READING BETWEEN THE LINES:
THREE INVESTIGATIONS OF USER GENERATED CONTENT
USING TEXT ANALYTICS

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

---

by
Ni (Nina) Huang
May 2017

Examining Committee Members:

Susan Mudambi, Advisory Chair, Marketing & Supply Chain Management Department
Paul Pavlou, Examining Chair, Management Information Systems Department
Gordon Burtch, Information & Decision Sciences Department, University of Minnesota
Nathan Fong, Marketing & Supply Chain Management Department
Eric Eisenstein, Marketing & Supply Chain Management Department
Hilal Atasoy, External Examiner, Accounting Department

ABSTRACT

User generated content (UGC) is a ubiquitous phenomenon on the Internet. UGC inform, entertain, and facilitate conversations among online users. The three essays of this dissertation examine different antecedents of UGC characteristics with text analytics. The first essay explored the effects of psychological distance on UGC positivity and found that spatial and temporal distance boost UGC positivity. The second essay investigates the effects of social media integration on the linguistic characteristic of UGC and showed that social media integration leads to increased review quantity, while more emotional, less rational and less negative language in UGC content. The third essay examines the impact of book-to-film adaptation on the rating and linguistic characteristics of UGC. The results suggest that, after the release of book-to-film adaptations, book ratings decline, and the use of language reflecting viewing, comparison and affective processes increase in book reviews. To summarize, the three essays in this dissertation contributes to research on UGC by improving our understanding on the various antecedents of UGC characteristics.

DEDICATION

My sincere gratitude goes to my dissertation advisory chair Professor Susan Mudambi and examining chair Professor Paul Pavlou for their continuous support of my Ph.D. study. I also thank my dissertation committee members: Professor Gordon Burtch, Professor Nathan Fong, Professor Eric Eisenstein, and Professor Hilal Atasoy for their insightful comments and constructive suggestions. Lastly, I dedicate this dissertation to my loving family: my husband Yili (Kevin) Hong and my son Evan N. Hong.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1.

INTRODUCTION

User generated content (UGC) has become an important aspect of the online experience of consumers (eMarketer 2009). It is reported that 1.6 billion Internet users spend an average of 5.4 hours per day with UGC (Caramela 2016). UGC involves in consumers actively sharing their experiences, opinions, and thoughts on the Internet (Ghose and Han 2011). Taking the forms of online reviews, online referral, online community discussions, social media interactions, and "other consumer-initiated contributions," UGC is an interesting research topic that has attracted tremendous attention from scholars in marketing (Chen and Xie 2008; Chen and Lurie 2013), information systems (Mudambi and Schuff 2010; Pavlou and Dimoka 2006), strategy (Kovács and Sharkey 2014), economics (Mayzlin et al. 2014) and psychology (Snefjella and Kuperman 2015).

Much work has focused on examining various consequences of UGC, such as product sales (Chevalier and Mayzlin 2006; Zhu and Zhang 2010), business intelligence (Lee and Bradlow 2011; Mallapragada et al. 2012), firm financial valuation (Nam and Kannan 2014), investment decisions (Park et al. 2014), consumer purchases (Chen et al. 2011; You et al. 2015), etc. However, limited research has considered the antecedents of UGC characteristics (Li and Hitt 2008; Goes et al. 2014). Further, with the growing need from both academia and industry to gain greater insights from UGC (Fader and Winer 2012), it is necessary for researchers to use new methodology, for instance, text analytics to better utilize UGC data (Ghose et al. 2012; Netzer et al. 2012). Bearing the above in mind, these three essays aim to investigate different antecedents of UGC characteristics using text analytics.

In the first essay, I study the effects of psychological distances, specifically spatial and temporal distance, on construal level and review positivity in the context of UGC. Building on construal level theory (CLT), I hypothesize the separate and joint effects of spatial distance and temporal distance on construal level and review positivity. I then empirically test the hypotheses by analyzing a large-scale panel data of 166,215 restaurant reviews that were collected from *TripAdvisor.com*. Employing content analysis and multiple econometric methods, I find that both spatial and temporal distance positively affect construal level and review positivity in UGC. Moreover, I find a distance boosting effect, in which simultaneously experiencing spatial and temporal distances would amplify consumers' construal level and review positivity. This study contributes not only to the CLT literature by presenting evidence of the distance boosting effect, but also to the UGC literature by demonstrating spatial and temporal distances as antecedents of UGC characteristics.

In the second essay, I examine the impact of social network integration on UGC characteristics. Grounding on social presence theory, I propose several hypotheses on the effects of social network integration on review volume and linguistic characteristics of review content. The data of this study contain online reviews for a matched set of restaurants across two leading online review platforms, namely *Yelp.com* and *TripAdvisor.com*. The linguistic characteristics of review content were measured by using Linguistic Inquiry and Word Count (LIWC), a text analytics software. I exploit two natural experiments of social network integration at the two online review platforms to test the hypotheses. Estimating a difference-in-difference (DID) model, I find that social network integration positively affects review volume, increases affective process, decreases cognitive process, and reduces negations in review content. These findings suggest that social network integration works as a double-edged sword, providing benefits in terms of review quantity, while at the cost of review quality. This study contributes to the literature on the design of online review systems as well as the literature on UGC by considering social network integration as a system design feature and an important antecedent that affects review volume and linguistic characteristics of review content.

The third essay empirically investigates the impact of book-to-film adaptations on the rating and linguistic characteristics of user content generation on books. We collected a unique dataset collected from *Amazon.com*, *Goodreads.com* and *IMDB.com*. Our results show that book-to-film adaptation would lead to decreases in book rating and increase the use of language reflecting viewing, comparative language and affective processes in book reviews. We attributed the decreases in book rating to the publicity effect of book-to-film adaptation. In addition, we argued that the changes in the viewing, comparative languages, and affective processes were driven by the viewing, comparative, and emotional spillover effect of book-to-film adaptation. We then conducted additional analyses on the robustness of the main effects as well as the impact of the review changes on review helpfulness. Our results show that book review might have become more helpful after the release of book-to-film adaptation because the decreased star rating, as well as increased viewing and affective language in book reviews after film adaptation is positively related to review helpfulness. Films adaptations, through multiple mechanisms, are making book reviews more helpful.

To conclude, the three essays of this dissertation proposal are connected through a common theme of investigating different antecedents of UGC characteristics using text analytics. With each essay filling a unique gap in the UGC literature, this research offers important contributions to research on consumer generated content and the practice of online marketing management.

CHAPTER 2.

ESSAY1: EFFECTS OF MULTIPLE PSYCHOLOGICAL DISTANCES ON CONSTRUAL AND CONSUMER EVALUATION: A FIIELD STUDY OF ONLINE REVIEWS

Abstract

Through a large-scale field study of 166,215 online restaurant reviews, we found evidence of a *distance boosting effect*, whereby experiencing spatial distance (i.e., authoring a review about a geographically distant restaurant, rather than proximate one) and temporal distance (i.e., authoring a review after a lengthy delay, rather than immediately) amplified consumers' high-level construals. Although past research has explored the relationship between spatial distance, temporal distance, and construal, these effects have only been considered in isolation (on a notable range of outcomes), yet never in tandem. Our research contributes to past work by testing the effects of experiencing two dimensions of psychological distance simultaneously on construal level, and on a downstream consequence thereof: positivity. Moreover, because our data contain naturalistic observations, our research includes a wide range of temporal and spatial distances. In all, we found that the effect of each distance increases the effect (on construal and positivity) of the other distance. Metaphorically speaking, the effect of one distance is boosted by another.

Introduction

Mark Twain is famous for saying, “distance lends enchantment to the view,” attesting, anecdotally, that faraway things will appear better than they are in reality. Broadly speaking, Mark Twain observed a pattern that has interested scholars in psychology and marketing for the last six decades – the effects of feeling something as closer or farther relative to the self (Lewin, 1951). In more recent times, research on feelings of (greater or lesser) closeness forms part of a large body of work known as construal level theory (CLT), which has theorized and found that psychological distance changes people’s mental representations of events (Liberman & Trope, 1998). Specifically, distant events (e.g., those taking place in the past, or in faraway places) are more likely to be represented in high-level terms of abstract and central features; whereas close events (e.g., events in places and at times that are here and now, respectively) are more likely to be represented in low-level terms of concrete and peripheral features (Trope & Liberman, 2000, 2003).

Informed by theories of psychological distance and CLT, the present research examined whether the passing of time and space *jointly* affect online review favorability. Notably, our focus is on two dimensions of psychological distance (time and space), which addresses a broad, theoretical question: Does experiencing multiple dimensions of psychological distance of an event, simultaneously, lead to different event-appraisals compared to when only a single dimension of psychological distance is experienced? Research on psychological distance has identified four different dimensions of distance: temporal-, spatial-, social-, and certainty-distance, yet despite the extensive research on these dimensions' role in a varied range of outcomes (e.g., in decision making, persuasion, negotiation, creativity, metacognition, self-control, and memory; for a review see Liberman, Trope, & Wakslak, 2007), little research has examined the influence of experiencing more than one dimension of distance simultaneously.

By adopting a multiple-dimensions approach to distance, the current research extends the literatures on psychological distance and CLT in both theoretical and empirical terms. Research has shown that distant events generate high-level construals. However, what if an event will take place in the *future*, and also in a *faraway* place? Would this scenario generate even *more* high-level construals? As a specific example, does the level of construal that results when considering an event that will take place 5,000 miles away grow even greater when a second distance is introduced, such as if the event were also to take place 5 years into the future (a case of two remote distances)?

It is this question that the current research helps address – what the effect is of experiencing multiple distances, at different levels on construal level and, subsequently, the evaluative judgments people generate of events. Our investigation enables us to examine the cross-dimensional effects of multiple distances (i.e., events that are close in time yet far in space; or far in time yet close in space) in a relatively nuanced and precise way: Owing to our novel dataset (containing a sample of 166,215 restaurant reviews) we are able to examine distance *continuously* (as opposed to categorically). By that we mean, in a departure from past research on psychological distance that typically compares one researcher-created discretionary, categorical level of distance with another level (e.g., comparing an event happening in "1 year" with an event happening "tomorrow"), we examined a wide range of naturalistic temporal and spatial distances, from 0 to 11 months and 0 to 11,910 miles respectively. Besides testing the separate effects of each distance across its range, our study design enables us to test our main focus: the effects of experiencing multiple distances in tandem – specifically, the interactive effect of different levels of temporal distance against a range of spatial distances, and vice versa.

We begin with a theoretical framework describing the effects of psychological distance on construal level, and argue that positive appraisals of past events are a result of generating high-level construals. Then, we shift our discussion from the effects of experiencing one dimension of distance to the effects of experiencing multiple dimensions, to advance our prediction that, compared with experiencing one distance, experiencing two distances fosters more high-level construals, and hence more positive event-appraisals.

Psychological Distance, CLT, and Positivity

CLT suggests that individuals process events at higher levels of construal when they perceive greater psychological distance between themselves and those events, where psychological distance is defined as an "egocentric" perception of "the different ways in which an object might be removed" from "the reference point of self in here and now" (Trope & Liberman, 2010, p. 440). Accordingly, theorizing on psychological distance has identified four means by which targets can drift in distance: in time, from now to earlier or later (Trope & Liberman, 2000); in space, from here to another place (Fujita, Henderson, Eng, Trope, & Liberman, 2006); in social contexts, from the self to others (Polman, 2012); and in probability, from certain to uncertain (Wakslak, Trope, Liberman,

& Alony, 2006). Built on the notion that abstraction is required for predicting and planning for what is not present, CLT proposes that psychological distance influences how individuals mentally construe events. Under high construal, individuals tend to think in an abstract, decontextualized manner; whereas under low construal, individuals tend to think concretely and focus on contextual details, which are often peripheral or incidental.

Several studies hint at the possibility of a positivity bias when evaluating under high-level construal. Research has found that the positive aspects of an experience are more salient with high-level construal, such that when led to think with a high-level construal mindset, people contemplate more pros in favor of an action than cons (Eyal, Liberman, Trope, & Walther, 2004). Consistent with this view, research has also found that pros are easier to think of when considering temporally distant versus close events (Herzog, Hansen, & Wänke, 2007). Similarly, Williams, Stein and Galguera (2014) found that past experiences felt more pleasant when they were described with a high-level construal, compared to a low-level construal. On the basis of these findings, we expect that psychologically distant (vs. closer) events foster more high-level construals (vs. low-level construals), which we predict makes appraisals of past events more postive. We test this hypothesis in the context of online restaurant reviews, and predict that (a) the more time between visiting a restaurant and reviewing it (temporal distance) and, separately, (b) the more space between the location of the restaurant and the location the reviewer (spatial distance), the higher the construal of the review, and hence the more positive the review.

Multiple Distances

Moving beyond the standard focus on psychological distance and its effects on construal level, we turn to research on multiple dimensions of distance and argue that experiencing more than one dimension of distance simultaneously will induce higher-level construal, and thus transform the downstream consequences of distance on positivity. Though each distance has been found to have a separate similar effect on construal – where more distance leads to higher levels of construal (Liberman, Sagristano, & Trope, 2002), no research has examined the level of construal that results under conditions of multiple distances. Research concerning multiple distances has thus far focused on two areas: (1) fit, which describes synchronizing levels of distance, such that when a remote (close) distance is aligned with another remote (close) distance, the choice feels "right" and people are faster to make choices and more susceptible to persuasion (Bar-Anan, Trope, Liberman, & Algom, 2007; Kim, Zhang, & Li, 2008; Zhao & Xie, 2011); and (2) distance-on-distance, which describes how a feeling of distance on one dimension serially influences how far another dimension of distance feels (Kim, Zauberman, & Bettman, 2012; Maglio & Polman, 2014; Maglio, Trope, & Liberman, 2013; Wakslak, 2012; Williams & Bargh, 2008; Yan, 2014; Zhang & Wang, 2009).

Our research question illuminates a third, new area: simultaneous distance, which describes how consumers behave when more than one distance is experienced at one time. We propose a *distance boosting effect*, whereby experiencing two distances amplifies consumers' high-level construals. That is to say, using restaurant reviews as an example, we predict that high-level construals will be greater when a restaurant patron reviews a faraway restaurant that she visited a long time ago, compared to solely (separately): a faraway restaurant or a restaurant she visited a long time ago. Although direct evidence for a distance boosting effect has not yet been reported in the literature, the findings from the two existing areas noted above (fit and distance-on-distance) are consistent with our prediction. Based on the fit literature, support for distance boosting comes from research that shows that participants preferred products that were described in positive high-level (vs. positive low-level) terms when they received cues to instantiate at least one remote level of either temporal or social distance (Kim et al., 2008). Specifically, participants' preferences aligned with what reliably follows from experiencing a remote level of distance: preference for the product described in high-level terms (cf. Liberman & Trope, 1998) – this is despite the fact that participants received cues with one remote and one close level of distance, as if a remote distance subsumes a closer one and prioritizes judgment. In other words, a remote distance does not appear to shrink (or get averaged) when it is misaligned (mismatched) with a close distance.

In fact, based on the distance-on-distance literature, the opposite is true – distance instead expands: feeling like one dimension of distance is remote subsequently makes all other dimensions feel remote (Yan, 2014). Indeed, the central thesis of research on psychological distance is that judgments of distance are subjective, and a range of distance-on-distance findings has found that priming participants with one remote (vs. close) level of distance makes other dimensions of distance feel more remote in kind (Kim et al., 2008, 2012; Wakslak, 2012; Zhang & Wang, 2009). Thus, previous research shows that judgments of distance grow according to how far a previous judgment of distance feels: the bigger the initial distance, the bigger the subsequent distance. While this past work has emphasized how experiencing a subjective level of distance can adjust judgment of another distance (in a serial fashion), the present investigation makes a distinct point about construal level to suggest that when two distances are configured together (experienced simultaneously), construal level will also increase – that is to say that in cases of two distances, each distance boosts the construal level of the other, leading to more high-level construals and consequences thereof.

Overview

To summarize, we aim to improve upon the understanding of the relationship between psychological distance and evaluations by exploring the joint effects of temporal and spatial distance on the content of online reviews (rating favorability). In this vein, we carried out a field study and examined the association between (spatial and temporal) distance and rating favorability in a large sample of *TripAdvisor* restaurant reviews. Further, to measure construal level, we coded the text of a sub-sample of the reviews using the Linguistic Categorization Model. The coded construal levels, in turn, were used in a mediation analysis to test CLT as the primary mechanism underlying the observed associations. In all, we predicted that construal level mediates the effects of temporal and spatial distance on rating favorability. Moreover, we predicted a positive interaction effect that would evidence our distance boosting account, whereby the effects of temporal and spatial distance (on construal level and rating favorability) amplify one another.

Our work attempts to replicate past work on psychological distance, using a high-powered large-scale field study. In light of the growing demand for independent direct replications of findings (Pham, 2013), our work not only heeds that demand, it also contributes to the extensive body of literature on psychological distance, a subject of substantial value and broad theoretical importance, by illuminating new theoretical aspects of experiencing multiple psychological distances on construal level, and the downstream consequences thereof. In this vein, our data, which contain 166,215 restaurant reviews, constitute what is to our knowledge the largest dataset concerning psychological distance and CLT ever assembled. Furthermore, we employed econometric analyses, which bring to bear a study design and methodological approach that is less common in consumer psychology. Because experimental methods dominate consumer psychology, a recent call by consumer psychologists has suggested that researchers test hypotheses with other methods (Lynch, Alba, Krishna, Morwitz, & Gürhan-Canli, 2012).

Methodology

Data

We combined two different data sources for our empirical analyses. First, we obtained publicly available data on restaurant reviews from *TripAdvisor*, spanning the period between 2003 and 2014. We constructed a panel, where each observation captures an online review, with a review date, the identity and characteristics of the reviewer, as well as the restaurant. Our data thus incorporate repeated observations across reviewers (who may author reviews about multiple restaurants) and restaurants. The body of reviews we consider pertains to a random set of restaurants located in seven major cities throughout the United States: Chicago, Houston, Los Angeles, New York, Phoenix, Philadelphia, and Seattle.

We next obtained geographic coordinates for each restaurant, based on its address; and for each reviewer, based on his or her self-reported city of residence. The coordinates (latitude and longitude) were obtained via geocoding, using the *Google Maps* Application Programming Interface (API). We manually cleaned the data to deal with some cases where acronyms of cities were used (for example, "LA" rather than Los Angeles), before performing the geocoding process (note: these observations constitute less than 1% of the sample).

Key Measures

Review Favorability. Review favorability is operationalized as the star rating of the review. This variable takes on positive integer values between 1 and 5. More stars indicate more favorability.

Temporal Distance. Temporal distance is defined as the delay between a dining experience and the consumer's submission of a review. On *TripAdvisor*, both the date of the review submission and the month in which the reviewer's relevant dining experience took place are publicly available. We therefore calculated the number of months between the month of the consumption date and the review date. We log-transformed this measure to achieve normality, because of the skewness of the variable (*skewness* = 74.626, $p < 0.001$).

Spatial Distance. Spatial distance is defined as the geographic distance between the location of the reviewed restaurant and the reviewer's place of residence (in miles), based on the geodesic distance between the pair of geographic coordinates. This distance was calculated using the equation for great-circle distance: the shortest distance between two points on the surface of a sphere (Vincenty, 1975). We log-transformed the resulting distance measure to achieve normality, again because of the skewed distribution ($skewness = 1.847$, $p < 0.001$).

As standard procedure, we mean-centered our two distance variables before constructing the interaction term (Cohen, Cohen, West, & Aiken, 2003). For variables that contain zeroes, we added a value of 1 (the lowest non-zero value) before applying the log-transformation (McCune, Grace, & Urban, 2002).

Control Variables. Control variables were included to account for time, reviewer, and restaurant fixed effects (Wooldridge, 2002, p. 265). The time effect was captured via the "review month" variable, a vector of dummy indicators reflecting the year and month in which a review was submitted. Our complete set of control variables is identified in Table 1. Specifically, we controlled for mobile device usage because consumers who submit reviews from mobile devices may systematically author reviews under lower temporal distance (given that they can author a review immediately, on the spot). Mobile reviews may also exhibit systematically different characteristics for other reasons (e.g., device-specific user interface differences), thus it is important to account for this variable (Burtch & Hong, 2014). We included a binary indicator of whether the user was an American Express (*Amex*) credit card holder because these individuals may be systematically different from other consumers in ways that may influence the content of their reviews (e.g., Frankel, 2014; Amex users may exhibit higher average levels of wealth, and thus exhibit lower price sensitivity). We controlled for the reviewer's average star rating across all the reviews they have provided, in order to account for the fact that individuals may be systematically positive or negative in their opinions, on average (Goes, Lin, & Yeung, 2014). We controlled for the reviewer's prior volume of reviewing activity for similar reasons, because past reviewing experience has the potential to influence subsequent evaluation behavior. We controlled for a restaurant's average rating as a proxy for restaurant quality, which influences the ratings it receives. We controlled for the lowest and highest item prices appearing on the restaurant's menu, again to account for the possible role of price sensitivity, and as an additional proxy for unobserved restaurant quality. We controlled for the restaurant's "dining style," because consumers may provide systematically different ratings in a city, depending on the local population's experience with, or preference for that dining style. Finally, we controlled for the restaurant's accessibility to public transportation, because accessibility may influence the customer segments who choose to frequent the restaurant, and thus may produce systematic differences in consumer ratings.

Table 1. Control variables

Variable	Description
Mobile	A binary variable that measures whether a consumer submitted his or her review from a mobile device.
Amex User	A binary variable that measures whether a consumer has associated an American Express (Amex) credit card with his or her *TripAdvisor* profile.
Reviewer Average Rating	The average of all past ratings of a reviewer.
Restaurant Average Rating	The average of all past ratings of a restaurant.
Lowest Price	The lowest price of the listed menu in a restaurant.
Highest Price	The highest price of the listed menu in a restaurant.
Reviewer Status	A categorical variable that indicates the levels of reviewer status on *TripAdvisor*.
Dining Style	A categorical variable that classifies restaurants into different dining styles.
Public Transportation	A binary variable that measures whether a restaurant is accessible to public transportation.

In all, we estimated multiple models to demonstrate the robustness of our results. In some models, reviewer or restaurant static characteristics are not identified because of the incorporated reviewer or restaurant fixed effects. The summary statistics and correlation matrix for all of our variables are reported in Table 32, in the Methodological Details Appendix (A).

Results

Our identification strategy relies on the application of three-way fixed effects: for the consumer, restaurant, and time of the review submission. Additionally, we employed propensity score matching (PSM) and exact covariate matching as robustness checks, in order to address any remaining endogeneity concerns. For our outcome variable (rating favorability), we examined the effects of spatial and temporal distance using alternative estimators (*viz*. ordinary least square regression, ordered logistic regression, one-way fixed effects, and two-way fixed effects). The various models demonstrate robustness to different model specifications. However, our three-way fixed effect regressions are the most conservative because these estimations control for static unobservable heterogeneity across time, reviewers, and restaurants, while simultaneously leveraging our entire sample. Following the procedure proposed by Cornelissen (2008), we employed the following econometric specification, which corresponds to the three-way fixed effects models with interactions and time dummies (see final column of Table 2):

(1) $$Rating\_favorability_{ijt} = \boldsymbol{\alpha} * TD_{ij} + \boldsymbol{\beta} * SD_{ij} + \boldsymbol{\gamma} * (TD * SD)_{ij} + \delta_i + \sum_J \lambda_j * R_j + \sum_T \tau_t * M_t + Control_{ijt} + \varepsilon_{ijt}$$

In Equation (1), TD and SD are our indicators of temporal and spatial distance, respectively. Subscript *i* indexes consumers, *j* indexes restaurants, and *t* indexes time. In addition, δ_i represents a vector of consumer fixed effects, R_j a vector of restaurant fixed effects, and M_t a vector of month fixed effects. Lastly, $Control_{ijt}$ represents our set of control variables.

Table 2 presents our regression results. In support of our predictions, we found that temporal distance is positively and significantly related to rating favorability (α = .024, $p < .01$), and that spatial distance is positively and significantly associated with rating favorability (β = .021, $p < .01$).

Next, we tested our main prediction concerning the distance boosting effect. As Table 2 shows, we found that temporal distance and spatial distance are positively and significantly related to rating favorability ($\gamma = .006$, $p < .05$). That is, we found that temporal and spatial distance positively moderate each other's effects on rating favorability, amplifying one another.

In sum, we found support for the role of two different dimensions of distance in influencing review favorability. Moreover, our test demonstrates a novel finding, by showing that when two distances are experienced *at the same time*, each distance boosts the effect of the other (on review favorability) – as evidenced by the positive, significant interaction. In other words, the influence of one dimension of psychological distance on review favorability increases not only when its *own* units increase (e.g., with temporal distance, when more time passes), but when an entirely *separate* dimension of distance also increases.

Table 2. Results for rating favorability (DV = star rating)

	(1) OLS		(2) Ordinal Logit		(3) Restaurant FE		(4) Reviewer FE		(5) Three-way FE	
	Main	Interaction	Main	Interaction	Main	Interaction	Main	Interaction	Main	Interaction
Temporal D	0.047***	0.047***	0.105***	0.104***	0.046***	0.045***	0.023***	0.032***	0.017**	0.024***
	(0.004)	(0.004)	(0.009)	(0.009)	(0.004)	(0.004)	(0.009)	(0.009)	(0.008)	(0.009)
Spatial D	0.013***	0.013***	0.026***	0.026***	0.016***	0.016***	0.017***	0.018***	0.020***	0.021***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)
Temporal D X Spatial D		0.004***		0.013***		0.005***		0.008***		0.006**
		(0.001)		(0.003)		(0.001)		(0.003)		(0.003)
Mobile	-0.023**	-0.022*	-0.023	-0.019	-0.029**	-0.027**	-0.077***	-0.075***	-0.086***	
	(0.012)	(0.012)	(0.026)	(0.026)	(0.012)	(0.012)	(0.026)	(0.026)	(0.027)	
Amex User	0.020**	0.021**	-0.010	-0.009	0.020**	0.020**				
	(0.008)	(0.008)	(0.019)	(0.019)	(0.008)	(0.008)				
Reviewer Average Rating	0.896***	0.896***	2.109***	2.109***	0.899***	0.899***				
	(0.006)	(0.006)	(0.014)	(0.014)	(0.009)	(0.009)				
Restaurant Average Rating	0.631***	0.631***	1.413***	1.414***			0.699***	0.699***		
	(0.007)	(0.007)	(0.016)	(0.016)			(0.012)	(0.012)		
Lowest Price	0.000	0.000	0.001	0.001			0.001***	0.001***		
	(0.000)	(0.000)	(0.001)	(0.001)			(0.000)	(0.000)		
Highest Price	0.000*	0.000	0.001***	0.001***			0.000***	0.000***		
	(0.000)	(0.000)	(0.000)	(0.000)			(0.000)	(0.000)		
Public Transportation	-0.014***	-0.015***	-0.056***	-0.057***			-0.004	-0.004		
	(0.004)	(0.004)	(0.010)	(0.010)			(0.008)	(0.008)		
Constant	-2.298***	-2.299***			0.337***	0.337***	1.147***	1.148***		
	(0.056)	(0.056)			(0.061)	(0.061)	(0.096)	(0.096)		
Reviewer Status Dummies	Yes	Yes	Yes	Yes	Yes	Yes	--	--	--	--
Dining Style Dummies	Yes	Yes	Yes	Yes	--	--	Yes	Yes	--	--
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reviewer Fixed Effects	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Restaurant Fixed Effects	No	No	No	No	Yes	Yes	No	No	Yes	Yes
R-squared	0.210	0.211	0.100	0.103	0.160	0.161	0.070	0.071	--	--
Observations	166,215	166,215	166,215	166,215	166,215	166,215	166,215	166,215	166,215	166,215
Number of reviewers	--	--	--	--	--	--	104,554	104,554	104,554	104,554
Number of restaurants	--	--	--	--	2,593	2,593	--	--	2,593	2,593

Notes: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$; Robust standard errors are in parentheses. Within R-squared is reported for fixed effect models.

Construal Level Content Analysis

To identify the theoretical mechanism underlying the distance boosting effect, we coded textual content from a sub-sample of online reviews. We employed the widely used Linguistic Categorization Model (Semin & Fiedler, 1988, 1991; Fujita et al., 2006; Trope, Liberman, & Wakslak, 2007) to code construal level. The Linguistic Categorization Model (LCM) is a five-level classification scheme that distinguishes between linguistic terms to measure the abstractness of content (Semin & Fiedler, 1991; Coenen, Hedebouw, & Semin, 2006). More abstract content indicates higher levels of construal (Semin & Smith, 1999; Fujita et al., 2006). The five levels of linguistic terms that measure increasing levels of abstractness include (Coenen et al., 2006, pp. 6-7): descriptive action verbs (DAV), interpretive action verbs (IAV), state action verbs (SAV), state verbs (SV) and adjectives (ADJ). DAV reflect the lowest level of abstract construal; this includes verbs that describe a specific action with at least one "physically invariant feature" (e.g., walk, speak, punch). IAV include verbs that are not physically invariant and contain an evaluative component (e.g., help, escape, praise). SAV include verbs that express the "emotional consequence of an action" (e.g., surprise, amuse, satisfy). SV include verbs that capture the "enduring cognitive or emotional state" (e.g., love, hate, appreciate). Finally, ADJ, which indicate the highest level of construal, include adjectives that describe a "characteristic or feature" of an entity (e.g., honest, nice, excellent).

