *disagreement* is the deviation in sentiment, which can come from either (1) a lack of consensus around



brand opinions or (2) from a diversity of posts instead of a single, viral message (Chen et al. 2011). When there is a low disagreement (i.e., consensus) online around a brand, then users will update their perceptions and be more likely to visit that brand's stores because there is less uncertainty. This is shown in the model by a dashed arrow that connects *disagreement* and *strength*, indicating a negative relationship. Conversely, high disagreement also indicates many users are sharing diverse opinions, providing more information, and signaling greater engagement with the brand (Cui et al. 2012). Heightened conversation can lead to more awareness for the *behavior of others*, more *feature* awareness, and—accordingly—more visits. Therefore, we hypothesize the cumulative effect of disagreement to be ambiguous.

H3 Subjectivity of brand tweets negatively impacts store visits.

We predict that fact-based conversations about a brand are more likely to discuss features of that retailer (Loria et al. 2014). If true, then more objective messages offer greater information about a retailer and increased sales. This compares to Archak et al. (2011) who found customer reviews that mention specific product features have a higher impact on sales than the customer reviews alone. We, therefore, represent *subjectivity* with a dashed arrow to source *strength* because of its negative expected relationship with store visits.

H4 Popularity of a brand positively impacts store visits.

H5 Number of likes for brand tweets positively impacts store visits.

H6 Number of followers for a user tweeting about a brand positively impacts store visits.

Popularity, likes, and followers all connect to *number of sources* with a solid arrow, indicating a positive anticipated impact on foot traffic. They follow the same mechanism by allowing more users to see a branded message and by making a single user see more likely to see multiple branded messages. Higher visibility of a brand on Twitter leads to greater attention—brining it front of mind to more consumers (Kim et al. 2019). Additionally, an increase in posts about a brand will signal to a user through observational learning and peer influence the value of visiting a store (Deloitte 2013; Joshi and Musalem 2021). In short, these three hypotheses contend that hype matters.

We could also connect *popularity, likes, and followers* to the source *strength*, meaning that more social media activity not only increases views but also signals a higher quality

user, higher quality content, or merely a higher impact on other users who opted to follow them (Araujo et al. 2017; Ismagilova et al. 2020). However, such a specification would not change any hypotheses, as the directional impact on retail visits remains positive.

H7 Recency in days from brand tweets positively impacts store visits.

Social media activity is more likely to impact consumer behavior in the short run, as documented by Lovett et al. (2019) and You et al. (2015). Social media platforms tend to be ephemeral and microblogging sites like Twitter especially emphasize trending content over historical posts. We expect stronger short-run effects here as well and represent a positive connection between *recency* and *immediacy*.

Methodology

We test the above hypotheses on how social media activity about brands impacts store visits. Our process, shown below, follows a similar approach to Liu et al. (2016) (Fig. 3).

Information extraction

The above evaluation begins with raw Twitter data of mentioned brands and undergoes natural language processing to define the variables of interest. It also takes the number of tweets, the number of likes each tweet gets, and the number of followers for the user who sent each tweet to create a system of weights for the social media variables. The data

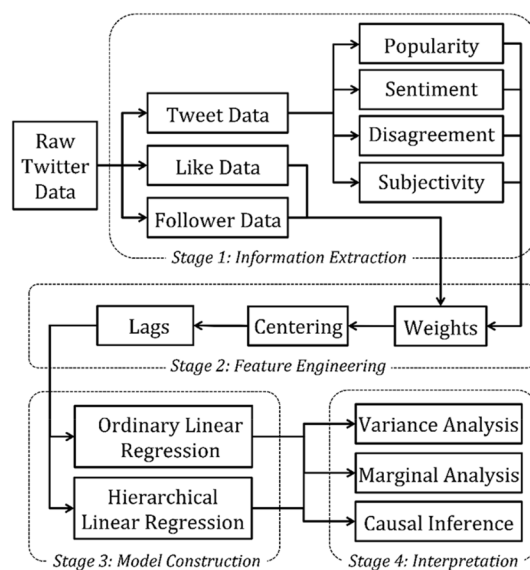


Fig. 3 System architecture



is then centered and represented as various lags and moving averages. Foot traffic data is centered but not lagged. We perform a baseline linear regression followed by hierarchical linear regression on different permutations of the data before evaluating model variance, marginal effects, and causal direction of those effects.

Feature engineering

We also consider separate ways to weigh social media activity. The primary purpose is to separately test if each measure of social media activity impacts foot traffic, as described in H4, H5, and H6. The secondary purpose of weights is to serve as a robustness check on whether our results are sensitive to a specific definition of social media activity. We define three distinct weights for our measures of social media activity. First, is our preferred model where we treat all tweets equally. The second representation is to consider the engagement a post receives by weighing it against accumulated likes. For example, a brand-day sentiment of 0.5 represents average sentiment per like received. Third, we consider each tweet's potential reach. Here we weigh measures according to the number of followers that could have seen each tweet.

The variables are then converted as the number of standard deviations from their own brand average on a particular day. This centering has the benefit of allowing for cross-brand comparisons by measuring social media activity in relative terms. Additionally, standardization improves interpretation in hierarchical linear models and facilitates convergence when estimating parameters by maximum likelihood (Hox et al. 2010, p. 63).

Next, we consider multiple temporal definitions of social media activity: same-day, previous-day, previous-3-days, and previous-7-days. To recap, social media activity is represented with one of three weights and as one of four temporal measures. We compare these twelve combinations to more widely view how Twitter impacts retail visits.

Baseline model

Consider the following model:

$$\widetilde{visits}_{ij} = \alpha + \beta' \tilde{X}_{\tau j} + \gamma date_t + \theta store_{ij} + \delta brand_j + R_{ij} \quad (1)$$

Observations occur at time t for store i of brand j . The dependent variable, \widetilde{visits}_{ij} , has a tilde representing it is standard deviations from a store's day-of-week average number of visits. Similarly, $\tilde{X}_{\tau j}$ is a vector of Twitter variables measured as standard deviations from each brand average, and β is their marginal effects on visits. The subscript τ is the different time lags for social media activity. R_{ij} is the error.

Since the dependent variable is centered around zero at the store level, we expect α , θ , and δ are zero.

We first test for a relationship between social media activity and foot traffic by estimating Eq. (1) with ordinary least squares (OLS) linear regression. One important assumption of OLS, with later implications, is the independent and identical distribution of the error term R_{ij} .

Full hierarchical linear model

Now we update the baseline model to a hierarchical specification including random effects, where U_j represents the brand-specific impact the social media variables have on visits, and V_{ij} represents the store-specific impact.

$$\widetilde{visits}_{ij} = \alpha + (\beta' + U_j' + V_{ij}') \tilde{X}_{\tau j} + \gamma date_t + \theta store_{ij} + \delta brand_j + R_{ij} \quad (2)$$

Note that fixed effects are represented as Greek letters and random effects are Latin. The random effects express store-specific and brand-specific ways social media activity influences foot traffic for each store. Fixed effects express the constant, expected impact social media activity has on any given store within the dataset. We are most interested in β , the global impact across all brands.

There are three primary benefits for estimating a hierarchical linear model (HLM). First, it addresses the correlation of error across stores by correcting for heterogeneous brand-level effects and improving estimates of standard errors (Hox et al. 2010, p. 3). Second, HLM guards against the ecological fallacy (interpreting aggregated data at the individual level) and atomistic fallacy (interpreting individual data on the aggregate) (ibid). Finally, HLM controls for within-group dependence of errors and within-group heteroskedasticity (De Leeuw et al. 2008, p. 14). Other variable-centered methods can also meet these criteria for grouped data, such as structural equation modeling. However, HLM performs better in use cases like this where (1) there are more than two levels, (2) groups are indistinguishable (e.g., Store 1 of Panera has no relation to Store 1 of Aldi), and there are many low-level observations (Huta 2014).

