The Review of Socionetwork Strategies (2020) 14:19–40
https://doi.org/10.1007/s12626-019-00041-5

ARTICLE

Topic Features in Negative Customer Reviews: Evidence Based on Text Data Mining

Zhen Li[1] · **Fangzhou Li**[2] · **Jing Xiao**[3] · **Zhi Yang**[3]

Received: 22 January 2019 / Accepted: 7 June 2019 / Published online: 2 July 2019

Abstract
Several studies have focused on the effects of online negative customer reviews on sales, especially pertaining to Internet shopping and e-retailing. However, there is mixed evidence and the theoretical studies have mainly focused on the volume and valence. To understand the effects of negative customer reviews on sales, the present study uses text data mining techniques to investigate how three factors, namely "content topic, proportion, and consistency," bout the textual content of negative customer reviews influence online sales. Relevant data were collected from a large-scale online shopping platform. The results of content association and topic extraction reveal four topics—product quality, delivery service, cost performance, and taste. A new econometric model proposed in this study shows that different topics have different effects on sales. Negative customer reviews with a higher percentage or consistency about these four topics significantly jeopardize product sales. Theoretical and managerial implications and future research directions are also presented.

Keywords Internet shopping · Negative customer reviews · Online sales · Text data mining · Content topic · Proportion · Consistency

✉ Jing Xiao
jingxiao@hust.edu.cn

Zhen Li
li@toyo.jp

Zhi Yang
zhiyang@hust.edu.cn

[1] Faculty of Business Administration, Toyo University, Tokyo 112-8606, Japan

[2] Graduate School of Business Administration, Kobe University, Kobe 657-8501, Japan

[3] School of Management, Huazhong University of Science and Technology, Wuhan 430074, Hubei, China

1 Introduction

Recent rapid proliferation of Internet technology has increased academics' and practitioners' interest. Customers are now empowered to easily share their experience or provide feedback on a product or service they have bought or experienced [1, 2], thereby affecting other customers' purchase decisions [3, 4]. Customers' statements are known as reviews (or product reviews), which are available to a multitude of people via the Internet [5]. Both academics and practitioners believe that the use of customer reviews strongly affects customer purchase process [6–8]. Previous studies have documented that customer review is a vital, trustworthy source of information for people to gather product-related information [9, 10] that helps them avoid laborious search efforts and mitigate the risk of decision making [6]. While studies on customer reviews are close to saturation, the stream of negative online customer reviews needs further exploitation, especially for its impact on sales outcomes.

Both researchers and managers are focusing more on negative customer reviews rather than on positive and neutral reviews. Although negative reviews are fewer in the total quantity of online reviews [11, 12], they are perceived as more diagnostic, trustworthy, and persuasive [13–16]. In addition, earlier studies indicated that negative reviews have stronger effects on sales outcomes [9, 17, 18]. However, empirical studies on how negative customer reviews influence sales performance were found to be mixed and equivocal. Floyd et al. [7] reviewed several studies and summarized that negative reviews are detrimental to product sales [9, 18, 19], whereas some researchers found that negative reviews did not jeopardize sales [6, 20, 21]. Surprisingly, there is evidence that negative reviews can even improve product evaluations and the trading volume [4, 22]. These conflicting results led to debates over the relationship between negative reviews and product performance and confusion between academics and practitioners on when and how to deal with negative reviews. Indeed, there is little empirical evidence to help explain what attributes of negative customer reviews work and how. Therefore, theoretical and managerial inputs are urgently needed to understand the mechanism of negative reviews.

Compared with numerous literatures analyzing the effect of individual negative reviews, there is scant work that emphasizes the importance of group features of negative reviews. In fact, customers usually form their opinion regarding a product by considering all negative reviews as a whole and evaluating the group attributes of negative reviews, and then they act accordingly. Moreover, regarding the impact of customer reviews on sales performance, most extant studies mainly highlighted the quantitative substitutes of customer reviews such as volume [21, 23], valence [24, 25], or both [26–29]. However, using these quantitative substitutes without considering the textual content and group attributes of reviews may be likely have led to the inconsistent results [18, 19, 21].

Some studies showed that the impact of customer reviews depends on the context, and customers read and respond to the textual content of reviews rather than depend only on summary statistics [9, 23]. Therefore, there is need for an in-depth analysis of the textual content of customer reviews considering all customer reviews of one product from one seller as a whole to elucidate its effect on sales. Although several researchers have focused on the textual content, a majority of them only used computational algorithms to distinguish negative from positive reviews, rather than mining the latent features of negative reviews [22, 30]. Recent research has proposed the text analysis using the Latent Dirichlet Allocation (LDA) model to capture the latent topics about customer reviews and the corresponding valence, validity, importance, dynamics, and the heterogeneity [31, 32]. However, the methods and results in those studies still need to be improved, and the relationship between the topics of negative reviews and sales outcomes has yet to be investigated.

To reconcile these issues, the present study applied text data mining techniques to explore the latent topics about negative customer reviews in Internet shopping and how the proportion and consistency of topics in negative reviews affect online sales outcomes. Data on customer reviews for snacks from a B2C online shopping platform in China were used for this study through web scraping, and this was matched with sales data at the product level. Following the content association and cluster analysis with the topic model, four principal latent topics were extracted, namely product quality, delivery service, cost performance, and taste, from the textual content of aggregated reviews, and the proposed statistical models were used to estimate the effects of negative reviews on sales in detail. Results suggested that different topics have different impacts on sales outcomes. The effect of negative reviews is the highest for the topic of taste, followed by the topics of product quality and cost performance, and the lowest for delivery service.

This study contributes to the literature in several ways: First, this study highlights the significance of group attributes of online reviews, which were previously neglected by researchers. Second, using the large-scale secondary data, this study reveals the underlying mechanism through which negative reviews affect sales outcomes. Third, this study proposes a model that considers both qualitative and quantitative substitutes of negative reviews. This provides some valuable insights not only for marketing academics but also for practitioners working for retailers or customer services to understand when and how to deal with online negative customer reviews.

In this background, Sect. 2 presents a literature review related to the discussion on customer reviews and proposes the conceptual foundations and corresponding hypotheses. Section 3 introduces the study design and data collection process and explores the topics that constitute negative reviews. Section 4 presents the influences of these topics on sales outcomes. Finally, Sect. 5 offers conclusions, implications, and future directions for research.

2 Conceptual Foundations and Hypotheses

Existing studies on customer reviews and Internet shopping have highlighted the effect of review volume and valence on customers' purchase decision-making process and sales outcomes [8, 26–29]. These studies confirm that the information on negative customer reviews is stronger and more influential than that on positive reviews [7, 33]. Although there are several studies on the effects of negative reviews on sales performance, the study results are mixed, and there is lack of information on the mechanism of the effects [6, 23, 34]. Earlier research acknowledges that customers appear more solicitous regarding the textual content of the reviews rather than the summary statistics, such as the total number of reviews [9]. Advancing this idea, the present study posits that the effects of customer reviews on sales depend on the textual content. Therefore, the literature reviewed in this study outlines the theory for the textual content of reviews and proposes corresponding hypotheses.

