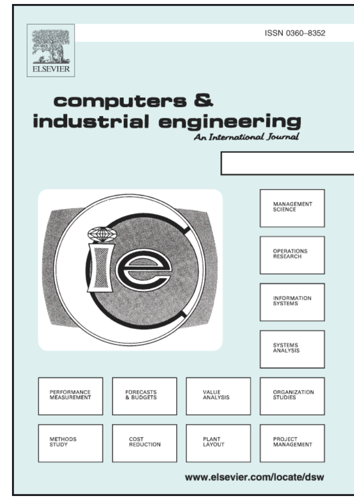
Accepted Manuscript



Predicting online e-marketplace sales performances: A big data approach

Boying Li, Eugene Ch'ng, Alain Yee-Loong Chong, Haijun Bao

PII: S0360-8352(16)30275-3
DOI: http://dx.doi.org/10.1016/j.cie.2016.08.009
Reference: CAIE 4437

To appear in: *Computers & Industrial Engineering*

Predicting online e-marketplace sales performances: A big data approach

Boying Li
Nottingham University Business School,
International Doctoral Innovation Centre,
University of Nottingham Ningbo China
199 Taikang East Road, Ningbo, 315100, China
Email: Boying.Li@nottingham.edu.cn

Eugene Ch'ng
School of Computer Science,
University of Nottingham Ningbo China
199 Taikang East Road, Ningbo, 315100, China
Email: Eugene.Chng@nottingham.edu.cn
Telephone: +86 (0) 574 88183049

Alain Yee-Loong Chong
Nottingham University Business School,
University of Nottingham Ningbo China
199 Taikang East Road, Ningbo, 315100, China
Email: Alain.Chong@nottingham.edu.cn
Telephone: +86 (0) 574 88180587

Haijun Bao
School of Urban-Rural Planning and Management
Zhejiang University of Finance & Economics
No. 18,Xueyuan Street, Hangzhou, 310018, China
Email: Baohaijun@zufe.edu.cn

Predicting online e-marketplace sales performances: A big data approach

Abstract

To manage supply chain efficiently, e-business organizations need to understand their sales effectively. Previous research has shown that product review plays an important role in influencing sales performance, especially review volume and rating. However, limited attention has been paid to understand how other factors moderate the effect of product review on online sales. This study aims to confirm the importance of review volume and rating on improving sales performance, and further examine the moderating roles of product category, answered questions, discount and review usefulness in such relationships. By analyzing 2,939 records of data extracted from Amazon.com using a big data architecture, it is found that review volume and rating have stronger influence on sales rank for search product than for experience product. Also, review usefulness significantly moderates the effects of review volume and rating on product sales rank. In addition, the relationship between review volume and sales rank is significantly moderated by both answered questions and discount. However, answered questions and discount do not have significant moderation effect on the relationship between review rating and sales rank. The findings expand previous literature by confirming important interactions between customer review features and other factors, and the findings provide practical guidelines to manage e-businesses. This study also explains a big data architecture and illustrates the use of big data technologies in testing theoretical framework.

Keywords

E-business, product reviews, moderation effect, big data architecture

Introduction

E-business is a crucial part of today's economy. Customers highly rely on e-business marketplaces for their daily purchases and organizations are competing to survive in e-business environment (Wu et al., 2014; Zhuang and Lederer, 2006). While the Internet and computer technologies allow organizations to reach their customers and run their businesses easily, the competitive environment also posts many challenges. One of the most important challenges in e-business operations is to understand sales performance so that the organization can manage its supply chain efficiently (Chong et al. 2015). With a good understanding of customer purchase decision and product sales performance, organizations can better arrange the procurement and develop strategies to improve sales, which may become the competitive advantage of the organization (Chong and Zhou, 2014).

Customer review is important to customers purchase decisions. Online customer reviews are perceived to be highly trustworthy by customers- ranked as the second most reliable review format (Chong, et al., 2015; Cheung, et al., 2008). Considering the importance of customer reviews on shaping customer purchase decision, the role of customer review in facilitating the understanding and management to e-business's sales should be given enough attention. Thus, in order to understand sales, it is crucial for organizations to understand how customer reviews influence sales. Previous studies identified review volume, review rating or valence, review helpfulness and review word count as representative features of customer review (e.g. Mudambi and Schuff, 2010; Ye et al., 2011). However, apart from the customer reviews, there are other important factors in e-business websites. For example, organizations can offer discounts and respond to customer's questions (Chong et al., 2015). The differences of product nature can also cause the fluctuation of sales.