The content coding was performed as follows (note: the complete details of the content coding procedure are provided in the Methodological Details Appendix). First, the various linguistic terms were identified in the text of each online review, and the total number of terms falling into each category (e.g., DAV, ADJ) was counted. We then assigned each category a weighting, based on the degree to which the category indicates high-level construal. In particular, consistent with the literature (Coenen et al., 2006), we assigned a weight of 1 to DAV terms; IAV and SAV terms were assigned a weight of 2; SV terms were assigned a weight of 3; and ADJ was assigned a weight of 4. Finally, based on the number of linguistic terms appearing in each category, we calculated an abstractness score for each review, using the weighted average equation proposed by Coenen et al. (2006, p. 15). This involved multiplying the number of terms appearing in the review from each category by the category's respective weight, to obtain the review's score on each level. The scores for each level were then summed, and the result was divided by the total number of linguistic terms appearing in the review, across all levels.

Due to the time-consuming nature of the coding task and the volume of reviews at our disposal, it was not feasible to code the entire sample. As such, we used a stratification sampling approach. We purposefully sampled 1,000 reviews, at random, conditional on the presence of high or low spatial or temporal distance (here, high or low refers to a distance value falling in the top or bottom 25% of the empirical distribution in our overall sample, respectively). Thus, we identified four sub-samples, each containing 250 reviews, which were separately characterized by (a) low spatial and low temporal distance (Group 1); (b) high spatial and low temporal distance (Group 2); (c) low spatial and high temporal distance (Group 3); and (d) high spatial and high temporal distance (Group 4).

We recruited eight research assistants (coders) to perform the content coding. The coders attended two, two-hour instructional sessions. In the first instructional session, the instructor explained the LCM to the assistants and each assistant then independently coded 10 reviews. The instructor and the eight assistants then compared results to ensure proper understanding of the instructions. The coders were asked to independently code 25 additional reviews before the second instruction session. In the second instruction session, the instructor revisited the LCM and discussed the assistants' coding of the 25 reviews to further ensure understanding and agreement (note: the 35 reviews considered in the instruction sessions did not contribute to the sample of 1,000 reviews that were considered in the subsequent analysis). Following the instruction sessions, the coders were asked to independently code the set of 1,000 reviews. Each review was examined by two coders. Each of the coders reported working between 15 and 25 hours to code 250 reviews over the course of 5 weeks. After the coding process was completed, we assessed the consistency of the results. Inter-rater reliability was above the commonly accepted threshold (Krippendorff's Alpha = 0.814). We then used the weighted average equation provided by Coenen et al. (2006, p. 15), described above, to calculate the construal level score.

The descriptive statistics for the coded data are reported in Table 3. We performed an ANOVA to examine both the main effects and the interaction effect of spatial and temporal distance on construal level. That is, we estimated an ANOVA using a 2 (spatial distance: high vs low) × 2 (temporal distance: high vs. low) design, with 250 observations in each of the cells. All the effects on construal level were at least marginally significant: temporal distance ($F = 8.28$, $p = 0.004$), spatial distance ($F = 4.03$,

$p = 0.045$) and their interaction ($F = 3.30$, $p = 0.070$). To decompose the interaction, we performed t tests to determine statistical differences in construal level among the four groups. In further support of the distance boosting effect, the highest level of construal level was observed when spatial and temporal distance were both high (Group 4). Specifically, we found a significant difference between the "high temporal, high spatial" group (Group 4; $M = 3.03$, $SD = 0.46$) and all other groups, ts > 4.65, ps < .001; as well as a significant difference between the "low temporal, low spatial" group (Group 1; $M = 2.74$, $SD = 0.59$) and all other groups, ts > 1.98, ps < .05. The cross-dimensional groups (Groups 2 and 3; $M = 2.83$, $SD = 0.41$ and $M = 2.87$, $SD = 0.52$), where one distance is high and the other is low, were not significantly different from each other, $t = 0.96$, $p = .34$; furthermore, each cross-dimensional group was significantly different from Groups 1 and 4, ts > 1.98, ps < .05. In essence, as more distance is felt, a higher level of construal is observed, and as more than one dimension of distance is felt, a higher level of construal is likewise observed.

Table 3. Descriptive statistics for content analysis

	Group 1 (low temporal, low spatial)		Group 2 (low temporal, high spatial)	
	Mean	Std. Dev.	Mean	Std. Dev.
Temporal Distance	0	0	0	0
Spatial Distance	6.84	7.84	4552.66	2237.45
Construal Level	2.74	.59	2.83	.41
	Group 3 (high temporal, low spatial)		Group 4 (high temporal, high spatial)	
	Mean	Std. Dev.	Mean	Std. Dev.
Temporal Distance	6.34	2.70	7.06	2.98
Spatial Distance	5.73	6.71	4686.96	2168.15
Construal Level	2.87	.52	3.03	.46

Notes:
For Group 1 vs. Group 2 & Group 3 vs. Group 4, spatial distance is significantly different, whereas temporal distance is not;
For Group 1 vs. Group 3 & Group 2 vs. Group 4, temporal distance is significantly different, whereas spatial distance is not;
For Group 2 vs. Group 3 & Group 1 vs. Group 4, both temporal distance and spatial distance are significantly different.

Mediation Analysis

To further demonstrate the role of construal level in the relationship between temporal distance, spatial distance, and rating favorability, we performed a mediation test, using a bias-corrected bootstrapping procedure (n = 10,000; results presented in Table 4). Specifically, we used the MEDIATE macro (Hayes & Preacher, 2014) in SPSS, employing construal level as the mediator. The independent variables were spatial distance, temporal distance, and their respective interaction.

Table 4. Mediation analysis results

Biased corrected indirect effects on review positivity through construal level			
	Estimate	Bootstrap S.E.	Bootstrap Confidence Interval
Spatial Distance	0.0318	0.0175	95% CI = [0.0003; 0.0697]
Temporal Distance	0.0457	0.0185	95% CI = [0.0137; 0.0861]
Spatial×Temporal Distance	0.0408	0.0240	90% CI = [0.0034; 0.0811]

Notes: Number of bootstrap samples for bias corrected bootstrap confidence intervals is 10,000 times.

We observed that the indirect effect of spatial distance on rating favorability, through construal level, yielded a 95% confidence interval (CI) that did not contain zero (95% CI = [0.0003, 0.0697]). The indirect effect of temporal distance on rating favorability through construal level, also yielded a 95% CI that did not contain zero (95% CI = [0.0137, 0.0861]). Finally, the indirect effect of the interaction between spatial and temporal distance, on rating favorability through construal level, yielded a 90% CI that did not contain zero (90% CI = [0.0034, 0.0811]). These results indicate that construal level mediates the effects of spatial distance, temporal distance, and (marginally) their interaction, on rating favorability. The mediation analysis also shows that the effects on rating favorability from temporal distance, spatial distance, and their interaction, are fully mediated by construal level (none of the respective direct effects were significant).

In sum, we found that experiencing more distance increased consumers' construal level, and that experiencing more than one dimension of distance further increased consumers' construal level – which we found to be a candidate underlying mechanism (albeit a marginal one) that drives the distance boosting effect on review favorability, reported earlier.

General Discussion

This study provides empirical evidence for the separate and interactive effects of temporal distance and spatial distance on construal level and consumer evaluations (rating favorability). Previous research on psychological distance has studied the effects of distance on construal level (and other outcomes) in relative isolation. In contrast, we examined two dimensions of psychological distance in tandem. Consistent with past research, we found that temporal and spatial distance each (individually) increase construal level and rating favorability. Moreover, we found that temporal and spatial distance exhibit an interactive effect, such that each distance increases the effect (on construal and positivity) of the other distance. Furthermore, we tested the process by which multiple psychological distances impact consumer evaluations, through construal level, by directly measuring construal level via a content analysis of the review text. We found construal level to fully mediate the effects of temporal distance, spatial distance, and their interaction on rating favorability – such that, similar to the effect on consumers restaurant evaluations, which grow increasingly positive, construal level also received a boost when more than one distance was experienced.

Theoretical and Empirical Contributions

Our research highlights the value of investigating multiple distances and contributes to the psychological distance, construal level theory, and online review literatures in several ways. First, our results extend an array of prior findings that show the effects of distance on construal level. While research has identified four kinds of psychological distance, only the effects of experiencing a single dimension on construal level have been investigated previously, on a notable range of outcomes. For instance, experiencing a single dimension of distance has been shown to affect important decisions involving gift-giving, negotiation, and creativity (Baskin, Wakslak, Trope, & Novemsky, 2014; Henderson & Wakslak, 2010; Polman & Emich, 2011). Our research suggests that there is *more* to experiencing one dimension of distance, such that outcomes like creativity (among others) could be further increased when people experience *more* than one dimension of distance, in tandem.

Second, our findings were drawn from data that included a wide range of distances which allowed us to investigate a continuum of distance (e.g., 0 to 11 months). Our data thus enable a very fine-grained analysis of psychological distance that is not only rare, but informative: We are able to examine construal level and review favorability based on a wide range of different points of distance – and can measure with precision the marginal effect of, say, one month or one mile (and the respective interaction) on review favorability. Quite possibly, past tests of distance have yielded null results (and then likely gone unpublished) because the categorical levels of distance chosen by researchers were not different enough to return significant findings (a problem that would exacerbate in tests involving more than one distance). However, with our data, we are able to show a reliable pattern for the effects of distance across gradually increasing swaths of distance – providing fidelity and thus rigorous internal validity.

Third, we extend the online review literature by applying theories of psychological distance and construal level in a context of significant relevance to marketing scholars. We therefore expand the boundary of CLT's application to a context that offers great potential for studying psychological distance and evaluation. The majority of prior studies on online reviews has focused on understanding the consequences of online reviews, such as facilitating consumer decision making (Hu, Liu, & Zhang, 2008), driving firm sales (Chevalier & Mayzlin, 2006; Zhu & Zhang, 2010) and affecting market competition (Kwark, Chen, & Raghunathan, 2014). Our study, in contrast, examines the antecedents of online review characteristics (e.g., Chen & Kirmani, 2015). In this regard, our study builds upon a relatively small body of work (Biyalogorsky, Gerstner, & Libai, 2001; Goes, Lin, & Yeung, 2014; Stephen & Lehmann, 2009), demonstrating the importance of consumer behavioral context (psychological distance) to consumer evaluations.

Future Directions

With an eye toward future research, we suggest a few questions that we think would be fruitful for investigation. First, does the effect of distance always receive a boost from another distance? While we have found evidence for a distance boosting effect, we must bear in mind the Weber-Fechner law (Fechner, 1966, 1860), which describes a diminished sensitivity to differences of magnitude as magnitude increases (i.e., decreasing marginal returns). For example, not unlike prospect theory, most people would consider the difference between $0 and $100 as greater than the difference between $1,000,000 and $1,000,100. In this vein, Maglio et al. (2013) found that among participants thinking of a more remote (vs. closer) location, a 2-year period in the future felt shorter. At first glance, our findings might appear inconsistent with that research.

However, past research has shown that a remote feeling on one distance fosters a remote feeling on another, when distance is egocentric (i.e., relative to self; Yan, 2014); whereas other research finds the opposite pattern when distance is non-egocentric (Maglio et al., 2013). Thus, whether a feeling of distance makes other distances larger (vs. smaller) might depend on whether the distances are egocentric (e.g., distances that are relative to here, now, me, reality).

Another related question concerns research that has observed a negative relationship between distance perception and emotional intensity (Williams & Bargh, 2008; Van Boven, Kane, McGraw, & Dale, 2010), possibly suggesting that we should expect consumers to review a restaurant less favorably as the months and miles increase from a long-ago visited restaurant and one's "here and now," respectively. Recent research, however, has theorized that the oft-found causal connection between psychological distance and construal level does not imply that they are the same theory (Van Boven & Caruso, 2015). In support of this perspective (that CLT and psychological distance are distinct), Williams et al. (2014) showed that the effects of psychological distance and construal level have "separable paths" (p. 1126) revealing simultaneously: (1) an inverse link between psychological distance and intensity and (2) a positive link between construal level and positivity. It is therefore possible to generate positive event-appraisals in spite of and because of increases in psychological distance.

CHAPTER 3.

ESSAY 2: SOCIAL NETWORK INTEGRATION AND USER CONTENT GENERATION: EVIDENCE FROM NATURAL EXPERIMENTS

Abstract

This study examines how social network integration can affect the characteristics of user-generated content (volume and linguistic features) in the context of online reviews. Building on the social presence theory, we propose a number of hypotheses on how social network integration affects review volume and linguistic features of review text. We consider two natural experiments at leading online review platforms (*Yelp.com* and *TripAdvisor.com*), wherein each implemented a social network integration with *Facebook*. Constructing a unique panel dataset of online reviews for a matched set of restaurants across the two review sites, we estimate a difference-in-differences (DID) model to assess the impact of social network integration. We find that integration with *Facebook* increased the production of user-generated content and positive emotion in review text, while simultaneously decreasing cognitive language, negative emotion and expressions of disagreement (negations) in review text. Our findings demonstrate that social network integration works as a double-edged sword. On the one hand, integration provides benefits in terms of increased review quantity. On the other hand, these benefits appear to come at the cost of reduced review quality, given past research which has found that positive, emotional reviews are perceived by users to be less helpful. We discuss the implications of these results as they relate to the creation of sustainable online social platforms for user content generation.

"Humans are different in private than in the presence of others. While the private persona merges into the social persona in varying degrees, the union is never complete. Something is always held back."
– Brian Herbert, House Corrino, 2001

Introduction

Many online platforms have sought to supplement their home-grown communities by integrating with prominent social networking sites like *Facebook*, *Twitter*, and *Google+*, a practice known as social network integration (Blanchard 2011). Examples of social network integration include social login (Kontaxis et al. 2012), Facebook Connect (Holliday 2009) and instant personalization (Kincaid 2010). Social login allows a new user to register an account with an online platform using an existing account at a social networking service – e.g., *Facebook* (Goel et al. 2011, Frutiger et al. 2014). Once a user grants a platform access to their existing social networking account, Facebook Connect enables automatic or user-controlled social sharing of the user's activities on the platform, back to the social networking site (for instance, sharing *Yelp* reviews on *Facebook* pages). *Facebook*'s instant personalization option enables even greater levels of integration, because it allows the partner platform to access and make use of *Facebook* profile information, including the user's name, geographic location and social connections (Rapp et al. 2013). In sum, social network integration facilitates convenient user account creation and login, provides for a more personalized user experience, and promotes a greater perception of social presence (e.g., cognizance of one's audience, typically their friends).

The objective of this study is to examine how social network integration affects the volume and characteristics of user content generation, particularly in the context of online reviews. We explore how social network integration influences review volumes and the linguistic features of review text. In terms of linguistic features, we place a specific focus on the occurrence of i) affective content, i.e., words indicative of feelings and emotion, such as 'happy', 'sad', or 'cried' (Epstein 1993; Gill et al. 2008), which may vary in both valence (positive versus negative) and intensity (Shiv and Fedorikhin 1999; Gilovich et al. 2002); ii) cognitive content, i.e., words suggestive of reasoning and information processing, such as 'cause', 'know', or 'ought' (Pennebaker and Francis 1996; O'Neill 2002); and iii) negative language, i.e., words related to negation and disagreement, such as 'no', 'not', or 'never' (Lasersohn 2005; Horn 2010). Formally, we seek to investigate the following research question:

How does social network integration affect user content generation (online reviews), in terms of volume, the exhibition of affective and cognitive language, and use of negative language?

We propose that social network integration may instigate changes in the volume and linguistic features of reviews by increasing social presence on a website. For example, in terms of the volume of content produced, there are two plausible countervailing mechanisms by which social network integration may cause differences. On the one hand, social network integration may result in more content production because it leads to increased social interaction, which provides a greater opportunity for individuals to gain social benefits from the content they produce[1] (Dellarocas 2003; Lampel and Bhalla 2007; Zhang and Zhu 2011). On the other hand, social integration may cause a decrease in content production, because it may lead users to tailor or even cease their content generation (Das and Kramer 2013; Sleeper et al. 2013), out of fear of social disapproval by the newly (socially) proximal audience.

Social network integration may also impact the characteristics of content that is produced, for at least two reasons. First, social network integration may drive a shift in the composition of the user base by inducing selection into the platform; by facilitating the entry of a new group of individuals, who might then produce systematically different content because they hold inherently different personal traits. Second, social network integration may cause existing users to change the nature of the content they produce. Specifically, social network integration, by increasing social presence, may trigger existing users to exhibit feelings and emotions with greater intensity when authoring reviews (Shiv and Fedorikhin 1999; Gilovich et al. 2002), at the expense of cognitive processing (De Martino et al. 2006; Kahneman 2011). Additionally, social integration, by increasing social presence, may reduce individuals' tendency to employ negations, which are indicative of negativity or disagreement (Davis et al. 2002; Moor et al. 2010).

Grounded in social presence theory (Short 1974; Short et al. 1976), we propose several hypotheses relating social network integration to the quantity and linguistic features of online reviews. We analyze a unique data set comprised of online reviews for

[1] It is also possible that social network integration may increase content production because it typically comes paired with a social network login feature (e.g., login via Facebook). Social login makes it easier and more convenient for new individuals to register and enter the community (Frutiger, 2015).

a set of matched restaurants across two comparable, leading online review websites. We code the linguistic features of the review text using Linguistic Inquiry and Word Count (LIWC), a tool we describe in greater detail in the methodology section. Our econometric identification strategy hinges on two natural experiments: temporally staggered social network integrations on *Yelp.com* and *TripAdvisor.com*. These natural experiments allow us to infer the causal effects of social network integration via a difference-in-differences model (Frohlich 2004; Fricke 2015).

Our results show that social network integration increases the volume of online reviews that are authored, due to a combination of more rapid user entry and an increase in average reviewing activity amongst existing users. Moreover, we find that integration leads to changes in the linguistic features of reviews; we observe that emotional, affective language increases while cognitive language declines. Further, we observe a decline in users' tendency to employ negation terms, indicative of disagreement. Finally, a series of subsequent user-level analyses demonstrate that the changes we observe are driven primarily by shifts in user behavior, rather than shifts in user composition (i.e., self-selection).

Our study contributes to the literature on user content generation and the design of online review systems. While past research has primarily focused on the consequences of online reviews, our study provides a pioneering effort in understanding how an important system design feature – social network integration – affects review volumes and linguistic features of review text, by increasing social presence. Given the recent trend of online platforms toward promoting social network integration, it is crucial that we improve our understanding of the collateral consequences.

The findings of our study also carry important implications for the design of IT platforms that host and heavily rely upon user generated content. On the one hand, social network integration appears to be a boon for online review sites. First, social network integration appears to increase content production, which is likely to be desirable to online review platforms, which are known to face an under-provisioning problem (Avery et al. 1999; Burtch et al. 2016). However, social network integration also has its downsides. Considering the past literature's observation that consumers' perceive negative reviews to be more helpful (Rozin and Royzman 2001; Chen and Lurie 2013), and emotional reviews as less helpful (Baumeister et al. 2001; Yin et al. 2014; Hong et al. 2016), the fact that we see i) a shift away from cognitive language toward emotional language, ii) that the latter manifests primarily as positive emotions, and iii) that consumers employ fewer negations, suggests that, despite the apparent benefits of greater review volumes, social network integration may lead to content that is perceived to be less helpful, and thus of lower quality. In sum, our findings demonstrate that social network integration, and thus the associated increases in social presence, can be a double-edged sword, providing benefits in terms of increased review quantity, possibly at the cost of perceived review quality.

Prior Literature

Social Presence & Anonymity

Social presence was originally defined as "the degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationships"

(Short et al. 1976, p. 65). Modern definitions in the context of computer-mediated social networks refer to social presence as individuals' awareness of their social connections in a communication interaction (Gunawardena 1995; Cobb 2009; Kehrwald 2008). The degree of social presence depends on the level of interpersonal interactions that a communication medium supports. For example, face-to-face communication tends to have the most social presence, whereas text-based communication has relatively less social presence (Lowenthal 2009; Cui et al. 2012). Social presence has been found to be a significant predictor of user behavior in computer-mediated interactions (Gunawardena and Zittle 1997; Richardson and Swan 2003). In online contexts, increased social presence has been shown to make individuals less divergent or disagreeable in their thinking (Sia et al. 2002), to facilitate 'deeper' information processing, to promote a lesser breadth of information sharing (Miranda and Saunders 2003), and to lead to more socially fulfilling experiences (Jiang et al. 2013).

Prior research on social presence focuses on outcomes that largely pertain to transactions, such as trust (Ou et al. 2014), purchase intentions (Animesh et al. 2010), and product choice (Rhue and Sundararajan 2013). In contrast, the present study examines users' possible shifts in consumers' online reviewing behavior (in terms of volume and linguistic features) as novel outcomes that may be driven by increases in social presence on online platforms. Recent developments in social media have created the potential to increase social presence on the Internet (Kaplan and Haenlein 2010) by connecting individuals in social networks (Kane et al. 2014). In particular, online platforms have begun to implement social network integration (integration with social media services) to improve social interaction (Wright-Porto 2011; Kontaxis et al. 2012). Social network integration leads to increases in social presence on the adopting platforms (Rhue and Sundararajan 2013). As a platform changes from a relatively anonymous environment to a social environment, users are likely to adapt their behaviors to their newly proximal, salient audience (Acquisti and Gross 2006; Daughety and Reinganum 2010; Jones and Linardi 2014).

Of course, the corollary of increased social presence is the loss of anonymity. Anonymity refers to a state in which identifying information for an acting party is unknown (Hoffman et al. 1999; Pfitzmann and Köhntopp 2001). There are two-sides to the argument about anonymity's role in the literature. On one hand, anonymity is an important element in preserving information privacy (Ba 2001; Ayyagari et al. 2011; Pavlou 2011; Acquisti et al. 2013). On the other hand, anonymity contributes to online incivility, producing behaviors ranging from racism and hatred (Reader 2012; Santana 2014) to Internet trolling (Hardaker 2010; Phillips 2011) and cyber bullying (Campbell 2005). Scott and Orlikowski (2014) have recently summarized these points, arguing that anonymity is likely to be an important issue in online reviews because it may lead users to feel more comfortable and secure, resulting in more frequent contribution, while at the same time raising concerns about user regulation.

Prior studies on anonymity indicate that the presence or absence of anonymity leads individuals to adjust their information sharing behavior. For instance, with the loss of anonymity, users are more likely to publicize socially desirable information (Huberman et al. 2005). When prompted to consider their anonymity, users may become self-conscious and subsequently more conservative in their information sharing (John et al. 2009; Burtch et al. 2015). Dissociative anonymity leads users to intensify their information sharing behavior, a phenomenon known as the online disinhibition effect (Suler 2004). Building on prior research, this study discusses how the loss of anonymity due to increased social presence may impact users' engagement with online platforms and, in particular, their content contributions.

Online Reviews & Social Interactions

The extensive literature of online reviews can be classified into two broad categories. One body of work has focused on the consequences of online reviews, conditional on their characteristics, namely volume, valence and linguistic features (e.g., Ba and Pavlou 2002; Chevalier and Mayzlin 2006; Dellarocas et al. 2007; Duan et al. 2008; Zhu and Zhang 2010; Mudambi and Schuff 2010; Kwark et al. 2014, Yin et al. 2014), whereas a second has focused on the antecedents of, and processes underlying review generation (e.g., Godes and Silva 2012; Luca and Zervas 2016; Huang et al. 2016). Our study aims to contribute to the latter category, considering the influence of social aspects on review generation.

Past work suggests that social factors significantly influence users' authorship of online reviews (e.g., Aral 2013; Muchnik et al. 2013; Wang 2010; Wang et al. 2015). First, Wang (2010) observed that, given the opportunity to establish social image, consumers tend to write more reviews and give less extreme ratings. This finding suggests that users do respond to an audience. Similarly, Zhang and Zhu (2011) leveraged a natural experiment at Chinese Wikipedia to demonstrate that a larger audience size incentivizes users to contribute public content, due to potential social benefits. Other work, by Chen et al. (2010), provides further evidence for an audience effect. Those authors conducted a randomized experiment and found that providing information about the average rate of review authorship in the community could significantly increase a subject's own rate of authorship, if they perceived that they were lagging behind that average. However, Chen et al. (2010) also found the opposite effect for individuals who were initially contributing above the average; high contributors became less engaged once they realized they were doing more than their fair share. Finally, other work suggests that individuals seek to maintain social approval, once it has been obtained. For example, Goes et al. (2014) showed that after attracting subscribers (or followers), individuals begin to author reviews more objectively, and with greater negativity and variance in valence; features that are known to be perceived as helpful.

The present study advances our understanding of the effects of social network integration, and thus social presence, on users' authorship of online reviews. Prior work has examined the impact of social factors on contribution quantity (Huberman et al. 2009; Chen et al. 2010), rating negativity and extremity (Wang 2010; Goes et al. 2014). Here, we begin by considering the impact of social presence on review volume, but we also go further, investigating the impact on linguistic features of review content. Specifically, we provide a first consideration of the impact of social presence implemented by platform social network integration on review authors' psychological processes (affect and cognition) and their use of negative language (negations).

Hypothesis Development

In this section, we hypothesize the effects of social network integration on online review production, in terms of volume and linguistic features. First, we focus on review volumes, noting that sustainable platforms require a healthy volume of content, and online reviews are subject to an under-provisioning problem (Avery et al. 1999; Burtch et al. 2016). Second, we consider reviewers' mental processes reflected in language, i.e., the use of affective (emotional) language and, conversely, cognitive language, noting prior research which has found that emotional expression can generally reduce the perceived helpfulness of online reviews (Yin et al. 2014; Hong et al. 2016). Third, and last, we consider individuals' use of negations, i.e., words such as 'no', 'not', 'never', which are indicative of negativity or disagreement.

Volume Effect

Review volumes reflect the level of user engagement and thus the sustainability of a content-based platform. Social network integration, by increasing users' social presence and decreasing anonymity, may affect whether a user chooses to write a review. In particular, we consider multiple possible mechanisms, which may countervail one another.

First, social network integration, in the form of *Facebook Connect* or *Instant Personalization*, may lead to greater social presence, exposing a user's reviews to his or her friends and thereby increasing the perceived relevance of the audience who may ultimately read the reviews. As such, because users are likely to be aware that their friends may benefit from their contributions to the review platform, they may believe that there is a potential for social benefits or reputational gains (Zhang and Zhu 2011). This, in turn, may stimulate users to contribute larger volumes of reviews.

However, we also must consider a second, countervailing mechanism. Users may fear social disapproval, given a relative loss of anonymity (Kang et al. 2013). This suggests that users' willingness to share their experiences on the review platform may decline, especially when they have had very negative experiences. This countervailing mechanism is supported by the findings of Leshed (2009), who observed that a loss of anonymity was associated with a decline in the number of comments users made in online discussion forums. Further, past research has noted that individuals often create and maintain an alternate identity in online spaces (Froomkin 1999). When individuals' online anonymity is compromised, they may lose the ability to maintain their alternate persona (Scott and Orlikowski 2014). Although it is possible that a user could simply construct a secondary user account, from which they could post their negative experience, this would require a new added cost of time and effort, which many users may not wish to absorb.

Second, social network integration typically comes with a social login feature (e.g., login via Facebook). Such features make registration and login less time consuming. Consequently, social integration may lead to greater user enrollment and user involvement in a platform (Drebes 2011; Kontaxis et al. 2012). In turn, this may produce an increase in the volume of reviews being authored on the platform.

To summarize, social network integration may lead to multiple countervailing effects. On the one hand, review volumes may increase i) because the greater social presence that results from social network integration may create a greater opportunity for users to pursue social image and reputational gains, and ii) because social login is likely to facilitate higher enrollment of new users and stimulate greater involvement of existing users. On the other hand, a relative loss of anonymity may cause individuals to fear social disapproval from their peers, driving them to contribute less. Bearing in mind that i) a majority of mechanisms (social presence and social login) speak to a likely increase in review volumes, and ii) social disapproval might be avoided by creating a throw-away user account, it appears more likely that the positive effects of social network integration on review volumes will dominate. Accordingly, we propose the following hypothesis.

H1: Social network integration leads to more reviews.

Mental Process Effects

The ability of a platform to support social connections and interactions heightens perceived social presence and users' awareness of other users – e.g., the audience for their reviews (Kehrwald 2008; Cobb 2009). Here, we argue that increased social presence deriving from social network integration is likely to affect how users author reviews, in terms of their reliance on affective versus cognitive mental processes. Affective (emotional) processes incorporate feelings associated with the entity being evaluated, whereas cognitive (rational) processes incorporate attributes and beliefs about the entity (Millar and Tesser 1986). In the social psychology literature, it has frequently been suggested that affect and cognition are negatively correlated (Briggs 1977; Pervin and John 1999); that when affect dominates, cognition recedes, and vice versa. Recent neurophysiological evidence supports this belief, having shown that affective processes and cognitive processes are supported by different areas of the brain (Finucane et al. 2003), i.e., the anterior insula supports emotion while the dorsolateral prefrontal cortex supports cognition (Sanfey et al. 2003). Because affective and cognitive mental processes are largely associated with two opposing neural systems, when people draw on affective mental processes, they are less likely to draw on cognitive mental processes, and vice versa (De Martino et al. 2006; Talmi and Frith 2007). Bearing the above in mind, when crafting reviews, users might be expected to rely on one type of mental process (affective or cognitive) at the expense of the other. Thus, when users exhibit affective mental processes in crafting their reviews, they are likely to express their emotions in the text of their reviews (e.g., words like *happy*, *sad*, *cried*). When this happens, we might expect to observe a decline in cognitive mental processes, and thus a reduction in users' application of logic and analytical thought (e.g., words like *because, therefore, think*). Similarly, when cognitive mental processes take hold, we might expect to observe an increase in words associated with logical and analytical thought, and a commensurate decline in words associated with emotions.