2.1 Effects of Customer Reviews on Sales

Earlier studies on Internet shopping suggested that online customer review is an important and trustworthy source for people to acquire product/service-related information [9]. The emergence of customer reviews helps other customers to make less efforts on information gathering and mitigate the risk of decision making, which influences customers' purchase decisions [35, 36]. Majority of earlier studies on customer reviews have highlighted the volume and valence. Many have showed that positive reviews typically increase customers' expectations of quality and attitudes toward the product, while negative reviews may involve product denigration, rumor, or complaints, and usually have an unfavorable effect on the attitudes toward a product [21]. Moreover, several experimental studies demonstrated that the information given in negative reviews is usually stronger, more influential and more predictive than that in positive reviews [7, 27, 33]. However, interestingly, there are conflicting and inconclusive empirical results on the effects of customer reviews on product sales performance. For instance, Chevalier and Mayzlin [9] and Ho-Dac et al. [19] found that negative reviews unfavorably impact product sales and attitudes toward the product, but the same has not been observed by other studies [6, 37]. Furthermore, surprisingly, some empirical data even substantiated that negative reviews can improve product evaluations and sales performance [4, 22]. Therefore, further exploration is essential to clarify how the negative customer reviews influence sales, if at all.

In comparison with using the volume and valence of online reviews to predict customers' decision making, recent studies proposed that the influence of reviews depends on the context and highlighted the importance of the textual content [22, 30, 32]. However, this attempt is still in the nascent stage, and the data mining process for latent clusters of customer reviews needs further improvement [32].

Existing studies recognize that customers face perceived risks, uncertainty, and potentially undesirable consequences in online markets [38, 39], and customer reviews are believed to help other customers save on the search efforts and mitigate the risk of decision making [6]. Therefore, the conceptual framework of the research on customer reviews is based on the theory of perceived risk. Perceived risks directly affect customers' willingness to buy and moderate the relationship between trust and consumer behavior in online markets [40]. Most perception studies stressed that given the risks in online markets, the areas of concern for customers are the perceived risk in a vendor, product performance, finance, and delivery [41]. Vendor risk refers to whether the seller is capable of delivering the items as described and providing a good service [42]. Product risk mainly represents customer anxiety regarding the product quality and expected performance. Delivery risk in online shopping refers to the problems such as packaging damage and timely delivery [43]. Financial risk refers to the possibility of monetary loss, though this risk is progressively attenuated by the improvement in information encryption technology and the extensive use of third-party payment [44]. Therefore, the present study posits that customer reviews reflect product-related information through several topics, which helps other customers to mitigate their decision-making risk. In this context, the following assumption is made:

Assumption Sales effects of customer reviews depend on their latent topics, and each topic influences sales outcome with a different impact.

As a pioneering research on textual content of reviews, Tirunillai and Tellis [32] adopted the text data mining method to capture the latent topics about product reviews and the corresponding valence, validity, importance, dynamics, and the heterogeneity. However, their studies failed to explore the sales effect of each topic. To improve upon the methods and results in the previous research, the present study discusses in depth the manner in which the topics underlying negative reviews impact sales outcomes.

2.2 Proportion

In comparison with positive/neutral reviews, negative reviews are minorities represented by a small group [45]. Existing studies have pointed out that the information in negative customer reviews is marginal but diagnostic, trustworthy, and persuasive [12]. Predictive models of social influence postulate that numbers convey social impact. Larger number of sources will have a greater ability to exert influence [46, 47]. These models predict that when the ratio of minority to majority is higher, the minority will have a greater influence, which concurs with the dual-process perspective proposed by Moscovici [48, 49]. Minority is usually considered to be erroneous, while a high proportion can weaken this heuristic thinking. Earlier studies discovered that the percentage of negative reviews negatively affects the sales performance [9, 18]. This study assumes that several topics underlying customer reviews draw people's attention. Moreover, in case of a fewer reviews, people discount these topics in negative reviews as idiosyncratic, biased perspective [49, 50]. Increasing the number of topics impedes customers' resistance by attributing the appeal of the negative statements to personal idiosyncrasies. In other words, it is much more difficult to reject the negative statements by attributing them to subjective factors. Under this condition, the conflict is between two judgments of fact rather than between value and preference. Consequently, something objective must be involved and people will be less likely to purchase under this condition. Hence, the following hypothesis is drawn:

Hypothesis 1 *Proportions of topics in negative customer reviews have negative effects on sales performance in online shopping.*

2.3 Consistency

Consistency in customer reviews is defined as the degree to which all reviewers of one product from one seller agree on the product or its performance [4]. The consistency of source is a primary factor for eliciting minority influence [45, 49, 51]. This result is applicable to the negative reviews. Empirical studies have found that people attribute the dispersion in online word of mouth to the heterogeneity of reviewer preference and not product itself when the products are perceived to be dissimilar [52]. Furthermore, opinion consistency in the online negative word-of-mouth exerts significant influence on the customers' change in attitude toward the firm. Similarly, with greater consistency, customers are more likely to attribute problem causality to the target company [53]. It is also indicated that consistency has a positive impact on the perceived information credibility, which enhances the intention to accept an online word of mouth [54]. In other words, consistent customer reviews indicate that the sources are valid and unbiased and are more likely to be trustworthy. Therefore, the present study considers that customers engaging in online shopping are more likely to adopt the message of consistent reviews in their purchase decisions, which leads to the following hypothesis:

Hypothesis 2 *Consistency of topics in negative customer reviews has a negative effect on product sales.*

3 The Data

Data collection process and the details of the variable measurements for the entire study are described as follows.

3.1 Datasets

To draw the latent topics in negative customer reviews and investigate how those topics influence sales outcomes, the present study combines data on customer online reviews and product sales performance, which are available from a large-scale e-commence platform through web scraping. This platform is one of the fastest growing online B2C markets in China, with nearly 300 million active users and 1.3 trillion RMB Yuan turnover in sales. To eliminate the influence of merchant and product category, a famous snack brand with a flagship store on this platform was selected, and data of its fastest-selling product range, dried fruits, and vegetables were extracted.[1] It is believed that as a fast-moving product sold at a flagship

[1] A kind of casual snack that can be eaten after opening the packing.

Fig. 1 Information of sales outcomes

Table 1 Basic information concerning the product sales

	Mean	SD	Min.	Max.
Price	17.47	7.69	4.90	42.80
Number of reviews	12739	12749.3	1100	59900
Praise degree	0.98	0.01	0.96	0.99

The number of products is 39

store on this platform, it draws a large volume of customer reviews in each specific product. Moreover, posting feedback or comments on this platform does not require heavy censorship, which ensures a certain number of reviews, especially negative. A Python-based web crawler was developed to retrieve the relevant data. All customer review data were collected based on the transaction record, and each product involved at least 1000 reviews. The data on customer reviews and product sales performance by product SKU[2] and the date of purchase were matched.

Data on product sales performance in this study include the information of product SKU, sales price, overall accumulated number of observable customer reviews, and the daily number of good feedbacks. There were overall 39 specific products. After referring to the relevant literature, aggregated data were made available, which included monthly sales performance for each relevant product [42]. Fig. 1 shows a product profile sampled to retrieve the sales performance information. Table 1 displays the basic statistical information of data used in this study.

