

What Happens When Word of Mouth Goes Mobile?

Completed Research Paper

Gordon Burtch

Carlson School of Management
University of Minnesota
gburtch@umn.edu

Yili (Kevin) Hong

W. P. Carey School of Business
Arizona State University
hong@asu.edu

Abstract

Individuals are likely to exhibit different behaviors between mobile and non-mobile devices, for a number of reasons. For example, mobile devices enable ubiquitous access, through their portability, yet they also constrain users because of the smaller form factor. Very little work to date has attempted to examine the impact of these differences holistically. Indeed, the work that does exist has generally focused on approaches to location-based advertising. One particularly important aspect of mobile usage behavior pertains to user content generation. Bearing this in mind, we aim here to improve our understanding of device-dependent user behavior by examining differences in content generated on mobile and non-mobile devices, in the context of electronic word of mouth. We demonstrate a variety of important differences in reviews that are submitted via mobile devices; they exhibit lower and more varied star ratings, contain more concrete and emotional text, and are generally perceived as more helpful. We discuss the implications for both service providers and the management of online review platforms.

Keywords: mobile, word of mouth, online reviews

Introduction

Mobile¹ devices enable ubiquitous access while constraining usage due to the smaller yet portable form factor, leading to different user behavior across mobile and non-mobile devices. The differences run deeper, however. Observed behavior may also differ because of individuals' selection into the mobile medium, whether as a result of their personal preferences and characteristics (e.g., habit of only using mobile devices or desktops), or contextual factors that can influence device accessibility (e.g., social appropriateness or travel time). Given the fast growing mobile industry², it is notable that a dearth of work to date has attempted to explore differences between mobile and non-mobile users' behavior. To the authors' knowledge, the sparse empirical work on the subject has generally focused on how best to target mobile users for advertisement, in a context-aware manner (Ghose et al. 2013; Ghose and Han 2011; Goh et al. 2009; Luo et al. 2014; Molitor et al. 2014; Noulas et al. 2011; Shankar and Balasubramanian 2009; Shankar et al. 2010; Sultan et al. 2009). Accordingly, in this work, we take an important first step toward broadening our understanding of device-dependent user behavior, in the context of electronic word of mouth.

Electronic Word-of-Mouth is now generally accepted as an integral component of firm marketing efforts (Chen and Xie 2008; Dellarocas 2003). Numerous online platforms now enable consumers to post ratings and text reviews about merchants and products of all kinds, effectively allowing them to share their experiences and opinions with others the world over. Recently, online reviews have been getting even easier. TripAdvisor, a leading review platform, has allowed consumers to author reviews from their mobile device in the last few years. In August 2013, *Yelp* also enabled a mobile review function. Traditionally, mobile reviews have been avoided by online review platforms, out of fear that allowing them will result in 'thinner' review content (e.g., lazy, uninformative, or short reviews), as well as 'rants' and 'raves' from emotional consumers in the moment. In fact, in the past, *Yelp* has publicly proclaimed such concerns on their corporate blog³.

Despite the obvious drawbacks, a mobile review channel offers a number of potential benefits, thanks to its increased convenience (Ghose and Han 2011). Rather than waiting until after he or she returns home from a restaurant, using a mobile device, a consumer can author a review on the spot. This has a number of interesting implications. First, given their *ad hoc* nature, mobile reviews are less likely to have been paid for (i.e., fake), which in turn may result in fewer 5-star reviews (note: although one can pay for fake negative reviews to damage the reputation of competition, this is certainly not the dominant type of fake reviews in the marketplace)⁴. Increased convenience also implies that the under-reporting bias that have previously been noted in the literature, which drives *J*-shaped review distributions, are likely to play a much smaller role in the mobile scenario (Anderson 1998; Dai et al. 2012; Hu et al. 2009). As a final example, given the increase in access and convenience, in general, mobile reviews are expected to exhibit indications of consumption recency. This is notable, because numerous studies in the word of mouth literature have examined the roles of consumption recency, timing and delays on word of mouth content (Berger and Schwartz 2011; Chen and Lurie 2013; Moe and Schweidel 2012), finding important implications around consumption recency cues.

The objective of this study is to broadly explore the implications for review characteristics from offering a mobile online review channel to consumers. We undertake our analysis by examining different aspects of online reviews across mobile and desktop reviews. The primary factors we focus upon include the standard measure of review valence. Additionally, we also examine the textual content of reviews in terms of review length, concreteness, emotion, textual indications of consumption recency and, pursuant to the above, whether other consumers find the review helpful.

¹ We define mobile devices as smart phones and tablets for the purposes of this study, given that mobile is identified in our data based on the use of a smartphone or tablet app.

² <http://www.businessinsider.com/mobile-is-growing-2013-11>

³ <http://officialblog.yelp.com/2009/12/ask-yelp-why-cant-i-write-reviews-from-my-mobile.html>

⁴ <http://www.nytimes.com/2013/09/23/technology/give-yourself-4-stars-online-it-might-cost-you.html>

To identify these effects, we draw on a panel dataset from *www.TripAdvisor.com* (*TripAdvisor*), a website that hosts online reviews for the service industry, with a focus on restaurants and hotels. *TripAdvisor* provides an ideal context because individual reviews are publicly flagged as having been entered via a mobile device (or not). In examining differences between mobile and non-mobile reviews, we bear in mind that self-selection into the mobile review channel is a potentially source of endogeneity. Accordingly, we employ purposeful sampling to only include reviewers who have posted reviews from both a mobile device and a desktop computer.

Our econometric analysis produces the following results. First, we find evidence that mobile reviews are more likely to be extreme, consistent with the idea that mobile users partake in ‘rants and raves’. Second, we demonstrate that mobile reviews are significantly shorter in length (i.e., word count), containing approximately 11 fewer words, on average. Third, we find that mobile reviews are significantly more likely to contain indications of recent consumption (i.e., shorter delay between consumption and review posting). Fourth, we find that mobile reviews are significantly lower in valence, on average, attributing this result jointly to greater review fidelity (i.e., mobile reviews are less likely to have been paid for) and evidence in the psychology literature that negative thoughts fade more quickly with time. Fifth, we find that, all else held equal, mobile reviews receive more helpful votes from other consumers. Sixth, we find that mobile reviews are significantly more likely to contain concrete information, specifically in terms consumers’ references to perceptions and sensory experiences. Lastly, we find that mobile reviews exhibit significantly greater levels of emotion.

In the following sections, we begin by reviewing the literatures pertaining to online WOM and the mobile Internet. We then proceed to formulate a series of hypotheses, motivated by our review of the literature. Next, discussing our methodological approach, we detail our study context, data and econometric specification, before presenting our estimation results. We then offer a discussion of the implications of our findings, particularly in terms of the policy implications for online review platforms. Finally, we conclude by discussing the limitations of this study, and suggesting a number of potentially fruitful avenues for future research.