The emotional broadcaster theory of social sharing argues that individuals have an intrinsic drive to share experiences in a psychologically arousing manner (Harber and Cohen, 2005). In a social environment, individuals' emotions are activated and, therefore, they are more likely to share their feelings and emotions, whether intentionally or unintentionally. This behavior is known as emotional leakage (Kraut 1982). Supporting this theory, Wagner and Smith (1991) and Buck et al. (1992) both found that closer social relationships (e.g., friends versus strangers) facilitate emotional expressiveness. It has also been found that, with respect to the expression of emotion, similar patterns emerge in both face-to-face communication and computer-mediated communication (Derks et al. 2008). In the context of online reviews, social network integration increases social presence on the platform, which can be expected to stimulate users' emotional expressiveness, or even trigger emotional leakage. As a result, following social network integration, users are more likely to draw on affective processes when authoring reviews, and less likely to rely on cognitive processes. Bearing in mind the above, we propose the following two hypotheses:

H2a: Social network integration leads to more language reflecting affective processes in review text.

H2b: Social network integration leads to less language reflecting cognitive processes in review text.

Inhibition Effect

Increases in social presence reduce user anonymity, which has both benefits and pitfalls. A variety of studies in the Group Decision Support Systems (GDSS) literature have consistently reported that anonymity can provide the conditions necessary for the production of innovative, creative ideas (Connolly et al. 1990), and that users may exhibit a decline in social desirability concerns, as well as higher levels of self-esteem (Joinson 1999). However, online anonymity has also been shown to produce an “online disinhibition effect" (Suler 2004; Cho et al. 2012), in which individuals exhibit a greater willingness to reveal their true, uncensored opinions, thoughts and preferences. Accordingly, individuals may be more critical, disagreeable and argumentative when they are in an anonymous environment, with a low level of social presence (Jessup et al.

1990). Under anonymity, individuals have also been known to engage in a variety of behaviors that would otherwise meet with social disapproval, ranging from free-riding (Andreoni and Bernheim, 2009) to racism (Reader 2012; Santana 2014), Internet trolling (Hardaker 2010; Phillips 2011) and cyber bullying (Campbell 2005).

In the context of online reviews, review authors, once subject to increased social presence, may become concerned about their audience (now more likely to be comprised of offline friends) disapproving of their tone. Individuals generally strive to achieve a positive social identity (Jackson et al. 1996; Oldmeadow and Fiske 2010) because they derive utility from being judged positively by others (Gneezy et al. 2012). As an anonymous interviewee reported to Kang et al. (2013): *"I posted a very bad review [of a restaurant]. And I guess I did that [anonymously]. I live in a small town, so I certainly didn't want to put my real name...."*.

With increases in social presence (i.e., a decline in anonymity), we therefore expect a decline in users' application of negative language and negations, due to an increased desire to establish a positive social identity. This leads us to our final formal hypothesis:

H3: Social network integration leads to fewer negations in review text.

Identification of Competing Mechanisms

From an identification standpoint, we must acknowledge that multiple mechanisms may exist which would produce the same pattern of results we have hypothesized. For example, social network integration, by increasing social presence, may stimulate greater user activity by increasing the potential for reputational gains. At the same time, it is plausible that social network integration might lead to an influx of new users, particularly those who are most active on, or comfortable with Facebook. In turn, such active new users might also contribute to the growth in review volumes, over and above any increases in the average contributions of existing users.

Similarly, existing users might change their language use patterns (e.g., using more affective language and expressing less disagreement, as reflected by a decline in negative language and negations) in response to greater social presence on the platforms. At the same time, newly entered users, arriving as a result of social network integration, may be systematically different from existing users in their language usage, and thus may also introduce changes in the linguistic features of reviews. Ultimately, distinguishing between these various mechanisms poses a difficulty; however, in Section 5, we report on a number of secondary analyses that enable us to unravel and eliminate some of the mechanisms (most notably those related to self-selection). We also report several falsification tests, which further strengthen the identification of our study.

Research Methodology

Background

Our study considers two comparable online review platforms: *Yelp.com* and *TripAdvisor.com*, both of which implemented social network integrations with Facebook, at different points in time. First, we consider *Yelp.com*'s adoption of the Facebook Connect feature on July 2, 2009 (O'Neill 2009; Holliday 2009). Facebook Connect allows users to log into a website using their Facebook account (also known as social login) and to share reviews with friends on Facebook. Facebook Connect is an "opt-in" feature, in that it is up to the user to decide whether he or she would like to adopt the feature. In other words, the implementation of Facebook Connect did not require that users share their reviews on Facebook; users could choose whether or not to share their reviews, and could readily adjust the review content conditional on that decision. Figure 1 shows the review page with the Facebook Connect feature enabled for a *Yelp.com* user.

Second, we consider *TripAdvisor.com*'s adoption of *Facebook*'s Instant Personalization feature on December 21, 2010 (Kincaid 2010; *TripAdvisor* 2010). With Instant Personalization, if a user visits *TripAdvisor*'s website while logged into *Facebook* (or having logged into *Facebook* at any time in the prior 30 days, with cookies enabled), *TripAdvisor* will gain access to the user's *Facebook* account information. Instant Personalization then automatically presents users with personalized website content on *TripAdvisor.com* that shows their *Facebook* friends' travel and reviewing activities, such as recently authored restaurant and hotel reviews, and a list of *Facebook* friends' most popular destinations. Instant Personalization is an "opt-out" feature, in that the feature is enabled by default and requires that users take a series of cumbersome actions to disable it. Although users can choose to opt out Instant Personalization through *Facebook's* privacy controls, this is reportedly challenging to do[2]. After Instant Personalization, users will be aware that their Facebook friends can read their reviews on *TripAdvisor*, and thus they are likely to change their reviewing behavior (e.g., modify their review content). Figure 2 illustrates the webpage with the Instant Personalization feature for a *TripAdvisor.com* user.

Both Facebook Connect and Instant Personalization are IT artifacts that can be considered instances of social network integration. Notably, however, due to the artifact design differences, these two integrations are likely to be heterogeneous in their impact upon the adopting platform. Specifically, Facebook Connect was an 'opt-in' (optional) feature, requiring that individual users explicitly choose to accept it, whereas Instant Personalization was a 'opt-out' (mandatory) feature, imposed on users by default. To reject Instant Personalization, users would have been required to go through a series of steps to disable the feature. For this reason, we expect that Instant Personalization, a mandatory integration, would have a greater impact on reviewing activity than Facebook Connect, an optional integration (Johnson et al. 2002; Kressel et al. 2007). Each social network integration constitutes a system change that was exogenous from the perspective of website users, enabling us to treat them as sequential natural experiments, with each platform serving as a control at the time of the other platform's integration event. Further,

[2] TripAdvisor Support Forum (June 9, 2012). "Facebook Link Not Wanted." Retrieved from https://www.tripadvisor.com/ShowTopic-g1-i12105-k5489397-Facebook_link_not_wanted-TripAdvisor_Support.html

the different treatment strengths resulted from the artifact design differences, help us properly achieve econometric identification, which we discuss in the next section.

Figure 1. Review Page with Facebook Connect Feature On *Yelp.com*

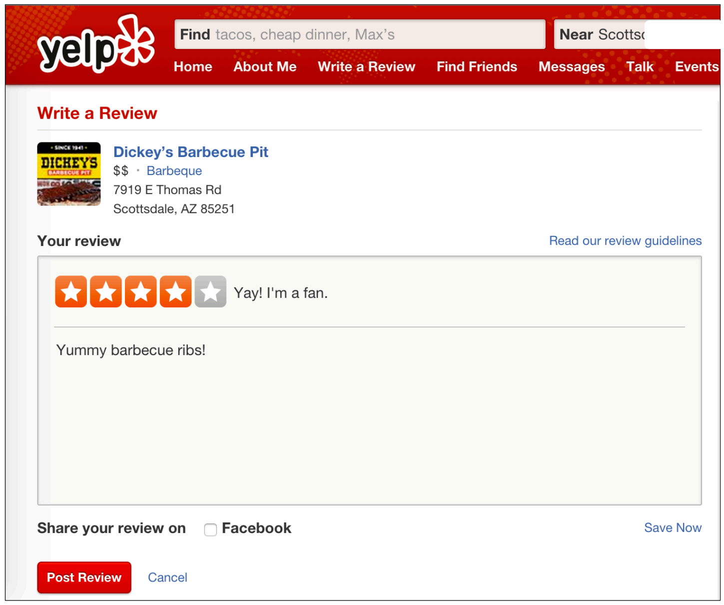
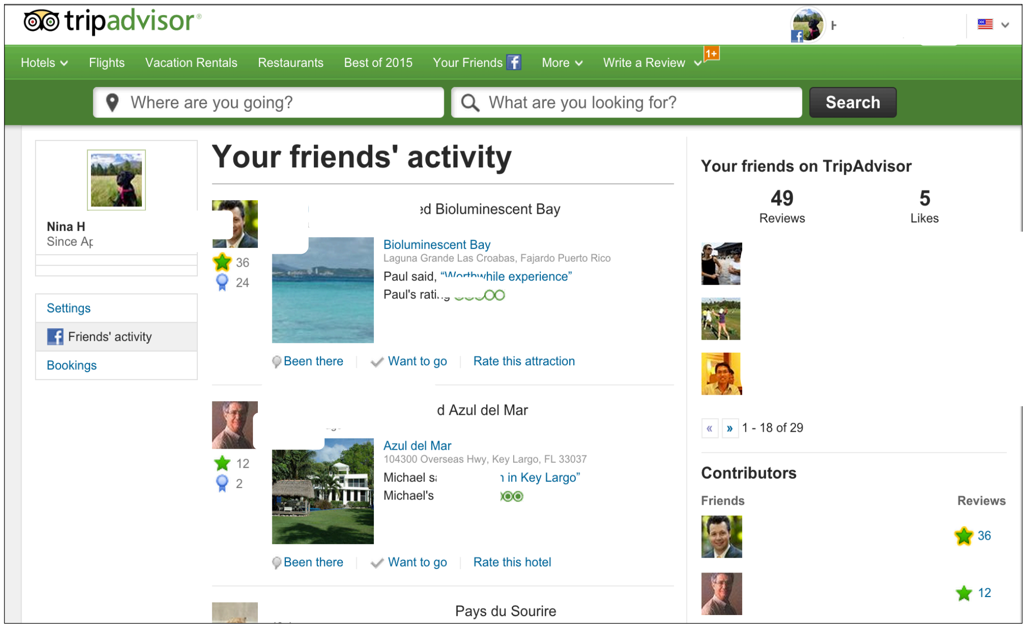


Figure 2. Webpage with Instant Personalization Feature on *TripAdvisor.com*



Data and Measures

We collected data on restaurant reviews authored between 2005 and 2014 from *Yelp.com* and *TripAdvisor.com* for a matched set of restaurants, selected at random, which are located in five major cities across the United States (New York City, Los Angeles, Chicago, Philadelphia and Phoenix). Notably, restaurant reviews have received considerable attention in the extant literature (Lu et al. 2013; Luca and Zervas 2016). The data contains time stamps and review content (ratings and text), in addition to reviewer profile and restaurant information. We created an indicator variable to mark reviews collected from *TripAdvisor.com* vs. those collected from *Yelp.com*, and we then pooled the data.

Review volume was measured as the monthly total volume of reviews submitted to the platform about a given restaurant. To construct measures of linguistic content, we leveraged the latest version of Linguistic Inquiry and Word Count (LIWC), a text analytics tool. LIWC calculates the prevalence of different categories of words in a text document based on the percentage of words that are matched to pre-defined keyword dictionaries (Pennebaker et al. 2007). LIWC has frequently been used in the psychology literature, and has also recently seen increased use in the Information Systems (Goes et al. 2014; Yin et al. 2014; Hong et al. 2016) and Marketing literature (Sridhar and Srinivasan 2012; Lurie et al. 2014). We focused on LIWC's measures of emotional, cognitive and negative language. Table 5 provides examples of words in the LIWC dictionaries for the linguistic categories we consider in this study. In addition, Table 6 presents examples of review text containing highly emotional, cognitive, or negative language on *TripAdvisor.com*.

Table 5. Sample Words in LIWC's Dictionaries

Language Characteristics	*Examples*	*Words in Category*
Affective Processes	Happy, cried, abandon	915
Positive Emotion	Love, nice, sweet	406
Negative Emotion	Hurt, ugly, nasty	499
Cognitive Processes	Cause, know, ought	730
Negation	No, not, never	57

Notes: Table 1a is adopted from the "LIWC2007 output variable information" table, retrieved from *http://liwc.net/descriptiontable1.php*. More information on LIWC and the entire list of words that are used for matching to obtain the linguistic measures can be obtained from *http://www.liwc.net*.

Table 6. Examples of Review Text

Language Characteristics	*Example Reviews*
Affective Processes	"The coffee is good a Cappuccino in this case, the place is super busy, super popular, super packed, a cute fun nice ambiance on the sidewalk in Zamalek. Cute"
Positive Emotion	"Love, love, love The Oinkster! Great patio, great burgers, yummy shakes and malts, relaxed atmosphere. My favorite place along a great strip of eateries."
Negative Emotion	"Terrible overall. Location right by the highway, terrible noise isolation, the worst unhealthy breakfast ever, outdated rooms, disgusting shower curtain, overpriced. Best to avoid all together."
Cognitive Processes	"Worn out place, trying to make it charming without really succeeding. Food was boring, sandwiches and chips - nothing to remember at all. Did not manage to make that little extra feel, neither with food, service or surroundings. The cottages looked as worn out as the restaurant, would not stay here."
Negation	"No big deal, I wouldn't highly recommend it unless you have nothing else to do. Not a lot of parking. Food, nothing special."

Notes: Table 1b provides examples of review text that are measured as having a high value in the corresponding linguistic category. The underlined text are the words matched to the LIWC dictionary for the respective categories.

We first measured the linguistic features of each review, then averaged them across reviews for each restaurant, aggregating to the monthly level, to avoid issues of sparsity – i.e., to ensure each observation included a reasonable amount of text. For the final measures used for analyses, we take two approaches and report results for them respectively. First, we use the raw data that results from LIWC, including those reviews where no words were matched with the LIWC dictionaries. This approach has the advantage of including reviews. At the same time, this approach has disadvantages, because it results in issues of sparsity. Accordingly, second, to establish robustness, we follow the approach of Snefjella and Kuperman (2015), trimming our data, and retaining only those monthly observations where at least one review comprising the observation contained at least one word that could be matched to an LIWC dictionary. Because there are no zeros in this second sample of data, it is straightforward to then log transform the dependent variables, to further address skewness in the variable distributions. Log transforming the dependent variables also has the benefit of enabling percentage interpretations of the parameter estimates. Table 7 presents the descriptive statistics of the variables in our raw data, and Table 8 provides a correlation matrix of the main variables.

Table 7. Descriptive Statistics

Variables	*Mean*	*S.D.*	*Min*	*Max*	*Median*
Review Volume	4.273	4.753	1	98	3
Rating	3.734	0.887	1	5	4
Words	123.782	57.929	27	281	116
Affective Processes	7.679	2.843	0	52.815	7.31
Positive Emotion	6.719	2.908	0	52.440	6.342
Negative Emotion	0.936	0.961	0	26.670	.772
Cognitive Processes	15.333	3.470	0	41.987	15.317
Negation	1.158	1	0	25.770	.989

Table 8. Correlation Matrix

Variables	**Review Volume**	**Rating**	**Words**	**Affective Processes**	**Positive Emotion**	**Negative Emotion**	**Cognitive Process**	**Negation**
Review Volume	1.000							
Rating	0.036	1.000						
Words	0.175	–0.135	1.000					
Affective Processes	0.042	0.230	–0.307	1.000				
Positive Emotion	0.039	0.351	–0.319	0.943	1.000			
Negative Emotion	0.007	–0.382	0.054	0.100	–0.233	1.000		
Cognitive Process	–0.012	–0.050	0.098	–0.070	–0.077	0.026	1.000	
Negation	–0.045	–0.230	–0.035	–0.062	–0.127	0.202	0.224	1.000

Econometric Identification

As noted above, our econometric identification hinges on two natural experiments related to social network integrations that occurred on *Yelp.com* and *TripAdvisor.com*, which we treat as exogenous shocks to the platform users, in the form of system changes. We employ difference-in-differences (DID) estimation to identify the effects of social network integration on the volume and linguistic features of online reviews on each platform. The DID estimator attempts to identify causal relationships by mimicking an experimental design in observational data (Angrist and Pischke 2008). DID is a common estimation approach, frequently used to establish causal relationships in data where experimental manipulation is generally difficult to implement (Card and Krueger 1994; Di Tella and Schargrodsky 2004).

Yelp.com introduced a Facebook Connect feature on July 2, 2009, and *TripAdvisor.com* implemented Instant Personalization on December 21, 2010. Because Facebook Connect is an “opt-in” feature while Instant Personalization is an “opt-out” feature, the two social network integrations are ordered in terms of increasing magnitude of their effects. The observation period spans from July 2008 to July 2012. We retain a 12-month pre-treatment period, in advance of *Yelp*’s integration event. Notably, the results we present in the following sections are not sensitive to this choice; expanding the window to 18 months or 24 months produces very similar results.

Yelp.com’s integration with Facebook constitutes the first treatment, with activity on *TripAdvisor.com* at the same point in time acting as the control group.

TripAdvisor.com's integration is the second treatment, with activity on *Yelp.com* at the same point in time acting as the control group. Though our work is not the first to employ a DID estimation in analyzing online reviews, prior research has typically employed a single-shock DID estimation (Chevalier and Mayzlin 2006; Zhang and Zhu 2011; Mayzlin et al. 2014; Chen et al. 2017). For example, Chen et al. (2017) employed a single-shock DID to identify the effect of introducing a multi-dimensional rating system on the characteristics of online reviews. Generally speaking, our two-treatment DID approach provides stronger identification and more conservative estimates (Choe et al. 2015). The assumption underlying the two-treatment DID identification strategy is that the exogenous shocks are ordered in time and vary in treatment strength. Our scenario of the two shocks exactly matches this assumption because, as discussed above, the first shock (Facebook Connect) is an optional 'opt-in' feature whereas the second shock (Instant Personalization) is a mandatory 'opt-out' feature, which is likely to have a stronger impact on user behavior.

Finally, it is worth noting that a key assumption underlying the validity of the DID specification, more generally, is the parallel trends assumption, i.e., that the trend of reviewing activity on the untreated platform can serve as a valid control for the trend of activity on the treated platform. Following Angrist and Pishke (2008), we assess this parallel trends assumption empirically in the Appendix B, employing a relative time specification around the *TripAdvisor* social network integration. We also report the results of separate, single-shock DID analyses in the for each of the natural experiments Appendix C.

Estimation Model

We estimate the two-treatment DID models reflected by Equations (1) through (4). Our estimation incorporates restaurant-level fixed effects via a within transformation, which allow us to effectively control for restaurant-level unobserved heterogeneity.

$$\ln(Review\ Volume)_{ipt} = \beta_0 Trip_p + \beta_1 Trip\_Change_t + \beta_2 Trip_p * Trip\_Change_t + \beta_3 Yelp\_Change_t + \beta_4 Yelp_p * Yelp\_Change_t + \beta_5 ln(words_{ipt}) + \beta_6 rating_{ipt} + \alpha_i + \varepsilon_{ipt} \tag{1}$$

$$Review\ Volume_{ipt} = \beta_0 Trip_p + \beta_1 Trip\_Change_t + \beta_2 Trip_p * Trip\_Change_t + \beta_3 Yelp\_Change_t + \beta_4 Yelp_p * Yelp\_Change_t + \beta_5 ln(words_{ipt}) + \beta_6 rating_{ipt} + \alpha_i + \varepsilon_{ipt} \tag{2}$$

$$\ln(Linguistic\ Characteristics)_{ipt} = \beta_0 Trip_p + \beta_1 Trip\_Change_t + \beta_2 Trip_p * Trip\_Change_t + \beta_3 Yelp\_Change_t + \beta_4 Yelp_p * Yelp\_Change_t + \beta_5 ln(words_{ipt}) + \beta_6 rating_{ipt} + \alpha_i + \varepsilon_{ipt} \tag{3}$$

$$Linguistic\ Characteristics_{ipt} = \beta_0 Trip_p + \beta_1 Trip\_Change_t + \beta_2 Trip_p * Trip\_Change_t + \beta_3 Yelp\_Change_t + \beta_4 Yelp_p * Yelp\_Change_t + \beta_5 ln(words_{ipt}) + \beta_6 rating_{ipt} + \alpha_i + \varepsilon_{ipt} \tag{4}$$

In these equations, *i* indexes restaurants, *p* denotes platforms and *t* indexes months. *Yelp* is a dummy variable which is equal to 1 if the observation pertains to reviews on *Yelp.com* and 0 if the observation pertains to reviews on *TripAdvisor.com*[3]. *Yelp_Change* is a dummy variable which is equal to 1 for observations that take place

[3] *Trip* is the reverse coding of *Yelp* (i.e., *Trip* = 1 – *Yelp*).

following the introduction of Facebook Connect on *Yelp.com*, and 0 for observations prior. *Trip_Change* is a dummy variable that is equal to 1 for observations that take place following the introduction of Instant Personalization on *TripAdvisor.com*, and 0 for observations prior. Because the linguistic features of a review might be affected by the reviewer's opinion about the quality of the restaurant, we also control for the average star rating (valence) of reviews in each restaurant-month. Additionally, because a lengthier review may provide greater opportunity for a reviewer to express certain linguistic patterns, we also control for the number of words appearing in the reviews.

The key parameters of interest in these models are β_2 and β_4, our DID estimates, which capture the effects of a policy or system change on user behaviors in the treatment group compared to user behaviors in the control group (Angrist and Pischke 2008). For example, in Equation (1), a positive coefficient for β_2 would suggest that *TripAdvisor*'s Instant Personalization has a positive effect on review volume, compared with the control site *Yelp*. Similarly, a positive coefficient for β_4 would imply that *Yelp*'s Facebook Connect has a positive effect on *Yelp*'s review volume, compared with the control site *TripAdvisor*, which had not yet implement social network integration.

Main Findings

In this section, we report the results of our two-treatment DID estimations. We test our hypotheses and then discuss the economic significance of the key estimates. Further, we provide additional specifications and estimations in our supplementary appendix, where we also report i) a relative time specification around the *TripAdvisor* integration, and ii) separate/single-shock DID estimates of each natural experiment. The relative time specification provides us with a sense of the dynamic effects of social network integration on our outcome variables, over the months that follow, as well as a means of evaluating the parallel trends assumption, which we are notably unable to reject.

Volume Effect

First, we tested the effect of social network integration on the volume of user-generated content (DV = ln (Review Volume) or Review Volume). Overall, our results reported in Table 9 suggest that social network integration is positively associated with the volume of user-generated content. We observe that both *Yelp*'s Facebook Connect feature and *TripAdvisor*'s Instant Personalization feature increased review volumes. This result indicates that both ease of use and reputational benefits seem to play a role in driving up review volumes and that their combined impacts dominate any possible negative effects deriving from the relative loss of anonymity.

In terms of effect sizes, based on the DID estimates of the raw review volume (Column 2 of Table 9), compared with the average monthly review volume of all restaurants in the sample (*mean* = 4.273), *TripAdvisor*'s Instant Personalization increased review volumes by 29.84% and *Yelp*'s Facebook Connect increased review volumes by 20.43%. These estimates are largely consistent with those we obtain in our log specification (Column 1 of Table 4). We therefore find evidence in support of Hypothesis 1 that the social integration led to increases in review volumes.

Interestingly, our result is different from Frutiger et al. (2014), who found that social login leads to decreases in user registration. We surmise that the different findings may be attributable to the differences in both the study context (Frutiger and his colleagues studied a virtual gaming platform, and it is possible users in that setting would not want their social connections to know that they are playing games) and the type of social network integration (Frutiger has solely focused on the social login feature). Moreover, in general, having others know that one is playing games does not provide the reputational or social benefits as in the case of a review platform, where a user's friends may observe that he or she have contributed a review to help others in their purchase decisions.

Table 9. Effect of Social Network Integration on Review Volume

Variables	(1) ln(Review Volume)	(2) Review Volume
Trip	–0.649***(0.019)	–2.543***(0.132)
Trip_Change	0.261***(0.006)	1.230***(0.037)
Trip * Trip_Change	**0.376***(0.014)**	**1.275***(0.111)**
Yelp_Change	0.043**(0.014)	0.209**(0.076)
Yelp * Yelp_Change	**0.194***(0.015)**	**0.873***(0.084)**
ln(words)	0.184***(0.004)	0.733***(0.023)
Rating	0.014***(0.003)	0.075***(0.012)
Constant	–0.075**(0.023)	–0.429**(0.135)
Observations	139,239	139,239
Within R-squared	0.220	0.146
Number of Restaurants	3,968	3,968
Restaurant Fixed Effect	Yes	Yes

Notes: Cluster-robust standard errors in parentheses; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

As noted earlier, social network integration also involves the introduction of a social login feature, making platform registration and login easier for users. In turn, an increase in the rate of new user entry may contribute to increases in review volumes, and possible self-selection on the part of heavy Facebook users. As such, for the moment, we interpret our results as the combined effect of new user entry and changes in user behavior. However, in Section 5, we will provide empirical evidence showing that the bulk of these results are in fact attributable to changes in user behavior, and not the entry of systematically different users.

Mental Process Effects

Table 10 reports our findings related to affective and cognitive processes. We observe that social network integration leads to an increase in the use of language related to affective mental processes, while leading to a decline in language related to cognitive mental processes, supporting Hypotheses 2a and 2b. Table 11 presents our findings related to positive and negative emotions. Note that although social network integration leads to more affective processes in general (more emotions), when we explore different types of emotion, we observe that integration leads to more positive emotions, yet fewer negative emotions. These results support the idea that social presence increases following social network integration, because users may not want their friends to perceive them as being overly negative. The fact that we observe increases in affective processes along with simultaneous declines in cognitive processes also supports our theory, in that the prior literature has suggested repeatedly that each mental process tends to receive focus at the expense of the other.

In terms of effect sizes, the DID estimates for the raw measures (Column 2 & 4 of Table 10; Column 2 & 4 of Table 11) indicate that, compared to the average occurrence of affective processes (*mean* = 7.679) and cognitive processes (*mean* = 15.333), amongst all reviews in our sample, *TripAdvisor*'s Instant Personalization increased language usage related to overall affective processes by 1.02%, while decreasing cognitive processes by 0.75%. Similarly, *Yelp*'s Facebook Connect increased language usage associated with overall affective processes by 9.98%, while decreasing cognitive processes by 2.52%. Further, compared to the average occurrence of positive emotions (*mean* = 6.719) and negative emotions (*mean* = 0.936) in our sample, *TripAdvisor*'s Instant Personalization increased language usage reflecting positive emotions by 1.91%, but decreased negative emotions by 5.02%. And *Yelp*'s Facebook Connect increased positive emotions by 12.07%, and decreased negative emotions by 4.06%, respectively.

Table 10. Affective and Cognitive Processes

Variables	(1) ln(Affective Process)	(2) Affective Process	(3) ln(Cognitive Process)	(4) Cognitive Process
Trip	–0.083***(0.010)	–0.184***(0.053)	0.026***(0.005)	0.342***(0.064)
Trip_Change	0.043***(0.002)	0.307***(0.018)	0.005***(0.001)	0.077***(0.022)
Trip * Trip_Change	**0.022***(0.006)**	**0.078*(0.038)**	**–0.007***(0.003)**	**–0.115*(0.046)**
Yelp_Change	–0.036***(0.011)	–0.489***(0.055)	0.022***(0.006)	0.374***(0.067)
Yelp * Yelp_Change	**0.075***(0.011)**	**0.766***(0.060)**	**–0.023***(0.006)**	**–0.386***(0.073)**
ln(words)	–0.220***(0.002)	–1.959***(0.015)	0.041***(0.001)	0.793***(0.018)
Rating	0.090***(0.001)	0.605***(0.009)	–0.012***(0.001)	–0.201***(0.011)
Constant	2.651***(0.013)	14.422***(0.086)	2.564***(0.007)	12.158***(0.104)
Observations	137,158	137,479	135,043	137,479
R-squared	0.146	0.162	0.016	0.019
Number of Restaurants	3,963	3,965	3,958	3,965
Restaurant Fixed Effect	Yes	Yes	Yes	Yes

Notes: Cluster-robust standard errors in parentheses; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 11. Positive and Negative Affective Processes

Variables	(1) ln(Positive Emotion)	(2) Positive Emotion	(3) ln(Negative Emotion)	(4) Negative Emotion
Trip	–0.102***(0.011)	–0.152**(0.052)	0.087***(0.021)	–0.022(0.018)
Trip_Change	0.048***(0.002)	0.303***(0.018)	–0.045***(0.005)	0.006(0.006)
Trip * Trip_Change	**0.035***(0.007)**	**0.128***(0.037)**	**–0.200***(0.014)**	**–0.047***(0.013)**
Yelp_Change	–0.043***(0.012)	–0.509***(0.054)	0.061**(0.021)	0.017(0.019)
Yelp * Yelp_Change	**0.093***(0.012)**	**0.811***(0.059)**	**–0.123***(0.022)**	**–0.038+(0.021)**
ln(words)	–0.255***(0.003)	–1.948***(0.015)	–0.257***(0.005)	–0.013*(0.005)
Rating	0.183***(0.002)	1.020***(0.009)	–0.283***(0.003)	–0.416***(0.003)
Constant	2.308***(0.015)	11.836***(0.084)	2.248***(0.029)	2.570***(0.029)
Observations	136,760	137,479	109,450	137,479
R-squared	0.233	0.221	0.124	0.127
Number of Restaurants	3,963	3,965	3,936	3,965
Restaurant Fixed Effect	Yes	Yes	Yes	Yes

Notes: Cluster-robust standard errors in parentheses; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Inhibition Effect

Finally, we consider the inhibition effect of social network integration (and thus social presence) on the use of words indicating disagreement or conflict, namely negations (i.e., no, not, never). Table 12 reports our findings. Overall, we observe that both Facebook Connect and Instant Personalization lead to fewer negations. This finding provides support for Hypothesis 3. In terms of effect sizes, the DID estimates for the raw measures (Column 2 of Table 7) show that compared to the average linguistic score of negation in all reviews (*mean* = 1.158), *TripAdvisor*'s Instant Personalization decreased negations by 7.43%, and *Yelp*'s Facebook Connect decreased negations by 5.01%.