Data

Our first step was to identify fifteen retailers with nationwide brand awareness ranging in market, size, and target customer. These brands represent diverse subsectors—namely—grocery/general, fast food, and specialty merchandise. We intentionally selected brands that met specific technical requirements, had available foot traffic data, and provided reliable search results with minimal false positives (e.g., searching for “Staples” stores but getting results for



Table 1 Exclusion criteria

Reason	Examples
Cumbersome number of tweets	Taco Bell, Starbucks, Walmart
Large non-American markets	McDonald's, Ikea, KFC
Highly regional	Tim Hortons, Waffle House, Albertsons
Apostrophe within name	Trader Joe's, O'Reilly Auto Parts, Lowe's
Commonly abbreviated	Dairy Queen, Panda Express, 7-Eleven
Commonly misspelled	T.J. Maxx, SuperValu, J.C. Penney
Proper nouns	Target, Gap, Staples

“staples” office supplies) or false negatives (e.g., searching for “T.J. Maxx” but not getting results for “TJ Maxx” or “T.J. Max”). The below table describes more specific exclusion criteria as which brands we removed from consideration. Ultimately, we as researchers had to make an intuitive decision as to which, of the brands of those made available to us, were most appropriate and representative for this analysis (Table 1).

Some of our selection criteria was driven by technical reasons, such as the Twitter API replacing punctuation with whitespaces in queries. For example, Dunkin' was included because our query of “Dunkin” alone would return “Dunkin” (with the apostrophe), “Dunkin Donuts,” and “Dunkin' Donuts.” The same applies to “Baskin Robbins” which returns both “Baskin Robbins” and “Baskin-Robbins.” We excluded the brand Popeyes,” which would not have returned tweets mentioning Popeye's.” We do not claim our results capture all tweets mentioning a brand. For example, we miss mentions of the abbreviation “DD,” which would have turned up more false positives than true positives of tweets discussing the brand Dunkin'. Instead, we content that looking at primary brand names captures a sufficient share of the conversation happening on Twitter about that brand. Also see “[Social media data](#)” and “[Centering](#)” sections for additional standardizations of the data to better represent deviations from typical behavior.

We collected 394,998 unique observations over 110 days, from November 12, 2019, to February 29, 2020—avoiding panic buying in March 2020 from the coronavirus pandemic.² We then created our variables of interest from the raw data, as summarized in the table below and further described within this section (Table 2).

² Despite this also covering the holiday season, only six out of fifteen brands had above-average daily visit levels in the month of December. We also calculated average daily foot traffic at the brand level and found they never exceeded three standard deviations, even on the retail event of Black Friday or when stores closed for a holiday.

Foot traffic data

Foot traffic data is superior to traditional surveys since the customer reveals their preference by incurring the travel cost to visit a store in person. Foot traffic data, which comes from SafeGraph, has been used to evaluate diverse areas of research such as the financial impact of Starbucks's open bathroom policy (Gurun et al. 2020).³ The data only measures a subset of store visitors and likely does not represent a random sample; however, an internal analysis by SafeGraph asserts users are representative of the broader country according to multiple observable characteristics (Squire 2019).

We received store visit data from the fifteen brands defined in Table 3, which included 60,295 U.S. stores. 13.6% were dropped for missing data or presumably being closed (e.g., having fewer than two average visits a day). The brand with the most remaining stores is Dollar General with 14,518, almost double the next highest of Walgreens with 7667. Finally, we randomly select 258 stores per brand to create a balanced dataset.

We calculate descriptive statistics from the 258-store samples and refer to average store-day values by brand. Costco has the highest average daily foot traffic at over 213 user visits. Home Depot follows with 73 user visits. AutoZone has the lowest average numbers of visits at less than eight. This small value reminds us that SafeGraph data measures only a subset of total visits collected from mobile phones. All brands have at least one store-day where there were no visitors, shown by the minimum column with all zeroes. The highest maximum daily foot traffic is 1473 user visits at a Dunkin' in a convention center.

Social media data

Social media data comes from the Twitter API with search terms defined in Table 4. We collected 2.7 million tweets—each between 2 to 8 days old so sufficient time passes for “likes” to accumulate. We find more than 90% of all likes received occur within the first 2 days of a tweet being posted. There were minor technical issues, resulting in several days with missing social media data; however, this involved less than 0.73% of the brand-days under consideration. Only public, English-language tweets that mention one of the fifteen brands are considered. Replies and retweets are treated the same as original tweets, since they are visible to that user's followers.

Recall that popularity is the daily number of tweets mentioning a brand. For sentiment we score the text of each

³ In December 2022, searching “SafeGraph” (with quotes) in Google Scholar returned 547 results.



Table 2 Construction of study variables

Data sources	Foot Traffic: SafeGraph collected November 12, 2019 to February 29, 2020 Social Media: Twitter API collected November 12, 2019 to February 29, 2020
Twitter activity	Popularity: Daily count of brand mentions Sentiment: Daily average sentiment of brand mentions Disagreement: Daily standard deviation of sentiment Subjectivity: Daily average subjectivity of brand mentions
Weights	Per Tweet: Each tweet counts equally Per Like: Tweets weighted by number of likes received Per Follower: Tweets weighted by number of followers the sender has
Centering	Foot Traffic: Measured in standard deviations from a store's day-of-week average number of visits Twitter Activity: Measured in standard deviations from a brand's average Twitter activity (by measure)
Time lags	Same Day: Compares foot traffic with Twitter activity from the same day Prior Day: Compares foot traffic with Twitter activity from the previous day 3-day average: Compare foot traffic with average Twitter activity from the same day and previous 2 days 7-day average: Compare foot traffic with average Twitter activity from the same day and previous 6 days
Unique observations	Unique brands: 15 Unique observations: 394,998 Unique days: 110 Unique Tweets: 2,747,353 Unique stores: 3870

Table 3 Foot traffic descriptive statistics

Brand	Stores (full)	Stores (sample)	Days	Mean	Store-day foot traffic (sample)			
					SD	Min.	Median	Max.
Aldi	1784	258	110	21.23	14.52	0	19	138
AutoZone	5246	258	110	7.81	6.34	0	6	220
Baskin-Robbins	517	258	110	213.80	101.06	0	205	942
Costco	2279	258	110	9.53	10.92	0	7	181
Dollar general	14,518	258	110	12.83	8.94	0	11	177
Dunkin'	7522	258	110	13.56	23.40	0	10	1473
Hobby lobby	860	258	110	48.80	42.72	0	39	501
Home depot	1919	258	110	73.41	39.39	0	67	378
Nordstrom	258	258	110	31.35	34.24	0	23	638
Panera	1915	258	110	25.28	28.46	0	21	504
Petco	1365	258	110	18.76	16.31	0	15	312
Pizza Hut	5274	258	110	8.66	9.70	0	7	258
Qdoba	602	258	110	14.66	15.46	0	11	262
Walgreens	7667	258	110	21.36	18.82	0	18	513
Whole foods	394	258	110	31.21	26.23	0	26	217
Total	52,120	3870	110	36.82	61.10	0	16	1473

tweet using three algorithms and average to the brand-day level.⁴ The sentiment of tweets ranges from -1 (completely negative) to 1 (completely positive). Disagreement is the standard deviation of a brand's daily sentiment. Finally,

subjectivity is also scored with *SentimentAnalyzer* and ranges from 0 (completely objective) to 1 (completely subjective). Popularity is summed to the brand-day level and the remaining variables of interest are averaged to the brand-day level. See the descriptive statistics below Table 4.