Literature Review

Online Word-of-Mouth

Online WOM is a very rich literature, dating back to at least the 1950s (Katz and Lazarfeld 1955). In general, WOM is now generally accepted as crucial to the success of a business (Mudambi and Schuff 2010). Of late, scholars have placed a significant emphasis on electronic WOM, as well as the impact of WOM on product demand (Chevalier and Mayzlin 2006; Godes and Mayzlin 2004; Godes and Mayzlin 2009), firm strategy (Chen and Xie 2008) and market competition (Kwark et al. 2014). A notable feature of online WOM is that it is typically characterized by reporting biases, namely in the form of positivity and underreporting (Anderson 1998; Dellarocas and Narayan 2006; Hu et al. 2009). Recently, scholars have also examined peer referrals in online settings (Burtch et al. 2014; Shi et al. 2013), as well as the role of variance and disagreement in online ratings (Hong et al. 2012; Nagle and Riedl 2014; Sun 2012)

Studies of online WOM pertaining to service providers (the focus of the present work) have also grown quite common over the last few years. Anderson and Magruder (2012) employ a regression discontinuity design (RD) to identify and quantify the relationship between online restaurant reviews and reservation availability. Luca (2011) employs a similar RD approach to identify the effect of online reviews from Yelp on restaurant revenue in the state of Washington. Byers et al. (2012) examine the effect of daily deal offers (e.g., Groupon) on service provider ratings at Yelp, finding evidence of a negative relationship. These authors argue that this negative relationship results from an influx of critical customers. However, these authors also discuss and attempt to rule out the possibility that this results from an increase in “real” reviews, which are less likely to be positive.

Most recently, a number of studies have begun to delve into novel aspects of reviews. In particular, scholars have recently examined the role of textual features in online reviews, considering such factors as emotion (Yin et al. 2014), concreteness (Li et al. 2013) and references to recent consumption experiences (Chen and Lurie 2013), and how these characteristics in turn dictate the perceived helpfulness of an online review by other consumers (Mudambi and Schuff 2010).

In this work, we build on much of this recent body of work by considering important differences between review valence and content submitted via mobile and non-mobile devices. In doing so, we consider differences in the user population, as well device effects, including increased access and opportunity as well as a more limited user interface, which in turn may have implications for many of the factors mentioned above.

Mobile Internet

Recent years have seen a growing interest in the mobile Internet among Information Systems and Marketing scholars (Shankar and Balasubramanian 2009; Shankar et al. 2010). Goh et al. (2009) offer a general study on the efficacy of mobile advertising in the context of automobile sales. Ghose et al. (2013) demonstrate that distance matters more to mobile Internet users than it does to users who work with a desktop. Ghose and Han (2011) examine the trade off and interplay between mobile user content generation and consumption. Molitor et al. (2014) employ randomized experimentation to evaluate the efficacy of location-based advertising, exploring the tradeoff between geographical distance, pricing discounts and screen position.

In general, the prior literature has placed a particular focus upon how mobile users can be better targeted to enhance advertising effectiveness. One notable exception to this is a recent working paper by Jung et al. (2013), which has sought to understand differences in user behavior in an online dating platform, following the adoption of a mobile application. We take a similar tact here, though we benefit from more fine-grained and precise individual-level data in our analysis. We observe device usage with each transaction on the website, whereas Jung and his colleagues only observe the initial download of a mobile app, and they then examine aggregate shift in average user behavior following app download by assuming mobile device usage.

Hypothesis Development

Review Valence

It has often been noted that a major benefit of the mobile Internet comes in the form of convenience. Busy users who are frequently in transit can use their mobile phones to view and also post content on the Internet during the course of their busy schedule, as they travel from place to place (Ghose and Han 2011). Shankar et al. (2010) refer to these users as “road warriors”. With increased convenience come increased opportunities to engage. Because mobile consumers have greater freedom to author a review, without the time constraints, the aforementioned extremity biases that have been shown to result in censored WOM (Anderson 1998; Hu et al. 2009) should be expected to wane. To clarify, whereas consumers may have previously required a very extreme experience in order to justify waiting and then posting their feedback, once they gained access to a desktop computer, increased opportunity to post feedback from anywhere, anytime, might reduce the threshold required to justify writing a review. In turn, this might be expected to result in more moderate reviews and, in general, a distribution that is more representative of the true preferences of the consumer population.

At the same time, however, the ability to post a review on the spot, immediately following a consumption experience, might enable consumers to “rant and rave.” This possibility is of particular note, given that the industry has previously expressed concerns about this very possibility⁹. If a consumer waits longer before posting their review, we might expect them to be in a calmer, cooler mindset, reflecting their experience more objectively. Given the above, it is difficult to anticipate whether mobile reviews will be more or less extreme in valence. Accordingly, we refrain from proposing any formal hypotheses about this relationship for the time being.

Next, it is interesting to consider the potential effect of mobile’s introduction on review valence as well. First, it is likely that mobile channels may impact the proportion of real versus paid reviews on a platform. Because mobile reviews are more likely to be *ad hoc* and spontaneous, they are in turn less likely to be

⁹ <http://officialblog.yelp.com/2013/10/android-users-prepare-for-a-thumb-workout-with-todays-addition-of-mobile-reviews.html>

compensated (i.e., paid for by a service provider), and thus less likely to be fake. This is important because recent work has noted the surprising prevalence of fake reviews on leading review platforms (Luca and Zervas 2013). Accordingly, mobile reviews might be expected to hold a lower valence, given that fake reviews will typically be highly positive. We therefore expect that mobile reviews should be of a lower valence. To this point, it is notable that a similar notion has been proposed and examined around the effect of service providers' daily deal offerings on Yelp reviews (e.g., Groupon). As noted previously, Byers et al. (2012) consider the possibility that daily deals drive a basic increase in traffic volumes at a service provider, resulting in a spike in real review content and thus a decline in average review valence. We consider an analogous result here.

There are also other reasons to expect that mobile reviews are likely to hold a lower valence. In particular, common sense dictates that memories fade with time. Interestingly, however, work in psychology has shown evidence of Fading Affect Bias (FAB); the notion that negative memories fade away more quickly than do positive memories (Ritchie et al. 2014). If we allow that mobile reviews should be entered more quickly following consumption (given greater access and opportunity, compared to desktops), on average, then they should also be more likely to contain negative thoughts and opinions. Taking all of the above together, we formalize our expectation in Hypothesis 1.

H1: Online reviews will have a lower valence when submitted via a mobile device.