The main objective of this study is to investigate the factors that moderate the influences of customer review on sales performance. Although previous research has studied various factors that influence online sales, most of the studies focused on testing the predictive powers of a list of factors (Chong et al., 2015, Marshall and Leng, 2002). There is a lack of research on interactions between different types of factors. This study thus is interested in examining whether other factors would moderate customer review's effect on sales performance. Review volume and rating are the two most recognized features of customer review (Berger et al., 2010; Forman et al., 2008). Therefore, volume and rating are the target variables for main effects, and four other factors are proposed as the potential moderators, specifically: product category, answered questions, discount value and helpful votes. Product category reflects the different natures between products (Cui et al. 2012), and answered questions provide a new dimension of user-generated content where customers can interact with each other or with the seller. Discount value is a type of one-way marketing strategy, and has been widely used to stimulate sales (Marshall and Leng, 2002). Helpful votes, although is also a feature of customer reviews, does not represent customer evaluations of products but instead indicate the review information usefulness. It would be interesting to test whether and how the interactions between customer review and those four variables from different dimensions would contribute to sales.

Traditional way of studying how certain factors influence sales or purchase behavior mainly uses historical sales data, test market data or survey data (e.g. Berbegal-Mirabent et al., 2016). Different from traditional approaches, this study applies a big data architecture to collect data. Big data technologies can extract objective data from reliable data source both in real time or from historical record. The data are of large volume, rapidly updated and of various types, and can provide valid information for understanding sales and corresponding supply chain management in online marketplaces (Chen et al., 2015). With our big data architecture, this study extracts data from product pages of Amazon.com, and use the data to test proposed hypotheses.

This study has made several contributions to the literature theoretically, methodologically and practically. Theoretically, although previous studies have widely accepted that online customer review volume and rating influence product sales significantly, very limited research has been done to investigate the factors that moderate such effects. This study examines the interactions between factors related to user-generated content, information usefulness, promotional marketing strategy and product nature. Methodologically, while many previous studies use questionnaires and surveys to study factors influencing sales, this study makes use of big data architecture to extract real data from online marketplaces. In this perspective, this study adopts big data technologies to facilitate theoretical research, illustrating a new approach to empirically test theories and models. Practically, the findings of this study can help practitioners to understand online customers and sales and perhaps stimulate sales.

In the next section, theories related to this study are reviewed. In section 3, we propose a research model that investigates the moderating roles of product category, answered questions, discount value and helpful votes. After that, the research methodology, including a detailed description of our big data architecture, is presented, followed by the results of analysis. Finally, we discuss about our findings, contributions and potential implications, and also comment on our limitations and future work directions.

Theory Review

Online customer review and related theories

Online customer review refers to the peer-generated evaluations about an organization's offerings posted online (Mudambi and Schuff, 2010). Previous research has studied the reviews on a company's website or an e-commerce platform and on third party online recommendation websites. Apart from product information provided by the sellers, prospective online customers can also refer to the customer reviews to make their purchase decisions.

The importance of online customer reviews can be explained with the theory of information economics. Prospective online customers often lack the information to evaluate the quality of a company's offering. Seeking additional information can be costly for customers in terms of the money and time, and thus customers must face the trade-off between their perceived costs and benefits from searching (Mudambi and Schuff, 2010). Giving more efforts in physical searching and cognitive processing, customers tend to increase the accuracy of their decisions. Online customer reviews can serve as an important source of information. The numerical indexes of customer reviews, including number of reviews and ratings, can contribute to the cognitive process and improve purchase decision accuracy while not increase the efforts too much (Poston and Speier 2005).

Apart from the theory of information economics, theories of social influence can also help understand the important roles of online reviews. Rooted in social psychology, theories of social influence argue that people can be influenced by others (Baek et al., 2015). Social influence can be divided into two processes, informational social influence where influence comes from informative evidence, and normative social influence where influence comes from the conformity (Deutsch and Gerard, 1955). According to Sridhar and Srinivasan (2012), social influence can come from peoples that the person does not know. Therefore, in the context of online marketplaces, the decision of consumers may conform to other people's opinions. While review rating and content provide informative materials, helpful votes of reviews reflect the consumer group norm. Customers' opinions are affected by their peers via these features related to online review. Therefore, via social influence, review features shape customer purchase decisions and thus influence the sales performance of products.