Data on customer online reviews include information of user id, feedback rating, product purchased, specific content of review, the number of helpfulness votes, and the date of purchase. A total of 366,430 customer personal reviews across the

[2] A stock keeping unit (SKU) is a 6–8 character long alphanumeric code used to identify a product and track its inventory. The SKU for each product and seller is a unique number.

Fig. 2 A sample of customer review

product range were collected up to July 2018. Figure 2 shows a sample of customer review.

3.2 Variables

Independent variables used in this study relate to the textual content of customer reviews. The characteristics of reviews are assessed through the latent topics on the negative customer reviews; the number of each topic in overall reviews; and the consistency of the topics in negative reviews. Latent topics in negative customer reviews are captured through natural language processing using text data mining methods. The proportions are identified by calculating the percentage of each topic accounting for the overall reviews of the given product. Consistency in the present study was measured using the sum of the squares of the topic shares within the whole negative reviews, where the topic shares are expressed as fractions such as Herfindahl–Hirschman Index (HHI). The result is proportional to the average topic share, which can range from 0 to 1, moving from scattering to centralizing. An increase in the consistency generally indicates common negative statements or customer feedback, whereas a decrease indicates the opposite. Proportion and consistency in the present study can be processed as follows:

$$p_{k,i,t} = \frac{T_{k,i,t}}{N_{i,t}} \times \frac{N_{i,t}}{R_{i,t}} \tag{1}$$

$$C_{i,t} = \sum_{k} \left(\frac{T_{k,i,t}}{N_{i,t}}, \right)^2 \tag{2}$$

where $p_{k,i,t}$ denotes the proportion of topic k for product i at period t; $T_{k,i,t}$ is the size of topic k (i.e., number of negative reviews that relate to topic k); and $N_{i,t}$ and $R_{i,t}$, respectively, indicate the accumulated number of negative and overall customer reviews, which can be easily aggregated from the dataset of customer online reviews. In the present study, negative reviews refer to those customer comments or feedbacks with a rating less than 2 stars, while those with a rating greater than 4 stars are classified as positive. $C_{i,t}$ represents the consistency whose value is always between 0 and 1.

Dependent variable in the present study is the sales outcome of each product. However, given the attributes of the product webpage on the platform used in this study, obtaining accurate information of sales is not easy. Therefore, the number of monthly customer reviews was combined and used as a substitute for sales, which is a practical approach, because previous studies have found a linear relationship between sales and the number of reviews [24, 25, 55]. Since the review can only be written by a customer who has bought the product, which means that the customer bought a product at period t, his or her review will be displayed at period $t + 1$. Therefore, the sales can be processed using the following equation:

$$S_{i,t} = R_{i,t+1} - R_{i,t}, \tag{3}$$

where $S_{i,t}$ is the difference between the accumulated number of overall reviews at period t and $t + 1$, and this indicator is used to refer to sales volume at period t.

Volume, valence, and the regular price for each product i at period t are considered the control variables. In particular, volume refers to the total number of all customer reviews, and the valence is denoted as the degree of praise that can be observed directly from the product webpage. Although the product page also provides the total number of customer reviews, the value is only the approximate number, such as 45,000+, which is not accurate (see Fig. 1). Consequently, in the present study, volume was determined by counting the number of customers' individual reviews that are stored in the database. In addition, regular price is denoted as the mode[3] of selling price during the relevant period.

4 Analysis

Natural language processing with the application of text data mining and LDA topic modeling is run to retrieve the latent cluster from the textual content of the accumulated negative customer reviews, followed by a detailed statistical estimation to find how the textual content of negative customer reviews impacts sales outcomes.

4.1 Text Analysis

After summarizing and aggregating the dataset of customer reviews, a total of 6260 negative reviews were obtained in the present study. Following is the pattern of analysis: process the keyword recognition, count their frequency, and identify the relationships among those keywords from the textual content and context. Figure 3 provides the co-occurrence network of keywords in negative customer reviews, in which the frequencies of their occurrence were at least 80, and the Jaccard indexes.[4] among them are more than 0.2 Figure 3 reveals that the high-frequency keywords in negative reviews appear to be concentrated on topics such as store service, product quality, cost performance, delivery service, packaging, and taste.

To further identify the constituent topics in negative customer reviews, which is one of the purposes of the present study, in addition to the aforementioned keyword

[3] The mode is the value that appears most often in the dataset.

[4] The Jaccard index is a statistic used for comparing the similarity and diversity of sample sets.

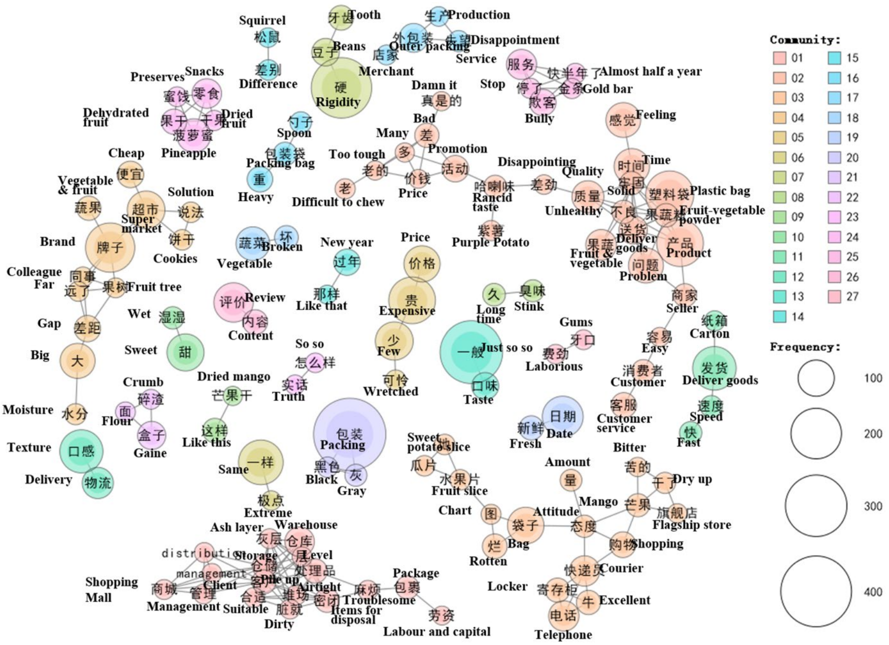

Fig. 3 Co-occurrence network of words in negative reviews

recognition, the study used three other methods: a self-organizing map, a hierarchical clustering analysis, and a topic model by LDA. Using these three methods, one can effectively and accurately analyze the words contained in the original review content to discover the common topics, the association between these topics, and how they change over time [32, 56]. In addition, these three methods were the preferred techniques of text analysis for the following two reasons: First, these approaches do not depend on the structure of the text, nor does it require strict syntactical or grammatical properties of the language posted on the Internet, which makes them ideal for processing the textual contents of customer reviews. Second, unsupervised methods can be used to complete several steps of the analysis so that little human intervention is involved, resulting in minimal bias or errors.

Figure 4 illustrates the result of the self-organizing map[5] with 1000 iterations, which is trained using an unsupervised learning approach similar to K-means. The figure clearly shows eight themes that can be concluded from customers' negative product reviews, including delivery service, service image, shipping, product quality, packing, store service, cost performance, and taste, respectively.