Review Length

We next consider the length of mobile reviews. Scholars have noted that certain tasks are more difficult to achieve using a mobile device. Specifically, a simplified user interface makes it significantly more challenging for users to locate information and parse content (Ghose et al. 2013), thereby increasing user search costs and effort. In the context of online reviews, this may translate into difficulty typing and navigating. Indeed, platforms supporting online reviews are clearly cognizant of these issues. This is evidenced by the past concerns they have voiced about 'thin' content, such as exceedingly short, uninformative reviews (noted previously). Indeed, some platforms, such as Yelp, have instituted a minimum word count in reviews to combat this. Nonetheless, we maintain an expectation of reduced length when it comes to mobile reviews.

H2: Online reviews will contain fewer words when submitted via a mobile device.

Review Textual Content

We consider three aspects of review textual content: indications of consumption recency, review concreteness and emotional reviews. Mobile apps and web browsers work on devices with different form factors: mobile phones and desktops, respectively. Typically, diners go to a restaurant with a mobile phone, but browse the Internet at home on a desktop. We expect the introduction of a mobile channel for online review posting to reduce the delay between consumption and posting, because this makes it possible for consumers to post their reviews immediately following consumption, rather than waiting until they return home, to a desktop. This is notable, because a number of studies in marketing have found the timing of WOM to be important. Studies have examined how temporal aspects of WOM are associated with product or service characteristics and consumer response to WOM. Berger and Schwartz (2011) consider product characteristics that drive immediate versus long term WOM, arguing that novel products will be more likely to spark immediate interests, but that such products will become attract less interest over time, and thus incite less sustained WOM, as consumers learn about them. Zhao and Xie (2011) discuss the temporal distance between consumers and their planned consumption, and how temporal distance affects their likelihood of relying on WOM. They find that a consumer is more likely to rely on WOM when it pertains to a consumption that is planned in the distant future. What is more, that effect is moderated by social distance between the WOM generator and the consumer, such that WOM pertaining to temporally proximate consumption plans are most influential when provided by a social proximate other, while WOM pertaining to temporally distant consumption activities is most influential when offered by social distant others.

Perhaps most relevant to the present study is the work of Chen and Lurie (2013), who report that, although consumers are more likely to place a greater weight on negative WOM in general, that focus is mitigated when the individual authoring the review indicates cues on "recent" consumption. Accordingly,

here, we look to identify any decline in posting delays when reviews are entered via a mobile device. To do so, we take the approach articulated by Chen and Lurie, and define a review as ‘recent’ if it contains textual references to recent consumption (e.g., ‘today’). We then hypothesize an increase in the probability that a review contains indications of consumption recency when it is entered via a mobile device.

H3: Online reviews are more likely to contain indications of consumption recency when submitted via a mobile device.

We also consider the textual content of reviews, as a function of their real-time nature. An extensive literature in psychology has argued and demonstrated repeatedly that temporal distance is associated with greater abstraction in thoughts and ideas due to the use of high level construal (Liberman et al. 2007), and above we have argued that mobile reviews should take place with less delay. That is, if two reviewing scenarios are identical in that they share the same author and pertain to the same restaurant, except that one review is posted on the spot while the other is posted following some delay, the above theory suggests that we would expect to observe differences in review content. The review posted on the spot (i.e., low temporal distance) would incorporate more detailed information, such as the taste of the food, the ambiance, and the helpfulness of the service staff (concrete information). In contrast, the review posted after some delay (i.e., high temporal distance) would contain less information, as fine-grained details fade from memory with time. We therefore anticipate greater levels of concreteness in the text of mobile reviews.

H4a: Online reviews are more likely to contain concrete text when submitted via a mobile device.

In addition to the above, we further expect that mobile reviews are more likely to prove helpful to other consumers, all else being equal. The reason for this is rather straight forward – in short, as noted above, memories generally fade with time (Ritchie et al. 2014). Accordingly, we might expect that mobile reviews should contain greater specifics and detail, and more importantly, we might expect that mobile reviews will provide an altogether more accurate representation of the author’s dining experience. We formalize this expectation below, in Hypothesis 4b.

H4b: Online reviews are more likely to be voted helpful when submitted via a mobile device.

Finally, we argue that mobile reviews likely contain greater emotional content than desktop reviews. This is fundamentally because mobile devices provide users with increased opportunity and access to the Internet, and thus enable impulsive, emotional actions, in the moment (otherwise known as visceral behavior), which would otherwise subside if the user were required to wait for a period before taking action (Loewenstein 1996; Loewenstein 2000). Accordingly, here, we anticipate that mobile reviews will be more likely to contain emotional textual content. We formalize this in Hypothesis 5, below.⁶

H5: Online reviews are more likely to contain emotional text when submitted via a mobile device.

Methods

Study Context & Identification

Our analyses consider online review data from the website *TripAdvisor.com*. *TripAdvisor* is a leading platform for online restaurant and hotel reviews. The site receives approximately 10 million visitors per month and is currently ranked 208th globally for web traffic (88th in the United States). Moreover, the site now hosts more than 150 million reviews, with a large proportion being contributed by users on mobile devices; the *TripAdvisor* mobile app has now been downloaded by more than 82 million people^{7,8,9}.

⁶ Note: the presence of emotional text need not coincide with extreme star ratings, and vice versa (Yin et al. 2014, footnote 2). It is for this reason that we offer separate discussions of the relationship between *Mobile*, star rating variance and emotional text.

⁷ <https://www.quantcast.com/tripadvisor.com>

Our identification strategy hinges on a purposeful sampling approach, which is intended to circumvent or account for self-selection biases around mobile device use, whether in terms of the reviewer's unobservable characteristics, or in terms of the service provider's unobservable characteristics. For example, it may be the case that a mobile review is more likely for restaurants where WiFi services are readily available. Similarly, a mobile review may be more likely when a consumer only has access to one type of device (e.g., the consumer does not own a mobile device, or a mobile device is their only source of Internet access).

To address these issues, we take two approaches. First, to address self-selection relating to device availability (i.e., the mobile- or desktop-only user), we begin by focusing our analyses solely on the subset of online reviews created by consumers exhibiting some variation in their device choice across reviews. That is, we exclude one-time reviewers, and we exclude reviews associated with consumers who limited their review authorship exclusively to one device type (i.e., mobile-only or desktop-only). This approach allows us to maintain some confidence that any effects identified truly pertain to cross-device usage behavior, and not to the population of users that prefer one device type.

Second, we incorporate three-way fixed effects, jointly addressing static unobservable heterogeneity across reviewers, restaurants and time periods. In doing so, we address the aforementioned influence of restaurant and user characteristics that remain relatively stable over time, such as WiFi availability or cuisine type. Moreover, time fixed effects allow us to address temporal trends, such as seasonality, in a non-parametric manner.