The first column shows the exact query used when searching for tweets of each brand. Next are the three measures of popularity on Twitter. It is unsurprising that Costco is mentioned the most in tweets and receives the most likes since

⁴ Specifically, we use *SentimentAnalyzer* from the *nlk* package in Python (Loper and Bird 2002), *TextBlob* in Python (Loria et al. 2014), and *sentimentr* in R (Rinker 2017).



Table 4 Average brand-day Twitter variables

Query	Popularity		Sentiment			Disagreement			Subjectivity						
	Mentions on Tweets	Likes on mentioned Tweets	Followers on mentioned Tweets	Per Tweet	Per	Like	Per Follower	Per Tweet	Per	Like	Per Follower	Per Tweet	Per	Like	Per Follower
“Aldi”	2418.6	9211.3	21,198,463.4	0.097	0.091	0.130	0.251	0.243	0.235	0.365	0.388	0.376	0.388	0.376	0.376
“AutoZone”	215.6	1064.1	2,626,215.1	0.010	0.117	0.190	0.243	0.221	0.202	0.355	0.367	0.358	0.367	0.358	0.358
“Baskin Robbins”	360.5	1401.3	1,239,685.5	0.095	0.086	0.112	0.214	0.202	0.188	0.339	0.364	0.374	0.364	0.374	0.374
“Costco”	4093.5	16,801.6	26,405,508.8	0.105	0.101	0.127	0.246	0.229	0.222	0.372	0.368	0.357	0.368	0.357	0.357
“Dollar General”	1320.2	3358.1	5,438,288.8	0.131	0.086	0.090	0.192	0.187	0.188	0.492	0.474	0.452	0.474	0.452	0.452
“Dunkin”	2998.2	12,610.1	17,096,187.8	0.074	0.081	0.154	0.258	0.235	0.230	0.344	0.361	0.367	0.361	0.367	0.367
“Hobby Lobby”	496.0	1980.1	1,435,496.7	0.125	0.120	0.140	0.241	0.223	0.218	0.339	0.354	0.347	0.354	0.347	0.347
“Home Depot”	2539.5	10,487.3	20,364,827.6	0.085	0.106	0.089	0.246	0.247	0.221	0.321	0.356	0.348	0.356	0.348	0.348
“Nordstrom”	844.9	2414.9	59,682,009.4	0.125	0.134	0.193	0.262	0.241	0.240	0.366	0.381	0.448	0.381	0.448	0.448
“Panera”	1241.0	6565.4	9,147,382.9	0.091	0.073	0.107	0.259	0.236	0.224	0.340	0.354	0.341	0.354	0.341	0.341
“Petco”	1709.5	6691.5	6,109,668.7	0.180	0.134	0.199	0.287	0.250	0.252	0.443	0.394	0.401	0.394	0.401	0.401
“Pizza Hut”	2043.9	9085.2	11,737,806.9	0.063	0.063	0.124	0.237	0.222	0.221	0.325	0.341	0.346	0.341	0.346	0.346
“Qdoba”	258.8	1204.9	1,224,522.9	0.113	0.137	0.153	0.265	0.242	0.245	0.372	0.405	0.365	0.405	0.365	0.365
“Walgreens”	1824.1	7905.9	34,421,050.3	0.065	0.074	0.243	0.246	0.230	0.277	0.357	0.373	0.443	0.373	0.443	0.443
“Whole Foods”	3049.4	14,340.8	22,104,151.3	0.121	0.124	0.164	0.215	0.205	0.217	0.437	0.439	0.432	0.439	0.432	0.432
Overall mean	1692.9	7,006.6	15,610,090.0	0.105	0.102	0.147	0.244	0.227	0.225	0.372	0.382	0.383	0.382	0.383	0.383



this is also the brand with the highest foot traffic. AutoZone is the least popular in terms of tweets and likes but Qdoba is the least popular in terms of followers. Tweets that mention Nordstrom reach the most followers (59.7 million)—more than double that of Costco (26.4 million).

The next section of variables describes our three measures of sentiment by brand. Positivity is more common on Twitter than negativity, shown by each brand with an average positive daily sentiment. Petco has the highest sentiment across all three dimensions. Turning to the overall mean for all brands, sentiment per like is approximately the same as sentiment per tweet, indicating users on average do not “like” tweets more when they are more positive. Conversely, tweets per follower have higher sentiment, indicating more influential users are also more positive.

We next see that disagreement on all three measures is rather similar across brands, except for Dollar General which has the greatest consensus (i.e., lowest disagreement). Finally, there is subjectivity with more heterogeneity across brands. Dollar General again is an outlier as having tweets with their brand mentioned with the highest degree of subjectivity.

Centering

We next center and standardize the variables of interest, providing multiple benefits. First, centering allows for cross-brand comparisons for stores of assorted sizes and lends more credence to out-of-sample interpretations. Second, foot traffic data only represents a share of the true numbers of visitors to each store. Since that share is also unknown, it would be inappropriate to interpret the dependent variable in terms of the absolute number of store visits. Centering overcomes this problem. Third, we can control for predictable fluctuations by centering on the mean of each individual store’s day-of-week average. Fourth, this reduces the overall range of observations, and accordingly, reduces the impact of outliers.⁵ Social media data, however, does not exhibit similar weekly functions, so we center Twitter variables at the brand average.

Results

Linear regression results

We first use Ordinary Least Squares (OLS) regression to estimate Eq. (1). These results, presented below, show the

⁵ After centering and standardizing foot traffic data, the dependent variable now ranges from -3.42 to 3.69 standard deviations from a store’s day-of-week average with 95.76% of observations within ± 2 standard deviations.

baseline relationship between social media activity and retail foot traffic (Table 5).

In this specification, virtually every measure of social media activity has a statistically significant relationship with visits. The coefficients are the same sign across different definitions and time dimensions of social media activity. These strong results provide preliminary evidence that word-of-mouth communication online impacts foot traffic to brick-and-mortar stores. Social media activity from the previous day, shown in Columns (4) through (6), also affects store visits the following day, indicating a causal relationship.

A brand’s popularity on social media has consistently positive and statistically significant effects on store visits of that brand. For example, consider Column (1) that measures social media activity on the same day of store visits. The coefficient indicates a one standard deviation (SD) increase in the average number of tweets about a brand is associated with a 0.047 SD increase in the number of visits a store expected for that day of the week. Column (5) tells a similar story, where a one SD increase in average number of *likes* received on tweets that mention a brand is associated with a 0.022 SD increase in average foot traffic *the next day* to that brand’s stores.

Increases in sentiment and disagreement about a brand are also both associated with higher expected foot traffic. Conversely, a decrease in subjectivity has a negative relationship with store visits. These results are true whether measuring on a per-tweet, per-like, or per-follower basis and whether social media activity is from the current day or previous 1, 3, or 7 days.

Unfortunately, there is one critical limitation. OLS regression assumes observations are independently and identically distributed, and any violation will lead to biased estimates. Given the nested structure of the data, it is unreasonable to presume daily visits do not fluctuate in a way that is store-dependent or brand-dependent. In other words, OLS assumes there is no random error unique to a store or brand (Snijders and Bosker 2011, p. 46). Additionally, OLS estimates global coefficients, meaning the relationship between social media and foot traffic is assumed constant across stores and brands. The fixed effects, shown in Eq. (1), control for the level of visits by day, store, and brand, but they do not control for the way brands react differently to online communication. This heterogeneity is evident below.