Table 12. Effect of Social Network Integration on Negation

Variables	(1) ln(Negation)	(2) Negation
Trip	0.394***(0.014)	0.523***(0.019)
Trip_Change	–0.018***(0.005)	0.028***(0.007)
Trip * Trip_Change	**–0.113***(0.010)**	**–0.086***(0.014)**
Yelp_Change	0.079***(0.015)	0.058**(0.020)
Yelp * Yelp_Change	**–0.106***(0.016)**	**–0.058**(0.022)**
lnwords	–0.235***(0.004)	–0.020***(0.006)
Rating	–0.192***(0.002)	–0.327***(0.003)
Constant	1.851***(0.023)	2.331***(0.031)
Observations	118,205	137,479
R-squared	0.135	0.094
Number of Restaurants	3,944	3,965
Restaurant Fixed Effect	Yes	Yes

Notes: Cluster-robust standard errors in parentheses; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Secondary Analysis: Separating Self-Selection from Behavioral Change

We next elaborate on our findings, with the objective of further identifying the multiple mechanisms underlying our observed effects. In particular, we consider that the increase in review volumes, as well as the shifts in language use, may have arisen from changes in the composition of the user base, changes in user behavior, or some combination of the two. In order to disentangle these three possibilities, we collected additional data at the user level from *TripAdvisor.com*. Specifically, we identified every user who contributed at least one review in our initial sample on *TripAdvisor*. We then collected every review ever written by these users, including the reviews they may have written about restaurants that did not appear in our initial sample. We then used these data to construct a user-level panel of reviewing activity.

User-Level Analysis

The user-level data enables us to examine i) any shifts in the rate of first-time reviewers (new user entry) over time, ii) any shifts in average reviewing volumes within users who had registered prior to the social integration event, and iii) whether new entrants and existing users exhibit any systematic differences in their language use in the post social integration period. Jointly, addressing these questions can help us to assess the degree to which self-selection or behavioral changes are driving our observed results.

Table 13 reports changes in user-level monthly reviewing volume and language characteristic over the two-year period around *TripAdvisor*'s social network integration (12 months before and after the event). For review volumes, we estimate the effect of *TripAdvisor* social network integration on users' average monthly number of reviews. For the linguistic features, due to limited scalability of the LIWC software to process large amounts of textual data, we randomly sampled a subset of users who jointly authored a total of approximately 750,000 reviews. Amongst these reviews, 96,356 were authored within our two year time window. Considering the results in Table 8, we observe that the social network integration is significantly associated with changes in all of our outcome variables, suggesting that our results are attributable to within-user changes in behavior.

Table 13. Effect of Social Network Integration on Within User Review Volume and Language Characteristics

Variables	(1) Review Volume	(2) Affective Process	(3) Positive Emotion	(4) Negative Emotion	(5) Cognitive Process	(6) Negation
Trip_Change	**0.792*** **	**0.103***	**0.139*** **	**–0.033***	**–0.165** **	**–0.036***
	(0.017)	**(0.044)**	**(0.041)**	**(0.013)**	**(0.219)**	**(0.014)**
Constant	2.288***	6.480***	5.501***	0.748***	7.927***	1.304***
	(0.012)	(0.036)	(0.034)	(0.010)	(0.044)	(0.011)
Observations	244,978	96,356	96,356	96,356	96,356	96,356
F-Statistic	2164.85***	5.44*	11.34***	6.64**	7.05***	6.63**
Number of Users	70,450	5,174	5,174	5,174	5,174	5,174
User Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Cluster-robust standard errors in parentheses; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Next we examine whether the users who entered *TripAdvisor* after the social network integration were systematically different from the pre-existing users in their language use, focusing on reviews authored *after* the integration had taken place. To answer this question, we conducted two additional analyses, assessing the relationship between language use and user tenure. We operationalized user tenure in two ways: i) in a continuous manner, as the number of month since the user registered on the platform, and iii) in a binary manner, via an indicator of whether a user registered before or after the social network integration. Table 14 and Table 15 report our findings on the relationship between language use and user tenure. The results show that the language characteristics of reviews (posted after social integration) do not significantly differ between pre-existing users and new users. In other words, the reviewers' language use does not depend on whether the authors had registered before or after social network integration took place. This implies, in turn, that the observed effects related to changes in language appear to be primarily caused by changes in user behavior, rather than self-selection.

Table 14. Effects of Continuous User Tenure on Review Language Characteristics

Variables	(1) Affective Process	(2) Positive Emotion	(3) Negative Emotion	(4) Cognitive Process	(5) Negation
ln(tenure)	0.019	0.029	–0.008	0.030	0.010
	(0.016)	(0.016)	(0.005)	(0.019)	(0.006)
ln(words)	–3.065***	–3.081***	0.010	0.935***	0.045**
	(0.047)	(0.048)	(0.012)	(0.043)	(0.014)
Rating	0.634***	1.036***	–0.406***	–0.185***	–0.444***
	(0.020)	(0.021)	(0.011)	(0.026)	(0.010)
Constant	18.270***	15.883***	2.386***	12.053***	3.088***
	(0.249)	(0.245)	(0.093)	(0.293)	(0.100)
Observations	46,341	46,341	46,341	46,341	46,341
R-squared	0.259	0.290	0.088	0.022	0.064
Number of Restaurants	2,755	2,755	2,755	2,755	2,755
Restaurant Fixed Effect	Yes	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes	Yes

Notes: Cluster-robust standard errors in parentheses; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 15. Effects of Binary User Tenure on Review Language Characteristics

Variables	(1) Affective Process	(2) Positive Emotion	(3) Negative Emotion	(4) Cognitive Process	(5) Negation
Reg_after_change	–0.062	–0.076	0.011	–0.104	–0.029
	(0.048)	(0.047)	(0.014)	(0.068)	(0.018)
ln(words)	–3.064***	–3.079***	0.009	0.936***	0.046**
	(0.047)	(0.048)	(0.012)	(0.043)	(0.014)
Rating	0.634***	1.037***	–0.406***	–0.185***	–0.444***
	(0.020)	(0.021)	(0.011)	(0.026)	(0.010)
Constant	18.327***	15.960***	2.368***	12.150***	3.116***
	(0.247)	(0.244)	(0.094)	(0.292)	(0.099)
Observations	46,355	46,355	46,355	46,355	46,355
R-squared	0.259	0.290	0.088	0.022	0.064
Number of Restaurants	2,755	2,755	2,755	2,755	2,755
Restaurant Fixed Effect	Yes	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes	Yes

Notes: Cluster-robust standard errors in parentheses; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Falsification Tests

We conclude our analyses with a number of falsification tests, intended to help rule out spurious correlation. Because there is no theoretical reason to expect that a website feature like social network integration would bear a relationship with the occurrence of common words, such as articles (e.g., 'a', 'an', 'the'), filler content (e.g., 'I mean', 'you know'), or numbers (e.g., 'second', 'thousand'), we would not expect to observe a significant effect on these measures. Thus, if we were to observe a significant effect, it would raise questions about the validity of our main results. Fortunately, as the results in Table 16 demonstrate, we observe no significant effects of social network integration on these outcome variables, lending further credence to our identification strategy.

Table 16. Falsification Tests

Variables	(1) Article	(2) Filler	(3) Numbers
Trip	0.354***(0.064)	–0.065***(0.010)	0.085***(0.022)
Trip_Change	–0.074***(0.014)	–0.013***(0.003)	–0.011*(0.005)
Trip * Trip_Change	0.054(0.041)	0.007(0.006)	–0.003(0.014)
Yelp_Change	–0.072(0.070)	–0.004(0.011)	–0.025(0.023)
Yelp * Yelp_Change	–0.012(0.072)	–0.013(0.011)	0.021(0.024)
ln(words)	0.103***(0.017)	0.018***(0.003)	0.102***(0.006)
Rating	0.150***(0.009)	–0.036***(0.002)	–0.022***(0.003)
Constant	7.468***(0.093)	0.303***(0.017)	0.316***(0.031)
Observations	137,479	137,479	137,479
R-squared	0.012	0.013	0.005
Number of Restaurants	3,965	3,965	3,965
Restaurant FE	Yes	Yes	Yes

Notes: Cluster-robust standard errors in parentheses; *** $p<0.001$, ** $p<0.01$, * $p<0.05$

Discussion

Implications

This study contributes to several important streams of IS research. First, this study adds to recent discussion on the potential value and impact of social media (Luo et al. 2013; Kane 2014; Kane et al. 2014; Huang et al. 2015). While prior studies examined the effect of social media on firm value (Luo et al. 2013), and the influence of social media on work-related content (vs. leisure-related content) sharing behavior within the firm's working environment (Huang et al. 2015), we unravel the effects of online review platforms' integration with social media websites (social network integration) on user content generation.

Second, this study also extends prior research on social presence, anonymity and online reviews in several ways. Past work on social presence has focused on limited outcomes such as trust (Ou et al. 2014), purchase intentions (Animesh et al. 2011), and consumer product choices (Rhue and Sundararajan 2013). We expand prior research by considering another outcome, users' authorship of online reviews, considering their volumes and linguistic features. In addition, prior studies on anonymity have mainly emphasized privacy implications (Ayyagari et al. 2011; Acquisti et al. 2013) and online disinhibition effects (Suler 2004; Reader 2012; Santana 2014). We present empirical evidence that anonymity (or its absence) can affect language use as well.

Lastly, previous research on the social aspects of online reviews, and user-generated content more broadly, have focused on outcomes such as contribution quantity (Huberman et al. 2009; Chen et al. 2010), evaluation negativity and extremity (Wang

2010; Goes et al. 2014). This study considers a novel aspect: the linguistic features of review content. Our study provides the first empirical evidence suggesting increases in the social elements of an online review platform may cause existing reviewers to write more often, with more emotional language and less cognitive language, as well as to employ less negative language.

This study carries important practical implications for firms operating IT platforms that host and heavily rely upon user generated content. Given the recent trend toward social network integration by various online platforms, it is crucial that we improve our understanding of the possible unintended consequences of social integration and, in turn, social presence and anonymity (Kane 2015). Specifically, we have found that social presence increases the contribution volume of online reviews. This result suggests that social network integration is likely to be most useful for online review websites (or websites that host other forms of user-generated content) that face challenges of under-provision (Avery et al. 1999; Burtch et al. 2016). In addition, our results also show that integration leads to decreases in cognitive language, along with increases in the use of (primarily positive) emotional language. Past work has noted that emotional, positive reviews are perceived by consumers to be less helpful (Chen and Lurie 2013; Yin et al. 2014; Hong et al. 2016). Thus, our results suggest that review platforms, if implementing social network integration, should consider pairing the system change with attempts to nudge users to be less emotional and more logical when authoring their reviews, in an effort to maintain or enhance review quality.

Limitations

Our work is subject to a number of limitations. First, our measures are relatively simplistic and the accuracy of the results depends on how comprehensive the LIWC dictionaries are. However, LIWC has recently been successfully applied by a number of scholars in Information Systems (Goes et al. 2014; Yin et al. 2014) and Marketing (Sridhar and Srinivasan 2012; Lurie et al. 2014). Second, due to the observational nature of our data, although multiple empirical tests suggest that self-selection is not a serious issue, we are not able to completely rule out selection as a partial explanation for our results. This is primarily because users had the ability to opt-in or opt-out of the social network integration. In turn, this may suggest that our results are driven in large part by a subset of users who chose to accept the integration, or users who had Facebook accounts. Although our data does not enable us to identify which users accepted the integration features, or which users have Facebook accounts, we do not believe this should be a significant concern. Again, as noted previously, opting out of the *TripAdvisor* integration, while feasible, was reportedly difficult for users to perform. Additionally, our user-level analyses suggest no evidence that users who entered the platform following *TripAdvisor*'s integration were systematically different from pre-existing users. Third, although we have tested the parallel assumption for the DID analysis, underlying differences that we have been unable to detect may still exist between the two platforms we studied. As such, similar to extant studies that use such a research design (Mayzlin et al. 2014), our results should be interpreted with some caution. Finally, even if our results might not to generalize to users who do not have Facebook accounts, Facebook now boasts more than 1.5 Billion users. Accordingly, such users are unlikely to be representative of the broader user population (that is, users who lack a Facebook account would quite possibly constitute a minority). Taken together, the above deliberation suggests that our results are likely to be both generalizable, and not driven predominantly by self-selection.

Future Research Directions

There is significant potential for future work. First, it may be fruitful to explore data mining techniques, such as natural language processing, to undertake a more nuanced textual analysis in a more granular fashion. Second, we have merely focused on a few aspects that social network integration may impact. Future research can extend our study by exploring other outcomes of social network integration. For example, researchers can examine the impact of integration on dimensions of psychological distance. Social network integration might change the perceived social distance among users on the online platforms, which in turn may affect word usage (Holtgraves 2003). Third, although we have leveraged natural experiments to identify the effects of social network integration on user-generated content, due to the observational nature of our data, these analyses may nonetheless suffer from issues of endogeneity (e.g., unobserved correlated shocks; self-selection). Because users in both cases had the option to reject the treatment, by failing to opt in, or choosing to opt out, our results may be largely attributable to individuals who were open to social network integration and choose to adopt it. Future studies could improve on our analysis via experimental manipulations of reviewer social presence or anonymity to better establish causal relationships. Alternatively, one might envision an experiment in which individuals are exogenously "friended" on a review platform, to examine whether the content of their reviews changes.

CHAPTER 4.

ESSAY 3. HOW BOOK-TO-FILM ADAPTATION AFFECTS USER CONTENT GENERATION

Abstract

This study empirically investigates the impact of book-to-film adaptations on books' user content generation. We propose that book-to-film adaptations not only influence evaluations of the adapted books (i.e. rating), but also affect the linguistic characteristics of book reviews (i.e. viewing, comparative, and affective processes). Using a combination of text analytics approaches and an econometric method to analyze a unique dataset of user reviews collected from *Amazon.com*, *Goodreads.com* and *IMDB.com*, we found that book-to-film adaptations would decrease book rating (publicity effect), but make greater parts of the reviews focus on reviewers' experiences viewing the film and comparing the adaptation and the original book (therefore increasing the viewing and comparative effects). Additionally, book-to-film adaptations would also increase the use of language in book reviews that reflects an affective process (emotional spillover effect). We then conducted additional analyses on the impact of the review changes on review helpfulness. The results suggests that decreased star rating, as well as increased viewing and affective language in book reviews after film adaptation is positively related to review helpfulness. Lastly, we explored the interaction effects of book-to-film adaptation with regard to book and movie rating differences, movie age-rating and book classifications.

"Never judge a book by its movie."
– *J. W. Eagan*

"Books and movies are like apples and oranges. They both are fruit,
but taste completely different."
– *Stephen King*

"The summer movies are coming out. My advice: just stay home and burn a good book."
– *Stephen Colbert*

Introduction

Book-to-film adaptation refers to the cinematographic transformation of a book into a feature film (Chatman 2002) — a practice that has been a popular in Hollywood for many years (Forbes 2015). It is estimated that a third of all films ever made are book-to-film adaptations (Public Broadcasting Service 2011). Successful film adaptations, such as *The Lord of the Rings*, *Jurassic Park*, and *Gone Girl*, have become household names, generating billions of dollars (Marlow 2015). There are clear benefits to book authors as well as film producers in adapting films from books. For instance, adapted books could provide films with more developed characters and coherent storylines; additionally, such films could capitalize on name recognition and inherit existing fan bases from adapted books (Green 2007).

With film adaptations flourishing, the books that are adapted to films also receive surges in name recognition and readership (Seger 2011). Numerous examples (*Harry Potter, The Girl on the Train, Hunger Games*) have shown that film adaptation significantly fuels book sales (Nielsen 2010; Lewis 2012; Cadden 2016). Meanwhile, there have been fierce debates fueled by comparisons between books and their film adaptations. Some people are inclined to think that books are almost always better their cinematic analogues because of the problems raised by the adaptation process (Leitch 2009; Santos 2013; Swanson 2016). In contrast, others argue that film adaptations sometimes are better than the books, as film adds more aesthetic and realistic elements to a book's story (Peitzman 2016; Weiss 2015).

The purpose of this study is to investigate the impact of book-to-film adaptations on consumer perception of book quality. Although past research on book-to-film adaptation has examined the nature of the theatrical translation and the financial calculations of film adaptations (Zatlin 2005), the reputational consequences (i.e. book reviews) of books that are brought about by book-to-film adaptation remain an open question with limited answers. This is a crucial issue for book publishers, because online reviews could significantly influence book sales (Chevalier and Mayzlin 2006), both positively and negatively. Furthermore, examining the impact of book-to-film adaptation on book reviews is important for academic researchers, because it could improve our understanding on the various antecedents of online review characteristics. Therefore, we seek to answer the following research questions:

How does the release of film adaptations influence consumer evaluations of the books that the films are based on? Specifically, how do the films affect the star ratings and linguistic features of book reviews?

Building on past research, we develop hypotheses that lead to an expectation that book-to-film adaptations have three types of effects on book reviews: the publicity effect, the viewing effect, and the emotional spillover effect. Analyzing a unique dataset of reviews collected from Goodreads.com, Amazon.com and IMDB.com, we found support for all three effects. Due to the publicity effect, the film publicity attracts a broader audience than the original book target market who have different tastes. As a result, book-to-film adaptations lead to a decrease in book star ratings. Since consumers had the opportunity to watch the movie trailer or see the movie, language related to viewing is more likely to be used in the reviews of the book. Thus, the viewing effect is indicated by an increase in reviewers' use of language reflecting their viewing experiences and comparing the original books with their cinematic antecedents following their release. In addition, we found evidence of an emotional spillover effect, since book-to-film adaptations lead to increases in the use of affective language in book reviews. Further, we conducted a series of additional analyses to understand the implications of the main findings. In particular, we found that a decrease in star rating, as well as increases in viewing and affective language, are positively associated with review helpfulness. Thus, our results suggest that book-to-film adaptation might be beneficial to book review helpfulness by shifting the star rating and certain linguistic characteristic of the review text. We lastly explored the interaction effects of book-to-film adaptations with regard to book and movie rating difference, movie age-rating and book classification. The results of the interaction effects show that the effects of book-to-film adaptation on viewing and comparative language are stronger when there are large book and movie rating differences. Moreover, the results suggest that the viewing effect of book-to-film adaptation is stronger for products with more mature audiences, while the emotional spillover effects of film adaptation are weaker for products with more mature audiences.

This study contributes to the recent trend in online review research that examines different antecedents of review characteristics (Goes et al. 2014). Prior literature has identified several antecedents, such as reviewer cultural background (Hong et al. 2016), reviewer psychological distance (Huang et al. 2016) and review system designs (Li and Hitt 2008; Huang et al. 2016; Chen et al. 2017). At the same time, related work has been looking at how to effectively stimulate online review contribution (Burtch et al. 2015; Huang et al. 2016). Although research on the antecedents of online reviews has been developing, this area still remains under-explored. In particular, there is limited research attention paid to the impact of film adaptation on book review characteristics. We aim to fill this void in the literature.

The rest of the paper is organized as follows. First, we summarize the related literature on book-to-film adaptation and online reviews. Second, we elaborate on the conceptual model and hypotheses development. Then, the data and methodology section provides details on the data and estimation equations. The results section follows, reporting our findings on hypotheses testing. There is also a series of additional analyses to help understand the implications and boundary conditions of the main effects. Lastly, the paper concludes with a discussion of this study's implications, limitations and possible future directions.

Related Literature

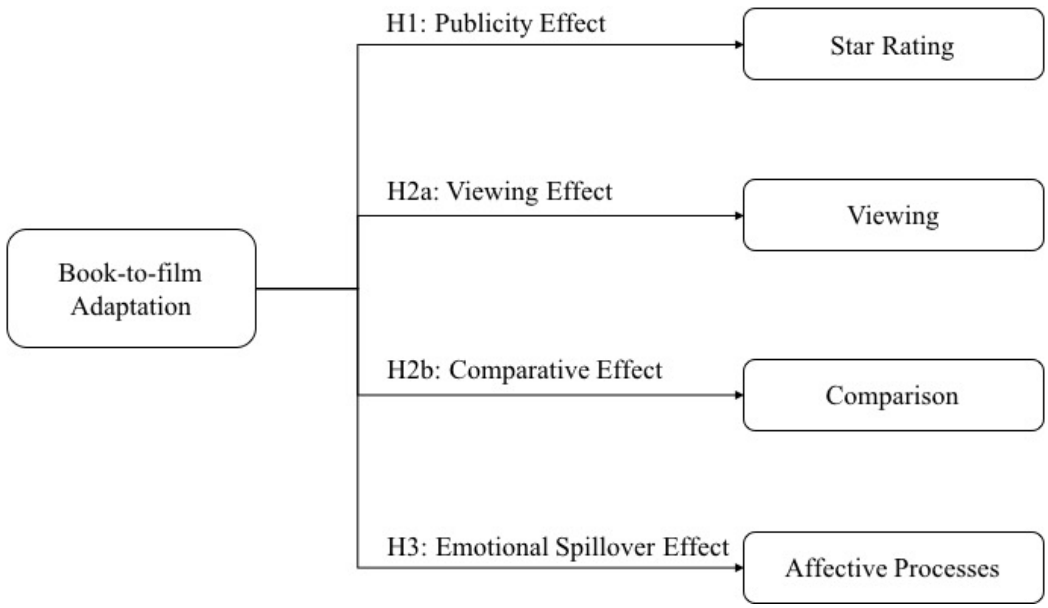
User content generation refers to the phenomenon that users create and share original content to online platforms (Ghose and Han 2011; Albuquerque et al. 2012). Examples of user content generation include online reviews (Mudambi and Schuff 2010), posts on online forums (Ba and Pavlou 2002) or other online communities (Ren et al. 2012), and social media interactions (Zeng and Wei 2013), to name a few. Research on user content generation resides in multiple disciplines, such as Marketing (e.g. Basuroy 2003; Chevalier and Mayzlin 2006), Information Systems (e.g. Dellarocas 2003; Clemons et al. 2006; Forman et al. 2008), and Psychology (e.g. Eichstaedt et al. 2015; Snefjella and Kuperman 2015). Prior literature has extensively examined the characteristics and consequences of user content generation, establishing its importance in business research and practice. For example, it has been shown that negative and longer reviews are perceived to be more helpful than shorter, positive ones (Mudambi and Schuff 2010; Chen and Lurie 2013). It has also been shown that user-generated content significantly influences product sales (Dellarocas et al. 2007), business strategy (Clemons et al. 2006), and market competition (Kwark et al. 2014). Meanwhile, review credibility (Lu et al. 2013) and review platform reputation (Gu et al. 2012) moderate the impact of user-generated content on business performance.

Less examined in the user content generation literature are antecedents of review characteristics. Prior research has primarily focused on users' characteristics, such as their popularity within a review platform (Goes et al. 2014). As a user becomes more popular in an online community, they tend to provide a higher volume of more objective reviews, and the ratings of their reviews become more negative with a higher variance. Cultural conformity of the users' country is identified as another important antecedent to review characteristics, such that users from countries that promote conformity as opposed to an individualistic culture tend to provide ratings that are in agreement with others and express fewer emotions in reviews (Hong et al. 2016). What we propose to examine in this study is how book-to-film adaptations would affect review characteristics, thereby extending this important line of research. Also less examined, yet steadily gaining attention, are the linguistic features of online reviews. Emerging research primarily uses two approaches to measure and quantify the linguistic features of online reviews as textual documents. While some researchers have begun to develop methods for extracting topics from short and noisy online reviews (Li et al. 2017), the majority of research focuses on matching lexicons to a word dictionary, such as the LIWC dictionary (Yin et al. 2014, Huang et al. 2016). Researchers have examined different linguistic features, such as anger, anxiety, negative and positive emotions, disagreements, etc. in different contexts. In this study, we propose to examine how book-to-film adaptations would affect different linguistic features in online reviews, which extends this line of research.

Conceptual Model

This study aims to test four kinds of effects of book-to-film adaptations on the characteristics of book UGC: the effects of book-to-film adaptations on book star rating (publicity effect), on the use of viewing (viewing effect) and comparative language (comparative effect) in review text, and on the prevalence of language reflecting affective processes (emotional spillover effect). The conceptual model of this study is shown in Figure 3.

Figure 3. Conceptual Model



Hypotheses Development

Publicity Effect

The release of book-to-film adaptations, along with complementary marketing efforts, can increase the publicity of the original books (Kim et al. 2016). The elevated publicity, either good or bad, could spur interest in consumers, leading them to purchase, read and even review the adapted books (Berger et al. 2010). As reported by the *Time* magazine, book sales often experience sale bumps with the release of film adaptations (Begley 2015). Meanwhile, prior research suggests that the consumers of experience goods, such as books, face the issue of product fit uncertainty (Hong and

Pavlou 2014). Purchasing, reading and reviewing a book requires consumers to invest money and time. Thus, consumers might employ external information during the book selection process, such as book publicity, which is derived from film adaptations, to help mitigate uncertainty (Winoto and Tang 2008).

However, as the audience for the original books expands, the diversity of consumer tastes also increases (Kovács and Sharkey 2014). If a consumer purchases a book because of its publicity rather than its features, a misfit between the book and the consumer's taste is likely to occur. In such a scenario, following a film adaptation, consumers might be less favorable to the book than before the adaptation because of the increased probability of mismatch between the book features and their tastes. Moreover, people intend to derive value from exclusivity, a phenomenon call "snob effects" (Becker 1991; Fionda and Moore 2009). As book-to-film adaptation makes a book more popular, increased audience size might drive consumers to devalue the now less exclusive books, leading to decreases in book ratings. Bearing the above arguments in mind, we propose that the increased publicity of books derived from book-to-film adaptations could lead to decreased evaluation of books. Formally, we hypothesize that:

H1: Book-to-film adaptations lead to a decrease in the star ratings of book reviews.

Viewing & Comparative Effect

Before the release of a book-to-film adaptation, consumers solely rely on the book itself when writing book reviews. In contrast, after the adaptation is released, consumers are offered access to both the film and the book. Thus, some book reviews posted after the release of an adaptation could be authored by reviewers who have seen the movie in addition to reading the book. Book-to-film adaptation typically involves borrowing the skeletal stories of the books and then building "stylistic equivalents" in the movies to the original book's "tone, values, imagery, and rhythm" (Zatlin 2005, p. 154). If a consumer has watched the movie and read the book, he or she would associate the book with its movie. Information integration theory suggests that individuals tend to combine information from different sources to form a global evaluation (Anderson 1981; 2014). As a result, when consumers write reviews for a book after having watched the movie, they might as well mention the movies in their review of the book. Hereby, we hypothesize that:

H2a: Book-to-film adaptation leads to increase in the use of language reflecting viewing in book reviews.

Consumers often compare books and their film adaptations, because films adapted from books are always not quite the same as the books themselves (Griffith 1997). Such differences can be attributed to factors such as storyline compression, a director or screenwriter's unique perspective, and film's production limitations, to name a few (Zatlin 2005). As Stephen King once said, "Books and movies are like apples and oranges. They both are fruit, but taste completely different." People look for differences between books and their movies, paying attention to disparities in everything ranging from characters and relationships to narrative themes. Social comparison theory suggests that individuals tend to compare things when striving to create an accurate evaluation (Suls 1977; Suls and Wheeler 2012). After the release of a book-to-film adaptation, consumers who have read the book and seen the movie might compare the two when evaluating the book. Thus, we hypothesize the following:

H2b: Book-to-film adaptation leads to increase in the use of language reflecting comparison in book reviews.

Emotional Spillover Effect

Emotion captures "mental experience with high intensity and high hedonic content" (Cabanac 2002). Media richness theory suggests that communication media differ in their capacity to deliver emotions (Daft and Lengel 1986; Dennis and Kinney 1998). Specifically, rich media (e.g. video) are capable of communicating more emotions than media that are low in richness (e.g. written documents) (Lengel and Daft 1989). To further elaborate, book authors mainly rely on verbal expressions to convey emotions, the communication of which depends on the readers' immersion and interpretations of the literal and figurative expressions (Fussell 2002). In contrast, filmmakers can employ a wide range of verbal, auditory and visual mechanisms to elicit emotions, such as acting, dialogue, sound effects, close-up scenes and special effects, among others (Plantinga and Smith 1999). Emotional experiences in films stem from the close association of the audience's attention and filmic stimulus (Smith 2003).

People report experiences of strong emotions caused by films: "films make us cry, flinch and cheer" (Pawlowski 2014). Recent neuroscience evidence has also reported the function of mirror neurons that helps people relate to emotions when seeing emotional content in a movie (Zacks 2014). For example, when a film presents a heartbreaking scene with characters crying and sad music playing, the audiences' mirror neurons are activated and people are likely to mirror the film's scene to their real life experiences when they themselves felt sad. Such a mirroring mechanism induces the audience to resonate with an emotional scene.

Bearing the above argument in mind, we propose that book reviewers who have watched the film adaptations and read the book might experience stronger emotions than reviewers who have only read the book. As a result, the reviewers' comparatively strong emotional experience with the film adaptation might spill over to the book, such that the book reviews that are authored by reviewers who have seen the movie and read the book show more emotional language than reviews that are written by reviewers who have solely read the book. Hereby, we hypothesize that:

H3: Book-to-film adaptation leads to increase in the use of language reflecting affective language in book reviews.