To ensure the accuracy and stability of the results of analysis, the present study also provides a hierarchical clustering analysis with Ward's method to reconfirm the topic classification for the study data. Through learning with the text data, the results of hierarchical clustering analysis are represented in Fig. 5, where seven

[5] A self-organizing map is a type of artificial neural network that uses unsupervised learning to build a two-dimensional map of a problem space.

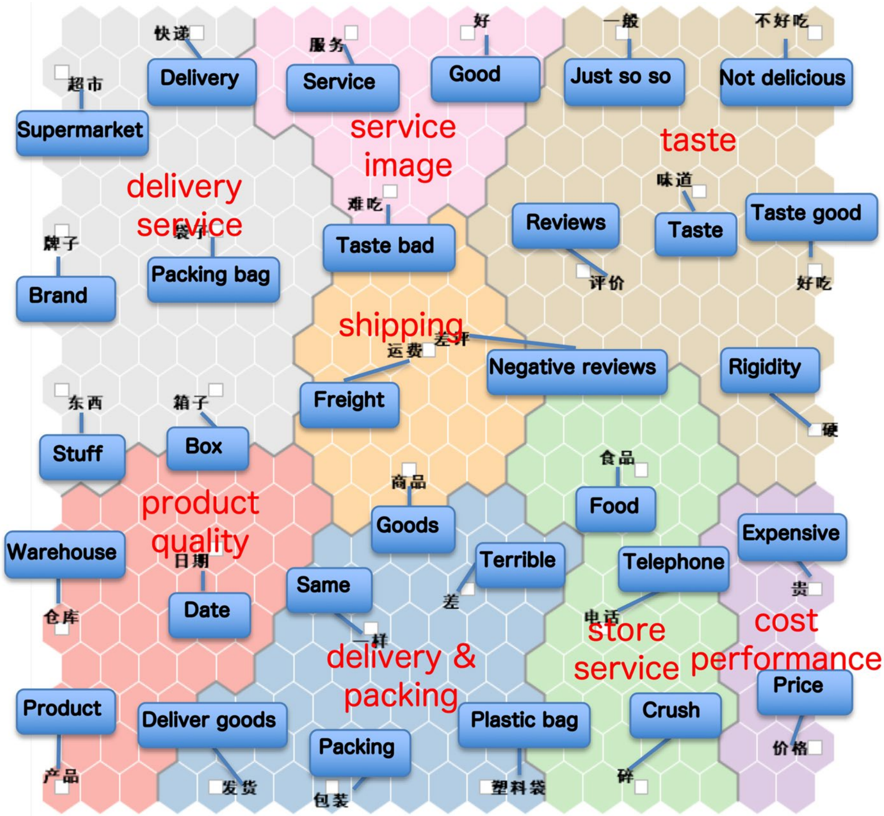

Fig. 4 Self-organizing map for negative reviews

Fig. 5 Clustering Result for Negative Reviews

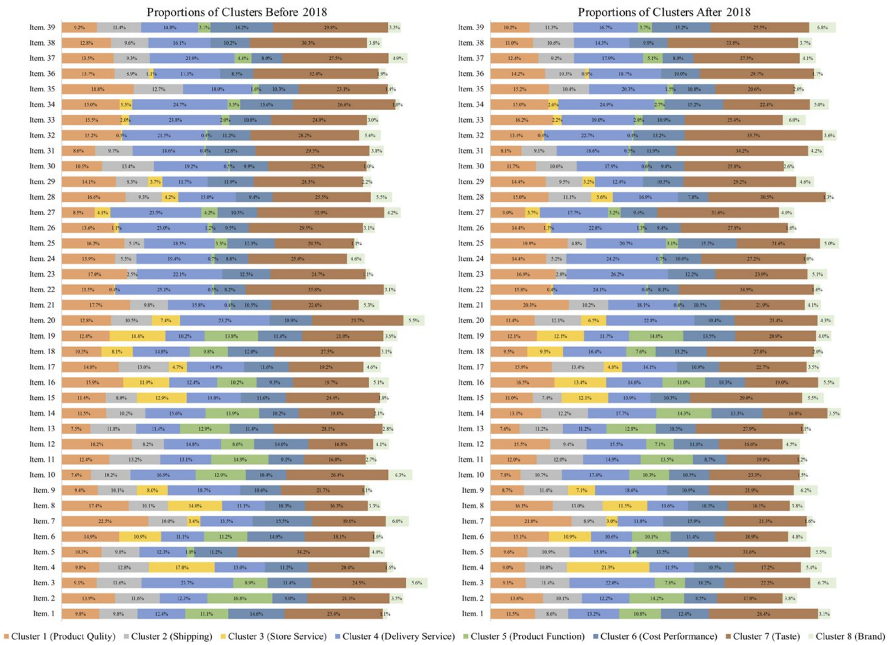

Fig. 6 Proportions of clusters for products

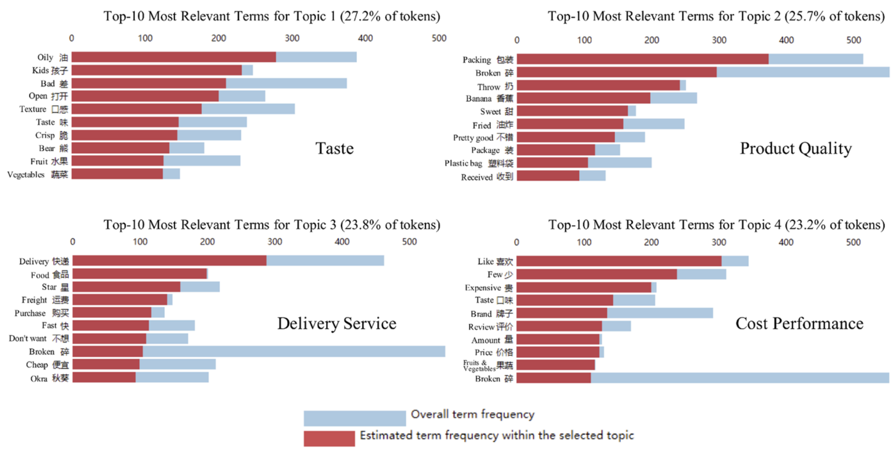

Fig. 7 Topics identified by LDA

clusters were created and classified as product quality, shipping, store service, delivery service, product function, brand image, and taste/ cost performance. These seven clusters show a similar classification as the self-organizing performed earlier.