Figure 4 shows scatter plots for select brands of daily store visits (as measured in standard deviations from each store’s day-of-week average) and daily popularity for a brand on social media (as measured same-day, per-tweet, and in standard deviations from the brand average) as well as the line of best fit for these two variables. The first store, Qdoba, shows a positive relationship between an increase in the number of tweets about Qdoba and an increase in foot traffic. Hobby Lobby also exhibits a positive relationship,



Table 5 OLS results

Twitter data	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)		(12)		
	Same day				Previous day				Previous 3 days				Previous 7 days												
	Per Tweet	Per Like	Per Follower	Per Tweet	Per Like	Per Follower	Per Tweet	Per Like	Per Follower	Per Tweet	Per Like	Per Follower	Per Tweet	Per Like	Per Follower	Per Tweet	Per Like	Per Follower	Per Tweet	Per Like	Per Follower	Per Tweet	Per Like	Per Follower	
Intercept	-0.396*** (0.015)	-0.375*** (0.014)	-0.394*** (0.014)	-0.367*** (0.014)	-0.379*** (0.014)	-0.379*** (0.014)	-0.376*** (0.014)	-0.376*** (0.014)	-0.372*** (0.014)	-0.385*** (0.014)	-0.378*** (0.014)	-0.384*** (0.014)													
Popularity	0.047*** (0.002)	0.021*** (0.002)	0.037*** (0.002)	0.035*** (0.002)	0.022*** (0.002)	0.003 [†] (0.002)	0.046*** (0.002)	0.038*** (0.002)	0.028*** (0.002)	0.052*** (0.002)	0.033*** (0.002)	0.028*** (0.002)													
Sentiment	0.008*** (0.002)	0.001 (0.002)	0.020*** (0.002)	0.010*** (0.002)	0.018*** (0.002)	0.024*** (0.002)	0.011*** (0.002)	0.013*** (0.002)	0.038*** (0.002)	0.005** (0.002)	0.009*** (0.002)	0.042*** (0.002)													
Disagreement	0.035*** (0.002)	0.010*** (0.002)	0.016*** (0.002)	0.032*** (0.002)	0.019*** (0.002)	0.008*** (0.002)	0.039*** (0.002)	0.024*** (0.002)	0.015*** (0.002)	0.043*** (0.002)	0.020*** (0.002)	0.019*** (0.002)													
Subjectivity	-0.020*** (0.002)	-0.018*** (0.002)	-0.011*** (0.002)	-0.006*** (0.002)	-0.004* (0.002)	-0.014*** (0.002)	-0.021*** (0.002)	-0.005** (0.002)	-0.022*** (0.002)	-0.024*** (0.002)	-0.003 [†] (0.002)	-0.034*** (0.002)													
Brand FE	True	True	True	True	True	True	True	True	True	True	True	True													
Store FE	True	True	True	True	True	True	True	True	True	True	True	True													
Date FE	True	True	True	True	True	True	True	True	True	True	True	True													
Adjusted R ²	0.1402	0.1390	0.1397	0.1394	0.1390	0.1389	0.1402	0.1396	0.1401	0.1409	0.1393	0.1405													
Observations	394,998	394,998	394,998	394,998	394,998	394,998	394,998	394,998	394,998	394,998	394,998	394,998													

The dependent variable is the number of standard deviations away from a store's day-of-week average visits. Standard errors are in parentheses. The associated significance is given as follows: ***0.001, **0.01, *0.05, [†]0.1



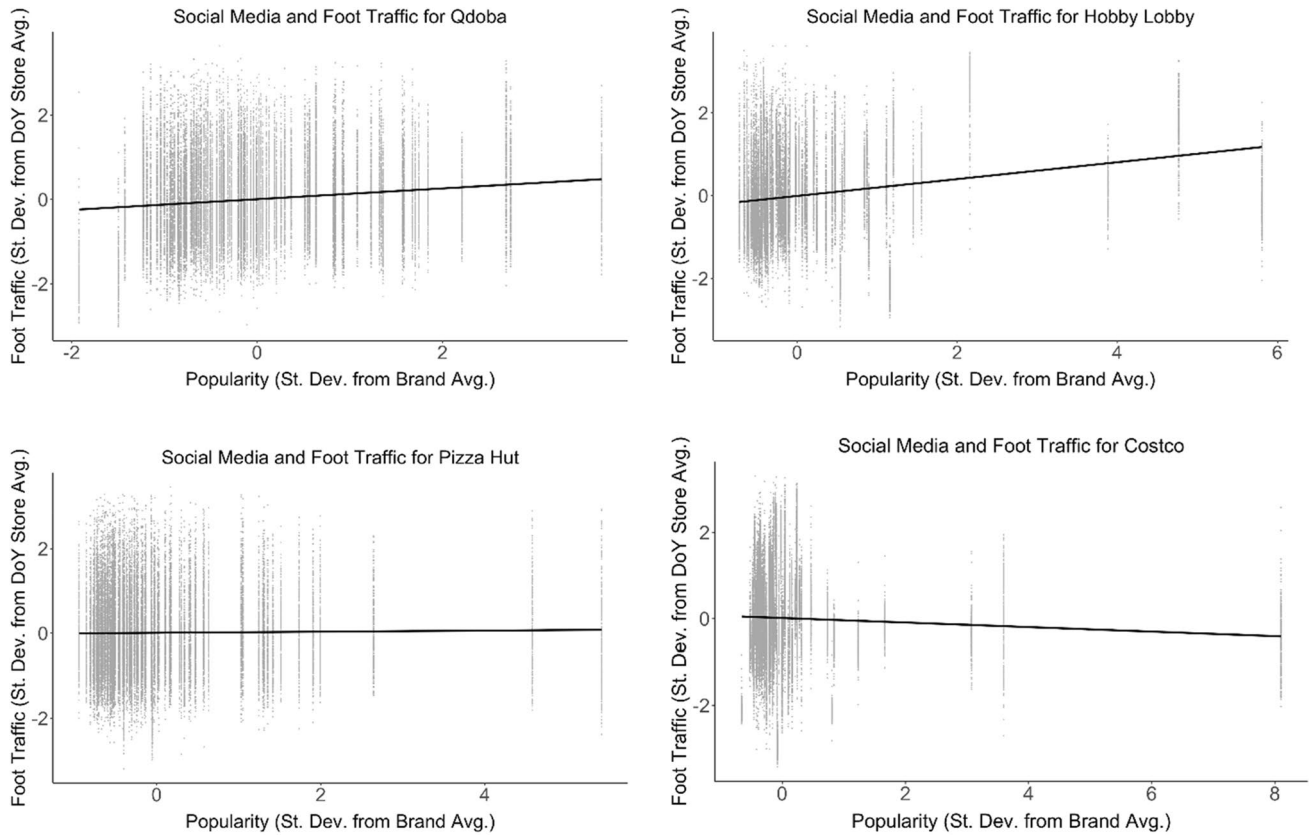


Fig. 4 Scatter plots of select brands

but the relationship appears stronger than that of Qdoba, shown by the steeper slope. There appears to be no relationship between popularity on social media and foot traffic to Pizza Hut, as seen by the flat line. Surprisingly, the scatter plot for Costco reveals a negative relationship visible in the downward sloping line. This is comparable to the heterogeneous impact of flyers on grocery store foot traffic observed in Gijbrecchts et al. (2003, p. 12). Overall, Fig. 4 provides compelling evidence of heterogeneity between brands, and that a random effects model can allow for the varying ways in which social media activity impacts foot traffic to stores and brands differently.