Dataset & Econometric Specification

Our dataset comprises all reviews from a random sample of 3,050 restaurants on TripAdvisor. For these restaurants, we collected the entire history of review content, from November of 2004 through February of 2014. The said sampling approach resulted in 23,827 reviews, authored by 6,021 reviewers. Our key independent variable across all of our regressions is a binary indicator of whether a review was entered via a mobile device, *Mobile*. Figure 1 provides a screenshot depicting a review from TripAdvisor.com that was entered on April 5th, 2014. Here, we see an icon indicating that this review was entered using a Mobile device.

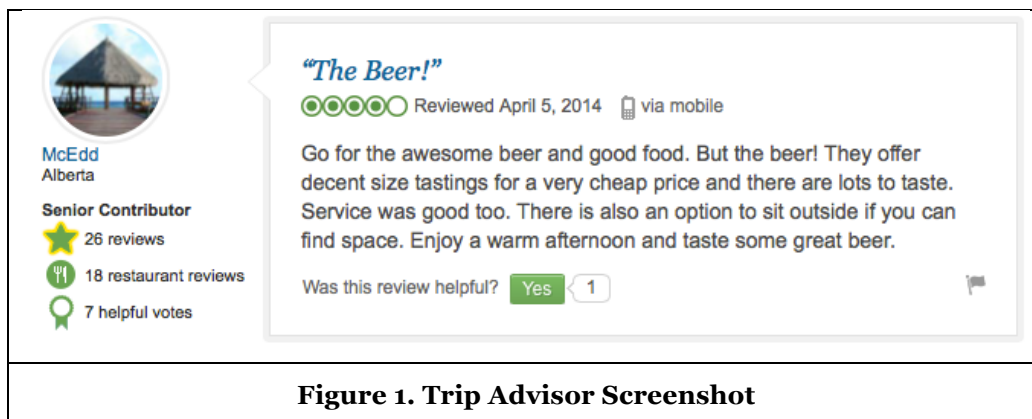


Figure 1. Trip Advisor Screenshot

Each of our hypotheses pertains to a different dependent variable. Hypothesis 1 refers to the valence of reviews. We measure review valence in terms of an ordinal variable, *Valence*, which can take whole integer values between 1 and 5. Hypothesis 2 refers to the length of online reviews, *Length*, in terms of character count of a text review. Hypothesis 3 refers to textual indications of consumption recency. To capture this, we construct the variable *Recent* based on the presence of recency-related keywords. We

⁸ <http://www.alexacom/siteinfo/tripadvisor.com>

⁹ http://www.tripadvisor.com/PressCenter-c4-Fact_Sheet.html

follow the prior literature (Chen and Lurie (2013)) to identify consumption recency through mention of any of the following keywords in the text of a review: ‘today’, ‘this morning’, ‘just got back’ and ‘tonight’. Notably, we expand upon this definition slightly, in that we also include mentions of ‘this evening’, as well. As per Chen and Lurie, we code reviews as containing indications of consumption recency if any of these keywords are present (1) or not (0).

Hypotheses 4a and 5 refer to the concreteness and emotional content of reviews’ textual content. In order to operationalize these variables, we leverage Linguistic Inquiry and Word Count (LIWC) - a text analysis application – to identify sentiment, emotion, etc. in textual content (Pennebaker et al. 2001). Notably, LIWC has seen frequent use in the psychology literature, and has recently begun to be used in the MIS and Marketing literature as well (Goes et al. 2014; Sridhar and Srinivasan 2012; Yin et al. 2014). Here, we leverage LIWC’s calculated measure of perceptions (sensory experiences), for the text of each review, to capture its “concreteness,” referring to this measure as *Concrete*. Perceptual words include references to having seen, heard, tasted, etc. Taking a similar approach, we define *Emotion* in term’s of LIWC’s calculated emotional score for the text of each review. Finally, Hypothesis 4b refers to how helpful other consumers perceive a particular review to be. We define our measure of helpfulness, *Helpful*, in terms of the number of helpful votes a review receives.

Beyond these outcome measures, we also incorporate a number of controls. Most notably, we include fixed effects for reviewers, restaurants and time (month) in all of our estimations. These controls respectively address unobserved heterogeneity across reviewers (e.g., personal characteristics or persistent preferences), restaurants (e.g., WiFi access, price range), and time (e.g., temporal trends in reviewing behavior or unobserved, exogenous shocks to review content over time). Additionally, in our regressions pertaining to review helpfulness we include controls for how long the review had been posted as of the date of our data collection, *Age*, as well as for review length (noted above).

We estimate our models employing a two-way fixed effects estimator with time period dummies (Cornelissen 2009). Equations 1 and 2, below, clarify our econometric specification regarding our hypotheses related to helpfulness and review extremity. We index reviewers by i , restaurants by j , and month by t . It is important to point out that *TripAdvisor* does not allow reviewers to enter multiple reviews for the same restaurant over time. Accordingly, the vast majority of our variables are not time varying, with the exception of *Helpful*, *Age* and of course our time fixed effects. *Helpful* is time varying because reviews can accrue helpful votes on an ongoing basis after they are initially posted. Because *Helpful* is our only time-varying dependent variable, it is the only model wherein we incorporate review *Age* as an additional control. Note: all of our other DVs are estimated using a specification of the form depicted in Equation 2 (i.e., our controls include our three-way fixed effects, and we focus on the *Mobile* effect.

$$Helpful_{ijt} = \beta_1 * Mobile_{ij} + \beta_2 * Length_{ij} + \beta_3 * Age_{ijt} + \beta_4 * Valence_{ij} + \lambda_i + \varphi_j + \delta_t + \varepsilon_{ijt} \quad (1)$$

$$Valence_{ij} = \beta_1 * Mobile_{ij} + \lambda_i + \varphi_j + \delta_t + \varepsilon_{ijt} \quad (2)$$

We provide definitions for all of our variables in Table 1. Further, in Table 2, we provide the overall distribution of our variables. Beyond the population-level statistics, we also break down the statistics between mobile and non-mobile reviews. For the latter two, we present t -tests for mean differences. Upon doing so, we already begin to see evidence that mobile reviews are significant different from non-mobile reviews; we observe statistically significant mean differences, generally in the anticipated direction.

Results

Main Results

We began by examining the extremity of mobile reviews, as compared to desktop reviews. To test this, we conducted a statistical test of variance equality between the two review distributions. Employing Levene's robust statistic, as well as that proposed by Brown and Forsythe. In both cases, we find that mobile reviews exhibit a significantly larger variance (Levene's $F = 4.392$, $p < 0.05$; Brown and Forsythe's $F = 4.554$, $p < 0.05$), indicating that they are generally more extreme in valence than desktop reviews.