Product categories

The search costs can be different across product categories. A simple yet widely accepted categorization of products is based on the information source that consumers rely on to evaluate product quality (Nelson 1970). Such categorization classifies products into search product and experience product. Search products, such as mobile phones and cameras, can be evaluated by search (Bei et al. 2004; Jiménez and Mendoza, 2013). For them, information about the quality can be obtained easily before making the consumption. Experience products, however, can be found costly to gather relevant information and are difficult to be evaluated until they are consumed (Mudambi and Schuff 2010). Thus, assessing experience product's quality highly relies on trials and experience.

Customers post reviews on both search and experience products and the reviews are used to support the purchase decisions on both product categories as well. Nevertheless, online customer reviews play different roles in making purchases of search and experience products. The perceived quality of search products depends on the objective attributes, while the perceived quality of experience products is based more on subjective and personal judgements. Therefore, for search products, information from the seller usually plays an important role in influencing consumer decisions. For experience products, the seller is not the most important information source for consumers; instead, purchase decision is made highly based on prior purchase experiences of other consumers. Therefore, when making purchase decisions on experience products, consumers tend to value the recommendations and word of mouth from others more than when they intend to buy search products.

Research Model and Hypotheses

Studies have illustrated that customer reviews are important in online marketplaces (Chevalier and Mayzlin 2006). Review volume and rating are two most widely accepted factors influencing sales (Clemons et al 2006). However, what are the factors that influence these relationships? This study identifies the factors that moderate the effect of online review volume and rating on online sales. Considering the differences in information-related natures between search and experience products, this study proposes that product category is one of the moderators. Taking common promotional marketing strategies into consideration, discount value is also included as a moderator. Numbers of answered questions and helpful votes are two other moderators proposed in this study. Our model is shown in Figure 1 and these factors and relationships are explained below.

Review volume and product sales

Online customer review volume is the total number of online customer reviews for an organization’s offerings. As existing research suggests, online review volume has a significant influence on online sales (Cheung and Thadani 2012). Large number of customer reviews is seen as a sign of publicity, attracting customer’s attention towards the product (Cheung and Thadani 2012). The awareness of product can further contribute to sales performance (Duan et al. 2008). Moreover, the high volume can also be seen as a cue for popularity, which often being associated with high product quality and trustworthy seller by consumers, and thus can lead to increase in sales. Empirical findings also support the importance of review volume for sales performance. Online customer review volume has been found to cause increase in restaurant sales (Lu et al. 2013) and online book sales (Chevalier and Mayzlin, 2006).

Based on both theoretical arguments and empirical evidences, it is proposed that:

H1: Online review volume influences online sales performance.

Overall review rating and product sales

Online customer review rating is usually the numerical rating that summarizes the level of positivity or negativity. Overall review rating reflects the general evaluation from previous reviewers. Numerical star rating ranging from one to five is the most common rating mechanism, where one to five stars representing a continuum from negative to positive evaluations of the product (Mudambi and Schuff 2010). The overall review rating can be seen as a representation of crowd intelligence. Because it is a highly summarized evaluation from different reviewers, the valence of the rating provides prospective customers important cognitive information. Positive rating reflects the positive attitude towards the product by the crowd, which could lead to purchase decisions that drive up the product sales. Moreover, overall review rating may not only influence the cognitive process of rational consumers; it can also influence irrational buyers with herding effect. The theory of herding suggests that people are likely to believe what others believe (Huang and Chen 2006). As the overall review rating represents the attitudes of others, when herding effect occurs, prospective customer tends to follow that attitude for decisions making. A positive rating can thus lead to high product sales. Previous research found that positive rating has positive influence on increasing product sales.

H2: Overall review rating influences online sales performance

Moderating role of product category

Product category moderates the relationship between online customer review volume and product sales performance. Online review volume has different influences on sales performances for experience products and search products (Cui et al. 2012). Volume itself provides relatively less information regarding the objective attributes of product; instead, it to some extent reflects the popularity of product. While consumers value objective features when making purchase decisions on search products, popularity is an extrinsic attribute that customers often rely on when evaluating experience products (Cui et al. 2012). Therefore, online review volume is proposed to have different influences on online sales performances for search and experience products. To be more specific, this study hypothesizes that:

H3a: Product category moderates the effect of online review volume on online sales performance.