Hierarchical linear regression results

We use an HLM regression to estimate Eq. (3) below. This is equivalent to Eq. (2) without the fixed effects for stores and brands. Excluding those factors makes it easier for the model to converge and should not impact the estimation due to the centered dependent variable.

$$\widetilde{visits}_{ij} = \alpha + (\beta' + U'_j + V'_{ij})\tilde{X}_{rj} + \gamma date_t + R_{ij} \quad (3)$$

We estimate Eq. (3) using the *lme4* package in R and present results in Table 6.⁶

Columns (13) through (15) evaluate the same-day impact social media has on retail foot traffic to stores while Columns (16) through (18) use social media activity from the previous day. Furthermore, each column measures social media activity differently, either per tweet, like, or follower. We first draw attention to the Likelihood Ratio (LR) Tests that show with a 0.001 level of significance how all HLM models are better specified than OLS.

We now turn to the random effects component of the model. The variance is small (< 0.01) for all Twitter coefficients compared to a residual variance of 0.78. This indicates that our Twitter variables are only capturing a small portion of the total variability in foot traffic, which intuitively makes sense given how social media is far from the dominant reason consumers visit a store.

⁶ Since our dependent variable is derived from foot traffic data with a natural lower bound of zero, it is possible for our results to be biased in the positive direction. We evaluate the residuals in all twelve models from Eq. (3) and find a slight positive skew ranging from 0.388 and 0.412 with an average of 0.404. This is sufficiently small to not pose any serious concern for our model specification.



Table 6 Full HLM results for same day and prior day social media activity

	(13)	(14)	(15)	(16)	(17)	(18)
Social media data						
Time measure	Same day	Same day	Same day	Prior day	Prior day	Prior day
Unit measure	Per Tweet	Per Like	Per Follower	Per Tweet	Per Like	Per Follower
Random effect variance						
Brand-level						
Popularity	0.0067	0.0018	0.0055	0.0037	0.0019	0.0022
Sentiment	0.0051	0.0011	0.0018	0.0062	0.0011	0.0200
Disagreement	0.0047	0.0028	0.0015	0.0031	0.0017	0.0018
Subjectivity	0.0011	0.0015	0.0008	0.0013	0.0009	0.0023
Store-level						
Popularity	0.0005	0.0002	0.0000	0.0018	0.0005	0.0013
Sentiment	0.0045	0.0006	0.0016	0.0051	0.0014	0.0010
Disagreement	0.0022	0.0002	0.0012	0.0037	0.0001	0.0012
Subjectivity	0.0003	0.0000	0.0005	0.0005	0.0000	0.0003
Date-level						
Residual	0.7880	0.8012	0.7969	0.7888	0.8002	0.7977
Fixed effect coefficients						
Intercept	− 0.3981*** (0.0147)	− 0.3865*** (0.0147)	− 0.4007*** (0.0149)	− 0.3612*** (0.0145)	− 0.3678*** (0.0145)	− 0.3676*** (0.0147)
Popularity	0.0451 [†] (0.0212)	0.0190 (0.0110)	0.0384 [†] (0.0192)	0.0364* (0.0158)	0.0243* (0.0113)	0.0041 (0.0122)
Sentiment	− 0.0044 (0.0186)	0.0001 (0.0086)	0.0172 (0.0112)	− 0.0009 (0.0204)	0.0185 [†] (0.0089)	0.0255* (0.0117)
Disagreement	0.0364 [†] (0.0179)	0.0074 (0.0138)	0.0183 [†] (0.0101)	0.0370* (0.0145)	0.0194 [†] (0.0109)	0.0107 (0.0110)
Subjectivity	− 0.0112 (0.0089)	0.0151 (0.0100)	− 0.0078 (0.0074)	0.0057 (0.0096)	− 0.0016 (0.0081)	− 0.0115 (0.0126)
Date fixed effects	True	True	True	True	True	True
Model performance						
Log likelihood	− 515,396	− 517,377	− 516,676	− 515,931	− 517,337	− 517,011
AIC	1,030,337	1,034,298	1,032,894	1,031,406	1,034,215	1,033,566
Degrees of freedom	129	129	129	129	129	129
LR test <i>p</i> value	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***

The dependent variable is the number of standard deviations away from a store's day-of-week average number of visits. The standard errors for fixed effect coefficients are in parentheses. The Likelihood Ratio (LR) Test, if statistically significant, rejects an OLS specification for the given HLM model with random effects. The associated significance is given as follows: ***0.001, **0.01, *0.05, [†]0.1

We next consider the fixed effect component of the model. The same widespread statistical significance from OLS in Table 5 is absent. The popularity of a brand on social media is associated with a positive and statistically significant (at the 0.1 level) increase in foot traffic to those brands stores. For example, consider Column (13) that shows a 1 standard deviation (SD) increase in the popularity (measured here as the number of tweets sent that same day) will lead to

a 0.0451 SD increase in average store visits.⁷ This effect, however, is not found in Column (14) when social media activity is measured on a per-like basis. Disagreement has a comparable effect. When disagreement per tweet increases by 1 SD, on the same day, average store visits are 0.0364 SD higher. On the other hand, measures of sentiment and

⁷ To better understand these magnitudes, note from Table 3 the average SD number of visits across all stores is 92.9% of their mean. Though imperfect, it is useful to think of 1 SD as about a 1% change.



subjectivity on social media do not have a statistically significant impact on foot traffic in any of the three specifications.

It is important to consider how endogeneity enters Columns (13) through (15). The model framing implies social media drives foot traffic to retail stores, but it is also plausible that foot traffic leads social media users to post about the brands they recently visited. To address this possible reverse causality, we also consider how the previous day's social media activity impacts store visits the following day. As shown in Columns (16) through (18), lagged social media activity still has a statistically significant relationship with foot traffic. We rule out concerns of reverse causality here since we believe it implausible that customers visit a store because they anticipate tweeting about it the next day.

Two of the three social media measures show a 1 SD increase in popularity leads to a 0.0245 to 0.0364 SD increase in store visits and is statistically significant at the 0.05 level. Unlike in the same-day measures, we now see sentiment having a positive and statistically significant effect on foot traffic in two of the three specifications. Finally, we see disagreement having a similar positive and significant impact and subjectivity still having no discernible impact.

We now consider multi-day averages of social media to discern how longer time horizons of online activity effect changes in store visits, shown below. Columns (19) through (21) reveal how a 3-day average of social media activity has a weaker impact on foot traffic than the previous day alone. Popularity is qualitatively the same, with two of the three measures being positive and statistically significant, but now sentiment and disagreement are only statistically significant (and positive) in one of the three measures. Furthermore, subjectivity is never statistically different from zero. Turning to Columns (22) through (24), only one of the twelve combinations of measures is statistically significant at the 0.01 level or below within a 7-day average. This demonstrates social media's effects are short-lived to only a few days, which is consistent with other research (Lovett et al. 2019) (Table 7).

The results are also meaningful in how different they are from the previous WoM communication literature. For example, our popularity elasticities and sentiment elasticities are approximately one-tenth and one-twentieth, respectively, similar elasticities measured in the meta-analysis by You et al. (2015). This also follows Pauwels et al. (2016) who find WoM elasticities between online reviews and *in-person* sales to be half the elasticities between online reviews and *online* sales. This makes intuitive sense that the further away online activity is from the point of sale, the lesser the impact it will have. Our results provide a sobering reminder that not all online activity is created equal and that we should expect estimated elasticity values to change between communication platforms.

Evaluating hypotheses

We now test the seven hypotheses from “Hypotheses” section against the hierarchical linear model results (Table 8).

Sentiment is not statistically significant in the preferred measures of social media on a per-tweet basis. However, when measuring social media activity on a per-like or per-follower basis, positive emotions lead to similarly positive increases in foot traffic to those brands' stores. The overall effect of sentiment is limited but positive in all instances where the variable is statistically significant. For this reason, we conclude there is modest, yet ultimately inconclusive, evidence H1 is true.