After summarizing the results in both self-organizing map and hierarchical clustering analysis, eight clusters were extracted, such as product quality, shipping, store service, delivery service, product function, cost performance, taste, and brand. The

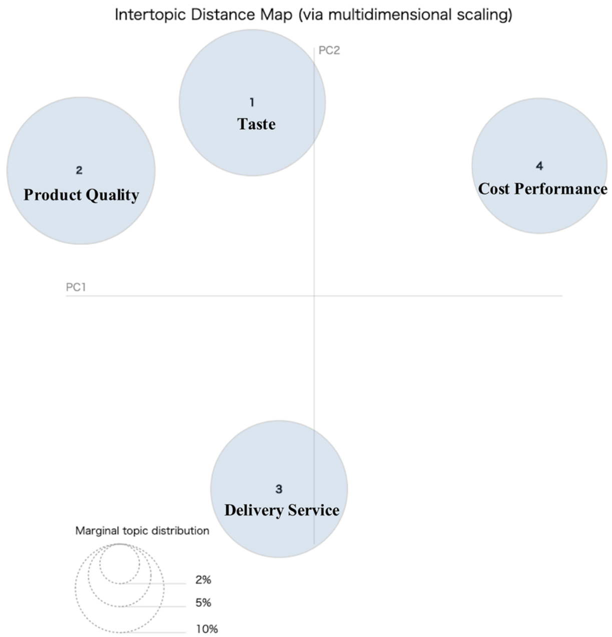

Fig. 8 Intertopic distance map

keywords for each cluster obtained were listed and the sizes and proportions of the eight clusters appearing in the negative customer reviews for each product were calculated (see Eq. 1). Meanwhile, the results at two different time periods are compared to ensure the robustness of the study results. Figure 6 shows the proportions of the aforementioned eight clusters for all 39 items at the time before and after 2018. Based on the results of the cluster sizes, four common topics for all the items were extracted, namely product quality, delivery service, cost performance, and taste (i.e., clusters 1, 4, 6, and 7).

Furthermore, to confirm the objectiveness and accuracy of the results for the four common topics, the present study adopted an LDA-based topic modeling to identify topics in customers' online negative reviews [57, 58]. Figures 7 and 8 provide the results of topic identification and intertopic distance map using LDA, which indicate the four common topics identified that are reliably proved and verified in a certain extent. Next, the details of the statistical estimation are elucidated on how much these four common topics affect the sales outcomes.

4.2 Sales Models

After capturing the common topics through text analysis, statistical models were proposed to analyze how the topics of negative customer reviews impact sales. The dataset was divided into two parts: training and test sets. Training set includes the data with sales outcomes and detailed original customer reviews before 2018, which

Table 2 Correlation matrix of model (A)

		(1)	(2)	(3)	(4)
(1)	x_1	1.00			
(2)	x_2	– 0.11	1.00		
(3)	x_3	– 0.03	– 0.01	1.00	
(4)	N/R	– 0.27	0.16	0.01	1.00

has in total 184,700 samples, while the test set included those after 2018, which has 181,730 samples.

A traditional approach to study the effects of negative reviews on sales without using textual topics of customer reviews is usually to include the variables of quantitative surrogates (i.e., volume and valence) and examine the relation between the proportion of negative reviews and sales. However, earlier studies present inconsistent results. To confirm this relationship, an approach called model (A) is used as a benchmark. Given that the independent variables in the study dataset are multidimensional and measured over time, the model is developed as follows, in which the time span and composite error term are introduced:

$$y_{i,t} = alpha_0 + \alpha_1 x_{1i,t} + \alpha_2 x_{2i,t} + \alpha_3 x_{3i,t} + \theta \frac{N_{i,t}}{R_{i,t}} + u_{i,t}, \tag{4}$$

$$u_{i,t} = w_i + v_{i,t}, \tag{5}$$

where i is the individual dimension and t is the time dimension, which identify product and period, respectively; y is the natural logarithm of sales; x_1 is the natural logarithm of price; x_2 is the log of accumulated number of overall reviews; and x_3 represents the degree of praise. $N_{i,t}$ and $R_{i,t}$ are defined the same as previously. The error term $u_{i,t}$ is assumed to be i.i.d. normally distributed and formed by two parts: w_i and $v_{i,t}$.[6] w_i is individual-specific and time-invariant effects, while $v_{i,t}$ represents the error that is dependent on both the individual and time. In addition, θ in this model is the core parameter capturing the relation between the proportion of negative reviews and sales, while the other parameters $\alpha_1, \alpha_2, \alpha_3$ relate to the impacts of control variables.

Literature related to the investigation of sales effect appended the preceding sales outcomes into prediction model [27, 59]. However, because the accurate sales are unobservable, we processed the volume of online customer reviews as the surrogate of sales. Therefore, the present study only uses the total number of overall reviews (i.e., $x_{2i,t} = \ln R_{i,t}$) as the independent variables without adding the preceding sales (i.e., $S_{i,t-1} = R_{i,t} - R_{i,t-1}$) to avoid the issue of multicollinearity,[7] which leads to

[6] In order to make the model developed from the training set reasonably applicable to test set, this study assumes that the data are independent identically distributed. What's more, the present study mainly highlights the sales effects of negative online reviews rather than sales forecast. Therefore, in the present study, the error term is assumed as i.i.d. distributed to simplify the model.

[7] The correlation coefficient between $x_{2i,t}$ and $S_{i,t-1}$ is 0.794.

Table 3 Parameter estimates of model (A)

	Before 2018			After 2018	
	Estimate	SE		Estimate	SE
α_1	0.233	0.173	α_1	0.733**	0.293
α_2	0.877***	0.089	α_2	0.833***	0.117
α_3	0.237	0.235	α_3	0.429	0.484
θ	− 205.97***	30.109	θ	– 106.91*	57.154
α_0	− 2.275*	1.330	α_0	– 4.949*	2.745
No. of groups	39		No. of groups	39	
R^2	0.627		R^2	0.459	
AIC	665.81		AIC	515.97	

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

inconsistent parameter estimates. In addition, before studying the impact of online customer reviews on sales, a correlation matrix is drawn to briefly confirm the multicollinearity among the explanatory variables. The results of correlation coefficients are displayed in Table 2.

The model describes a case where no lag of explained variable is used as regressor; therefore, the estimation result is compared with fixed effects (FE) and random effects (RE). In the present study, a Durbin–Wu–Hausman test was used to differentiate between FE and RE models. This test verifies whether there is a correlation between the unique errors and the regressors in the model. The null hypothesis here is that there is no correlation between the two, and the preferred model is RE, while the alternative hypothesis is that the model is FE.[8] The testing result indicated that the RE model is preferred ($\chi^2(1) = 1.15$, $p = 0.28$) to obtain an accurate assessment of the effect of negative online product customer reviews on sales outcomes. Table 3 (left) reports estimation results of model (A) with the data before 2018. This result indicates that after controlling for the impacts of volume, valence, and regular price, a significant and negative estimate was obtained, which relates to the effect of negative review on sales at the 0.01 level (i.e., 99% credible interval does not cover zero). Parameter θ represents this effect. In addition, only the estimate of the parameter α_2 is significant. The coefficient of determination for the overall training data is 0.627, and Akaike's information criterion (AIC) is 665.81. The value of sales is predicted using these estimation results and compared with the actual sales outcome after 2018. Figure 9 shows an example of the comparison results. The horizontal axis in the figure represents the aggregate of overall customer reviews, while the vertical axis indicates the sales. As shown in Fig. 9, the degree of fit was not encouraging. Next, the parameter was analyzed using the test set. The estimation results with the

[8] Random effects (RE) model is preferred under the null hypothesis because of higher efficiency, while under the alternative fixed effects (FE) model is at least consistent and thus preferred.

test set appear in Table 3 (right). Although the directions of all the parameters were estimated as earlier, there is a significant change in the estimate and asymptotic significance, especially for parameter θ.