Compared to search products, for experience products, the value of additional information in online customer reviews can make more contributions to purchase decision making (Cui et al. 2012). The overall customer review rating reflects the actual evaluations of product quality after previous customers experience the product. However, the evaluation of experience products is highly subjective and based on personal tastes. Instead, the assessment of search product is objective. Because the judgement of search product is more objective than experience product, the evaluation can follow a set of criteria that are more uniform, and the evaluation results are more consensus. Therefore, compared to that of experience products, the overall review rating of search products provides more valuable and consensus evaluation that contributes to purchase decisions. Based on the above, this study proposes that:

H3b: Product category moderates the effect of online overall review rating on online sales performance.

Moderating role of number of answered questions

Apart from online reviews where customers post their evaluations and experiences after making the purchase, in Amazon.com, customers also seek information before the purchase by posting questions. Both customers and the seller can offer answers to the questions, allowing customers to communicate with the sellers and interact with each other beyond reading reviews (Chong et al., 2015). Everyone who views the product page can have access to the questions and answers. The number of answered questions represents the total number of questions posted by customers with answers from sellers or/and customers. This feature reflects the level of interactions among customers or between customers and the seller (Chong et al., 2015), and it also provides additional information which prospective customers are interested in.

The number of answered questions reflects the demand for information. If customers post a large number of questions and seek answers very proactively, the product is likely to be a high involvement product that customers rely on external search to support their purchase decisions (Richins and Bloch, 1986). Therefore, customers tend to require large amount of information to support their purchase decisions. Large volume of customer reviews in general contains more information towards the product than low volume of reviews. The number of answered questions reflects customers' information demands. When there are large number of answered questions, customers tend to post many queries, and thus indicate their high demands regarding information. Therefore, number of reviews tends to play a more important role in influencing sales when there are more answered questions. In addition, when the number of answered questions is large, there is likely to be a higher level of social interactions. When customers are highly interactive, the number of reviews may have a better reflection on product popularity which influences sales performance. Therefore, it is hypothesized that:

H4a: Number of answered questions moderates the effect of online review volume on online sales performance.

When the number of answered questions is large, there is likely to be high level of social interactions which further leads to high level of trust (Ou et al., 2014; Gefen and Straub, 2004). With trust, prospective customers tend to perceive the existing customers' reviews as credible and value existing customers' evaluations. However, when the number of answered questions is small, there is relatively low level of social interactions and trust. In such case, customers may not value other customers' reviews and thus overall customer rating can be less important than when there is a large number of answered questions. Therefore, this study hypothesizes that:

H4b: Number of answered questions moderates the effect of online overall review rating on online sales performance.

Moderating role of discount value

Offering discount is a type of promotional marketing strategy in online marketplace (Chong et al., 2015). Although studies have explored the effects of discount and other marketing strategies on product demands and sales (McNeill, 2013; Lichtenstein et al., 1990), limited attention has been paid to the interaction between the one-way marketing approach and customer reviews in online marketplace. Because of the monetary savings, customers tend to perceive the purchases with discount offerings as of good value (Chandon et al., 2000).

With high discount value, the cost of purchase perceived by customers is relatively lower than with low or no discount offering (Marshall and Leng, 2002). Therefore, discount value offers customers the incentive to buy products without too much consideration on customer reviews. Therefore, this study proposes that:

H5a: Discount value moderates the effect of online review volume on online sales performance.

H5b: Discount value moderates the effect of online overall review rating on online sales performance.

Moderating role of number of helpful votes

Customers' evaluations on whether reviews are helpful reflect quality of reviews (Mudambi and Schuff, 2010; Cheung et al., 2008). Compared to low level of helpfulness, when the review is perceived as highly helpful, the review is evaluated as better facilitating consumer's purchase decision (Mudambi and Schuff, 2010). Although the proportion of helpful votes in total votes is one of the most commonly used ways to measure helpfulness, this approach is sometimes inaccurate. For example, if 1 out of 1 people think a review is helpful, the percentage of helpful vote is 100%, while when 90 out of 100 people think a review is helpful, the percentage of helpful vote is 90%. It may be biased to conclude that the quality of the first review is better than the second review because there are far more people vote for the second review than the first. Thus, the number of helpful votes is an important yet often overlooked indicator of review helpfulness. In this study, the number of helpful votes of the most helpful review is used because it tends to have higher exposure to customers and may influence customers' first impression of review usefulness in general (Chong et al., 2015).