Data and Methodology

The data of this study were collected from *Goodreads.com*, *Amazon.com* and *IMDB.com*. The data collecting process involved several steps. First, we collected book information (e.g. book title, book author and book's *Amazon* product page link) on a set of randomly selected books from the list of "popular books made into movies" on *Goodreads.com.*[3] It is notable that there are several "books made into movies" lists on Goodreads.com. We selected the list that seemed most comprehensive. Second, we matched each book with its corresponding book-to-film adaptation webpage on

[3] https://www.goodreads.com/shelf/show/books-made-into-movies

IMDB.com, creating a set of book-film pairings. Lastly, we collected book review information (e.g. rating, review text, verified purchase) from *Amazon.com* and movie information (e.g. movie release year, movie age-ratings, movie review ratings) from *IMDB.com* for the set of book-film adaptation pairings. During the data cleaning process, we excluded graphic books or books that were introduced after movies. We also dropped observations that showed missing values. After data cleaning and trimming, we observed 149 book-movie pairings with 142,601 book reviews. Table 17 lists all the books, book authors, average book ratings and movie ratings in our sample.

Table 17. List of Book-Movie Pairings

Amazon ASIN	*Book Title*	*Book Author*	*Book Rating* (1-5)	*Movie Rating* (1-10)
7173040	How the Grinch Stole Christmas!	Dr. Seuss	4.693	6.0
000720230X	Prince Caspian	C.S. Lewis	4.339	6.6
7442912	Insurgent	Veronica Roth	4.421	6.7
60512628	My Friend Flicka	Mary O'Hara	4.611	6.1
60744871	The Pursuit of Happyness	Chris Gardner	4.018	7.9
006075995X	Divine Secrets of the Ya-Ya Sisterhood	Rebecca Wells	4.002	5.9
60764899	The Lion, the Witch, and the Wardrobe	C.S. Lewis	4.647	6.9
006081926X	Million Dollar Baby: Stories from the Corner	F.X. Toole	4.078	8.1
60838582	Fast Food Nation: The Dark Side of the All-American Meal	Eric Schlosser	4.336	6.3
61142026	Stardust	Neil Gaiman	4.263	7.7
61173002	Amazing Grace: William Wilberforce and the Heroic Campaign to End Slavery	Eric Metaxas	4.670	7.5
62024035	Divergent	Veronica Roth	4.454	6.8
62107062	American Sniper	Chris Kyle, Scott McEwen, Jim DeFelice	4.451	7.4
62245538	The Bling Ring: How a Gang of Fame-Obsessed Teens Ripped Off Hollywood and Shocked the World	Nancy Jo Sales	3.432	5.6
64400565	Stuart Little	E.B. White, Garth Williams	4.115	5.9
99453452	Point of Impact	Stephen Hunter	4.615	7.2

Table 17, continued

140275363	The Iliad	Homer, Robert Fagles, Bernard Knox	4.355	7.2
141003774	The Mortdecai Trilogy	Kyril Bonfiglioli	3.833	5.5
141182245	Dream Story	Arthur Schnitzler	4.045	7.3
141439661	Sense and Sensibility	Jane Austen	4.277	7.7
014240165X	Stormbreaker	Anthony Horowitz	4.584	5.1
142437204	Jane Eyre	Charlotte Brontë, Michael Mason	4.346	7.4
152050167	The Whale Rider	Witi Ihimaera	4.308	7.7
156031191	Winter's Tale	Mark Helprin	3.322	6.2
307592731	Wild: From Lost to Found on the Pacific Crest Trail	Cheryl Strayed	4.230	7.2
307948684	Headhunters	Jo Nesbo	3.996	7.6
312305060	The Hours	Michael Cunningham	4.018	7.6
312421990	What Was She Thinking? [Notes on a Scandal]	Zoë Heller	4.230	7.4
312426089	The Good German	Joseph Kanon	3.966	6.1
031605755X	Winter's Bone	Daniel Woodrell	4.253	7.2
316068047	The Host	Stephenie Meyer	4.368	5.9
031610969X	Julie and Julia: 365 Days, 524 Recipes, 1 Tiny Apartment Kitchen: How One Girl Risked Her Marriage, Her Job, and Her Sanity to Master the Art of Living	Julie Powell	2.729	7.0
316159794	Cross	James Patterson	3.889	5.1
316706000	The Ice Storm	Rick Moody	3.615	7.5
330398121	Tiger Eyes	Judy Blume	4.503	6.4
330419641	Absolute Power	David Baldacci	4.315	6.7
340707429	Tomorrow Never Dies	Raymond Benson	4.211	6.5
340818670	Hearts in Atlantis	Stephen King	4.110	6.9
345453743	The Ultimate Hitchhiker's Guide to the Galaxy	Douglas Adams	4.668	6.8
345465083	Seabiscuit: An American Legend	Laura Hillenbrand	4.750	7.3
345498127	Starter for Ten	David Nicholls	4.269	6.8
345803922	Anna Karenina	Leo Tolstoy, Louise Maude, Alymer Maude	4.455	6.6
374165270	Gomorrah	Roberto Saviano	4.008	7.0
374533571	The Silver Linings Playbook	Matthew Quick	4.331	7.8
375408266	The Reader	Bernhard Schlink, Carol Brown Janeway	3.703	7.6
375415122	The Snow Queen	Hans Christian Andersen	4.556	7.7

Table 17, continued

375703314	The Last King of Scotland	Giles Foden	4.049	7.7
375704027	Norwegian Wood	Haruki Murakami	4.257	6.4
375706674	No Country for Old Men	Cormac McCarthy	4.011	8.1
037570910X	The Feast of Love	Charles Baxter	3.879	6.7
375727434	Hateship, Friendship, Courtship, Loveship, Marriage: Stories	Alice Munro	4.295	6.1
375826696	Eragon	Christopher Paolini	3.888	5.1
385319037	Animal Husbandry	Laura Zigman	3.381	6.1
393324818	Moneyball: The Art of Winning an Unfair Game	Michael Lewis	4.493	7.6
039915731X	Locked On	Tom Clancy, Mark Greaney	3.844	6.2
425098087	Zodiac	Robert Graysmith	4.072	7.7
425242064	Conan the Barbarian	Michael A. Stackpole	3.500	5.2
425259358	Tom Clancy Presents: Act of Valor	Dick Couch, George Galdorisi	4.439	6.5
439851165	The Ant Bully	John Nickle	3.875	5.9
440421705	Hoot	Carl Hiaasen	4.369	5.6
446547654	The Best of Me	Nicholas Sparks	4.060	6.6
446584975	Life Itself	Roger Ebert	4.259	7.9
446675369	Bless Me, Ultima	Rudolfo Anaya	4.096	6.4
446677388	Kiss the Girls	James Patterson	4.212	6.5
446692638	Along Came a Spider	James Patterson	4.255	6.3
450542882	Four Past Midnight	Stephen King	4.142	6.6
451184424	Unstrung Heroes: My Improbable Life with Four Impossible Uncles	Franz Lidz	4.828	6.8
451933028	The Green Mile	Stephen King	4.755	8.5
525478817	The Fault in Our Stars	John Green	4.680	7.9
545010225	Harry Potter and the Deathly Hallows	J. K. Rowling	4.722	8.1
553376055	How I Live Now	Meg Rosoff	3.819	6.5
553384155	Flags of Our Fathers	James D. Bradley, Ron Powers	4.699	7.1
553813951	Nights in Rodanthe	Nicholas Sparks	3.690	5.9
553816713	The Notebook	Nicholas Sparks	4.178	7.9
571199976	Proof	David Auburn	4.364	6.8
571212921	A Beautiful Mind	Sylvia Nasar	4.251	8.2
057506708X	The Pianist: The Extraordinary Story of One Man's Survival in Warsaw, 1939-45	Anthea (translator) Szpilman Wladyslaw	4.791	8.5
057507681X	A Scanner Darkly	Philip K. Dick	4.394	7.1

Table 17, continued

059011848X	Freaky Friday	Mary Rodgers	4.219	6.1
618346260	The Two Towers	J.R.R. Tolkien	4.610	8.7
670026220	Argo: How the CIA & Hollywood Pulled Off the Most Audacious Rescue in History	Antonio J. Mendez, Matt Baglio	4.245	7.8
670061107	That Summer	Sarah Dessen	3.956	5.7
671522817	Braveheart	Randall Wallace	3.528	8.4
679751319	Dead Man Walking: The Eyewitness Account of the Death Penalty That Sparked a National Debate	Helen Prejean	4.290	7.6
679879242	The Golden Compass	Philip Pullman	4.472	6.1
689040342	The Spiderwick Chronicles Box Set	Holly Black, Tony DiTerlizzi	4.565	6.6
715628488	Iris: A Memoir of Iris Murdoch	John Bayley	4.000	7.1
739416383	The Prize Winner of Defiance, Ohio: How My Mother Raised 10 Kids on 25 Words or Less	Terry Ryan	4.504	7.3
074326455X	Prince of Thieves	Chuck Hogan	4.235	7.6
743271327	Brokeback Mountain	Annie Proulx	4.421	7.7
743442849	Addicted	Zane	4.101	5.2
743454537	My Sister's Keeper	Jodi Picoult	4.102	7.4
743477545	A Midsummer Night's Dream	William Shakespeare	4.464	6.5
743482751	Much Ado About Nothing	William Shakespeare	4.183	7.2
743482832	The Tempest	William Shakespeare	4.152	5.4
743484258	Big Fish	Daniel Wallace	4.048	8.0
074349282X	Between a Rock and a Hard Place	Aron Ralston	3.911	7.6
743495667	In Her Shoes	Jennifer Weiner	3.926	6.5
747545251	Psycho	Robert Bloch	4.472	4.6
747560595	The Virgin Suicides	Jeffrey Eugenides	4.035	7.2
751525677	Apt Pupil	Stephen King	4.282	6.7
752863886	The Bourne Supremacy	Robert Ludlum	3.902	7.8
755102800	A Shine Of Rainbows	Lillian Beckwith	5.000	7.2
771068719	The English Patient	Michael Ondaatje	4.021	7.4
785122133	X-Men: The Dark Phoenix Saga	Chris Claremont, John Byrne	4.687	6.8
785144404	Winter Soldier, Vol. 1: The Longest Winter	Ed Brubaker, Butch Guice	3.727	7.8
786849002	The True Meaning of Smekday	Adam Rex	4.491	6.7
786869038	I Love You Phillip Morris: A True Story of Life, Love, & Prison Breaks	Steve McVicker	4.526	6.7
786889373	The Big Picture	Douglas Kennedy	4.279	6.7
786891076	Shopgirl	Steve Martin	3.847	6.4

Table 17, continued

786891084	Love, Rosie	Cecelia Ahern	4.015	7.2
802118585	Night Train to Lisbon	Pascal Mercier, Barbara Harshav	3.857	6.8
802142842	Cold Mountain	Charles Frazier	3.771	7.2
802775802	Alan Turing: The Enigma	Andrew Hodges, Douglas R. Hofstadter	3.422	8.1
080507516X	Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life	Paul Ekman	4.524	8.7
805080236	Derby Girl	Shauna Cross	4.087	6.9
806512768	The Collected Stories of Philip K. Dick, Volume 4: The Minority Report	Philip K. Dick	4.167	7.7
812968069	Snow Flower and the Secret Fan	Lisa See	4.500	6.1
812982428	The Best Exotic Marigold Hotel	Deborah Moggach	3.961	7.3
816638624	A Single Man	Christopher Isherwood	4.439	7.6
820323381	The Dangerous Lives of Altar Boys	Chris Fuhrman	4.520	7.1
1400075637	I'm Not Scared	Niccolo Ammanit	4.060	7.5
1405206233	Red Rackham's Treasure	Herge	4.636	7.4
1416524304	Carrie	Stephen King	4.329	6.0
141694740X	He's Just Not That Into You: The No-Excuses Truth to Understanding Guys	Greg Behrendt, Liz Tuccillo	3.940	6.4
1419701762	Me and Earl and the Dying Girl	Jesse Andrews	3.676	8.3
1426309376	Stolen into Slavery: The True Story of Solomon Northup, Free Black Man	Judith B. Fradin, Dennis Brindell Fradin	4.421	8.1
157008789X	The Other Side of Heaven	John H. Groberg	4.778	6.5
1577656938	Little Women (Great Illustrated Classics)	Louisa May Alcott	4.000	7.3
1582404976	Wanted	Mark Millar	3.410	6.7
1590514505	The Secret in Their Eyes	Eduardo Sacheri	4.222	8.3
1593070942	Hellboy, Vol. 1: Seed of Destruction	Mike Mignola, John Byrne	4.239	6.8
1593070950	Hellboy, Vol. 2: Wake the Devil	Mike Mignola	4.594	7.0
1594481938	A Long Way Down	Nick Hornby	3.507	6.4
1595543414	Thr3e	Ted Dekker	4.226	5.0
1619698366	Struck By Lightning: The Carson Phillips Journal	Chris Colfer	4.586	6.4
1780810296	Diary of a Wimpy Kid	Jeff Kinney	5.000	6.2
1840187166	Catch Me If You Can: The True Story Of A Real Fake	Frank W. Abagnale, Stan Redding	4.340	8.0
1841954802	Under the Skin	Michel Faber	3.752	6.3
1844080439	A Mighty Heart: The Brave Life and Death of My Husband, Danny Pearl	Mariane Pearl	3.500	6.7

Table 17, continued

1854594672	The Winslow Boy	Terence Rattigan	4.714	7.3
1858820618	Executioner: Pierrepoint. Albert Pierrepoint	Albert Pierrepoint	4.167	7.5
1860469620	Crimson Rivers	Jean-Christophe Grange	4.000	6.9
1860499473	Riding in Cars with Boys: Confessions of a Bad Girl Who Makes Good	Beverly Donofrio	4.284	6.5
2226131906	Dreamcatcher	Stephen King, William Olivier Desmond	2.750	5.5
2253140872	L'cume des jours	Boris Vian	5.000	6.5
8129104598	Five Point Someone: What Not to Do at IIT	Chetan Bhagat	3.418	8.5
B0074670VA	The Grey	Ian Mackenzie Jeffers	3.896	6.8
B007TXT9XA	Touchback	Don Handfield	4.656	6.5

Notes: The book ratings measures the average star ratings of each book across all reviews, and the movie ratings are the observed *IMDB* ratings at the time of data collection in December 2016.

For each review, we collected data on available review characteristics on *Amazon*, such as star rating, review text, review time stamps, number of images in a review, whether a review was based on an *Amazon* verified purchase and whether a review containing videos. Our panel was constructed by merging the reviews of each book with its book-to-film adaptation information and then ordering the observations by book and by review time stamps. We then created the key independent variable (i.e. review_after_movie) to indicate whether a book review was authored before or after its book-to-film adaptation's release. Due to the skewness of the variable distributions, we log-transformed the variables on the number of words and the number of images in reviews.

We leveraged the Linguistic Inquiry and Word Count (LIWC) dictionary to construct the linguistic measures of book reviews. For each linguistic measure, the LIWC dictionary identifies a set of key-words that reflects certain characteristics in a textual content. Text analytics using the LIWC dictionary is a lexicon-based method that calculates the percentage of words being matched to predefined key-word categories (Pennebaker et al. 2007). This method has been accepted and used in the fields of Psychology, Marketing, and Information Systems (Sridhar and Srinivasan 2012; Yin et al. 2014; Mudambi et al. 2016; Huang et al. 2016). In this study, we focus on the linguistic measures of viewing, comparative and affective processes in review text. Table 18 reports examples of words that are categorized as reflecting each linguistic characteristic, as well as examples of book reviews that have high values in certain linguistic measures.

Table 18. Linguistic Characteristics and Examples

Linguistic Characteristics	*Word Examples*	*Book Review Examples*
Viewing	"view, saw, seen"	"First time we saw the movie, it was suspense. After we saw the movies more than 2 times, we are still looking forward to read the book." "So glad I read this before watching the movie. It was incredibly more deep, beautiful, and tragic. Thanks John Green, needed this read."
Comparative	"greater, best, after"	"Better than I expected. Much better than the film adaptation. If you've only seen the film, it is more than worth reading this tiny, yet huge, book."
Affective Processes	"happy, cried"	"I loved it because it was sad happy romantic heartwarming interesting book but I loved it so much it was a really good book." "I cried my eyes out!! such a beautiful book, best book. great story, great writing, great characters, great passion. perfect, with flaws and all"

Notes: Example words of the linguistic characteristics are from Pennebaker et al. (2007).

To summarize, the data of this study contain book review information from *Amazon*, book classification by manual coding, linguistic characteristics of book reviews using the LIWC dictionary, and movie information from *IMDB.com*. When generating the variable that measures book-movie rating difference, we standardized the two variables on average book rating and movie rating to deal with the issue of different rating scales. We then measured the absolute differences of the standardized book vs. movie ratings. Lastly, we created a variable measuring book classifications based on books' target audience. Two researchers independently coded the books and then reached consensus on book classifications. Table 19 lists the key variables and their definitions in this study. In addition, Table 20 reports the descriptive statistics of all variables, and Table 21 presents the correlation matrix of the variables.

Table 19. Variables and Definitions

Variables	*Definitions*
Book_rating	Ratings of book reviews on Amazon, on a 5 point scale.
Movie_rating	Ratings of movies on IMDB, on a 10 point scale.
Viewing	A variable measuring the percentage of words suggesting viewing of information by the reviewer in a book review.
Comparative	A variable measuring the percentage of words reflecting comparatives in a book review.
Affective_process	A variable measuring the percentage of words expressing the affective process of the reviewer in a book review.
Cognitive_process	A variable measuring the percentage of words reflecting the cognitive process of the reviewer in a book review.
Review_after_movie	A dummy variable measuring whether a book review was written before or after its book-to-film adaptation's release, with 0 marking before and 1 indicating after.
Movie_age_rating	A categorical variable measuring the age ratings of the films adapted from books, with 1 standing for G, 2 denoting PG, 3 marking PG-13, and 4 representing R.
Book_classification	A dummy variable measuring the classification of books, with 0 marking books for children and young adult, and 1 indicating books for adults.
lnwords	A log transformed variable, measuring how many words a review contains.
Amz_verified	A dummy variable indicating whether or not a review was written by a reviewer who has purchased the book, with 0 indicating without verified purchase and 1 meaning with verified purchase.
lnimages	A log transformed variable, measuring how many images a review incorporates.
Has_video	A dummy variable indicating whether or not a review includes a video, with 0 meaning without video and 1 indicating with video.
Review_helpfulness	A variable measuring the number of helpful votes accumulated on a review by users on Amazon

Table 20. Descriptive Statistics

Variables	*Obs.*	*Mean*	*S.D.*	*Min*	*Max*
Book_rating	142,601	4.389	1.054	1.000	5.000
Movie_rating	142,601	7.187	0.845	4.600	8.700
Viewing	142,601	0.899	1.452	0.000	20.830
Comparative	142,601	2.536	2.481	0.000	29.170
Affective_process	142,601	7.879	4.354	0.000	88.660
Cognitive_process	142,601	11.687	5.153	0.000	56.520
Review_after_movie	142,601	0.437	0.496	0.000	1.000
Book_movie_diff	142,601	0.748	0.714	0.004	5.187
Movie_age_rating	142,601	3.178	0.551	1.000	4.000
Book_classification	142,601	0.525	0.499	0.000	1.000
lnwords	142,601	4.222	0.857	3.178	8.050
Amz_verified	142,601	0.568	0.495	0.000	1.000
lnimages	142,601	0.000	0.026	0.000	4.905
Has_video	142,601	0.000	0.006	0.000	1.000
Review_helpfulness	142,587	1.774	12.878	0.000	904.000

Table 21. Correlation Matrix

	1	2	3	4	5	6	7	8	9	10	11	12
1. Book_rating	1.000											
2. Affective_process	0.110	1.000										
3. Cognitive_process	-0.161	-0.050	1.000									
4. Visualization	0.057	0.013	-0.066	1.000								
5. Review_after_movie	-0.016	0.028	-0.007	0.075	1.000							
6. Book_movie_diff	-0.428	-0.088	0.029	-0.067	-0.030	1.000						
7. Movie_classification	-0.060	-0.006	0.032	0.026	0.026	-0.059	1.000					
8. Book_classification	-0.136	-0.106	0.002	-0.013	0.167	0.067	0.526	1.000				
9. lnwords	-0.164	-0.297	0.041	-0.031	-0.100	0.049	-0.018	0.152	1.000			
10. Amz_verified	0.120	0.169	0.013	0.054	0.143	-0.090	0.071	-0.172	-0.463	1.000		
11. lnimages	-0.002	-0.005	-0.002	0.016	0.013	-0.003	-0.004	0.007	0.010	0.002	1.000	
12. Has_video	-0.001	-0.006	-0.005	0.009	-0.003	-0.004	-0.002	-0.004	0.003	-0.002	0.000	1.000

In the empirical analyses, we employed panel-data regressions with two-way fixed effects at the review level. Specifically, we identified the effects of book-to-film adaptations on outcome variables (i.e. book rating, affective process, viewing and comparative language) by examining intra-book variance in book reviews via within transformation, while controlling for time effects by using monthly dummy variables. Meanwhile, we also control for review-level heterogeneity, using observable review information (i.e. number of words, number of images, *Amazon* verified purchase, and whether a review has video). Equations 1 and 2 delineate the estimation specifications for hypotheses testing.

(1) $\text{Book\_rating}_{ijt} = \boldsymbol{\alpha_1} \times \text{Review\_after\_movie}_{ij} + \gamma_1 \times \text{controls}_{ijt} + \delta_j + \sum_T \tau_t \times M_t + \varepsilon_{ijt}$

(2) $\text{Linguistic\_characteristics}_{ijt} = \boldsymbol{\alpha_2} \times \text{Review\_after\_movie}_{ij} + \gamma_2 \times \text{controls}_{ijt} + \delta_j + \sum_T \tau_t \times M_t + \varepsilon_{ijt}$

In the above equations, i indexes reviews, j indicates books or movies, and t denotes time. The notation δ_j stands for the book-fixed effect that controls for time-invariant book-level heterogeneity, and $\sum_T \tau_t \times M_t$ represents a vector of time dummies. In equation 2, the outcome variable, $\text{Linguistic\_characteristics}_{ijt}$, encompasses the four linguistic measures: viewing, comparative, affective and cognitive processes. In Equations 1 and 2, the key parameters of interests are $\boldsymbol{\alpha_1}$ and $\boldsymbol{\alpha_2}$. A significant and positive (or negative) coefficient of $\boldsymbol{\alpha_1}$ or $\boldsymbol{\alpha_2}$ would suggest that on average there are significant increases (or decreases) in book review characteristics (i.e. book rating, viewing, comparative, affective and cognitive processes) before vs. after book-to-film adaptation.

Results

Publicity Effect

We first tested the hypothesis on the publicity effect, which expects book-to-film adaptations to lead to a decrease in book star rating. In particular, we estimated three alternative specifications (without fixed effects, with a one-way fixed effect and with two-way fixed effects) to test the effect of book-to-film adaptation on star rating. Table 22 reports the results on star rating. Our findings are consistent under all three specifications, which show that the release of book-to-film adaptation is negatively related to book ratings. We also proposed that the negative relationship of book-to-film adaptation and star rating might be driven by the publicity effect of film release. After the release of an adaptation, the adapted book's increased publicity might attract readers with different tastes, resulting in a misfit of readers and book. To shed some light on this mechanism, we compared the rating deviations of reviews (i.e. the absolute difference between a rating of the book and its prior average ratings) before vs. after the release of the adaptation. The results of a *t*-test show that the rating deviations of reviews after film release are significantly higher than those before the film release ($p < 0.05$), indicating that rating variance increased after the release of film adaptation. To conclude, the present evidence shows that star ratings of book reviews decline after book-to-film adaptation, supporting H1.

Table 22. Effect of Book-to-film Adaptation on Star Rating

Variables	(1) Book_rating	(2) Book_rating	(3) Book_rating
Review_after_movie	**-0.034*** **	**-0.109*** **	**-0.059* **
	(0.006)	**(0.024)**	**(0.024)**
lnwords		-0.162***	-0.160***
		(0.020)	(0.022)
Amz_verified		0.126***	0.144***
		(0.029)	(0.029)
lnimages		-0.111	-0.095
		(0.114)	(0.112)
Has_video		-0.156	-0.116
		(0.431)	(0.405)
Constant	4.404***	5.050***	5.239***
	(0.004)	(0.086)	(0.070)
Observations	142,601	142,601	142,601
R-squared	0.000	0.023	0.029
Number of Books	No	Yes	Yes
Book Fixed Effect	No	No	Yes
Time Fixed Effect		149	149

Notes: Robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Viewing & Comparative Effects

We then tested the hypotheses on the viewing effects of book-to-film adaptations. Specifically, we proposed that after the release of book-to-film adaptation, book reviewers are likely to watch the film adaptation and compare the book to the movie. Thus, we expect that book-to-film adaptation leads to an increase in the use of viewing and comparative language in book reviews. Table 23 and Table 24 report the effects of book-to-film adaptations on viewing and comparative review language respectively. Our results show that reviews authored after (vs. before) film adaptation release show higher percentages of viewing and comparative language, supporting H2a and H2b.

Table 23. Effects of Book-to-film Adaptation on Viewing

Variables	(1) Viewing	(2) Viewing	(3) Viewing
Review_after_movie	**0.219***** **(0.008)**	**0.161***** **(0.024)**	**0.137***** **(0.022)**
Book_rating		0.063*** (0.014)	0.064*** (0.014)
lnwords		0.003 (0.025)	-0.002 (0.025)
Amz_verified		0.032+ (0.019)	0.041+ (0.022)
lnimages		0.893*** (0.094)	0.897*** (0.095)
Has_video		2.032+ (1.221)	2.029+ (1.222)
Constant	0.803*** (0.005)	0.520*** (0.068)	-0.253*** (0.035)
Observations	142,601	142,601	142,601
R-squared	0.006	0.005	0.007
Number of Books	No	Yes	Yes
Book Fixed Effect	No	No	Yes
Time Fixed Effect		149	149

Notes: Robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 24. Effects of Book-to-film Adaptation on Comparative Language

Variables	(1) Comparative	(2) Comparative	(3) Comparative
Review_after_movie	**0.153*** **	**0.134*** **	**0.094*** **
	(0.013)	**(0.023)**	**(0.022)**
Book_rating		-0.079***	-0.079***
		(0.016)	(0.016)
lnwords		0.056**	0.060**
		(0.020)	(0.019)
Amz_verified		0.059**	0.070***
		(0.021)	(0.019)
lnimages		0.260	0.224
		(0.230)	(0.228)
Has_video		-0.487	-0.535+
		(0.336)	(0.294)
Constant	2.469***	2.551***	4.231***
	(0.008)	(0.116)	(0.100)
Observations	142,601	142,601	142,601
R-squared	0.001	0.002	0.004
Number of Books	No	Yes	Yes
Book Fixed Effect	No	No	Yes
Time Fixed Effect		149	149

Notes: Robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Emotional Spillover Effects

Lastly, we tested the effects of book-to-film adaptations on the use of language reflecting affective process. We hypothesized that book reviewers might become more emotional and thus less rational towards a book after seeing its film adaptation. Table 25 presents the hypothesis testing results on affective processes. The evidence suggests that book-to-film adaptation leads to an increase in affective language in book reviews, supporting H3.

Table 25. Effects of Book-to-film Adaptation on Affective Processes

Variables	(1) Affective_ process	(2) Affective_ process	(3) Affective_ process
Review_after_movie	**0.242***** **(0.023)**	**0.159**** **(0.061)**	**0.102*** **(0.045)**
Book_rating		0.219*** (0.044)	0.220*** (0.045)
lnwords		-1.278*** (0.131)	-1.285*** (0.139)
Amz_verified		-0.278*** (0.042)	-0.317*** (0.060)
lnimages		-0.223 (0.290)	-0.206 (0.287)
Has_video		-4.170** (1.373)	-4.336** (1.387)
Constant	7.774*** (0.015)	12.404*** (0.439)	12.608*** (0.273)
Observations	142,601	142,601	142,601
R-squared	0.001	0.063	0.065
Number of Books	No	Yes	Yes
Book Fixed Effect	No	No	Yes
Time Fixed Effect		149	149

Notes: Robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Additional Analyses

Placebo Tests

We want to rule out the alternative explanation that the outcome variables generally change over time. Thus, we validated our results on the main effects using a set of placebo tests. The purpose of the placebo tests is to rule out the possibility that there are pre-treatment time trends driving the main results. In the main analyses, we relied on book and month two-way fixed effects to identify the effects of book-to-film adaptation on outcome variables. Here, the idea behind the placebo tests is that if the main results were not driven by pre-treatment trends, we should not observe significant effects of book-to-film adaptation on outcome variables in the pre-film book reviews. Thus, we created an artificial film adaptation date (i.e. placebo treatment) that is 12 months before the actual film release date. Then we consider the observation window before the actual film release dates and re-estimated the equations used in the main analyses. As reported in Table 26, the results of the placebo tests show that the effects of book-to-film adaptation are consistently insignificant across all outcome variables, ruling out pre-treatment time trend as an alternative explanation.