We next present the primary regression results for each of our hypothesis tests in Table 3. Each column in the table reflects a separate regression with a different dependent variable, evaluating a different hypothesis. Column one pertains to Hypothesis 1, regarding review valence. As noted previously, we hypothesized that all else being equal, mobile reviews are expected to be more negative in valence, primarily because they are less likely to be fake. This is exactly what we observe. Mobile reviews exhibit a significantly lower star rating, on average, compared to non-mobile reviews.

The second column provides results that pertain to Hypothesis 2, around the length of reviews. We find that mobile reviews are significantly shorter. In particular, we observe that, on average, mobile reviews contain approximately 11 fewer words. This is important, because this finding serves as a form of sanity check, ensuring that mobile device usage does indeed drive very clear, very simple differences in review content that derive directly from the nature of the medium (in this case, a limited user interface).

Column three demonstrates the effect of mobile device usage on the probability that review text contains keyword indications of consumption recency (Hypothesis 3). As hypothesized, we observe that mobile reviews are significantly more likely to contain text of this sort. In particular, mobile reviews are approximately 1% more likely to contain such indications. Although this effect size appears small in isolation, when we consider the baseline rate of recency indicators amongst desktop reviews (1.4%), this effect represents nearly a one third increase over the baseline.

Variable	Definition
Mobile	A binary indicator of whether the review was entered via a mobile device.
Valence	A positive integer value between 1 and 5 representing the star rating of the review.
Length	A positive integer value representing the number of characters in the body of the review.
Recent	A binary indicator of whether the review body contains keyword indicators of recent consumption (as per Chen and Lurie 2013).
Concrete	The <i>percept</i> score from LIWC for the body of the review, capturing mentions of perceptual (sensory) keywords.
Helpful	A positive integer value representing the number of times the review has been voted as helpful by other TripAdvisor users.
Emotion	The <i>affect</i> score from LIWC for the body of the review, capturing mentions of emotional keywords.
Age	The number of days that have elapsed from when the review was initially posted until the date of data collection.
Weekend	A binary indicator of whether the review was entered on a weekend.

Table 2. Descriptive Statistics									
	Population				Mobile		Non-Mobile		
Variable	Mean	St. Dev.	Min	Max	Mean	St. Dev.	Mean	St. Dev.	t-test
Mobile	0.40	0.49	0.00	1.00	--	--	--	--	--
Valence	4.16	0.95	1.00	5.00	4.15	0.97	4.17	0.93	-2.10*
Length	85.63	70.70	11.00	1,769.00	78.26	60.96	90.49	76.22	-13.10***
Recent	0.016	0.13	0.00	1.00	0.020	0.14	0.014	0.12	3.63***
Concrete	2.61	2.41	0.00	25.00	2.66	2.49	2.57	2.35	2.80**
Helpful	0.25	0.81	0.00	42.00	0.25	0.93	0.24	0.72	1.38+
Emotion	8.40	4.58	0.00	66.67	8.62	4.67	8.25	4.52	5.77***
Age	466.62	329.40	13.00	3,389.00	433.70	259.01	488.34	366.95	-12.57***

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$; $N_{tot} = 23,827$; $N_{mob} = 9,472$; $N_{nmob} = 14,355$.

Table 3. Regression Results: OLS /w Three-Way FE						
	Dependent Variable					
Variable	Valence	Length	Recent	Concrete	Helpful	Emotion
Mobile	-0.06*** (0.01)	-10.61*** (0.85)	0.01** (0.00)	0.10* (0.04)	0.02** (0.01)	0.36*** (0.08)
Length	--	--	--	--	0.00*** (0.00)	--
Valence	--	--	--	--	-0.05*** (0.01)	--
Age	--	--	--	--	-0.00 (0.00)	--
Time Effects	Included	Included	Included	Included	Included	Included
Reviewer Effects	Included	Included	Included	Included	Included	Included
Restaurant Effects	Included	Included	Included	Included	Included	Included
Observations	23,827	23,827	23,827	23,827	23,827	23,827

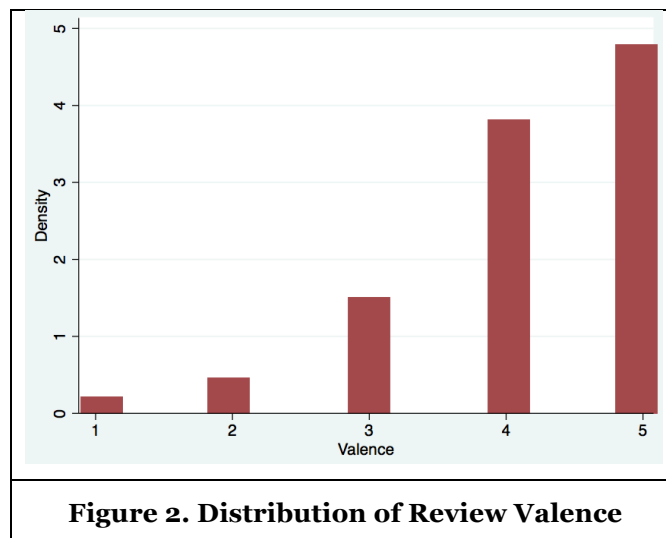
Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$; Estimates of fixed effects are jointly significant in all models, $p < 0.001$; Robust standard errors reported in brackets for coefficients; Models estimated with Stata's *felsdureg* command (Cornelissen 2009).

Finally, considering columns four, five and six, which pertain to hypothesis 4a, 4b and 5 (i.e., concreteness, helpfulness and emotion), we again find results that support our expectations. We find that mobile reviews exhibit significantly greater amounts of concrete and emotional text. Further, we find that, all else held equal, mobile reviews are significantly more likely to prove helpful to other consumers. Taken together, the above results suggest that the nature and content of mobile reviews are quite different from their non-mobile counterparts.

If we consider these effects in more concrete terms, it is readily apparent that the differences are economically important. For example, the effect of mobile on *Helpful* indicates that mobile reviews typically draw 8% more helpful votes, on average, all else held equal. Similarly, *Mobile* reviews are half a star rating lower than desktop reviews, on average. If we refer to the work by Luca on the relationship between Yelp reviews and restaurant revenues, a one-half star decrease in average star rating translates into an approximate 5% loss in revenue.