To improve the model and the accuracy of prediction, a new model is proposed. This model includes an analysis of the textual content of the reviews to capture the latent topics and then introduces the indicators for those latent topics instead of the single proportion. Therefore, the present study progresses to model (B), in which the proportions and consistency of the four topics obtained from the text analysis earlier are introduced as the new indicators. Model (B) is specified as follows:

$$y_{i,t} = \beta_0 + \beta_1 x_{1i,t} + \beta_2 x_{2i,t} + \beta_3 x_{3i,t} + \sum \gamma_k p_{k,i,t} + \delta C_{i,t} + e_{i,t}, \tag{6}$$

where all the variables are defined the same as previously defined. γ is a $(k \times 1)$ vector, which captures the effects of topic proportions on sales outcomes. Parameter δ reflects the impact of topic consistency on sales. To avoid the multicollinearity, the

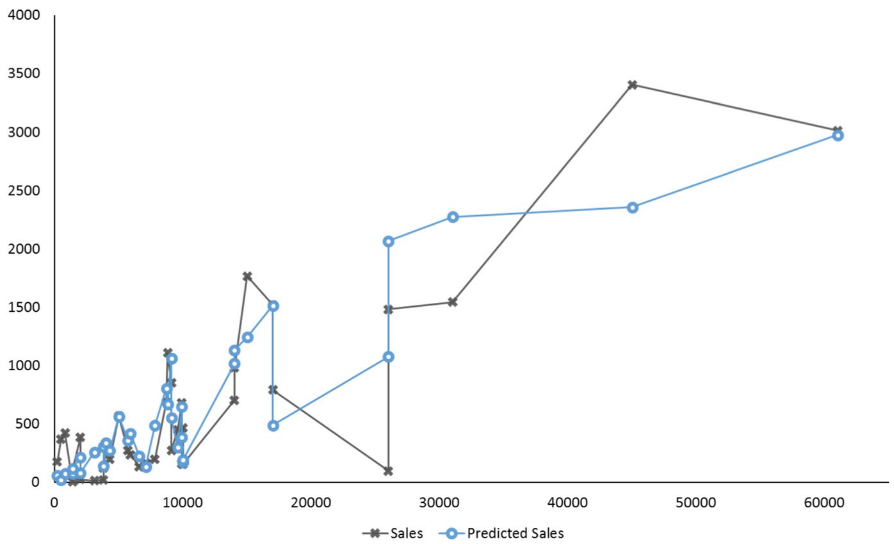

Fig. 9 Predicting outcomes of sales (1)

Table 4 Correlation matrix of model (B)

		(1)	(2)	(3)	(4)	(5)
(1)	$p_{k=1}$	1.00				
(2)	$p_{k=2}$	0.18	1.00			
(3)	$p_{k=3}$	0.12	− 0.09	1.00		
(4)	$p_{k=4}$	− 0.30	0.31	− 0.17	1.00	
(5)	C	0.07	0.38	0.08	0.35	1.00

Note: Four topics, in turn, are product quality, delivery service, cost performance, and taste correlation among the variables of proportions and consistency was studied. Table 4 shows the matrix of the correlation.

Table 5 (left) shows the estimation results of model (B) using a training set. Consistent with the estimation results in model (A), the estimation result of β_2 is significant, proving the existence of a bandwagon effect in online shopping. Hypothesis 1 (H1) predicts that topics in a higher number of negative customer reviews have a stronger negative effect on sales. As shown in the table, the estimates of the parameters from γ_1 to γ_4 are negative, and all of them are significant at least at the 0.05 level. Therefore, H1 is supported. In particular, as the aforementioned assumption, different topics have different effects on sales. In particular, for snacks in the study dataset, taste is the most detrimental to sales, which is consistent with the studies by He and Bond [52] that similarity in taste plays an important role for the effects of word of mouth on sales and other related outcomes. Product quality and cost performance almost have the same effect on product sales. In comparison, customers have a highest level of tolerance to delivery service.

In Hypothesis 2 (H2), consistency of topics in negative customer reviews is assumed to have a negative effect on product sales. This hypothesis can be verified by parameter δ. The estimate of δ is negative and significant at the 0.1 level. Therefore, H2 is supported.

In particular, the coefficient of determination for the overall training data in model (B) is 0.883, and AIC here is only 227.37. The accuracy of approach adopted in the present study is well illustrated by the change of these two indicators. In addition, Fig. 10 provides the plots of the predicted and actual sales value. The result of this comparison indicates that the fitting degree of the proposed model is better than that of the traditional model. Similarly, the parameter estimates using the data after 2018 are listed in Table 5 (right).

Table 5 Parameter Estimates of Model (B)

	Before 2018			After 2018	
	Estimate	Std.Err		Estimate	Std.Err
β_1	0.436	0.284	β_1	0.451	0.295
β_2	0.780***	0.034	β_2	0.845***	0.044
β_3	0.308***	0.097	β_3	0.429***	0.131
γ_1	– 138.144**	62.415	γ_1	– 135.155*	72.003
γ_2	– 84.927***	30.124	γ_2	– 67.361**	31.604
γ_3	– 128.936**	58.996	γ_3	– 139.062*	74.497
γ_4	– 254.338**	110.916	γ_4	– 331.923**	160.115
δ	– 29.613*	15.235	δ	– 26.054**	11.426
β_0	– 2.749	1.680	β_0	– 3.819*	2.214
No. of groups	39			No. of groups	39
R^2	0.883			R^2	0.833
AIC	227.37			AIC	52.92

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

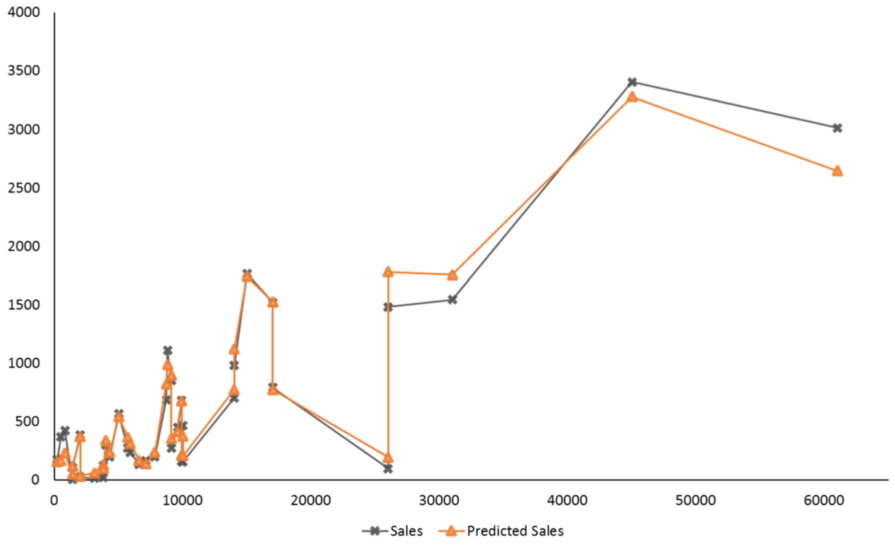

Fig. 10 Predicting outcomes of sales (2)

5 Conclusions

While negative customer reviews have attracted significant academic attention in the marketing field, there is lack of information on why negative customer reviews work at the group level. The present study reveals how negative customer reviews work as a whole to influence product sales by exploring the latent topics about these negative reviews by applying several text data mining techniques.