Table 26. Placebo Tests

Variables	(1) Book_rating	(2) Viewing	(3) Comparative	(4) Affective_process
Review_after_movie2	-0.006	0.015	-0.052	-0.062
	(0.028)	(0.023)	(0.036)	(0.048)
Book_rating		0.051***	-0.080***	0.150***
		(0.012)	(0.022)	(0.041)
lnwords	-0.176***	0.018	0.094***	-1.225***
	(0.024)	(0.031)	(0.026)	(0.169)
Amz_verified	0.179***	-0.001	0.059*	-0.200***
	(0.037)	(0.014)	(0.023)	(0.048)
lnimages	0.335***	0.567	0.643	0.254
	(0.088)	(0.470)	(0.464)	(0.763)
Has_video	-0.420	2.376	-0.356	-5.281***
	(0.357)	(1.500)	(0.345)	(1.400)
Constant	5.361***	-0.444***	4.087***	12.490***
	(0.111)	(0.055)	(0.170)	(0.414)
Observations	80,270	80,270	80,270	80,270
R-squared	0.035	0.006	0.005	0.060
Number of Books	134	134	134	134
Book Fixed Effect	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Falsification Tests

We also conducted a series of falsification tests to help rule out spurious correlations. There is no theoretical reason that book-to-film adaptation would significantly influence outcome measures such as informal language, filler words (e.g. "I mean," "you know"), nonfluencies (e.g. er, hm, umm), and netspeak words (e.g. btw, lol, thx). Thus, we would expect that book-to-film adaptation has no significant effect on the appearance of those words in book reviews. As shown in Table 27, the falsification tests did not find significant relationships between book-to-film adaptation and the outcome variables, including informal language, filler words, nonfluencies or netspeak words, lending further validity to the main results.

Table 27. Falsification Tests

Variables	(1) Informal	(2) Filler	(3) Nonfluencies	(4) Netspeak
Review_after_movie	-0.009	0.000	0.010	-0.006
	(0.013)	(0.001)	(0.009)	(0.010)
Book_rating	0.007	-0.004***	0.013***	0.010*
	(0.005)	(0.001)	(0.004)	(0.004)
lnwords	-0.149***	0.002*	-0.056***	-0.047***
	(0.012)	(0.001)	(0.008)	(0.008)
Amz_verified	-0.082***	0.001	0.009+	-0.049**
	(0.022)	(0.001)	(0.005)	(0.016)
lnimages	-0.066	-0.007*	-0.012	-0.018
	(0.081)	(0.003)	(0.043)	(0.048)
Has_video	-0.174	-0.012**	-0.253***	0.288
	(0.198)	(0.005)	(0.037)	(0.306)
Constant	0.362***	0.009	0.163***	-0.038+
	(0.041)	(0.007)	(0.038)	(0.021)
Observations	142,601	142,601	142,601	142,601
R-squared	0.008	0.003	0.005	0.004
Number of Books	149	149	149	149
Book Fixed Effect	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Review Helpfulness

To further explore the implications of our results for sites hosting reviews of books and other products that can be adapted into other media forms, we therefore conducted additional analyses, exploring the effects of star rating and linguistic characteristics on review helpfulness. Following Mudambi and Schuff (2010), review helpfulness in this study is measured as the number of votes a review has received from other users. Specifically, we estimated the following equation:

(3) $\text{Review\_helpfulness}_{ijt} = \alpha_3 \times \text{Book\_rating}_{ij} + \beta_3 \times \text{Viewing}_{ijt} + \gamma_3 \times \text{Comparative}_{ijt} + \eta_3 \times \text{Affective\_process}_{ijt} + \theta_3 \times \text{Cognitive\_process}_{ijt} + \zeta_3 \times \text{controls}_{ijt} + \delta_j + \sum_T \tau_t \times M_t + \varepsilon_{ijt}$

Table 28 reports the correlational evidence on the relationships between rating and linguistic characteristics and review helpfulness. The results show that book rating is negatively associated with review helpfulness, consistent with Mudambi and Schuff (2010). Meanwhile, the increase of viewing language in book reviews is positively related to review helpfulness. Moreover, we find that affective language is positively associated with review helpfulness, while language reflecting a cognitive process is negatively associated with review helpfulness. It is notable that the positive relationship between the affective process and review helpfulness contradicts prior findings on this relationship by Hong et al. (2016). We suspect that the differences in findings might be attributed to the variations in study contexts. Hong et al. (2016) examined the relationship between affective language and review helpfulness in the context of restaurant reviews, a utilitarian product/service. In contrast, this study is conducted in the context of book reviews, a hedonic product. It is possible that users find emotional reviews are more helpful if the product is for hedonic rather than utilitarian purpose. Future research could definitely explore product heterogeneity in the relationship between emotional language and review helpfulness.

To summarize, the main findings have shown that book-to-film adaptation decreases star rating and increases viewing and affective language. Here, additional analyses on review helpfulness suggest that decreased star rating, as well as increased viewing and affective language are positively associated with review helpfulness. Thus, it is possible that reviews become more helpful in accordance with the changes in star rating and linguistic characteristics of book reviews after book-to-film adaptation.

Table 28. Effects of Rating & Linguistic Characteristics on Review Helpfulness

Variables	(1) Review_helpfuness	(2) Review_helpfuness	(3) Review_helpfuness
Book_rating	**-1.919***** **(0.297)**	**-1.927***** **(0.297)**	**-1.974***** **(0.299)**
Viewing		**0.109*** **(0.048)**	**0.093*** **(0.047)**
Comparative		-0.030+ (0.017)	-0.019 (0.018)
Affective_process			**0.058***** **(0.017)**
Cognitive_process			**-0.057***** **(0.011)**
lnwords	2.638*** (0.414)	2.639*** (0.413)	2.731*** (0.419)
Amz_verified	0.853* (0.348)	0.855* (0.347)	0.867* (0.344)
lnimages	4.195 (3.310)	4.153 (3.329)	4.155 (3.331)
Has_video	-3.223* (1.312)	-3.492** (1.255)	-3.430* (1.353)
Constant	5.193** (1.957)	5.364** (1.964)	5.749** (1.963)
Observations	80,275	80,275	80,275
R-squared	0.049	0.049	0.049
Number of Books	136	136	136
Book Fixed Effect	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Interaction Effects

Lastly, to better understand the interactions of types of movies and types of books, we explored possible interaction effects on the linguistic characteristics of review text. In particular, we examined the interaction effects of book-to-film adaptation with book and movie rating difference, movie age-rating or book classification. In doing so, we estimated the following equations:

(4) $\text{Linguistic\_characteristics}_{ijt} = \alpha_3 \times \text{Review\_after\_movie}_{ij} + \eta_3 \times \text{Book\_movie\_diff}_j + \boldsymbol{\theta_3} \times \text{Review\_after\_movie}_{ij} \times \text{Book\_movie\_diff}_j + \beta_3 \times \text{Book\_rating}_{ijt} + \gamma_3 \times \text{controls}_{ijt} + \delta_j + \sum_T \tau_t \times M_t + \varepsilon_{ijt}$

(5) $\text{Linguistic\_characteristics}_{ijt} = \alpha_4 \times \text{Review\_after\_movie}_{ij} + \eta_4 \times \text{Movie\_age\_rating}_j + \boldsymbol{\theta_4} \times \text{Review\_after\_movie}_{ij} \times \text{Movie\_age\_rating}_j + \beta_4 \times \text{Book\_rating}_{ijt} + \gamma_4 \times \text{controls}_{ijt} + \delta_j + \sum_T \tau_t \times M_t + \varepsilon_{ijt}$

(6) $\text{Linguistic\_characteristics}_{ijt} = \alpha_5 \times \text{Review\_after\_movie}_{ij} + \eta_5 \times \text{Book\_classification}_j + \boldsymbol{\theta_5} \times \text{Review\_after\_movie}_{ij} \times \text{Book\_classification}_j + \beta_5 \times \text{Book\_rating}_{ijt} + \gamma_5 \times \text{controls}_{ijt} + \delta_j + \sum_T \tau_t \times M_t + \varepsilon_{ijt}$

In Equations 4, 5 and 6, the key parameters of interest are $\boldsymbol{\theta_3}$, $\boldsymbol{\theta_4}$ and $\boldsymbol{\theta_5}$, the coefficient of which estimates the significance and direction of the interaction effects. A significant and positive (or negative) coefficient of $\boldsymbol{\theta_3}$ would suggest that book and movie rating difference reinforce (or attenuate) the effects of book-to-film adaptations on book reviews, such that the higher the book and movie rating difference, the stronger (or weaker) the impact of book-to-film adaptations on book review characteristics. The interpretation of $\boldsymbol{\theta_4}$ and $\boldsymbol{\theta_5}$ is similar to that of $\boldsymbol{\theta_3}$.

When estimating the interaction effects, we followed the procedure suggested by Cohen et al. (2013) and mean-centered the variable on book-movie rating difference, movie age-rating and book classification before generating the product terms in Equation 4, 5 and 6. In addition, we measure the variables book-movie rating difference, movie age-rating and book classification at the book level. Thus, the main effects of such variables are omitted during the estimations with book-fixed effects. Our findings on the interaction effects are reported in the following tables.

Table 29 presents the results on the interaction effects of book-to-film adaptation with book and movie rating difference. The books are rated on *Amazon.com*, and the movies are rated on *IMDB.com*. People often argue about which was better the book or the movie, and sometimes there can be a big gap in the star rating of the book and the star rating of the movie. Sometimes a highly rated book is not adapted into a highly rated movie, while some very highly rated movies were adapted from books that were not that well rated. Our results show that book and movie rating difference reinforces the positive relationships between book-to-film adaptations and viewing and comparative processes, such that these relationship are stronger when there is a high rating difference between book and movie. These results suggest that book reviewers are more likely go see the movie and compare the book with its movie when there are large discrepancies between the book and the film adaptation.

Table 29. Interaction Effects: Book and Movie Rating Difference

Variables	(1) Viewing	(2) Comparative	(3) Affective_ process
Review_after_movie	0.137***	0.094***	0.102**
	(0.012)	(0.021)	(0.035)
Book_movie_diff	-	-	-
Review_after_movie * Book_movie_diff	**0.031***	**0.037+**	-0.005
	(0.012)	**(0.021)**	(0.035)
Book_rating	0.064***	-0.079***	0.220***
	(0.004)	(0.007)	(0.011)
lnwords	-0.002	0.060***	-1.285***
	(0.005)	(0.009)	(0.015)
Amz_verified	0.040***	0.069***	-0.317***
	(0.011)	(0.019)	(0.031)
lnimages	0.897***	0.223	-0.206
	(0.147)	(0.255)	(0.419)
Has_video	2.027**	-0.538	-4.336*
	(0.637)	(1.103)	(1.808)
Constant	-0.243	4.237+	12.603**
	(1.423)	(2.464)	(4.039)
Observations	142,601	142,601	142,601
R-squared	0.007	0.004	0.065
Number of Books	149	149	149
Book Fixed Effect	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes

Notes: Standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Table 30 reports the interaction effects of book-to-film adaptation with movie age-rating. Interestingly, when a movie is rated G ($\text{Movie\_age\_rating}_{G} = 1$), the effect of book-to-film adaptation on viewing is negative. When movie age-rating becomes PG ($\text{Movie\_age\_rating}_{PG} = 2$), PG-13 ($\text{Movie\_age\_rating}_{PG\text{-}13} = 3$) or R ($\text{Movie\_age\_rating}_{R} = 4$), the impact of book-to-film adaptation on viewing turns positive. In addition, we found that movie age-rating negatively interacts with book-to-film adaptations in influencing affective language in book reviews, such that the positive relationship between book-to-film adaptation and affective language is attenuated as movies are rated for more mature audiences. The present results suggest that more mature audiences are more responsive to viewing effects and less sensitive to emotional spillover effects.

Table 30. Interaction Effects: Movie Age-rating

Variables	(1) Viewing	(2) Comparative	(3) Affective_ process
Review_after_movie	0.139***	0.094***	0.099**
	(0.012)	(0.021)	(0.035)
Movie_age_rating	-	-	-
Review_after_movie * Movie_age_rating	**0.086*** (0.016)**	-0.008 (0.028)	**-0.080+ (0.047)**
Book_rating	0.064***	-0.079***	0.220***
	(0.004)	(0.007)	(0.011)
lnwords	-0.002	0.060***	-1.284***
	(0.005)	(0.009)	(0.015)
Amz_verified	0.040***	0.070***	-0.316***
	(0.011)	(0.019)	(0.031)
lnimages	0.901***	0.223	-0.209
	(0.147)	(0.255)	(0.419)
Has_video	2.024**	-0.534	-4.331*
	(0.637)	(1.103)	(1.808)
Constant	-0.262	4.232+	12.617**
	(1.422)	(2.463)	(4.039)
Observations	142,601	142,601	142,601
R-squared	0.007	0.004	0.065
Number of Books	149	149	149
Book Fixed Effect	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes

Notes: Standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Table 31 shows the results of testing the interaction effect with regard to book classification. About 78.94% of the books are classified as adult books, and 83.56% of the film adaptations are rated PG-13 or R. Thus, the interaction effects of book-to-film adaptation with book classification are largely consistent with those of book-to-film adaptation and movie age-rating.

Table 31. Interaction Effects: Book Classification

Variables	(1) Viewing	(2) Comparative	(3) Affective_ process
Review_after_movie	0.139***	0.093***	0.098**
	(0.012)	(0.021)	(0.035)
Book_classification	-	-	-
Review_after_movie * Book_classification	**0.086*** **	**-0.054+**	**-0.188*** **
	(0.019)	**(0.032)**	**(0.053)**
Book_rating	0.064***	-0.079***	0.220***
	(0.004)	(0.007)	(0.011)
lnwords	-0.003	0.061***	-1.283***
	(0.005)	(0.009)	(0.015)
Amz_verified	0.038***	0.072***	-0.310***
	(0.011)	(0.019)	(0.031)
lnimages	0.896***	0.224	-0.204
	(0.147)	(0.255)	(0.419)
Has_video	2.018**	-0.528	-4.312*
	(0.637)	(1.103)	(1.808)
Constant	-0.220	4.208+	12.536**
	(1.423)	(2.464)	(4.039)
Observations	142,601	142,601	142,601
R-squared	0.007	0.004	0.065
Number of Books	149	149	149
Book Fixed Effect	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes

Notes: Standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

To conclude, the results of additional analyses have shown several significant interaction effects of book-to-film adaptation with regard to book and movie rating difference, movie age-rating and book classification on certain linguistic characteristics. One caveat of interpreting and generalizing the interaction effects is that these are mere correlational evidence without adequate information in order to establish causality. Furthermore, we did not rely on theoretical foundations for proposing the interaction effects. Although we have provided plausible explanations for the interaction effects, there could still be alternative reasons in play.

Discussions

Theoretical Implications

By empirically examining the effect of book-to-film adaptation on users' content generation behaviors, this research provides a deeper understanding of this interesting phenomenon. We leverage two main mechanisms to explain the findings: the publicity effect and the emotional spillover effect. The publicity effect suggests that book-to-film adaptation may lead to a wider audience of readers for the book, even though those readers may not like the book. In this respect, our research extends prior research on the publicity effect on online ratings (Kovács and Sharkey 2014). The emotional spillover effect suggests that readers who have experienced both mediums of the content (text and visual) are likely to express more emotions when providing an evaluation. In this respect, our study provides empirical evidence for the media richness theory (Daft and Lengel 1986) in the unique context of book-to-movie adaptations, which predicts that rich media is more likely to stimulate users' emotions.

Second, this research also provides implications to the review helpfulness literature (Mudambi and Schuff 2010). While prior research has found mixed findings with regard to the effect of emotion on review helpfulness in different contexts (Yin et al. 2014; Hong et al. 2016; Jensen et al. 2013; Li and Zhan 2011), we find that emotional reviews are perceived to be more helpful, whereas cognitive reviews are perceived to be less helpful. We surmise that the effect of review emotion on review helpfulness could well be context dependent. For products that have many objective criteria for evaluation (utilitarian goods), review emotion may be seen as detrimental, because the authors of the reviews may be considered biased or not objective in evaluation. However, for hedonic goods for which different consumers have idiosyncratic preferences, emotions may be seen as a positive signal because they provide an indication of the reviewers' strong attitude and active involvement.

Managerial Implications

This study provides several implications for online review platform managers, as well as book authors and publishers. First, given that book-to-movie adaptations negatively affect readers' evaluation of a movie, which will likely affect the book's future demand, it is important for the review platform to provide additional metadata on the book reader's source of referral. For example, a review platform could ask a reader one additional question in the review interface: where did you hear about this book? Alternatively, the review platform may be able to get around this issue by providing metadata on the focal reader's (who provided the book review) book preference.

Secondly, because we found that book-to-movie adaptations increase book reviews' emotional qualities (likely a spillover from the movie), the authors may be able to identify what plots or scenes are most likely to engage readers, which could further inspire the authors when they write sequels or related books.

Limitations and Future Research

This study has several limitations, while offering ample opportunities for future research. First, due to the observational nature of this study, we could only use econometric methods to show correlational evidence. Future studies might consider employing experimental methods to further infer causality. For example, there can be a between-subjects experiment, comparing user evaluations on textual information alone versus textual information accompanied by a video. Second, although we have supported the possible mechanisms with theoretical arguments, we did not directly measure or test them. There could therefore be alternative explanations for the observed relationships in our study. We encourage future studies to examine possible alternative mechanisms of our findings. For instance, the decrease in book rating after the release of book-to-film adaptation might also be explained by expectation-disconfirmation (Oliver 1980; Hossain and Quaddus 2012). Specifically, consumers might form certain expectations about a book after seeing its film adaptation without reading the book. If the consumers ever find that the book doesn't meet the expectations they derived from the movie, expectation-disconfirmation occurs. As a result, consumers are likely to feel dissatisfied, resulting in lower book ratings. Moreover, even though we were able to include 149 books that represented a range of types, subject matter and authors in our data, future work could expand our sample by including other books that have been adapted to enhance the external validity of our findings. Third, due to the limitation of our data, we did not investigate the role of book/movie genre in the effect of book-to-film adaptations on book evaluations. Future research can explore the role of other between-book variations beyond the book and movie rating difference and book classification in this study. For example, a romantic comedy film may have a different effect on the book reviews than a horror film. Lastly, due to the nature of our data and observational research design, although we have used multiple econometric identification strategies, readers should be cautious in any causal interpretations, because the results may be correlational. It will be interesting for future research to examine some of the findings in a research lab environment. To conclude, this study provides a good starting point for better understanding how the adaptation of a book into a film affects consumer evaluations of the book. The findings of our study can lay a foundation for future investigations into consumer perceptions of cross-media adaptations.

CHAPTER 5.

CONCLUSIONS

As an important research stream across multiple disciplines, UGC is a ubiquitous phenomenon that has generated tremendous value for businesses and consumers. In three related essays, I employ different methodologies to deepen the hitherto limited understanding of various textual aspects of UGC. The three essays are largely empirical in nature, informed by a number of important theories, such as construal level theory and social presence theory, and contributes back to these theories.

My first essay focuses on two psychological distances that naturally occurs in the context of writing online reviews – spatial and temporal distance, and uses econometric analysis and content analysis to find a reinforcement effect between spatial and temporal distances on review positivity. When it comes to recounting past experiences – like writing online reviews about restaurants once-visited – how do consumers construe and subsequently evaluate those past events? As time and space increase (from here and now to then and there), are they more or less likely to construe events at a high-level and evaluate them more favorably? Further, does the commingling of distance in both time and space have its own effect on judgment, separate from the individual effects of time and space? As revealed in the current field research, the answers to these questions lie at the intersection of theories of psychological distance and construal level. By adopting a simultaneous, cross-dimensional approach to psychological distance, we found that an increase in one distance increased the effect (on construal and review favorability) of the other. A famous anonymous poet once said, “distance makes the heart grow fonder” – we might say, based on our research, that "distance*s* make the heart grow fonder*er*" (albeit less poetically).

The second essay focuses on how a significant system change on online review platforms – social network integration – would affect how users write online reviews. First, we find that social network integration with *Facebook* increased the volume of reviews on both *Yelp.com* and *TripAdvisor.com*. Second, we find that social network integration generally increased the prevalence of emotional language in review text, but a parallel decline in cognitive language. Breaking down emotional language into positive and negative, we further observe that the increase is largely attributable to an increase in positive emotions. Third, and last, we observed a significant decline in the use of negations, indicating a decline in disagreement or negativity. Subsequent analyses at the user-level demonstrate that the observed effects of social network integration on both volume and language use are driven at least in part by changes in user-behavior, rather than mere self-selection, as we observe significant changes in average reviewing activity and language characteristics within users who had registered prior to the social integration event on *TripAdvisor.com*. Taken together, although social network integration delivers apparent benefits, in terms of increases in review volumes and a decline in disagreement, when we consider that past work has found that emotional, positive, and conforming reviews tend to be perceived as less helpful (Chen and Lurie 2013; Yin et al. 2014; Hong et al. 2016), it seems that social network integration constitutes a double-edged sword, providing some benefits in terms of review quantity, possibly at the cost of perceived quality.

In the third essay, we empirically examined the impact of book-to-film adaptation on characteristics of book reviews. This essay found that book-to-film shows four types of effects on book reviews: publicity effect, viewing and comparative effects, as well as emotional spillover effect. Specifically, publicity effect suggests that book-to-film adaptation leads to decrease in book star ratings. We attribute the negative publicity effect of film adaptation on star ratings to the increased likelihood of product misfit after the release of the film. In addition, consistent with the predictions based on information integration theory and social comparison theory, we found evidence of viewing and comparative effects revealing that book-to-film adaptation increases the use of language reflecting viewing and comparison in book review text. Meanwhile, emotional spillover effect based on media richness theory indicates that book-to-film adaptation leads to increases in the use of affective language in book reviews. Further, we conducted a series of additional analyses to understand the implications of the main findings. In particular, we found that the decrease in star rating, as well as increases in viewing and affective language are positively associated with review helpfulness. Thus, our results suggest that book-to-film adaptation might be beneficial to book review helpfulness by shifting the star rating and certain linguistic characteristic of review text. We lastly explored the interaction effects of book-to-film adaptation with regard to book and movie rating difference, movie age-rating and book classification. The results of the interaction effects show that the effects of book-to-film adaptation on viewing and comparative language are stronger when there are large book and movie rating difference. Lastly, our results suggest that the viewing effect of book-to-film adaptation are stronger, while the emotional spillover effects of film adaptation are weaker for more mature audiences.

In conclusion, user generated content has increasingly become an important element of firm strategy and consumer decision making. The three essays that comprise my dissertation are connected through a common theme of investigating different antecedents of UGC characteristics using various methods of text analytics. In particular, the first essay examines psychological distances as antecedents of rating positivity using content analysis; the second essay investigates how social media integration affects the volume and linguistic characteristics of UGC, and the third essay explores the impact of book-to-film adaptation on rating and linguistic features of book reviews using semantic text analytics approach. This dissertation is intended to contribute to research related to UGC in the fields of Marketing and Information Systems. The three essays of my dissertation shed light on the possible antecedents of UGC characteristics, which not only improves our theoretical understanding on UGC as an important phenomenon, but also provides practical implications for business managers, consumers and policy makers.

BIBLIOGRAPHY

Abadie, A., & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica, 74*(1), 235-267.

Acquisti, A., & Gross, R. (2006). Imagined communities: Awareness, information sharing, and privacy on the Facebook. *Lecture Notes in Computer Science*, 36-58.

Acquisti, A., John, L., & Loewenstein, G. (2013). What is privacy worth? *The Journal of Legal Studies, 42*(2), 249-274.

Albuquerque, P., Pavlidis, P., Chatow, U., Chen, K. Y., & Jamal, Z. (2012). Evaluating promotional activities in an online two-sided market of user-generated content. *Marketing Science, 31*(3), 406-432.

Anderson, N. H. (1981). *Information integration theory*. New York.

Anderson, N. H. (2014). *Contributions to information integration theory: Volume 1: Cognition.* Psychology Press.

Andreoni, J., & Bernheim, B. D. (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica, 77*(5), 1607-1636.

Angrist, J. D. & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion.* Princeton University Press.

Animesh, A., Pinsonneault, A., Yang, S. B., & Oh, W. (2011). An odyssey into virtual worlds: Exploring the impacts of technological and spatial environments on intention to purchase virtual products. *MIS Quarterly, 35*(3), 789-810.

Aral, S. (2013, December 19). *The problem with online ratings.* Retrieved from: http://sloanreview.mit.edu/article/the-problem-with-online-ratings-2/

Austin, P. C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics*, *10*(2), 150-161.

Avery, C., Resnick, P., & Zeckhauser, R. (1999). The market for evaluations. *The American Economic Review, 89*(3), 564-584.

Ayyagari, R., Grover, V., & Purvis, R. (2011). Technostress: Technological antecedents and implications. *MIS Quarterly, 35*(4), 831-858.

Ba, S., (2001). Establishing online trust through a community responsibility system. *Decision Support Systems, 31*(3), 323-336.

Ba, S. & Pavlou, P. (2002), Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS Quarterly, 26*(3), 243-268.

Bar-Anan, Y., Liberman, N., Trope, Y., & Algom, D. (2007). Automatic processing of psychological distance: evidence from a Stroop task. *Journal of Experimental Psychology: General*, *136*(4), 610.

Baskin, E., Wakslak, C. J., Trope, Y., & Novemsky, N. (2014). Why feasibility matters more to gift receivers than to givers: a construal-level approach to gift giving. *Journal of Consumer Research*, 41(1), 169-182.

Basuroy, S., Chatterjee, S., & Ravid, S. A. (2003), How critical are critical reviews? The box office effects of film critics, star power, and budgets, *Journal of Marketing, 67*(4), 103-117.

Baumeister, Roy F., Bratslavsky, E., Finkenauer C., & Vohs K. D. (2001), Bad is stronger than good. *Review of General Psychology, 5*(4), 323-70.

Becker, G. S. (1991). A note on restaurant pricing and other examples of social influences on price. *Journal of Political Economy, 99*(5), 1109-1116.

Begley, S. (2015, July 9). *How movies give new life to old classics.* Retrieved from: http://time.com/3837328/classics-publishers-sales-movies/

Berger, J., Sorensen, A. T., & Rasmussen, S. J. (2010). Positive effects of negative publicity: When negative reviews increase sales. *Marketing Science, 29*(5), 815-827.

Biyalogorsky, E., Gerstner, E., & Libai, B. (2001). Customer referral management: Optimal reward programs. *Marketing Science, 20*(1), 82-95.

Blanchard, O. (2011). *Social media ROI: Managing and measuring social media efforts in your organization*. Pearson Education.

Briggs, Katharine C. (1977). *Myers-Briggs type indicator*. Palo Alto, CA: Consulting Psychologist Press, Inc.

Buck, R., Losow, J. I., Murphy, M. M., & Costanzo, P. (1992). Social facilitation and inhibition of emotional expression and communication. *Journal of Personality and Social Psychology, 63*(6), 962-968.

Burtch, G., & Hong, Y. (2014). What happens when Word of Mouth goes mobile? *Proceedings of the International Conference on Information Systems (ICIS)*, Auckland, New Zealand.

Burtch, G., Hong, Y., Bapna, R. & Griskevicius, V., (2016). What are social incentives worth? A randomized field experiment in user content generation, *Management Science*, forthcoming.

Burtch, G., Carnahan, S., & Greenwood, B. N. (2016). Can you gig it? An empirical examination of the gig-economy and entrepreneurial activity. *Working Paper*.

Burtch, G., Ghose, A., & Wattal, S. (2015). The hidden cost of accommodating crowdfunder orivacy preferences: A randomized field experiment. *Management Science, 61*(5), 949-962.

Cabanac, M. (2002). What is emotion?. *Behavioural processes, 60*(2), 69-83.

Cadden, M. (2016, October 19). *Book buzz: Films continue to foster book sales.* Retrieved from: http://www.usatoday.com/story/life/books/2016/10/19/book-buzz-films-continue-foster-book-sales/92371576/

Campbell, M. A. (2005). Cyber bullying: An old problem in a new guise? *Australian Journal of Guidance and Counselling, 15*(1), 68-76.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys, 22*(1), 31-72.

Card, D. & Krueger, A. B. (1994). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. *American Economic Review, 84*(4), 772-793.

Chan, J., & Ghose, A. (2014). Internet's dirty secret: Assessing the impact of online intermediaries on HIV transmission. *MIS Quarterly, 38*(4), 955-976.

Chatman, S. (2002). Film adaptation. *Film Quarterly, 56*(1), 57-60.

Chen, Y., Harper, F. M., Konstan, J., & Li, S. X. (2010). Social comparisons and contributions to online communities: A field experiment on MovieLens. *American Economic Review*, *100*(4), 1358-1398.

Chen, Z., & Lurie, N. H. (2013). Temporal contiguity and negativity bias in the impact of online Word of Mouth. *Journal of Marketing Research, 50*(4), 463-476.

Chen, Z., & Berger, J. (2013). When, why, and how controversy causes conversation. *Journal of Consumer Research, 40*(3), 580-593.

Chen, Y. J., & Kirmani, A. (2015). Posting strategically: The consumer as an online media planner. *Journal of Consumer Psychology*, *25*(4), 609-621.

Chen, P., Hong, Y. & Liu, Y. (2017), The effect of rating system design on negativity bias, *Management Science*, forthcoming.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of Word of Mouth on sales: Online book reviews. *Journal of Marketing Research, 43*(3), 345-354.

Cho, D., Kim, S., & Acquisti, A. (2012). Empirical analysis of online anonymity and user behaviors: the impact of real name policy. *the 45th Hawaii International Conference on System Science (HICSS)*, *IEEE*.

Choe, C., Flores-Lagunes, A., & Lee, S. J. (2015). Do dropouts with longer training exposure benefit from training programs? Korean evidence employing methods for continuous treatments. *Empirical Economics, 48*(2), 849-881.

Clemons, E., Guodong G., & Hitt, L. (2006), When online reviews meet hyper differentiation: A study of the craft beer industry, *Journal of Management Information Systems, 23*(2), 149-171.