Secondary Analyses & Robustness Checks

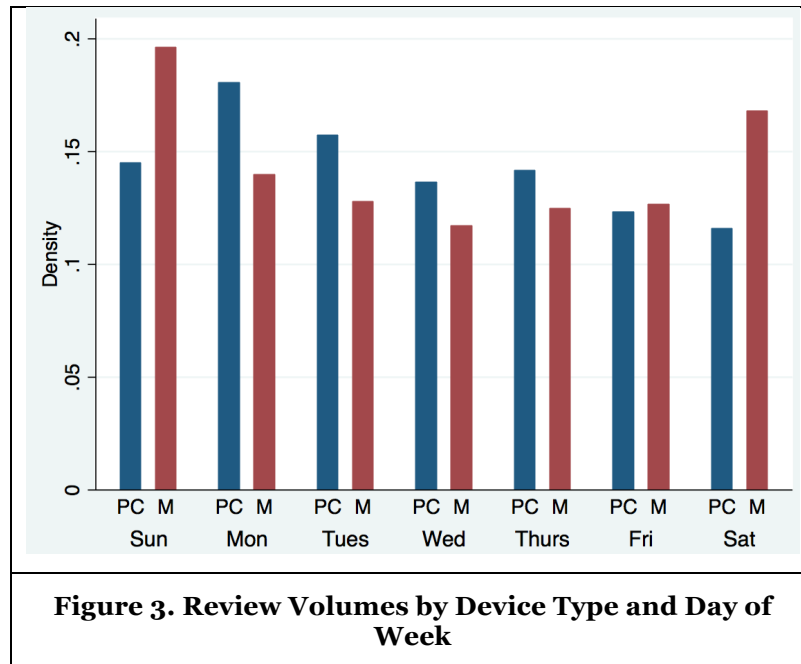
We began by looking at the overall distribution of review valence in our data, which is depicted as a histogram in Figure 2. Upon doing so, we observed something interesting. We saw an exponential distribution, rather than the J-shaped distribution that is frequently noted in the literature. In an effort to determine whether this observation was an aberration or a persistent feature of restaurant reviews, we examined the distribution of reviews for these same restaurants, on *Yelp* and *OpenTable*. Upon doing so, we observed the same pattern. Although not a focus of the present study, this finding is interesting in and of itself, as it suggests that the J-shaped distribution, which is frequently taken for granted in the literature, does not necessarily manifest for all types of products and services. The distribution appears to depend at least in part on the nature of the product or service in question, or perhaps on platform characteristics. More to the point, this finding suggests that extremity biases are less apparent in restaurant reviews. This provides some indication for why we find no evidence for a reduction in extremity biases, while we do find evidence of “ranting and raving” in the data – i.e., more emotional text and higher variance.



We next made an effort to understand whether our results are attributable to self-selection versus a causal effect of mobile device usage. To elaborate, our *Valence* model might be suffering from potential endogeneity (and reverse causality in particular). To clarify, the notion of ‘rants and raves’ suggests that a customer may be more likely to employ a mobile device following a very negative or positive experience. If the level of satisfaction of a consumption experience were to drive mobile use, while mobile use simultaneously has a causal effect on content, then reverse causality would be at play, resulting in a biased estimate of *Mobile’s* effects. Bearing this in mind, we next sought to address the potential endogeneity issue via instrumental variable regression. We focus upon our *Valence* model in these re-estimations.

We leverage available data on the date of the review and construct a binary indicator of whether the review was entered on a weekday or weekend (i.e., Saturday or Sunday). Our logic here is that mobile device usage is more likely on the weekend compared to PC usage, because consumers are more likely to be in transit at any given time on the weekend, as they go “out and about” to conduct errands, etc. Assuming this is the case, mobile usage will be more likely for weekend reviews, yet the day of the week should be unrelated to the valence or extremity of a restaurant dining experience (within reason).

First, we sought to establish that mobile device usage is indeed more common than PC device usage on weekends, and that the reverse is true on Weekdays. Figure 3 provides a clustered histogram by day of week for the volume of reviews arriving via each device type. We see quite strong support for our expectation, as mobile reviews are significantly more common on weekends than weekdays, and the reverse is true of PC reviews.



Next, we sought to directly test the notion that weekend reviews should be no different in terms of valence, helpfulness, concreteness or emotional content, once we control for *Mobile*. This is a falsification test of sorts, given that, if we identify a significant relationship between *Weekend* and any of our DVs when controlling for *Mobile*, this would provide direct evidence that our instrument is in fact invalid. However, upon executing these models (Table 4), we find no significant relationship between *Weekend* and the different dependent variables. This lends credence to our argument that the instrument is indeed exogenous, and thus valid.

Given this descriptive support for our instrument, we next proceed to our estimations. Table 5 presents a re-estimation of our *Valence* model employing our binary indicator, *Weekend*, as an instrument. Here, our estimate of the effect of *Mobile* on *Valence* remains significant and negative; mobile reviews continue to be more negative in star ratings than their desktop counterparts. Indeed, the estimate becomes larger in magnitude. This analysis suggests that our results are not severely biased by the presence of any endogeneity.

Discussion & Implications

These results provide an important first step toward understanding the role of mobile devices and mobile users in online WOM, in that they demonstrate stark differences in both content and influence. In a broad sense, it appears that offering mobile review channels is beneficial, in that mobile reviews appear more credible and likely to prove helpful to others (conditional, of course, on mobile users actually taking the time to craft informative content).

Table 4. Robustness Check: Exogeneity of Weekend Instrument				
Dependent Variable				
Variable	Valence	Concrete	Helpful	Emotion
Weekend	-0.004 (0.016)	-0.04 (0.05)	-0.01 (0.01)	0.09 (0.08)
Mobile	-0.06*** (0.01)	0.10* (0.04)	0.02* (0.01)	0.36*** (0.07)
Length	--	--	0.00*** (0.00)	--
Valence	--	--	-0.05*** (0.01)	--
Age	--	--	-0.00 (0.00)	--
Time Effects	Included	Included	Included	Included
Reviewer Effects	Included	Included	Included	Included
Restaurant Effects	Included	Included	Included	Included
Observations	23,827	23,827	23,827	23,827

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$; Estimates of fixed effects are jointly significant in all models, $p < 0.001$; Robust standard errors reported in brackets for coefficients; Models estimated with Stata's *felsdureg* command (Cornelissen 2009).

Table 5. Robustness Checks: 2SLS /w Three-Way FE	
Variable	Valence
Mobile	-0.18** (0.18)
Time Effects	Included
Reviewer Effects	Included
Restaurant Effects	Included
Observations	23,827

Notes: ** $p < 0.01$, + $p < 0.10$; Estimates of fixed effects are jointly significant in all models, $p < 0.001$; Robust standard errors reported in brackets for coefficients, corrected for two-stage estimation; Models estimated using Stata's *felsdureg* command (Cornelissen 2009); Weak instrument concern is alleviated as the Cragg-Donald Wald F statistic surpasses Stock and Yogo (2005)'s critical value of 16.38 (with Stata's *xtivreg2* command) of the first stage regression.