5.1 The Main Findings

First, a co-occurrence network was examined based on high-frequency keywords to macroscopically analyze the knowledge structure of negative customer reviews. The results show that negative customer reviews mainly focus on six topics, such as store service, product quality, cost performance, delivery service, packaging, and taste.

Furthermore, a self-organizing map and a hierarchical clustering analysis were employed to identify the specific topics underlying the textual content of negative customer reviews. Satisfyingly, there is a similar analysis result in both self-organizing map and hierarchical clustering analysis in which a total of eight clusters were found, that is, product quality, shipping, store service, delivery service, product function, cost performance, taste, and brand. After comparing the proportion of each cluster in online product reviews at two different time periods, four common topics, that is product quality, delivery service, cost performance, and taste, were determined. In addition, the finding of these four common topics is also supported by the LDA-based topic modeling.

Moreover, a new econometric model including the four topics to estimate the effects of negative customer reviews on product sales was proposed. Topics with greater negative reviews exert a stronger negative impact on sales outcomes, while each topic is found to influence sales outcomes with a different effect. In particular, taste accounts for the largest proportion, which exerts most impact on sales outcomes. Topics of product quality and cost performance have an almost equal effect, while the influence of delivery service is weakest. Furthermore, as assumed, negative customer reviews with a higher level of consistency regarding these four topics significantly jeopardizes product sales. In other words, the study findings can have potential important contributions at the theoretical and managerial level alike.

5.2 Theoretical Contributions

Although the research on customer reviews is very popular in marketing literature, there are relatively few studies about the textual content of reviews. This study has two academic contributions. First, text data mining techniques help in finding the latent topics of negative customer reviews and reveal the underlying mechanism through which negative customer reviews function and the extent to which they help explain product sales at the level of group. Although there are several factors of negative customer reviews that affect sales outcomes, the present study suggests that topic proportion and consistency may be important considerations. The study findings give other researchers a new perspective to interpret and utilize negative customer reviews.

Second, the study contributes methodologically by showing how the results of text data mining techniques can be reliably applied to an econometric model to predict the sales outcomes. The present study shows that including textual attributes of negative customer reviews in the proposed forecasting model significantly improves the forecasting performance, but not in the benchmark model.

5.3 Managerial Implications

Having recognized the excessively detrimental effect of negative customer reviews on brand equity or product sales, retailers or customer services marketers face challenges to draw various marketing strategies based on modern technological tools. Several managerial implications can be drawn from the findings of this research. First, this study suggests that retailers or customer services should provide smart information about customer reviews. To help customers process numerous reviews, more and more platforms provide topics underlying customer reviews. When the consistency of topics in reviews is very high, it may be a bad idea for retailers or customer services to provide topic labels that are detrimental for sales outcomes.

Second, as the negative customer review can be a useful measure to help forecast product sales, marketers should monitor negative customer reviews as a part of their routine work. Good forecasts can help marketers perceive the potential problems or customers' discontent and respond to negative customer reviews precisely. Different group attributes of negative customer reviews have different impacts on product sales, so marketers should take different actions. The proportion directly indicates the extent to which these topics differ. The topic with a largest proportion represents customers' biggest concern. It also indicates that there might be some problems with the product. The level of consistency of latent topics in the negative customer reviews plays a similar role. Marketers should focus more on topics with high proportion or consistency and take effective action to weaken the consequent negative effect.

5.4 Limitations and Future Research

As with any study, there are several limitations that present opportunities for future research. First, although the study data included about 39 kinds of products, these products are categorized as experience goods. It is well known that customers process the information about experience or search goods quite differently. Experiential customer reviews are likely to be subjective and emotional and reflective of personal taste, while reviews of search goods are likely to be objective and factual and reflective of product functionality. Future studies could sample both experience and search goods in order to check the generalizability of the study findings. Second, this study conceptualizes topic proportion and consistency independently, while there is likely to be interaction between these two variables. For example, customers are less likely to adopt reviews with high consistency but with very low proportion. There is a lack of research on the interplay of these two variables, and it is likely to be an interesting stream for future research. Moreover, sales data were not forthcoming in this study. If possible, future studies should use such data to analyze the direct effect of negative customer reviews. Finally, the data used in the present study are from only one platform, so the impact of the platform itself cannot be ruled out. In the future, researchers could expand data sources and compare the similarities and differences between different platforms.

Acknowledgements We thank the editor and the anonymous reviewers for their thoughtful reviews and constructive suggestions during the review process. This work was supported by JSPS KAKENHI Grant numbers JP15H06747, JP17K18152, and the National Natural Science Foundation of China NSFC71572065.

References

1. Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, *49*(10), 1407–1424.
2. Godes, D., & Mayzlin, D. (2004). Using online conversations to study word of mouth communication. *Marketing Science*, *23*(4), 545–560.
3. Chu, S. C., & Kim, Y. (2011). Determinants of consumer engagement in electronic word-of-mouth (eWOM) in social networking sites. *International Journal of Advertising*, *30*(1), 47–75.
4. Doh, S. J., & Hwang, J. S. (2009). How consumers evaluate eWOM (electronic word-of-mouth) messages. *CyberPsychology & Behavior*, *12*(2), 193–197.
5. Hennig-Thurau, T., Walsh, G., & Walsh, G. (2003). Electronic word of mouth: Motives for and consequences of reading customer articulations on the internet. *International Journal of Electronic Commerce*, *8*(2), 51–74.
6. Duan, W., Gu, B., & Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *Journal of Retailing*, *84*(2), 233–242.