Cobb, S. C. (2009). Social presence and online learning: A current view from a research perspective. *Journal of Interactive Online Learning, 8*(3), 241-254.

Coenen, L. H., Hedebouw, L., & Semin, G. R. (2006). *The Linguistic Category Model (LCM). Manual.* Free University Amsterdam.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences.* Routledge.

Connolly, T., Jessup, L. M., & Valacich, J. S. (1990). Effects of anonymity and evaluative tone on idea generation in computer-mediated groups. *Management Science*, *36*(6), 689-703.

Cornelissen, T. (2008). The Stata command felsdvreg to fit a linear model with two high-dimensional fixed effects. *Stata Journal*, *8*(2), 170-189.

Cui, G., Lockee, B., & Meng, C. (2012). Building modern online social presence: A review of social presence theory and its instructional design implications for future trends. *Education and Information Technologies, 18*(4), 661-685.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*(5), 554-571.

Das, S., & Kramer, A. (2013). Self-censorship on Facebook, *Poceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, ICWSM, 120-127.

Daughety, A. F., & Reinganum, J. F. (2010). Public goods, social pressure, and the choice between privacy and publicity. *American Economic Journal: Microeconomics, 2*(2), 191-221.

Davis, John P., Farnham, S., & Jensen, C. (2002). Decreasing online "Bad" behavior. *CHI'02 Extended Abstracts on Human Factors in Computing Systems*, ACM, 718-719.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics, 84*(1), 151-161.

Dellarocas, C. (2003). The digitization of Word of Mouth: Promise and challenges of online feedback mechanisms. *Management Science, 49*(10), 1407-1424.

Dellarocas, C., X. M. Zhang, & Awad, N. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing, 21*(4), 23-45.

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science, 313*(5787), 684-687.

Dennis, A. R., & Kinney, S. T. (1998). Testing media richness theory in the new media: The effects of cues, feedback, and task equivocality. *Information Systems Research, 9*(3), 256-274.

Derks, D., Fischer, A. H., & Bos, A. E. (2008). The Role of emotion in computer-mediated communication: A Review. *Computers in Human Behavior, 24*(3), 766-785.

Di Tella, R., & Schargrodsky, E. (2004). Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack. *The American Economic Review, 94*(1), 115-133.

Drebes, L. (2011, October 26). *Social login offers new ROI from social media.* Retrieved from http://blogs.hbr.org/cs/2011/10/social_login_offers_new_roi_fr.html

Duan, W., Gu, B., & Whinston, A. B. (2008). Do online reviews matter? – An empirical investigation of panel data. *Decision Support Systems, 45*(4), 1007-1016.

Eichstaedt, J. C., Schwartz, H. A., Kern, M. L., Park, G., Labarthe, D. R., Merchant, R. M., & Weeg, C. (2015). Psychological language on Twitter predicts county-level heart disease mortality. *Psychological Science, 26*(2), 159-169.

Epstein, S. (1993). Emotion and Self-Theory, *Hand-book of emotions,* ed. Michael Lewis and Jeannette M. Haviland, New York: Guilford, 313-326.

Eyal, T., Liberman, N., Trope, Y., & Walther, E. (2004). The pros and cons of temporally near and distant action. *Journal of Personality and Social Psychology, 86*(6), 781-795.

Fechner, G. T. (1966). *Elements of psychophysics* (Vol. 1, H. E. Adler, Trans.). New York, NY: Holt, Rinehart & Winston. (Original work published 1860)

Finucane, M.L., Peters, E., & Slovic, P. (2003). Judgment and decision making – The dance of affect and reason. *Emerging Perspectives on Judgment and Decision Research*, Cambridge University Press, 327-357.

Fionda, A. M. & Moore, C. M. (2009). The anatomy of the luxury fashion brand. *Journal of Brand Management, 16*(5-6), 347-363.

Forbes (2015, May 28). *How books dominate the box office*. Retrieved from http://www.forbes.com/sites/natalierobehmed/2015/05/28/how-book-adaptations-dominate- thebox-office/

Forman, C., Ghose, A., & Wiesenfeld B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research, 19*(3), 291-313.

Frankel, M. (2014, June 16). *Here's why American Express can charge more than Visa or MasterCard.* Retrieved from http://www.fool.com/investing/general/2014/06/16/heres-why-american-express-can-charge-more-than-vi.aspx

Fricke, H. (2015). Identification based on difference-in-differences approaches with multiple treatments. *University of St. Gallen, School of Economics and Political Science Economics Working Paper Series No. 1510.*

Frolich, M. (2004). Programme evaluation with multiple treatments, *Journal of Economic Surveys, 18*(2), 181-224.

Froomkin, M. A. (1999). Legal issues in anonymity and pseudonymity. *Information Society, 15*(2), 113-127.

Frutiger, M., Overby, E., & Wu, D. (2014). Is social network platform integration valuable for an online service? A randomized field experiment and archival data analysis, *International Conference on Information Systems*.

Fujita, K., Henderson, M. D., Eng, J., Trope, Y., & Liberman, N. (2006). Spatial distance and mental construal of social events. *Psychological Science*, *17*(4), 278–282.

Fussell, S. R. (Ed.). (2002). *The verbal communication of emotions: Interdisciplinary perspectives.* Psychology Press.

Ghose, A. & Han, S. P. (2011). An empirical analysis of user content generation and usage behavior on the mobile Internet. *Management Science, 57*(9), 1671-1691.
Gill, A. J., French, R. M., Gergle, D., & Oberlander, J. (2008). The language of emotion in short blog texts. *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*. ACM.

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press.

Gneezy, A., Imas, A., Brown, A., Nelson, L. D., & Norton, M. I. (2012). Paying to be nice: Consistency and costly prosocial behavior. *Management Science, 58*(1), 179-187.

Godes, D. & Silva J. C. (2012). Sequential and temporal dynamics of online opinion. *Marketing Science, 31*(3), 448-473.

Goel, L., Johnson, N.A., Junglas, I., & Ives, B. (2011). From space to place: Predicting users' intentions to return to virtual worlds. *MIS Quarterly*, *35*(3), 749-771.

Goes, P. B., Lin, M., & Yeung, C. M. A. (2014). Popularity effect in user-generated content: Evidence from online reviews. *Information Systems Research, 25*(2), 222-238.

Green, J. (2007, March 15). *Is adaptation the way to make a film?* Retrieve from http://www.firstshowing.net/2007/is-adaptation-the-way-to-make-a-film/

Greenwood, B. N., & Wattal, S. (2016). Show me the way to go home: An empirical investigation of ride-sharing and alcohol related motor vehicle fatalities. *MIS Quarterly*, Forthcoming.

Griffith, J. J. (1997). *Adaptations as imitations: Films from novels*. University of Delaware Press.

Gu, B., Park J., & Prabhudev K. (2012). Research note - the impact of external Word-of-Mouth sources on retailer sales of high-involvement products. *Information Systems Research, 23*(1), 182-196.

Gunawardena, C. N. (1995). Social presence theory and implications for interaction collaborative learning in computer conferences. *International Journal of Educational Telecommunications. 1*(2/3), 147-166.

Gunawardena, C. N., & Zittle, F. J. (1997). Social presence as a predictor of satisfaction within a computer mediated conferencing environment. *American Journal of Distance Education, 11*(3), 8-26.

Guo, S., & Fraser, M. W. (2010). *Propensity score analysis: Statistical methods and applications*. Menlo Park, CA: Sage.

Harber, K. D., & Cohen, D. J. (2005). The emotional broadcaster theory of social sharing. *Journal of Language and Social Psychology, 24*(4), 382-400.

Hardaker, C. (2010). Trolling in asynchronous computer-mediated communication: from user discussions to theoretical concepts. *Journal of Politeness Research, 6*(2), 215-242.

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology*, *67*(3), 451-470.

Henderson, M. D., & Wakslak, C. J. (2010). Over the hills and far away: The link between physical distance and abstraction. *Current Directions in Psychological Science*, *19*(6), 390-394.

Herzog, S. M., Hansen, J., & Wänke, M. (2007). Temporal distance and ease of retrieval. *Journal of Experimental Social Psychology*, *43*(3), 483-488.

Hoffman, D., Novak, T., & Peralta, M. A. (1999). Information privacy in the market-space: Implications for the commercial use of anonymity on the Web. *Information Society, 15*(4), 129-139.

Holliday, M. (2009, July 3). *Yelp integrates Facebook Connect, online reviews becoming increasingly social*. Retrieved from http://www.adweek.com/socialtimes/*Yelp*-integrates-facebook-connect-online-reviews-becoming-increasingly-social/224427

Holtgraves, T. M. (2013). *Language as social action: Social psychology and language use*. Psychology Press.

Hong, Y., Huang, N., Burtch, G., & Li, C. (2016). Culture, conformity and emotional suppression in online reviews. *The Journal of the Association for Information Systems,* forthcoming.

Hong, Y., and Pavlou, P. A. (2014). Product fit uncertainty in online markets: Nature, effects, and antecedents. *Information Systems Research, 25*(2), 328-344.

Horn, L. R. (Ed.). (2010). *The expression of negation* (4). Walter de Gruyter.

Hossain, M. A., & Quaddus, M. (2012). Expectation–confirmation theory in information system research: A review and analysis. *Information Systems Theory*, 28, 441-469.

Huang, N., Burtch, G., Hong, Y. & Polman. E. (2016). Effects of multiple psychological distances on construal and consumer evaluation: A field study of online reviews. *Journal of Consumer Psychology, 26*(4), 474-482.

Huang, N., Y. Hong, & Burtch. G. (2016). Social media integration and user content generation: Evidence from Natural Experiments. *MIS Quarterly*, forthcoming.

Huang, Y., Singh, P. V., & Ghose, A. (2015). A structural model of employee behavioral dynamics in enterprise social media. *Management Science, 61*(12), 2825-2844.

Huberman, B. A., Adar, E., & Fine, L. (2005). Valuating privacy. *IEEE Security & Privacy, 3*(5), 22-25.

Huberman, B. A., Romero, D. M., & Wu, F. (2009). Crowdsourcing, attention and productivity. *Journal of Information Science, 35*(6), 758-765.

Hu, N., Liu, L., & Zhang, J. J. (2008). Do online reviews affect product sales? The role of reviewer characteristics and temporal effects. *Information Technology and Management, 9*(3), 201-214.

Jackson, L. A., Sullivan, L. A., Harnish, R., & Hodge, C. N. (1996). Achieving positive social identity: Social mobility, Social creativity, and permeability of group boundaries. *Journal of Personality and Social Psychology, 70*(2), 241-254.

Jensen, M. L., Averbeck, J. M., Zhang, Z., & Wright, K. B. (2013). Credibility of anonymous online product reviews: A language expectancy perspective. *Journal of Management Information Systems, 30*(1), 293-324.

Jessup, L. M., Connolly, T., & Galegher, J. (1990). The effects of anonymity on GDSS group process with an idea-generating task. *MIS Quarterly, 14*(3), 313-321.

Jiang, Z., Heng, C. S., & Choi, B. C. (2013). Research note – privacy concerns and privacy-protective behavior in synchronous online social interactions. *Information Systems Research, 24*(3), 579-595.

John, L. K., Acquisti, A., & Loewenstein, G. (2009). The best of strangers: Context dependent willingness to divulge personal information. Available at SSRN: http://ssrn.com/abstract=1430482.

Johnson, E. J., Bellman, S., & Lohse, G. L. (2002). Defaults, framing, and privacy: Why ppting in≠opting out. *Marketing Letters, 13*(1), 5-15.

Joinson, A. (1999). Social desirability, anonymity, and Internet-based questionnaires. *Behavior Research Methods, Instruments, & Computers, 31*(3), 433-438.

Jones, D., & Linardi, S. (2014). Wallflowers: Experimental evidence of an aversion to standing Out. *Management Science, 60*(7), 1757-1771.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kane G. (2014). Enterprise social media: Current capabilities and future possibilities. *MIS Quaterly Executive, 14*(1), 1-16.

Kane, G. C., Alavi, M., Labianca, G., & Borgatti, S. (2014). What's different about social media networks? A framework and research agenda. *MIS Quarterly, 38*(1), 275-304.

Kane G. (2015, October 6). *Digital transparency and performance.* Retrieved from http://sloanreview.mit.edu/article/digital-transparency-and-permanence/?utm_source=linkedin&utm_medium=social&utm_campaign=sm-direct

Kang, R., Brown, S., & Kiesler, S. (2013). Why do people seek anonymity on the Internet? Informing policy and design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons, 53*(1), 59-68.

Kehrwald, B. (2008). Understanding social presence in text-based online learning environments. *Distance Education*, *29*(1), 89-106.

Kim, S. H., Yi, J. S., & Ohm, J. (2016). Do sales of an extended brand affect popularity of its original? *Asia Marketing Journal, 18*(2), 45-61.

Kim, K., Zhang, M., & Li, X. (2008). Effects of temporal and social distance on consumer evaluations. *Journal of Consumer Research, 35*(4), 706-713.

Kim, B. K., Zauberman, G., & Bettman, J. R. (2012). Space, time, and intertemporal preferences. *Journal of Consumer Research*, *39*(4), 867-880.

Kincaid, J. (2010, December 21). *Facebook launches instant personalization on TripAdvisor.* Retrieved from http://techcrunch.com/2010/12/21/facebook-launches-instant-personalization-on-*TripAdvisor*/

Kollock, P. (1997). *Design principles for online communities*. Peter Kollock.

Kontaxis, G., Polychronakis, M., & Markatos, E. P. (2012). Minimizing information disclosure to third parties in social login platforms, *International Journal of Information Security, 11*(5), 321-332.

Kovács, B., & Sharkey, A. J. (2014). The paradox of publicity: How awards can negatively affect the evaluation of quality. *Administrative Science Quarterly, 59*(1), 1-33.

Kraut, R. E. (1982). Social presence, facial feedback, and emotion. *Journal of Personality and Social Psychology, 42*(5), 853-863.

Kressel, L. M., Chapman, G. B., & Leventhal, E. (2007). The influence of default options on the expression of end-of-life treatment preferences in advance directives. *Journal of General Internal Medicine, 22*(7), 1007-1010.

Kwark, Y., Chen, J., & Raghunathan, S. (2014). Online product reviews: Implications for retailers and competing manufacturers. *Information Systems Research, 25*(1), 93-110.

Lampel, J., & Bhalla, A. (2007). The role of status seeking in online communities: Giving the gift of experience. *Journal of Computer-Mediated Communication, 12*(2), 434-455.

Lasersohn, P. (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy*, *28*(6), 643-686.

Leitch, T. (2009). *Film adaptation and its discontents: from Gone with the Wind to The Passion of the Christ.* JHU Press.

Lengel, R. H., & Daft, R. L. (1989). The selection of communication media as an executive skill. *The Academy of Management Executive*, *2*(3), 225-232.

Leshed, G. (2009). Silencing the clatter: Removing anonymity from a corporate online community. *Online Deliberation: Design, Research, and Practice*, CLSI Publications, 243-251.

Leuven, E., & Sianesi, B. (2012). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. *Statistical Software Components*.

Lewin, K. (1951). *Field theory in social science: selected and theoretical papers*: Harpers, Oxford.

Lewis, A. (2012, March 28). *Hunger Games' movie fuels sharp rise in book sales.* Retrieved from http://www.hollywoodreporter.com/news/hunger-games-twlight-book-sales-versus-jennifer-lawrence-josh-hutcherson-305457

Li, Ji., Ge, Y., Hong, Y., Cheema A., & Gu, B. (2017). Review dimensionality and helpfulness: Evidence from a supervised topic modeling, an empirical analysis, and a randomized experiment. *working paper.*

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

Li, J., & Zhan, L. (2011). Online persuasion: How the written word drives WOM. *Journal of Advertising Research, 51*(1), 239-257.

Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: a test of temporal construal theory. *Journal of Personality and Social Psychology*, *75*(1), 5.

Liberman, N., Sagristano, M. D., & Trope, Y. (2002). The effect of temporal distance on level of mental construal. *Journal of Experimental Social Psychology, 38*(6), 523-534.

Liberman, N., Trope, Y., & Wakslak, C. (2007). Construal level theory and consumer behavior. *Journal of Consumer Psychology, 17*(2), 113-117.

Lowenthal, P. R. (2009). The evolution and influence of social presence theory on online learning. *Social Computing: Concepts, Methodologies, Tools, and Applications: Concepts, Methodologies, Tools, and Applications*, 124-139.

Lu, X., Ba, S., Huang, L. & Feng, Y., (2013). Promotional marketing or word-of-mouth? Evidence from online restaurant reviews. *Information Systems Research. 24*(3), 596-612.

Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and *Yelp* review fraud. *Management Science,* Forthcoming.

Luo, X., Zhang, J., & Duan, W. (2013). Social media and firm equity value. *Information Systems Research, 24*(1), 146-163.

Lurie, N., Ransbotham, S., & Liu, H. (2014). The content and impact of mobile vs. desktop reviews. *Association for Consumer Research, Chicago, IL.*

Lynch, J. G., Alba, J. W., Krishna, A., Morwitz, V., & Gurhan-Canli, Z. (2012). Knowledge creation in consumer research: Multiple routes, multiple criteria. *Journal of Consumer Psychology*, *22*, 473-485.

Maglio, S. J., Trope, Y., & Liberman, N. (2013). Distance from a distance: Psychological distance reduces sensitivity to any further psychological distance. *Journal of Experimental Psychology: General, 142*(3), 644-657.

Maglio, S. J., & Polman, E. (2014). Spatial orientation shrinks and expands psychological distance. *Psychological Science, 25*(7), 1345-1352.

Marlow, J. R. (2015). *Box office.* Retrieved from http://makeyourbookamovie.com/box-office/

Mayzlin, D., Dover, Y. & Chevalier, J. (2014). Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review, 104*(8), 2421-2455.

McCune, B., Grace, J. B., & Urban, D. L. (2002). *Analysis of ecological communities* (Vol. 28). Gleneden Beach, OR: MjM software design.

Millar, M. G. & Tesser, A. (1986). Effects of affective and cognitive focus on the attitude–behavior relation. *Journal of Personality and Social Psychology, 51*(2), 270-276.

Miranda, S. M., & Saunders, C. S. (2003). The social construction of meaning: An alternative perspective on information sharing. *Information Systems Research. 14*(1), 87-106.

Moor, P. J., Heuvelman, A., & Verleur, R. (2010). Flaming on YouTube. *Computers in Human Behavior, 26*(6), 1536-1546.

Muchnik, L., Aral, S., & Taylor, S. J. (2013). Social influence bias: A randomized experiment. *Science, 341*(6146), 647-651.

Mudambi, S. M., & Schuff, D. (2010). What makes a helpful review? A study of customer reviews on Amazon.com. *MIS Quarterly, 34*(1), 185-200.

Mudambi, S., Schuff, D., & Zifla, E. (2016), What's "funny" about technology adoption? Humorous appropriation of online review platforms. *The 2016 International Conference on Information Systems (ICIS)*, Dublin, Ireland.

Nielsen Holdings PLC (2010, November 17), *Deathly Hallows' film breathes life into Harry Potter book sales.* Retrieved from http://www.nielsen.com/us/en/insights/news/2010/deathly-hallows-film-breathes-life-into-harry-potter-book-sales.html.

Norris, P. (2004). *The bridging and bonding role of online communities*. Thousand Oaks, California: Sage Publications, 31-43.

Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, *17*(4), 460-469.

Oldmeadow, J. A., & Fiske, S. T. (2010). Social status and the pursuit of positive social identity: Systematic domains of intergroup differentiation and discrimination for high- and low-status groups. *Group Processes & Intergroup Relations, 13*(4), 425-444.

O'Neill, H. (2002). Cognitive – cognitive what?. *The British Journal of Occupational Therapy, 65*(6), 288-290.

O'Neill, N. (2009, July 2). *Yelp launches Facebook Connect.* Retrieved from http://www.adweek.com/socialtimes/*Yelp*-launches-facebook-connect/309713

Ou, C. X., Pavlou, P. A., & Davison, R. (2014). Swift Guanxi in online marketplaces: The role of computer-mediated communication technologies. *MIS Quarterly, 38*(1), 209-230.

Pavlou, P. A. (2011). State of the information privacy literature: Where are we now and where should we go? *MIS Quarterly, 35*(4), 977-988.

Pawlowski, A. (2014, November 2). *Your brain on movies: Why films make us cry, flinch and cheer*, Retrieved from http://www.today.com/health/your-brain-movies-why-films-make-us-cry-flinch-cheer-1D80258284

Peitzman, L. (2016, May 31). *39 movies that are, in fact, better than the book.* Retrieved from https://www.buzzfeed.com/louispeitzman/movies-that-are-better-than-the-book?utm_term=.gt0Qz78Qo#.ebJwQRYwD

Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition & Emotion, 10*(6), 601-626.

Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007). The development and psychometric properties of LIWC2007 [LIWC manual]. Austin, TX: LIWC.net.

Pervin, L. A., & John, O. P. (1999). *Handbook of personality: Theory and research*. Elsevier.

Pfitzmann, A., & Köhntopp, M. (2001). Anonymity, unobservability, and pseudonymity – A proposal for terminology. *Designing Privacy Enhancing Technologies,* Springer Berlin Heidelberg. 1-9.

Pham, M. T. (2013). The seven sins of consumer psychology. *Journal of Consumer Psychology*, *23*(4), 411-423.

Phillips, W. (2011). LOLing at tragedy: Facebook trolls, memorial pages and resistance to grief online. *First Monday,* 16(12).

Plantinga, C. R. & Smith, G. M. (Eds.). (1999). *Passionate views: Film, cognition, and emotion.* Johns Hopkins University Press.

Polman, E., & Emich, K. J. (2011). Decisions for others are more creative than decisions for the self. *Personality and Social Psychology Bulletin*, *37*(4), 492-501.

Polman, E. (2012). Self-other decision making and loss aversion. *Organizational Behavior and Human Decision Processes* (119), 141-150.

Public Broadcasting Service (2011). *Adaptation: From novel to film.* Retrieved from http://d2buyft38glmwk.cloudfront.net/media/cms_page_media/11/FITC_Adaptation_1.pdf

Rapp, A., Beitelspacher, L. S., Grewal, D., & Hughes, D. E. (2013). Understanding social media effects across seller, retailer, and consumer interactions. *Journal of the Academy of Marketing Science, 41*(5), 547-566.

Reader, B. (2012). Free press vs. Free speech? The rhetoric of "civility" in regard to anonymous online comments. *Journalism & Mass Communication Quarterly, 89*(3), 495-513.

Ren, Y., Harper, F. M., Drenner, S., Terveen, L., Kiesler, S., Riedl, J., & Kraut R. E. (2012). Building member attachment in online communities: Applying theories of group identity and interpersonal bonds. *MIS Quarterly, 36*(3), 841-864.

Rhue, L. & Sundararajan, A. (2013) *How digital social visibility shapes consumer choice.* Available at SSRN: http://ssrn.com/abstract=2347793

Richardson, J. C., & Swan, K. (2003). Examining social presence in online courses in relation to students' perceived learning and satisfaction. *Journal of Asynchronous Learning Networks, 7*(1), 68-88.

Rishika, R., Kumar, A., Janakiraman, R., & Bezawada, R. (2013). The effect of customers' social media participation on customer visit frequency and profitability: An empirical investigation. *Information Systems Research*, *24*(1), 108-127.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (Vol. 2): McGraw-Hill, New York.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5*(4), 296-320.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Scienc*e, *300*(5626), 1755-1758.

Santana, A. D. (2014). Virtuous or vitriolic: The effect of anonymity on civility in online newspaper reader comment boards. *Journalism Practice, 8*(1), 18-33.

Santos, R. (2013, October 2). *6 reasons the book is (almost always) netter than the movie.* Retrieved from http://www.barnesandnoble.com/blog/6-reasons-the-book-is-almost-always-better-than-the-movie/

Scott, S. V., & Orlikowski W. J. (2014). Entanglements in practice: Performing anonymity through social media. *MIS Quarterly, 38*(3), 873-893.

Seger, L. (2011). *The art of adaptation: Turning fact and fiction Into film*. Macmillan.

Semin, G. R., & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology*, *54*(4), 558.

Semin, G. R., & Fiedler, K. (1991). The linguistic category model, its bases, applications and range. *European Review of Social Psychology*, *2*(1), 1-30.

Semin, G. R., & Smith, E. R. (1999). Revisiting the past and back to the future: Memory systems and the linguistic representation of social events. *Journal of Personality and Social Psychology,* 76, 877-892.

Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research*, *26*(3), 278-292.

Short, J.A. (1974). Effects of medium of communication on experimental negotiation. *Human Relations, 27*(3), 325-334.

Short, J.A., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications.* New York: John Wiley & Sons.

Sia, C. L., Tan, B. C., & Wei, K. K. (2002). Group Polarization and Computer-mediated Communication: Effects of Communication Cues, Social Presence, and Anonymity. *Information Systems Research, 13*(1), 70-90.

Sleeper, M., Balebako, R., Das, S., McConahy, A. L., Wiese, J., & Cranor, L. F. (2013). The post that wasn't: Exploring self-censorship on Facebook. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, ACM.

Smith, G. M. (2003). *Film structure and the emotion system*. Cambridge University Press.

Snefjella, B., & Kuperman, V. (2015). Concreteness and psychological distance in natural language Use. *Psychological Science, 26*(2), 159-169.

Sridhar, S., R & Srinivasan. (2012). Social influence effects in online product ratings. *Journal of Marketing, 76*(5), 70-88.

Stephen, A. T., & Lehmann, D. R. (2009). Why do people transmit word-of-mouth? The effects of recipient and relationship characteristics on transmission behaviors. *Working Paper, Columbia University.*

Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & Behavior, 7*(3), 321-326.

Suls, J. M. (1977). *Social comparison theory and research*. Hemisphere Publishing. 1-9.

Suls, J., & Wheeler, L. (2012). Social comparison theory. *Handbook of Theories of Social Psychology* (1), 460-482.

Swanson A. (2016, January 5). *The book really is better than the movie.* Retrieved from https://www.washingtonpost.com/news/wonk/wp/2016/01/05/the-book-really-is- better-than-the-movie/

Talmi, D., & Frith, C. (2007). Neurobiology: Feeling right about doing right. *Nature*, *446*(7138), 865-866.

TripAdvisor (2010). *TripAdvisor launches Facebook® Integration, making travel planning more social for millions.* Retrieved from http://www.*TripAdvisor*.com/PressCenter-i4470-c1-Press_Releases.html

Trope, Y., & Liberman, N. (2000). Temporal construal and time-dependent changes in preference. *Journal of Personality and Social Psychology, 79*(6), 876-889.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review, 110*(3), 403-421.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*(2), 440-463.

Trope, Y., Liberman, N., & Wakslak, C. (2007). Construal levels and psychological distance: Effects on representation, prediction, evaluation, and behavior. *Journal of Consumer Psychology, 17*(2), 83-95.

Van Boven, L., Kane, J., McGraw, A. P., & Dale, J. (2010). Feeling close: Emotional intensity reduces perceived psychological distance. *Journal of Personality and Social Psychology, 98*(6), 872-885.

Van Boven, L., & Caruso, E. M. (2015). The tripartite foundations of temporal psychological distance: Metaphors, ecology, and teleology. *Social and Personality Psychology Compass*, *9*(11), 593-605.

Vincenty, T. (1975). Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. *Survey Review*, *23*(176), 88-93.

Wagner, H. L., & Smith, J. (1991). Facial expression in the presence of friends and strangers. *Journal of Nonverbal Behavior, 15*(4), 201-214.

Wakslak, C. J., Trope, Y., Liberman, N., & Alony, R. (2006). Seeing the forest when entry is unlikely: probability and the mental representation of events. *Journal of Experimental Psychology: General*, *135*(4), 641.

Wakslak, C. J. (2012). The where and when of likely and unlikely events.*Organizational Behavior and Human Decision Processes*, *117*(1), 150-157.

Wang, Z. (2010). Anonymity, Social image, and the competition for volunteers: a case study of the online market for reviews. *The BE Journal of Economic Analysis & Policy*, *10*(1), 44-44.

Wang, A., Zhang, M., & Hann, I.-H. (2015). Socially nudged: a quasi-experimental study of friends' social influence in online product ratings. *Information Systems Research*, Forthcoming.

Weiss, A. (2015, September 11). *9 movies that were even better than the nooks they were based on.* Retrieved from: http://www.bustle.com/articles/108987-9-movies-that-were-even- better-than-the-books-they-were-based-on

Williams, L. E., & Bargh, J. A. (2008). Keeping one's distance the influence of spatial distance cues on affect and evaluation. *Psychological Science*, *19*(3), 302-308.

Williams, L. E., Stein, R., & Galguera, L. (2014). The distinct affective consequences of psychological distance and construal level. *Journal of Consumer Research*, *40*(6), 1123-1138.

Winoto, P. & Tang, T. (2008). If you like the Devil Wears Prada the book, will you also enjoy the Devil Wears Prada the movie? a study of cross-domain recommendations. *New Generation Computing, 26*(3), 209-225.

Wooldridge, J. (2002). *Econometric analysis of cross-section and panel data.* The MIT Press: Cambridge, MA.

Wright-Porto, H. (2011). Social network integration. *Creative Blogging,* 207-237.

Yan, D. (2014). Future events are far away: Exploring the distance-on-distance effect. *Journal of Personality and Social Psychology, 106*(4), 514-525.

Yin, D., Bond, S., & Zhang, H. (2014). Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *MIS Quarterly, 38*(2), 539-560.

Zacks, J. M. (2014). *Flicker: Your brain on movies.* Oxford University Press, USA.