Disagreement on social media, as measured by the standard deviation of sentiment, is also positive and statistically significant in some specifications. This result is contrary to H2, where we expected lower disagreement corresponds to message strength about a brand. The finding is like Cui et al. (2012) who show that divergence in sentiment on social media better predicts new product sales than sentiment itself. Subjectivity is not statistically different from zero in any model—a stark lack of evidence to support H3. In testing H1–H3, we find that source strength of the Social Impact Theory is meaningful for our research question, indicating that the effect is weakly tied to these measures and overall difficult to measure.

Next, popularity broadly has a positive impact on foot traffic, as speculated in H4. This follows Cui et al. (2012) who show popularity is more important than sentiment for experience products, which would include shopping experiences such as the retailers considered here. In testing H5, we find that the number of likes—at times—is associated with higher levels of retail visits, but the effects are less than the per-tweet impact from H4. Finally, there is virtual no effect on store visits when measuring social media activity on a per-follower basis. Thus, we fail to reject the null hypothesis in favor of H6. Overall, our results show a statistically significant effect across most, but not all, specifications of measuring the number of sources within the Social Impact Theory. We, thus, give a cautious interpretation that online popularity generally leads to more foot traffic; however, context can shape how online popularity matters.

In H7, we speculated that more recent tweets have a stronger effect on foot traffic than later tweets. We find a more nuanced result where social media has weak effects on the same-day decision to visit a store, stronger effects in the next 1 to 3 days, and virtually no effect by 7 days. This is a clear signal that the immediacy of a source, as stated in the Social Impact Theory, corresponds to higher foot traffic; however, it may take a day or two for an action to occur.



Table 7 Full HLM results for 3- and 7-day average social media activity

	(19)	(20)	(21)	(22)	(23)	(24)
Social media data						
Time measure	3-day avg.	3-day avg.	3-day avg.	7-day avg.	7-day avg.	7-day avg.
Unit measure	Per Tweet	Per Like	Per Follower	Per Tweet	Per Like	Per Follower
Random effect variance						
Brand-level						
Popularity	0.0059	0.0024	0.0054	0.0147	0.0031	0.0062
Sentiment	0.0082	0.0020	0.0030	0.0184	0.0054	0.0052
Disagreement	0.0048	0.0033	0.0025	0.0186	0.0036	0.0031
Subjectivity	0.0027	0.0017	0.0028	0.0109	0.0040	0.0069
Store-level						
Popularity	0.0000	0.0031	0.0035	0.0031	0.0090	0.0107
Sentiment	0.0093	0.0057	0.0027	0.0129	0.0120	0.0107
Disagreement	0.0077	0.0033	0.0065	0.0138	0.0107	0.0140
Subjectivity	0.0030	0.0007	0.0040	0.0071	0.0044	0.0108
Date-level						
Residual	0.7754	0.7871	0.7818	0.7546	0.7624	0.7492
Fixed effect coefficients						
Intercept	- 0.3611*** (0.0144)	- 0.3784*** (0.0145)	- 0.3691*** (0.0146)	- 0.3619*** (0.0143)	- 0.3884*** (0.0144)	- 0.3704*** (0.0145)
Popularity	0.0443* (0.0200)	0.0284* (0.0128)	0.0272 (0.0191)	0.0280 (0.0315)	0.0159 (0.0146)	0.0091 (0.0205)
Sentiment	- 0.0036 (0.0236)	0.0139 (0.0117)	0.0349* (0.0142)	0.0126 (0.0351)	0.0130 (0.0191)	0.0356 [†] (0.0188)
Disagreement	0.0452* (0.0181)	0.0181 (0.0149)	0.0157 (0.0131)	0.0408 (0.0354)	0.0086 (0.0158)	0.00732 (0.0147)
Subjectivity	- 0.0060 (0.0136)	- 0.0011 (0.0107)	- 0.0159 (0.0139)	- 0.0220 (0.0271)	- 0.0046 (0.0164)	- 0.030 (0.0216)
Date fixed effects	True	True	True	True	True	True
Model performance						
Log likelihood	- 513,896	- 515,893	- 515,158	- 510,248	- 512,586	- 510,248
AIC	1,027,337	1,031,329	1,029,860	1,020,035	1,024,713	1,020,036
Degrees of freedom	129	129	129	129	129	129
LR test <i>p</i> value	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***

The dependent variable is the number of standard deviations away from a store's day-of-week average number of visits. The standard errors for fixed effect coefficients are in parentheses. The Likelihood Ratio (LR) Test, if statistically significant, rejects an OLS specification for the given HLM model with random effects. The associated significance is given as follows: ***0.001, **0.01, *0.05, [†]0.1

Discussion

The overall results point to a modest but economically meaningful relationship between social media activity and retail foot traffic. In a world where customers have thousands of reasons to visit a store, these results show a sliver of motivation comes from word-of-mouth communication on social media. We qualify these outcomes according to the limitations of the study. The data comes from fifteen high-recognition brands with stores located around the United States.

Therefore, these results point to the impact social media activity has on established, well-known brands.

The results also demonstrate how the definition of social media activity matters. We recommend focusing on the per-tweet models, as that is the most widely used in the literature. The results are dampened when measuring on a per-like or per-follower basis; however, the values are often either directionally the same or virtually indistinguishable from zero. This broader consistency shows the results in the paper reflect a global impact that is present but modest. By comparing the weaker hierarchical results against the stronger OLS results, we conclude that the predominant



Table 8 Summary of hypotheses and results

Hypothesis	Result
H1 Sentiment of brand tweets <i>positively</i> impacts store visits	Brand sentiment has an inconclusive impact on foot traffic with no influence in the preferred model but a positive relationship in other models
H2 Disagreement of brand tweets <i>negatively</i> impacts store visits	More disagreement on social media leads to more foot traffic, indicating that greater disagreement represents higher brand engagement through diverse posts
H3 Subjectivity of brand tweets <i>negatively</i> impacts store visits	The subjectivity of social media posts has no measurable impact on retail foot traffic
H4 Popularity of a brand <i>positively</i> impacts store visits	More tweets about a brand leads to more store visits
H5 Number of likes for brand tweets <i>positively</i> impacts store visits	There is higher foot traffic when a brand get more likes; however, the effects are not as large compare to the per-tweet effects in H4
H6 Number of followers for a user tweeting about a brand <i>positively</i> impacts store visits	As the number of followers per tweet mentioning a brand increases, there is virtually no impact on retail visits
H7 Recency in days from brand tweets <i>positively</i> impacts store visits	Time has a nonlinear effect where social media activity is less impactful in the same day, more impactful after a few days, and insignificant within a week

impact social media has on retailers is specific to the brand and the store. Companies, therefore, should be skeptical of external research and instead quantify the unique ways social media impacts their own business.

This brand salience relates to Luca (2016), who finds a higher rating on Yelp.com will raise a restaurant's revenue, unless it is a national chain. Online reviews also have a lower impact on video game sales (Zhu and Zhang 2010) and DVD/Blu-ray sales (Ho-Dac et al. 2013) when the product is a so-called "strong brand." Importantly, these results focus on a generalizable impact of social media, but there is still wide heterogeneity between brands and even stores of the same brand. Much of the previous literature showed how social media influences product sales, but we were able to show that social media about a retail brand can impact foot traffic. Our results also demonstrate how these empirical effects are sensitive in how social media is measured.

Since we find wide evidence of store and brand heterogeneity, marketing managers should evaluate the distinct, and sometimes opposing, ways social media activity influences their own customers and stores. We also suggest marketing managers can use publicly available social media data to forecast retail foot traffic. Our results show that activity on Twitter can serve as a leading indicator to future store visits, suggesting a direction of causality. The demand signal is most effective for measuring next-day activity; however, it fully dissipates by 7 days. Finally, brand managers should monitor the less-common measure of sentiment spread (i.e., disagreement), which we find is more important than sentiment itself.