7. Floyd, K., Freling, R., Alhoqail, S., Cho, H. Y., & Freling, T. (2014). How online product reviews affect retail sales: A meta-analysis. *Journal of Retailing*, *90*(2), 217–232.
8. Gu, B., Park, J., & Konana, P. (2012). The impact of external word-of-mouth sources on retailer sales of high-involvement products. *Information Systems Research*, *23*(1), 182–196.
9. 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.
10. Dellarocas, C. (2006). Strategic manipulation of internet opinion forums: Implications for consumers and firms. *Management Science*, *52*(10), 1577–1593.
11. Chiou, J. S., & Cheng, C. (2003). Should a company have message boards on its web sites? *Journal of Interactive Marketing*, *17*(3), 50–61.
12. Fowler, G. A., & Avila, J. D. (2009). On the internet, everyone's a critic but they're not very critical. *The Wall Street Journal*.
13. Chatterjee, P. (2001). Online reviews: Do consumers use them? *Advances in Consumer Research*, *28*, 129–133.
14. Kusumasondjaja, S., Shanka, T., & Marchegiani, C. (2012). Credibility of online reviews and initial trust: The roles of reviewer's identity and review valence. *Journal of Vacation Marketing*, *18*(3), 185–195.
15. Papathanassis, A., & Knolle, F. (2011). Exploring the adoption and processing of online holiday reviews: A grounded theory approach. *Tourism Management*, *32*(2), 215–224.
16. Zhang, J. Q., Craciun, G., & Shin, D. (2010). When does electronic word-of-mouth matter? A study of consumer product reviews. *Journal of Business Research*, *63*(12), 1336–1341.
17. 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.
18. Chen, Y., Wang, Q., & Xie, J. (2011). Online social interactions: A natural experiment on word of mouth versus observational learning. *Journal of Marketing Research*, *48*(2), 238–254.
19. Ho-Dac, N. N., Carson, S. J., & Moore, W. L. (2013). The effects of positive and negative online customer reviews: Do brand strength and category maturity matter? *Journal of Marketing*, *77*(6), 37–53.
20. Rosario, A Babić, Sotgiu, F., De Valck, K., & Bijmolt, T. H. (2016). The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *Journal of Marketing Research*, *53*(3), 297–318.
21. Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing*, *70*(3), 74–89.
22. Tirunillai, S., & Tellis, G. J. (2012). Does chatter really matter? Dynamics of user-generated content and stock performance. *Marketing Science*, *31*(2), 198–215.
23. Li, X., & Hitt, L. M. (2008). Self-selection and information role of online product reviews. *Information Systems Research*, *19*(4), 456–474.
24. Ogut, H., & Tas, B. K. Onur. (2012). The influence of internet customer reviews on the online sales and prices in hotel industry. *The Service Industries Journal*, *32*(2), 197–214.
25. Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behavior*, *27*(2), 634–639.
26. Cui, G., Lui, H. K., & Guo, X. (2012). The effect of online consumer reviews on new product sales. *International Journal of Electronic Commerce*, *17*(1), 39–58.
27. Li, Z., & Shimizu, A. (2018). Impact of online customer reviews on sales outcomes: An empirical study based on prospect theory. *The Review of Socionetwork Strategies*, *12*(2), 135–151.
28. Yang, J., Kim, W., Amblee, N., & Jeong, J. (2012). The heterogeneous effect of WOM on product sales: Why the effect of WOM valence is mixed? *European Journal of Marketing*, *46*(11/12), 1523–1538.
29. Zhang, Z., Li, X., & Chen, Y. (2012). Deciphering word-of-mouth in social media: Text-based metrics of consumer reviews. *ACM Transactions on Management Information Systems*, *3*(1), 1–23.
30. Sonnier, G. P., McAlister, L., & Rutz, O. J. (2011). A dynamic model of the effect of online communications on firm sales. *Marketing Science*, *30*(4), 702–716.
31. Cho, V., & Chan, A. (2017). A study on the influence of eWOM using content analysis: How do comments on value for money, product sophistication and experiential feeling affect our choices? *Enterprise Information Systems*, *11*(6), 927–948.
32. Tirunillai, S., & Tellis, G. J. (2014). Mining marketing meaning from online chatter: Strategic brand analysis of big data using Latent Dirichlet allocation. *Journal of Marketing Research*, *51*(4), 463–479.
33. Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370.

34. Chen, P.Y., Dhanasobhon, S., & Smith, M. D. (2008). All reviews are not created equal: The disaggregate impact of reviews and reviewers at amazon. com. Working Paper. https://doi.org/10.2139/ssrn.918083.
35. Gruen, T. W., Osmonbekov, T., & Czaplewski, A. J. (2006). eWOM: The impact of customer-to-customer online know-how exchange on customer value and loyalty. *Journal of Business Research*, *59*(4), 449–456.
36. King, R. A., Racherla, P., & Bush, V. D. (2014). What we know and don't know about online word-of-mouth: A review and synthesis of the literature. *Journal of Interactive Marketing*, *28*(3), 167–183.
37. 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.
38. Dowling, G. R., & Staelin, R. (1994). A model of perceived risk and intended risk-handling activity. *Journal of Consumer Research*, *21*(1), 119–134.
39. Taylor, J. W. (1974). The role of risk in consumer behavior. *The Journal of Marketing*, *38*(2), 54–60.
40. Stewart, K. J. (2003). Trust transfer on the world wide web. *Organization Science*, *14*(1), 5–17.
41. Girard, T., & Dion, P. (2010). Validating the search, experience, and credence product classification framework. *Journal of Business Research*, *63*(9), 1079–1087.
42. Ye, Q., Xu, M., Kiang, M., Wu, W., & Sun, F. (2013). In-depth analysis of the seller reputation and price premium relationship: A comparison between eBay US and TaoBao China. *Journal of Electronic Commerce Research*, *14*(1), 1–10.
43. Liao, T. H., & Keng, C. J. (2013). Online shopping delivery delay: Finding a psychological recovery strategy by online consumer experiences. *Computers in Human Behavior*, *29*(4), 1849–1861.
44. Forsythe, S. M., & Shi, B. (2003). Consumer patronage and risk perceptions in internet shopping. *Journal of Business Research*, *56*(11), 867–875.
45. Wood, W., Lundgren, S., Ouellette, J. A., Busceme, S., & Blackstone, T. (1994). Minority influence: A meta-analytic review of social influence processes. *Psychological Bulletin*, *115*(3), 323–345.
46. Latané, B., & Wolf, S. (1981). The social impact of majorities and minorities. *Psychological Review*, *88*(5), 438–453.
47. Tanford, S., & Penrod, S. (1984). Social influence model: A formal integration of research on majority and minority influence processes. *Psychological Bulletin*, *95*(2), 189–225.
48. Moscovici, S., & Personnaz, B. (1980). Studies in social influence: V. Minority influence and conversion behavior in a perceptual task. *Journal of Experimental Social Psychology*, *16*(3), 270–282.
49. Moscovici, S. (1985). *Social Influence and Conformity*. Handbook of Social Psychology.
50. Turner, J. C. (1991). *Social Influence*. Pacific Grove: Thomson Brooks/Cole Publishing Co.
51. Nemeth, M. C., Swedlund, M., & Kanki, B. (1974). Patterning of the minority's response and their influence on the majority. *European Journal of Social Psychology*, *4*(1), 53–64.
52. He, S. X., & Bond, S. D. (2015). Why is the crowd divided? Attribution for dispersion in online word of mouth. *Journal of Consumer Research*, *41*(6), 1509–1527.
53. Lee, C. H., & Cranage, D. A. (2014). Toward understanding consumer processing of negative online word-of-mouth communication: The roles of opinion consensus and organizational response strategies. *Journal of Hospitality & Tourism Research*, *38*(3), 330–360.
54. Chang, H. H., & Wu, L. H. (2014). An examination of negative e-WOM adoption: Brand commitment as a moderator. *Decision Support Systems*, *59*, 206–218.
55. Zhu, M., & Lai, S. (2009). A study about the WOM influence on tourism destination choice. In *2009 International conference on electronic commerce and business intelligence*, pp. 120–124.
56. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.
57. Chuang, J., Manning, C. D., & Heer, J. (2012). Termite: Visualization techniques for assessing textual topic models. In *Proceedings of the international working conference on advanced visual interfaces* (pp. 74–77). ACM.
58. Sievert, C., & Shirley, K. (2014). LDAvis: A method for visualizing and interpreting topics. In *Proceedings of the workshop on interactive language learning, visualization, and interfaces*, pp. 63–70.
59. Zhang, J., Wedel, M., & Pieters, R. (2009). Sales effects of attention to feature advertisements: A Bayesian mediation analysis. *Journal of Marketing Research*, *46*(5), 669–681.