Our study contributes to a number of streams of literature in MIS and marketing. First, we build on the nascent body of work dealing with the mobile Internet. Our work takes the literature a step beyond leveraging geo-location data for the purposes of targeted advertising. We identify a number of important differences in the content contributed by mobile users (from mobile devices). Our findings therefore suggest that not only should marketers leverage location-based data on mobile users to improve advertising; rather, they should go a step further, attempting to infer context or user-level characteristics and preferences based on a user's historical device usage patterns and their present device choice.

Our work also offers insights for the literature dealing with electronic WOM. While some of our findings are merely confirmatory (e.g., shorter review lengths, consumption recency), others are not immediately apparent. In particular, our finding that, all else held equal, mobile reviews are more likely to prove helpful to other consumers has important implications. While on the surface one might expect that limitations of the mobile interface would result in thinner contributions, the increased access afforded by mobile devices appears to result in less delay prior to posting. Given this results, and prior findings that explicit indications of consumption recency are associated with greater perceived helpfulness and impact when it comes to reviews (Chen and Lurie 2013), our work demonstrates that the value of mobile reviews is a complex story. Although the user interface can certainly be limiting, once those limitations are overcome, significant value can be derived from mobile contributions.

This study provides implications for practitioners as well. First, to date, many online review platforms have taken the basic step of enforcing character or word count minimums to avoid uninformative content. *TripAdvisor* currently enforces a 100-character minimum, and *Yelp* takes a similar approach. Going forward, however, platforms might do more to encourage thoughtful content by providing heuristics and guidelines. For example, platforms might coach users by suggesting aspects of the restaurant experience for them to comment upon. More specifically, users might be prompted to discuss how the food tasted, the ambience of the restaurant, how noisy it was, etc. *Yelp* has taken some steps in this regard already, briefly displaying an example of a “good” review to mobile users before they are allowed to author their own reviews¹⁰.

Further, platforms might consider prompting mobile users to enter a review immediately after completing their dining experiences (e.g., leveraging location-based data to identify when a user departs a restaurant following check-in). Doing so could encourage references to consumption recency in review content. In addition, given our findings that ratings submitted over the mobile channel typically are of lower valence and are also likely to be more credible, online review platforms might look to implement appropriate algorithms to place different weights to ratings submitted from different devices, in order to maximize review accuracy and fairness.

Limitations & Opportunities

As with all studies, our work is subject to some limitations. Most notably, our data is observational. It therefore remains possible that a portion of our results actually derive from unobserved dynamic factors that our econometric specifications cannot address. Going forward, additional data could be leveraged to construct or identify more instrumental variables explaining mobile device usage, in order to better identify causal effects. Further, future research could expand upon the analyses herein to explicitly consider self-selection and population-based differences across certain segments of consumers. For example, it would be interesting to consider behavioral differences that manifest between mobile-only and desktop-only users, relative to those users who truly have a device choice. As noted earlier, an estimated 31% of American Internet users rely solely on a mobile device. It would therefore be interesting to explore this new segment of the consumer population in greater detail.

It is also worth pointing out that there is a nuanced distinction between whether mobile device use ‘causes’ different content to emerge, or whether mobile device use takes place when consumers have something different to say. This speaks again to the important potential role of self-selection, and thus the interpretation of our results. Although our data at present do not allow us to piece these two mechanisms apart, the short-term implications are essentially identical in each scenario. In short, in either case, the logical implication is to incentivize or encourage mobile use, or not, depending on whether the content consumers produce is desirable. That being said, the long-term implications may vary. If mobile device use causes the differences in content, it is entirely possible that these effects might grow weaker or stronger with time, as users learn and gain experience on the mobile device. However, if the observed differences were instead simply enabled by mobile access, then we would expect them to persist into the future and perhaps grow stronger as users become more accustomed to using the mobile channel.

¹⁰ <http://www.fastcompany.com/3027249/lessons-learned/how-yelp-encourages-users-to-write-more-thoughtful-reviews-even-on-mobile>

Conclusion

According to Pew, mobile Internet use is growing at a rapid pace relative to other channels¹¹. Understanding consumers' behavioral differences across mobile and non-mobile devices is of paramount importance for firms looking to formulate appropriate mobile strategies. This work presents what is to our knowledge a first step down the path toward an understanding of the behavioral differences between mobile and non-mobile users, in the context of online WOM. Our findings suggest that, on mobile devices, reviewers tend to write shorter reviews, and they post reviews more quickly after the consumption experience. Reviews authored via mobile devices are more likely to be extreme and emotional. At the same time, reviews entered via mobile devices are more likely to contain concrete text as well. Perhaps most importantly, we find that, all else held equal, mobile reviews prove more helpful to other consumers. This last result is the most striking, as it suggests that the content produced on mobile devices can actually be more valuable, if online review platforms can find feasible ways to encourage mobile reviewers to expend sufficient effort when they are contributing content. Our study calls for a more nuanced understanding of the mobile Internet and mobile user behavior.

References

- Anderson, E.W. 1998. "Customer Satisfaction and Word of Mouth," *Journal of Service Research* (1:1), pp. 5-17.
- Anderson, M., and Magruder, J. 2012. "Learning from the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database," *The Economic Journal* (122:563), pp. 957-989.
- Berger, J., and Schwartz, E.M. 2011. "What Drives Immediate and Ongoing Word of Mouth?," *Journal of Marketing Research* (48:5), pp. 869-880.
- Burch, G., Ghose, A., and Wattal, S. 2014. "An Empirical Examination of Peer Referrals in Online Crowdfunding " *International Conference on Information Systems (ICIS)*, Auckland, New Zealand.
- Byers, J.W., Mitzenmacher, M., and Zervas, G. 2012. "The Groupon Effect on Yelp Ratings: A Root Cause Analysis," *Working Paper*.
- Chen, Y., and Xie, J. 2008. "Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix," *Management Science* (54:3), pp. 477-491.
- Chen, Z., and Lurie, N.H. 2013. "Temporal Contiguity and Negativity Bias in the Impact of Online Word of Mouth," *Journal of Marketing Research* (50:4), pp. 463-476.
- Chevalier, J., and Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research* (43), pp. 345-354.
- Cornelissen, T. 2009. "The Stata Command Felsdsvreg to Fit a Linear Model with Two High-Dimensional Fixed Effects," *The Stata Journal* (8:2), pp. 170-189.
- Dai, W., Jin, G., Lee, J., and Luca, M. 2012. "Optimal Review Aggregation of Consumer Ratings: An Application to Yelp.Com," *NBER Working Paper*.
- Dellarocas, C. 2003. "The Digitization of Word-of-Mouth: Promise and Challenges of Online Feedback," *Management Science* (49:10), pp. 1407-1424.