This article also contributes to diverse academic fields. In economics, the research shows how social media provides a demand signal to monopolistically competitive firms. In marketing, the article demonstrates that using hierarchical data in a nonhierarchical model can lead to overstated claims

in how the predictors are associated with the outcome variable. In communications, we describe and formulate methods for considering distinct ways to quantify social media activity in terms of measures, weights, and time horizons. We also believe hierarchical methods deserve more attention in the economics and marketing literature, where there are many opportunities to measure how variables on one level affect the outcome on another level. Some laudable exceptions, however, include Aiello and Bonanno (2018) assessing the profitability of banks nested within local markets and Keller et al. (2019) assessing the effect of a promotion event on sales of brands nested within retailers.

Conclusion

The first aspect of our research was to identify which measures of social media activity led to changes in retail foot traffic. Here we identified that the disagreement of brand mentions on Twitter has a moderate impact on store traffic, but sentiment has no discernable effect. Although brand disagreement is rather consistent, the magnitude and significance of brand sentiment becomes positive and meaningful when defined on a per-like or per-follower basis. Subjectivity has no discernible influence. We show that the strongest effect of social media activity on foot traffic comes from the number of tweets that mention a brand, and next from the number of likes received by tweets mentioning a brand. However, increases in the number of followers of users that mention a brand also has no discernible effect. Finally, the recency of posts matter. Same-day social media activity around retail brands has minimal impact on same-day store visits, but the effect is greatest within 1 to 3 days and negligible within a week. These results support the Social Impact Theory, showing how the number sources, the immediacy of a source,



and—to a lesser extent—the strength of a source all lead to greater changes in consumer behavior.

The second aspect of our research question was to estimate the marginal impact of changes in social media activity on changes to the number of retail visits. We first identify the wide heterogeneity between the ways traffic to individual brands and individual stores responds to changes in online chatter about a brand. We then control for that heterogeneity using a hierarchical linear model (HLM). The resulting elasticities provide better estimates than could be obtained from ordinary linear regression, which is based on overly strong assumptions of error independence among nested observations.

For example, we find that a one standard deviation increase in brand mentions or disagreement will increase foot traffic the next day to stores of those brands by about 0.04 standard deviations (3–4%). Although this may appear small, it is economically significant when considering all stores of a national brand. Overall, the results point to a modest but meaningful relationship between social media activity and foot traffic to retailers.

We provide improved global estimates for recognizable brands outside of our sample; however, the diversity we observe should provide a sober reminder that all measures represent average effects. Therefore, we recommend that more applications of hierarchical linear models where there is nested data, such as product sales within stores. At the same time, we recognize there are limitations in how average effects can be applied to business situations with wide heterogeneity, such as our research problem. Because of this, we recommend reproducing the methodology outlined here to estimate the unique way social media can impact other organizations. This paper shows how Twitter, being publicly available, is a useful source of large-scale, online word-of-mouth communication that can signal near-future consumer demand.

There are multiple ways to expand this paper. First is to test if these results hold across multiple platforms, each with their own unique set of consumers and rules on how users interact and share information. It would be of note if the magnitudes outlined here apply to a larger, more private virtual space like Facebook. Second, it would be meaningful to see whether the described elasticities compare to lower-recognition retail brands. Luca (2016) suggests how more local brands are more susceptible to online reviews—future research could test if this also applies to other online formats. Third, although we did not find any long-term effect of social media on retail foot traffic, our analysis was limited to daily measures. An expanded model could consider the impact of accumulated social media impressions on store visits. Finally, we built a framework based on the Social Impact Theory to measure the directional influence social

media. Additional research could relax this perspective and consider if retail visits lead to more activity on Twitter.

Funding This research did not receive any specific Grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

- Aiello, F., and G. Bonanno. 2018. Multilevel Empirics for Small Banks in Local Markets. *Papers in Regional Science* 97 (4): 1017–1037.
- Antweiler, W., and M.Z. Frank. 2004. Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. *Journal of Finance* 59 (3): 1259–1294.
- Archak, N., A. Ghose, and P.G. Ipeirotis. 2011. Deriving the Pricing Power of Product Features by Mining Consumer Reviews. *Management Science* 57 (8): 1485–1509.
- Argo, J.J., D.W. Dahl, and R.V. Manchanda. 2005. The Influence of a Mere Social Presence in a Retail Context. *Journal of Consumer Research* 32 (2): 207–212.
- Araujo, Theo, Peter Neijens, and Rens Vliegenthart. 2017. Getting the Word out on Twitter: The Role of Influentials, Information Brokers and Strong Ties in Building Word-of-Mouth for Brands. *International Journal of Advertising* 36 (3): 496–513.
- Bollen, J., H. Mao, and X. Zeng. 2011. Twitter Mood Predicts the Stock Market. *Journal of Computational Science* 2 (1): 1–8.
- Chen, Y., Q.I. Wang, and J. Xie. 2011. Online Social Interactions: A Natural Experiment on Word of Mouth versus Observational Learning. *Journal of Marketing Research* 48 (2): 238–254.
- Chenga, J.H., and L.W. Linb. 2020. The Impact of 5G Online Social Media Interaction on Consumer Purchase Intentions: A Social Impact Theory Perspective. *Revista Argentina De Clínica Psicológica* 29 (5): 1985–2001.
- Cheung, C.M.K., and D.R. Thadani. 2012. The Impact of Electronic Word-of-Mouth Communication: A Literature Analysis and Integrative Model. *Decision Support Systems* 54 (1): 461–470.
- Cui, G., H.K. Lui, and X. Guo. 2012. The Effect of Online Consumer Reviews on New Product Sales. *International Journal of Electronic Commerce* 17 (1): 39–58.
- Deloitte. 2013. Tweets for Sales: Gaming.
- Gijsbrechts, E., K. Campo, and T. Goossens. 2003. The Impact of Store Flyers on Store Traffic and Store Sales: A Geo-Marketing Approach. *Journal of Retailing* 79 (1): 1–16.
- Godes, D., and D. Mayzlin. 2009. Firm-Created Word-of-Mouth Communication: Evidence from a Field Test. *Marketing Science* 28 (4): 721–739.
- Gurun, U., J. Nickerson, and D.H. Solomon. 2020. The Perils of Private Provision of Public Goods. SSRN.
- Ho-Dac, N.N., S.J. Carson, and W.L. Moore. 2013. The Effects of Positive and Negative Online Customer Reviews: Do Brand Strength and Category Maturity Matter? *Journal of Marketing* 77 (6): 37–53.
- Hovland, C.I. 1948. Social Communication. *Proceedings of the American Philosophical Society* 92 (5): 371–375.
- Hox, J.J., M. Moerbeek, and R. de Schoot. 2010. *Multilevel Analysis: Techniques and Applications*. London: Routledge.