Zatlin, P. (2005). *Theatrical translation and film adaptation: A practitioner's view (Vol. 29)*. Multilingual Matters.

Zeng, X. & Wei, L. (2013). Social ties and user content generation: Evidence from Flickr. *Information Systems Research, 24*(1), 71-87.

Zhang, M., & Wang, J. (2009). Psychological distance asymmetry: The spatial dimension vs. other dimensions. *Journal of Consumer Psychology, 19*(3), 497-507.

Zhang, M., & Zhu, F. (2011). Group size and incentives to contribute: A natural experiment at Chinese Wikipedia. *The American Economic Review, 101*(4), 1601-1615.

Zhao, M., & Xie, J. (2011). Effects of social and temporal distance on consumers' responses to peer recommendations. *Journal of Marketing Research*, *48*(3), 486-496.

Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing, 74*(2), 133-148.

APPENDIX A

METHODOLOGICAL DETAILS APPENDIX

Descriptive Statistics for Regression Variables

We report descriptive statistics for the variables entering our regressions in Table 32. These descriptive statistics include the mean, standard deviation and correlation matrix of all variables.

Table 32. Descriptive statistics

	Mean	S. D.	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1. Rating Favorability (star rating)	4.216	0.964	1									
2. Temporal Distance (lndelay_months)	0.335	0.533	0.041	1								
3. Spatial Distance (lndistance_miles)	5.620	2.613	0.043	-0.004	1							
4. Mobile	0.039	0.193	-0.010	-0.062	0.043	1						
5. Amex User	0.063	0.243	0.005	-0.002	-0.081	-0.003	1					
6. Reviewer Average Rating	4.168	0.417	0.405	0.034	0.025	-0.010	0.013	1				
7. Restaurant Average Rating	4.167	0.334	0.254	0.014	0.009	0.010	0.003	0.080	1			
8. Lowest Price	27.136	11.257	0.016	0.021	-0.005	-0.011	0.023	-0.007	0.021	1		
9. Highest Price	51.372	30.401	0.031	0.038	0.042	-0.009	0.004	-0.002	0.075	0.333	1	
10. Public Transportation	0.575	0.494	-0.007	0.008	0.087	-0.008	0.017	0.011	-0.037	0.193	0.120	1

Content Coding Procedure to Measure Construal Level

The Linguistic Categorization Model (LCM) classifies five levels of linguistic terms that increase in abstractness (Semin & Fiedler, 1991). The five levels are descriptive action verbs (DAV), interpretive action verbs (IAV), state action verbs (SAV), state verb (SV) and adjective (ADJ). Coenen, Hedebouw, and Semin (2006) provided a comprehensive manual on how to use the LCM to measure language abstractness. And we conducted our content coding of construal level by following their guidelines.

We identified the five levels of linguistic terms in each review. We adopted the definitions, characteristic features, and classification criteria from Coenen et al. (2006, p. 6 and p. 14). Specifically, we define DAV as “verb that refers to a single specific action with a clear beginning and end, and with a physically invariance feature" (e.g., walk, speak, punch); IAV as “verb that refers to a multitude of different actions with a clear beginning and end that have the same meaning but don’t share a physically invariant feature” (e.g., help, escape, praise); SAV as “verb refers to a behavioral event but expresses the emotional consequence of an action rather than referring to an action as such" (e.g., surprise, amuse, satisfy); SV as “verb that refers to an enduring cognitive or emotional state with no clear definition of beginning and end" (e.g., hate, love, appreciate); and ADJ as “adjective that refers to a characteristic or feature qualifying a person" (e.g., honest, nice, excellent).

Furthermore, we counted the number of linguistic terms at each level in a review. We then give each level a numerical score (Semin & Fiedler, 1988, 1991). Particularly, DAV is given a score of 1; IAV and SAV are assigned a score of 2; SV is given a score of 3; and ADJ is assigned a score of 4. Lastly, based on the number of linguistic terms in each level and its corresponding coding weights, we obtain the abstract construal level for each review using the equation below (Coenen et al., 2006, p. 15):

$$c = \frac{n_1(1) + n_2(2) + n_3(2) + n_4(3) + n_5(4)}{n_1 + n_2 + n_3 + n_4 + n_5}$$

Here, *c* is the calculated abstract construal level; n_1 denotes the number of DAVs in a review; n_2 denotes the number of IAVs; n_3 denotes the number of SAVs; n_4 denotes the number of SVs; and n_5 denotes the number of ADJs. The calculated abstract construal level is a continuous variable ranging from 1 to 4. A larger number of *c* indicates higher construal level.

Propensity Score Matching (PSM)

Even though this study has econometrically identified the effects using the three-way fixed effects approach, it is possible that some unobserved factors might influence both psychological distance and review characteristics. For example, consumers might choose to visit higher quality restaurants when travelling, which could explain the positive association between spatial distance and rating favorability. Unobserved factors like natural manifestation of consumer preference would constitute selection bias. To address this potential selection issue, we used Propensity Score Matching (PSM; Abadie & Imbens, 2006; Dehejia & Wahba, 2002; Rosenthal & Rosnow, 1991).

PSM is a quasi-experimental approach that tests a causal treatment effect by controlling for the covariates that could affect the probability of subjects receiving the treatment – i.e., sources of selection (Angrist & Pischke, 2008). In our context, the goal is to identify a sample of the control group (low temporal / low spatial distance) that is extremely similar in observable characteristics, and thus comparable to, a treatment group (high temporal / high spatial distance). Here, matching is based on a subject's "propensity score," which refers to the conditional probability of a subject belonging to the treatment group. The conditional probability is determined based on the observed characteristics of the subject. "Matching" is achieved by pairing a treated subject with an untreated subject, based on similar prior propensity to receive treatment (Rosenbaum & Rubin, 1983). This matching reduces "sample selection bias in non-experimental settings" (Dehejia & Wahba, 2002, p. 151) and self-selection bias in nonrandomized experimental settings. In short, PSM is more appropriate than standard regression in the presence of selection bias; it entails identifying a subset of the control group that is directly comparable to the treatment group, and estimating the weighted average effect of the "treatment-on-the-treated." (Angrist & Pischke, 2008, pp. 51-56).

Following the procedure suggested by Caliendo and Kopeinig (2008) and Guo and Fraser (2010), two PSM analyses were performed, variably taking spatial distance or temporal distance as the treatment condition. To begin, we estimated the propensity score for each observation (subject) via Logit, regressing "treatment" on observable covariates, including the different characteristics of reviewers and restaurants from our models. These estimated propensity scores represent the conditional probability of a reviewer writing a review in the presence of high spatial or high temporal distance, given his or her observable characteristics, as well as the observable characteristics of the restaurant.

Next, we matched the observations (subjects) between the treatment and the control groups based on the estimated propensity scores. The matched sample, which is comprised of observations having equivalent propensity scores, is assumed to be homogeneous in terms of the distribution of observed reviewer and restaurant characteristics. Multiple matching algorithms were used, including nearest neighbor matching, kernel matching, radius matching, etc. (e.g., Guo & Fraser, 2010; Rishika, Kumar, Janakiraman, & Bezawada, 2013). Following the suggestion of Austin (2011), we further specified a matching threshold – i.e., a maximum deviation in propensity (caliper distance = 0.01), to improve the precision of the matching process.

We checked the assumptions of PSM using methods suggested in the prior literature (Leuven & Sianesi, 2012; Rosenbaum & Rubin, 1983). Besides visually checking the box-plots and kernel density plots of the propensity score distributions before and after matching for the control group and treatment group, the results of a balance check show that the mean differences of the covariates between the treatment and the control groups after matching are not significantly different from zero, suggesting that the two groups are comparable.

After propensity score matching, the treatment effects of spatial and temporal distance on rating favorability remain significant (see Table 33). Additionally, one-on-one matching, nearest neighbor matching, and kernel matching yield estimates that are almost identical, in terms of statistical significance, directionality and magnitude. Thus, the PSM results demonstrate the robustness of our main results.

Table 33. Propensity score matching

ATT of temporal distance		ATT of spatial distance	
Matching Algorithm	Rating	Matching Algorithm	Rating
NN Match (n=4, caliper=0.1) Sample Size on support = 134,214	.018*** (.006)	NN Match (n=4, caliper=0.1) Sample size on support = 135,711	.135*** (.010)
1-1 Matching (Caliper=0.001) Sample Size on support =122,628	.012** (.0059)	1-1 Matching (Caliper=0.001) Sample size on support=59,310	.100*** (.008)
Kernel Matching Sample Size on support =134,202	.016** (.0056)	Kernel Matching Sample size on support=135,551	.134*** (.009)

Notes: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$; ATT: average treatment effect on the treated.

Exact Covariate Matching

To further identify the effect of spatial and temporal distance on rating favorability, we used exact covariate matching. Our goal is to identify the differences between pairs of reviews that are authored by the same consumer, about the restaurants with exactly the same observable characteristics. In short, we achieve identification by focusing upon reviews that only vary in spatial or temporal distance. For example, to identify the effect of spatial distance, we compare reviews authored by the same consumer after they dined at restaurants that are the same in observed aspects, except for their geographic distance from the consumer. The variables we use in exact matching include consumer identifier, restaurant average rating, and the availability of public transit. This matching procedure resulted in 7,169 pairs of reviews for spatial distance, and 6,090 pairs of reviews for temporal distance.

Tables 34 present the results from exact covariate matching. Consistent with the results from our fixed effect and propensity score matching (PSM) estimations, we found that spatial and temporal distance increase rating favorability.

Table 34. Exact covariate matching

Effects of temporal distance

Variable	Mean (STD)		Difference (S.E.)	*t*-test	
	Temporal – high	Temporal – low		*t*	*p*-value
Rating	4.126 (.967)	4.097 (.969)	.029** (.017)	1.66	.048

Effects of spatial distance

Variable	Mean (STD)		Difference (S.E.)	*t*-test	
	Spatial – high	Spatial – low		*t*	*p*-value
Rating	4.188 (.924)	4.116 (.974)	.071*** (.016)	4.476	.0001

Notes: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

APPENDIX B

ASSESSING PRE-TREATMENT TRENDS UUSING

A RELATIVE TIME MODEL

A key identifying assumption of the DID specification is the existence of parallel trends between the treatment and control group leading up to the treatment. Under a relative-time specification, it is possible to test the assumption of parallel trends explicitly. The test is to check if the observed effects in the DID analyses were due to a pre-treatment trend that continued after the treatment. The DID assumption is jeopardized if we observe a pre-treatment trend that is in the same direction as the post-treatment effect, as this would imply that the effect began to manifest prior to the treatment).

As we discuss in the main text of the paper, *TripAdvisor*'s Instant Personalization is an "opt-out" feature and the effect is presumably more salient than *Yelp*'s Facebook Connect (an "opt-in" feature), therefore we use the time window around *TripAdvisor*'s exogenous shock to examine the pre-treatment and post-treatment differences in trends. We implement the approach suggested by Angrist and Pishke (2009, p. 177, Equation 5.2.6), interacting our platform dummy, *Trip*, with monthly dummies, to examine differences in trends around reviewing behaviors on each platform in the months before and after the *TripAdvisor*'s Instant Personalization treatment. Notably, this general approach has seen extensive use in recent IS work (Burtch et al. 2016, Chan and Ghose 2014, Greenwood and Wattal 2015).

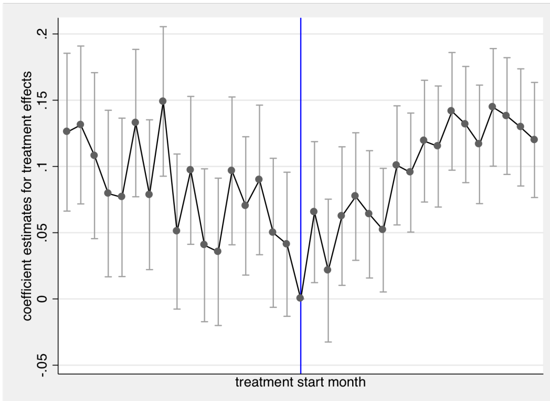
Specifically, we incorporate a platform fixed effect, *Trip*, a set of absolute time (monthly) dummies τ_t (e.g., January 2011, February 2011), their interactions, and a vector of restaurant fixed effects. Our econometric specification is detailed in Equation A1. We plot the coefficients associated with each relative month, taking the month of integration (December 2010) as the reference period.

$$DV_{ipt} = Trip_p + \tau_t + Trip_p * \tau_t + \gamma_1 \ln(words_{ipt}) + \gamma_2 rating_{ipt} + \alpha_i + \varepsilon_{ipt}$$

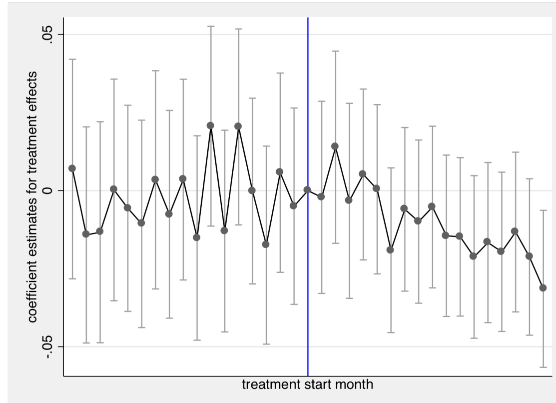
(A1)

In the above equation, *i* denotes restaurants, p indexes platforms and *t* indicates months. Figure 4 presents visualizations the coefficient estimates associated with our time dummy interactions for our DVs. As shown, we observe no apparent pre-treatment trends that is in the same direction as the after-treatment trend. Accordingly, we are unable to reject the parallel trends assumption. Moreover, we do not observe a peak in any of the effects, suggesting that the effects continue to progress in magnitude beyond our window of observation.

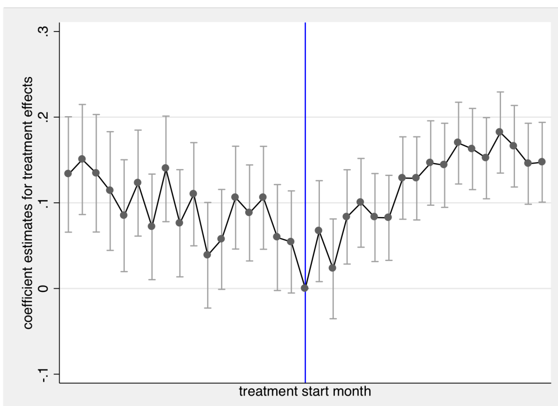
Figure 4. Visualization of Treatment Effects Over Time on DVs



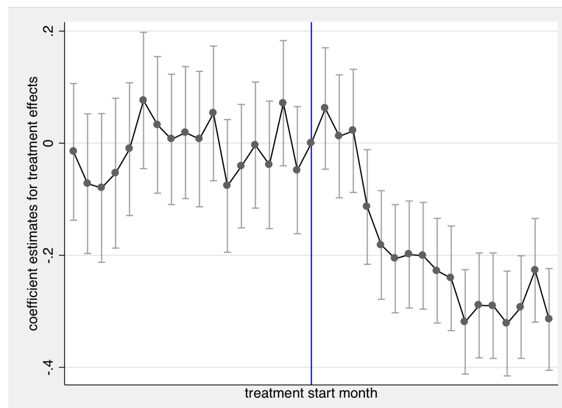
(a) Affective Process



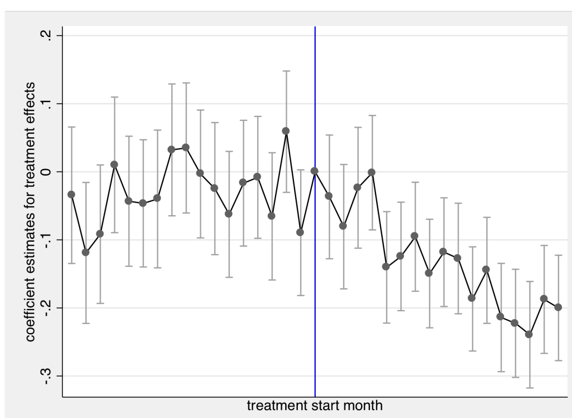
(b) Cognitive Process



(d) Positive Emotion



(c) Negative Emotion



(e) Negation

APPENDIX C

SEPARATE DID ANALYSES

As an additional robustness check, we report separate/single-shock DID analyses for the two exogenous shocks, to evaluate the robustness of our main findings, which were obtained via a double DID specification. We estimate the following models, where the parameter of interest (i.e., the DID estimate) is β_2.

$$\ln(Review\ Volume)_{ipt} = \beta_0 Trip_p + \beta_1 Trip\_Change_t + \beta_2 Trip_p * Trip\_Change_t + \beta_3 \ln(words_{ipt}) + \beta_4 rating_{ipt} + \alpha_i + \varepsilon_{ipt} \quad (1)$$

$$Review\ Volume_{ipt} = \beta_0 Trip_p + \beta_1 Trip\_Change_t + \beta_2 Trip_p * Trip\_Change_t + \beta_3 \ln(words_{ipt}) + \beta_4 rating_{ipt} + \alpha_i + \varepsilon_{ipt} \quad (2)$$

$$\ln(Review\ Volume)_{ipt} = \beta_0 Yelp_p + \beta_1 Yelp\_Change_t + \beta_2 Yelp_p * Yelp\_Change_t + \beta_3 \ln(words_{ipt}) + \beta_4 rating_{ipt} + \alpha_i + \varepsilon_{ipt} \quad (3)$$

$$Review\ Volume_{ipt} = \beta_0 Yelp_p + \beta_1 Yelp\_Change_t + \beta_2 Yelp_p * Yelp\_Change_t + \beta_3 \ln(words_{ipt}) + \beta_4 rating_{ipt} + \alpha_i + \varepsilon_{ipt} \quad (4)$$

$$\ln(Linguistic\ Characteristic)_{ipt} = \beta_0 Trip_p + \beta_1 Trip\_Change_t + \beta_2 Trip_p * Trip\_Change_t + \beta_3 \ln(words_{ipt}) + \beta_4 rating_{ipt} + \alpha_i + \varepsilon_{ipt} \quad (5)$$

$$Linguistic\ Characteristic_{ipt} = \beta_0 Trip_p + \beta_1 Trip\_Change_t + \beta_2 Trip_p * Trip\_Change_t + \beta_3 \ln(words_{ipt}) + \beta_4 rating_{ipt} + \alpha_i + \varepsilon_{ipt} \quad (6)$$

$$\ln(Linguistic\ Characteristic)_{ipt} = \beta_0 Yelp_p + \beta_1 Yelp\_Change_t + \beta_2 Yelp_p * Yelp\_Change_t + \beta_3 \ln(words_{ipt}) + \beta_4 rating_{ipt} + \alpha_i + \varepsilon_{ipt} \quad (7)$$

$$Linguistic\ Characteristic_{ipt} = \beta_0 Yelp_p + \beta_1 Yelp\_Change_t + \beta_2 Yelp_p * Yelp\_Change_t + \beta_3 \ln(words_{ipt}) + \beta_4 rating_{ipt} + \alpha_i + \varepsilon_{ipt} \quad (8)$$

First, we report the separate DID analyses results for review volume in Table 35 and Table 36, where we observe that, compared with *Yelp*, the review volume of *TripAdvisor* increased by 38.8% after implementing Instant Personalization. Similarly, compared to *TripAdvisor*, the review volume of *Yelp* increased by 18.2% after integrating Facebook Connect.

Table 35. *TripAdvisor* DID Volume Effect

Variables	(1) ln(Review Volume)	(2) Review Volume
Trip	–0.841***(0.018)	–3.482***(0.132)
Trip_Change	0.257***(0.006)	1.210***(0.037)
Trip * Trip_Change	**0.388***(0.014)**	**1.377***(0.116)**
ln(words)	0.247***(0.005)	0.952***(0.027)
Rating	0.014***(0.003)	0.082***(0.015)
Constant	–0.124***(0.026)	–0.353*(0.143)
Observations	112,262	112,262
R-squared	0.220	0.140
Number of restaurants	3,964	3,964
Restaurant Fixed Effect	Yes	Yes

Notes: Cluster-robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Table 36. *Yelp* DID Volume Effect

Variables	(1) ln(Review Volume)	(2) Review Volume
Yelp	0.612***(0.021)	2.117***(0.119)
Yelp_Change	0.006(0.015)	–0.010(0.067)
Yelp * Yelp_Change	**0.182***(0.016)**	**0.858***(0.078)**
ln(words)	0.149***(0.006)	0.573***(0.029)
Rating	0.006(0.003)	0.026*(0.013)
Constant	–0.462***(0.041)	–1.570***(0.231)
Observations	47,151	47,151
R-squared	0.195	0.150
Number of restaurants	3,178	3,178
Restaurant Fixed Effect	Yes	Yes

Notes: Cluster-robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Second, we present the separate DID results for mental processes. In Tables 37 and Table 38, we observe that, compared with *Yelp*, affective processes on *TripAdvisor* increased by 1.8%, whereas cognitive processes decreased by 0.9% after implementing Instant Personalization. Additionally, positive emotion on *TripAdvisor* increased by 3% while negative emotion decreased by 21%. According to Table 39 and Table 40, compared with *TripAdvisor*, affective processes on *Yelp* increased by 6.4%, while cognitive processes declined by 2.4% after implementing Facebook Connect. Further, positive emotion on *Yelp* increased by 7.8% but negative emotion decreased by 10.2%.

Table 37. *TripAdvisor* DID Affective and Cognitive Processes

Variables	(1) ln(Affective Process)	(2) Affective Process	(3) ln(Cognitive Process)	(4) Cognitive Process
Trip	–0.160***(0.006)	–0.967***(0.033)	0.051***(0.003)	0.739***(0.056)
Trip_Change	0.042***(0.002)	0.294***(0.017)	0.004**(0.001)	0.067**(0.022)
Trip * Trip_Change	**0.018**(0.006)**	**0.095*(0.037)**	**–0.009*(0.003)**	**–0.115+(0.059)**
ln(words)	–0.235***(0.003)	–1.797***(0.017)	0.039***(0.002)	0.855***(0.028)
Rating	0.094***(0.002)	0.593***(0.009)	–0.012***(0.001)	–0.198***(0.015)
Constant	2.745***(0.015)	13.971***(0.096)	2.571***(0.009)	11.860***(0.154)
Observations	110,337	110,669	108,368	110,669
R-squared	0.152	0.143	0.015	0.020
Number of restaurants	3,958	3,961	3,953	3,961
Restaurant Fixed Effect	Yes	Yes	Yes	Yes

Notes: Cluster-robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Table 38. *TripAdvisor* DID Positive and Negative Affective Processes

Variables	(1) ln(Positive Emotion)	(2) Positive Emotion	(3) ln(Negative Emotion)	(4) Negative Emotion
Trip	–0.196***(0.006)	–1.016***(0.043)	0.225***(0.014)	0.025(0.015)
Trip_Change	0.047***(0.003)	0.292***(0.018)	–0.044***(0.005)	0.006(0.007)
Trip * Trip_Change	**0.030***(0.007)**	**0.096*(0.047)**	**–0.210***(0.015)**	**–0.053**(0.016)**
ln(words)	–0.270***(0.003)	–2.154***(0.024)	–0.300***(0.007)	–0.020*(0.008)
Rating	0.188***(0.002)	1.101***(0.011)	–0.295***(0.003)	–0.436***(0.006)
Constant	2.410***(0.016)	12.815***(0.126)	2.425***(0.033)	2.656***(0.054)
Observations	109,966	110,669	86,307	110,669
R-squared	0.236	0.228	0.129	0.128
Number of restaurants	3,958	3,961	3,929	3,961
Restaurant Fixed Effect	Yes	Yes	Yes	Yes

Notes: Cluster-robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Table 39. ***Yelp*** **DID Affective and Cognitive Processes**

Variables	(1) ln(Affective Process)	(2) Affective Process	(3) ln(Cognitive Process)	(4) Cognitive Process
Yelp	0.076***(0.011)	0.083(0.089)	–0.013*(0.006)	–0.389***(0.087)
Yelp_Change	–0.035**(0.012)	–0.443***(0.098)	0.025***(0.007)	0.346***(0.104)
Yelp * Yelp_Change	**0.064***(0.013)**	**0.650***(0.100)**	**–0.024**(0.007)**	**–0.354***(0.107)**
ln(words)	–0.193***(0.004)	–1.620***(0.030)	0.063***(0.003)	0.646***(0.036)
Rating	0.076***(0.002)	0.474***(0.015)	–0.011***(0.001)	–0.172***(0.021)
Constant	2.498***(0.024)	13.174***(0.181)	2.458***(0.015)	13.151***(0.206)
Observations	46,767	46,821	46,807	46,821
R-squared	0.103	0.114	0.022	0.015
Number of restaurants	3,174	3,174	3,174	3,174
Restaurant Fixed Effect	Yes	Yes	Yes	Yes

Notes: Cluster-robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Table 40. ***Yelp*** **DID Positive and Negative Affective Processes**

Variables	(1) ln(Positive Emotion)	(2) Positive Emotion	(3) ln(Negative Emotion)	(4) Negative Emotion
Yelp	0.097***(0.012)	0.063(0.088)	–0.052*(0.022)	0.036(0.022)
Yelp_Change	–0.040**(0.013)	–0.433***(0.096)	0.049*(0.024)	0.038(0.023)
Yelp * Yelp_Change	**0.078***(0.014)**	**0.662***(0.098)**	**–0.102***(0.025)**	**–0.047+(0.026)**
ln(words)	–0.226***(0.005)	–1.622***(0.029)	–0.231***(0.009)	–0.024**(0.009)
Rating	0.171***(0.003)	0.862***(0.014)	–0.253***(0.005)	–0.434***(0.006)
Constant	2.118***(0.027)	10.768***(0.178)	2.054***(0.050)	2.650***(0.048)
Observations	46,670	46,821	38,299	46,821
R-squared	0.191	0.175	0.112	0.101
Number of restaurants	3,173	3,174	3,108	3,197
Restaurant Fixed Effect	Yes	Yes	Yes	Yes

Notes: Cluster-robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Third, we show separate DID analyses results for the inhibition effect in Table 41 and Table 42. We observe that, compared with *Yelp* as the baseline control group, the use of negations on *TripAdvisor* decreased by 11.9% after implementing Instant Personalization. Similarly, compared with *TripAdvisor*, language references to negation on *Yelp* decreased by 8.1% after integrating with Facebook Connect.

Table 41. *TripAdvisor* DID Inhibition Effect

Variables	(1) ln(Negation)	(2) Negation
Trip	0.509***(0.010)	0.576***(0.018)
Trip_Change	–0.018***(0.005)	0.026***(0.007)
Trip * Trip_Change	**–0.119***(0.011)**	**–0.068***(0.019)**
ln(words)	–0.265***(0.006)	–0.006(0.009)
Rating	–0.200***(0.003)	–0.351***(0.005)
constant	1.993***(0.030)	2.355***(0.053)
Observations	93,870	110,669
R-squared	0.143	0.093
Number of restaurants	3,937	3,961
Restaurant Fixed Effect	Yes	Yes

Notes: Cluster-robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Table 42. *Yelp* DID Inhibition Effect

Variables	(1) ln(Negation)	(2) Negation
Yelp	–0.389***(0.017)	–0.543***(0.030)
Yelp_Change	0.061***(0.017)	0.042(0.033)
Yelp * Yelp_Change	**–0.081***(0.018)**	**–0.038+(0.024)**
ln(words)	–0.231***(0.008)	–0.047***(0.012)
rating	–0.170***(0.004)	–0.285***(0.007)
constant	2.138***(0.044)	2.855***(0.071)
Observations	40,877	46,821
R-squared	0.121	0.098
Number of restaurants	3,139	3,174
Restaurant Fixed Effect	Yes	Yes

Notes: Cluster-robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

APPENDIX D

PEMISSION LETTER FOR INCLUSION OF COPYRIGHTED MATERIAL

ELSEVIER LICENSE
TERMS AND CONDITIONS

Feb 17, 2017

This Agreement between Ni Huang ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4051560550118
License date	Feb 17, 2017
Licensed Content Publisher	Elsevier
Licensed Content Publication	Journal of Consumer Psychology
Licensed Content Title	Effects of multiple psychological distances on construal and consumer evaluation: A field study of online reviews
Licensed Content Author	Ni Huang,Gordon Burtch,Yili Hong,Evan Polman
Licensed Content Date	October 2016
Licensed Content Volume	26
Licensed Content Issue	4
Licensed Content Pages	9
Start Page	474
End Page	482
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	Reading Between the Lines: Three Investigations of User Generated Content Using Text Analytics
Expected completion date	Apr 2017
Estimated size (number of pages)	160
Elsevier VAT number	GB 494 6272 12
Requestor Location	Ni Huang 832 W Wagner Dr GILBERT, AZ 85233 United States Attn: Ni Huang
Publisher Tax ID	98-0397604
Total	0.00 USD

MIS Quarterly
Carlson School of Management
University of Minnesota
Suite 4-339 CSOM
321 19th Avenue South
Minneapolis, MN 55455

March 15, 2017

Nina Huang
Fox School of Business
Temple University
1801 Liacouras Walk
Philadelphia, PA 19122

Permission to use article from
MIS Quarterly in doctoral dissertation

Permission is hereby granted for Ni Huang to use the article "Social Network Integration and User Content Generation: Evidence from Natural Experiments" forthcoming in *MIS Quarterly* as part of her doctoral dissertation, *Reading Between the Lines: Three Investigations of User Generated Content Using Text Analysis* being completed at Temple University.

The citation information for the work should include

> Copyright © 2017/2018, Regents of the University of Minnesota. Used with permission.

The permission extends to publication of the dissertation by UMI. Any other usage or publication will require specific permission and may require payment of a fee.

Janice I. DeGross
Manager