¹¹ <http://www.pewinternet.org/2013/09/19/cell-phone-activities-2013/>

- Dellarocas, C., and Narayan, R. 2006. "A Statistical Measure of a Population's Propensity to Engage in Post-Purchase Online Word-of-Mouth," *Statistical Science* (21:2), pp. 277-285.
- Ghose, A., Goldfarb, A., and Han, S.P. 2013. "How Is the Mobile Internet Different? Search Costs and Local Activities," *Information Systems Research* (24:3), pp. 613-631.
- Ghose, A., and Han, S. 2011. "An Empirical Analysis of User Content Generation and Usage Behavior on the Mobile Internet," *Management Science* (57:9), pp. 1671-1691.
- Godes, D., and Mayzlin, D. 2004. "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science* (23:4), pp. 545-560.
- Godes, D., and Mayzlin, D. 2009. "Firm-Created Word-of-Mouth Communication: Evidence from a Field Test," *Marketing Science* (28:4), pp. 721-739.
- Goes, P., Lin, M., and Yeung, C. 2014. "'Popularity Effect' in User-Generated Content: Evidence from Online Product Reviews," *Information Systems Research*, Articles in Advance.
- Goh, K.Y., Chu, C., and Soh, W. 2009. "Mobile Advertising: An Empirical Study of Advertising Response and Search Behavior," in: *30th International Conference on Information Systems (ICIS)*. Phoenix, AZ.
- Hong, Y., Chen, P.-Y., and Hitt, L.M. 2012. "Measuring Product Type with Dynamics of Online Product Review Variance," *Proceedings of the 33rd International Conference on Information Systems (ICIS)*, Orlando, Florida.
- Hu, N., Zhang, J., and Pavlou, P.A. 2009. "Overcoming the J-Shaped Distribution of Product Reviews," *Communications of the ACM* (52:10), pp. 144.
- Jung, J., Umyarov, A., Bapna, R., and Ramaprasad, J. 2013. "Love Unshackled: The Causal Effect of Mobile App Adoption in Online Dating," in: *Workshop on Information Systems and Economics (WISE)*. Milan, Italy.
- Katz, E., and Lazarsfeld, P.F. 1955. *Personal Influence: The Part Played by People in the Flow of Mass Communications*. New York: Free Press.
- Kwark, Y., Chen, J., and Raghunathan, S. 2014. "Online Product Reviews: Implications for Retailers and Competing Manufacturers," *Information Systems Research* (25:1), pp. 93-110.
- Li, M., Huang, L., Tan, C., and Wei, K. 2013. "Helpfulness of Online Product Reviews as Seen by Consumers: Source and Content Features," *International Journal of Electronic Commerce* (17:4), pp. 101-136.
- Lieberman, N., Trope, Y., and Wakslak, C. 2007. "Construal Level Theory and Consumer Behavior," *Journal of Consumer Psychology* (17:2), pp. 113-117.
- Loewenstein, G. 1996. "Out of Control: Visceral Influences on Behavior," *Organizational Behavior and Human Decision Processes* (65:3), pp. 272-292.
- Loewenstein, G. 2000. "Emotions in Economic Theory and Economic Behavior," *The American Economic Review* (90:2), pp. 426-432.
- Luca, M. 2011. "Reviews, Reputation, and Revenue: The Case of Yelp.Com," *HBS Working Paper*.

- Luca, M., and Zervas, G. 2013. "Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud," *HBS Working Paper*. Luo, X., Andrews, M., Fang, Z., and Phang, C. 2014. "Mobile Targeting," *Management Science*, Forthcoming.
- Moe, W.W., and Schweidel, D.A. 2012. "Online Product Opinions: Incidence, Evaluation, and Evolution," *Marketing Science* (31:3), pp. 372-386.
- Molitor, D., Reichhart, P., Spann, M., and Ghose, A. 2014. "Measuring the Effectiveness of Location-Based Advertising: A Randomized Field Experiment," *Working Paper*.
- Mudambi, S.M., and Schuff, D. 2010. "What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.Com," *MIS Quarterly* (34:1), pp. 185-200.
- Nagle, F., and Riedl, C. 2014. "Online Word of Mouth and Product Quality Disagreement," *HBS Working Paper*.
- Noulas, A., Scellato, S., Mascolo, C., and Pontil, M. 2011. "An Empirical Study of Geographic User Activity Patterns in Foursquare," in: *AAAI Conference on Weblogs and Social Media*. Barcelona, Spain: pp. 570-573.
- Pennebaker, J.W., Francis, M.E., and Booth, R.J. 2001. *Linguistic Inquiry and Word Count: Liwc 2001*. Mahwah, NJ: Lawrence Erlbaum.
- Ritchie, T., Batteson, T., Bohn, A., Crawford, M., Ferguson, G., Schrauf, R., Vogl, R., and Walker, W.R. 2014. "A Pancultural Perspective on the Fading Affect Bias in Autobiographical Memory," *Memory*, pp. 1-13.
- Shankar, V., and Balasubramanian, S. 2009. "Mobile Marketing: A Synthesis and Prognosis," *Journal of Interactive Marketing* (23:2), pp. 118-129.
- Shankar, V., Venkatesh, A., Hofacker, C., and Naik, P. 2010. "Mobile Marketing in the Retailing Environment: Current Insights and Future Research Avenues," *Journal of Interactive Marketing* (24:2), pp. 111-120.
- Shi, N., Hong, Y., Wang, K., and Pavlou, P.A. 2013. "Social Commerce Beyond Word of Mouth: Role of Social Distance and Social Norms in Online Referral Incentive Systems," *Proceedings of the 34th International Conference on Information Systems (ICIS), Milan, Italy*.
- Sridhar, S., and Srinivasan, R. 2012. "Social Influence Effects in Online Product Ratings," *Journal of Marketing* (76:5), pp. 70-88.
- Stock, J.H., and Yogo, M. 2005. "Testing for Weak Instruments in Linear Iv Regression," In D.W.K. Andrews and J.H. Stock, eds. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge: Cambridge University Press, 2005, pp. 80-108.
- Sultan, F., Rohm, A.J., and Gao, T. 2009. "Factors Influencing Consumer Acceptance of Mobile Marketing: A Two-Country Study of Youth Markets," *Journal of Interactive Marketing* (23:4), pp. 308-320.
- Sun, M. 2012. "How Does the Variance of Product Ratings Matter?," *Management Science* (58:4), pp. 696-707.
- Yin, D., Bond, S., and Zhang, H. 2014. "Anxious or Angry? Effects of Discrete Emotions on the Perceived Helpfulness of Online Reviews," *MIS Quarterly*, Forthcoming.

Zhao, M., and Xie, J. 2011. "Effects of Social and Temporal Distance on Consumers' Responses to Peer Recommendations," *Journal of Marketing Research* (48:3), pp. 486-497.