- Huta, V. 2014. When to Use Hierarchical Linear Modeling. *The Quantitative Methods for Psychology* 10 (1): 13–28.
- Ismagilova, E., E. Slade, N.P. Rana, and Y.K. Dwivedi. 2020. The Effect of Characteristics of Source Credibility on Consumer Behaviour: A Meta-Analysis. *Journal of Retailing and Consumer Services* 53: 101736.
- Java, A., X. Song, T. Finin, and B. Tseng. 2009. *Why We Twitter: An Analysis of a Microblogging Community*. Lecture Notes in Computer Science 5439 LNAI, 118–38. New York: Springer.
- Joshi, Y.V., and A. Musalem. 2021. When Consumers Learn, Money Burns: Signaling Quality via Advertising with Observational Learning and Word of Mouth. *Marketing Science* 40 (1): 168–188.
- Jun, S.P., and H.D. Park. 2016. Consumer Information Search Behavior and Purchasing Decisions: Empirical Evidence from Korea. *Technological Forecasting and Social Change* 107: 97–111.
- Kalyanam, K., J. McAteer, J. Marek, J. Hodges, and L. Lin. 2018. Cross Channel Effects of Search Engine Advertising on Brick & Mortar Retail Sales: Meta Analysis of Large Scale Field Experiments on Google.com. *Quantitative Marketing and Economics* 16 (1): 1–42.
- Keller, W.I.Y., B. Deleersnyder, and K. Gedenk. 2019. Price Promotions and Popular Events. *Journal of Marketing* 83 (1): 73–88.
- Kim, N., and W. Kim. 2018. Do Your Social Media Lead You to Make Social Deal Purchases? Consumer-Generated Social Referrals for Sales via Social Commerce. *International Journal of Information Management* 39: 38–48.
- Kim, K., S. Yoon, and Y.K. Choi. 2019. The Effects of EWOM Volume and Valence on Product Sales—An Empirical Examination of the Movie Industry. *International Journal of Advertising* 38 (3): 471–488.
- Kwahk, K.Y., and X. Ge. 2012. The Effects of Social Media on e-Commerce: A Perspective of Social Impact Theory. The 2012 45th Hawaii International Conference on System Sciences, IEEE Chicago, 1814–1823.
- Latané, B. 1981. The Psychology of Social Impact. *American Psychologist* 36 (4): 343.
- Leeuw, J., De., E. Meijer, and H. Goldstein. 2008. *Handbook of Multilevel Analysis*. New York: Springer.
- Liu, T., X. Ding, Y. Chen, H. Chen, and M. Guo. 2016. Predicting Movie Box-Office Revenues by Exploiting Large-Scale Social Media Content. *Multimedia Tools and Applications* 75 (3): 1509–1528.
- Loper, E., and S. Bird. 2002. NLTK: The Natural Language Toolkit.
- Loria, S., P. Keen, M. Honnibal, R. Yankovsky, D. Karesh, E. Dempsey, et al. 2014. Textblob: Simplified Text Processing.
- Lovett, M.J., R. Peres, and L. Xu. 2019. Can Your Advertising Really Buy Earned Impressions? The Effect of Brand Advertising on Word of Mouth. *Quantitative Marketing and Economics* 17 (3): 215–255.
- Luca, M. 2016. Reviews, Reputation, and Revenue: The Case of Yelp.com. Harvard Business School, NOM Unit Working Paper, no. 12–016.
- Mainardes, E.W., and M. Vieira Cardoso. 2019. Effect of the Use of Social Media in Trust, Loyalty and Purchase Intention in Physical Stores. *The International Review of Retail, Distribution and Consumer Research* 29 (4): 456–477.
- Naeem, M., and W. Ozuem. 2021. Understanding the Social Consumer Fashion Brand Engagement Journey: Insights about Reputed Fashion Brands. *Journal of Brand Management* 28 (5): 510–525.
- Pauwels, K., K. Aksehirli, and A. Lackman. 2016. Like the Ad or the Brand? Marketing Stimulates Different Electronic Word-of-Mouth Content to Drive Online and Offline Performance. *International Journal of Research in Marketing* 33 (3): 639–655.
- Pew Research Center. 2020. Social Media Fact Sheet.
- Rinker, T. 2017. Package 'Sentimentr'.
- Sakaki, T., M. Okazaki, and Y. Matsuo. 2010. Earthquake Shakes Twitter Users: Real-Time Event Detection by Social Sensors. Proceedings of the 19th International Conference on World Wide Web, 851–860.
- Sanchez, J., J. Abril, and M. Haenlein. 2020. Competitive Spillover Elasticities of Electronic Word of Mouth: An Application to the Soft Drink Industry. *Journal of the Academy of Marketing Science* 48 (2): 270–287.
- Squire, R.F. 2019. What About Bias in the SafeGraph Dataset?. SafeGraph Blog. 17 October.
- Smith, A.N., E. Fischer, and C. Yongjian. 2012. How Does Brand-related User-generated Content Differ across YouTube, Facebook, and Twitter? *Journal of Interactive Marketing* 26 (2): 102–113.
- Snijders, T.A.B., and R.J. Bosker. 2011. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Thousand Oaks: Sage.
- Statista. 2022. Daily time spent on social networking by internet users worldwide from 2012 to 2022.
- Stricker, G. 2014. The 2014 #YearOnTwitter, Twitter Blog, 10 December.
- Suddath, C. 2019. Popeyes Chicken Sandwich, Internet Superstar. *Bloomberg Businessweek*, 4 December.
- U.S. Census Bureau. 2021. Quarterly Retail E-Commerce Sales: 3rd Quarter 2022. 18 November.
- You, Y., G.G. Vadakkepatt, and A.M. Joshi. 2015. A Meta-Analysis of Electronic Word-of-Mouth Elasticity. *Journal of Marketing* 79 (2): 19–39.
- Zhang, Z., X. Li, and Y. Chen. 2012. *Deciphering Word-of-Mouth in Social Media: Text-Based Metrics of Consumer Reviews. Transactions on Management Information Systems*. New York: ACM.
- Zhu, F., and X. Zhang. 2010. Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. *Journal of Marketing*.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Thomas J. Weinandy is a Research Economist at Upside, a digital promotions marketplace driving incremental, attributable profit for brick-and-mortar retail through personalized consumer offers. He received his B.A. in Social Entrepreneurship and Spanish from John Carroll University, his M.B.A from Wheeling Jesuit University, and his Ph.D. from Western Michigan University with the dissertation “Applied Microeconomics and Business Intelligence in the Digital Age.” His research looks at how digital technologies impact business and consumer decision making, particularly in retail. Dr. Weinandy also does pro bono consulting, from conference organizing to nonprofit program evaluation.

Kuanchin Chen is a Professor of Computer Information Systems, Director of the Center for Business Analytics, and John W. Snyder Fellow at Western Michigan University. His research interests include electronic business, data analytics, social networking, privacy & security, online behavioral issues, and human computer interactions. Dr. Chen has published in many reputable academic journals, including *Information Systems Journal*, *Decision Support Systems Information & Management*, and many others. He has been an editor, associate editor, and editorial member of several scholarly journals. He has also received research and teaching awards from scholarly journals & conferences, department, college, university, and U.S. Fulbright program.



Susan Pozo is a professor of Economics at Western Michigan University. She received her undergraduate degree from Barnard College, Columbia University and a Ph.D. in Economics from Michigan State University. Her research areas are in asset accumulation, currency crises, exchange rate uncertainty, immigrants' remittances, immigration policy, and returns to international human capital. She has held visiting positions at Oxford University, Universidad de Salamanca, and Universidad de Montevideo. Her research has been funded by the United Nations, the Inter-American Development Bank, the National Institutes of Health, the Fulbright Foundation and the Banco Bilbao Vizcaya Argentaria.

Michael J. Ryan is the director of undergraduate studies and a professor in the Department of Economics at Western Michigan University. He received his Ph.D. in Economics from Indiana University. Dr. Ryan's primary research interests include theoretical and empirical studies on foreign direct investment and multinational enterprise behavior. Recent studies focus on affiliate financing, location, and ownership structure choice. He has served as a Visiting Professor at the universities of Kiel (Germany), Bamberg (Germany), and Otago (New Zealand). He was recently also a visiting scholar at Chuo University (Japan).