The large number of helpful votes of the most helpful review may be associated with high quality of customer reviews in general (Mudambi and Schuff, 2010). If so, customer review features including the review volume and overall review rating may be thought as helpful for purchase decisions and of great importance to sales. In contrary, small number of helpful votes for the most helpful review may adversely affect prospective customers' impression and trust towards existing reviews. Prospective customers thus may less likely to rely on review volume and overall rating. Therefore, review volume and overall review rating tend to be less influential for sales performance when number of helpful votes is smaller. However, when certain reviews receive large number of helpful votes, those reviews reflect the group norm of consumers, and have great social influence on prospective consumers. As a result, the number of helpful votes may be perceived as more important than the total review volume, and the influence of reviews with large number of helpful votes may outperform the overall review rating. Considering the above, this study proposes that:

H6a: Number of helpful votes moderates the effect of online review volume on online sales performance.

H6b: Number of helpful votes moderates the effect of online overall review rating on online sales performance.

To summarize, the research model in this study (Figure 1) proposes four factors that moderate the effects of review volume and overall review rating on product sales performance in online marketplaces, namely product category (search vs. experience product), number of questions, discount value and number of helpful votes.

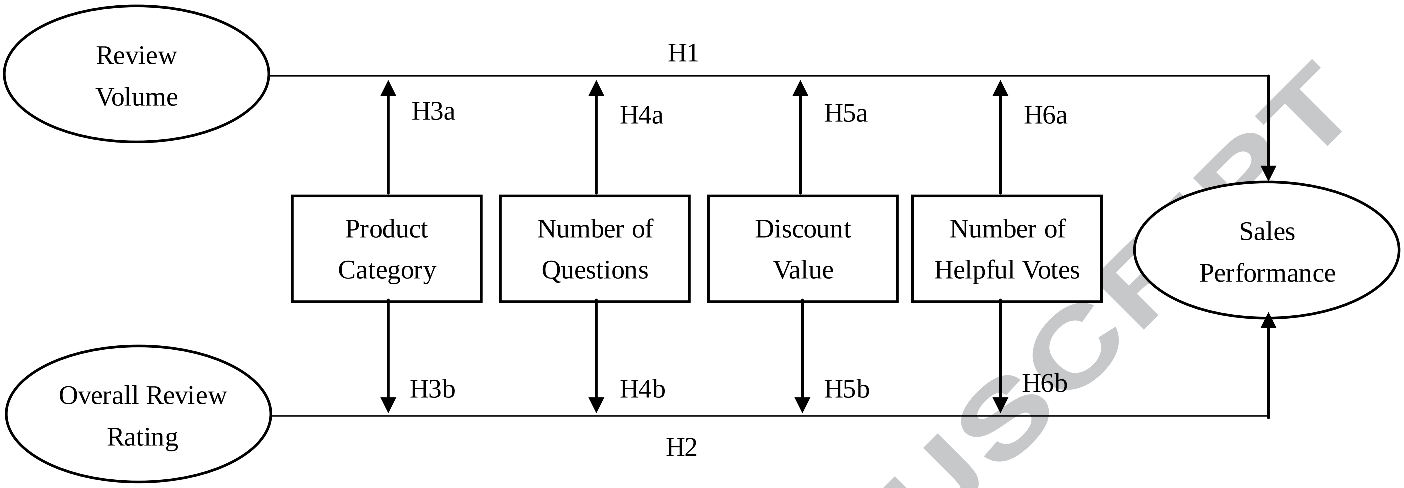


Figure 1 Research Model

Research Methodology

Data collection and big data architecture

The data for this study was collected from Amazon.com. Amazon.com is chosen because it is one of the most influential e-commerce platform and is also used as the research context by many other studies (e.g. Lu et al. 2013).

Product page and information is extracted using our big data architecture. Our big data architecture is a generic and fundamental technical system which is set up to serve the general data capturing, storage and processing purposes. The system is able to access tens of thousands of webpages, pick up the various type of content that we require, and collect and clean the data in real-time before storing the records. These functions of our system fulfil the volume, variety, velocity and value characteristics of big data, and thus can be seen as a big data architecture (Chong et al. 2015).

With 2x HP DL388p physical servers which can be scaled horizontally to fit the need and the 6x Linux Ubuntu 64bit Virtual Machines (VM) which set the need, a scalable system is built up. This system works within the Web and Social Media Big Data client-server architecture to assimilate many large organizations' various open-source server technologies (Ch'ng, 2015).

For the data collection and pre-processing of this study, we first coded asynchronous I/O algorithms server side in JavaScript using Node.js. The algorithms took a list of product paging links of Amazon.com as the input and crawled all product pages from the paging links. Then, in each product page, our asynchronous agents extracted the values we need with JavaScript's DOM and scattered HTML tags. After that, regular expressions were used to specify data patterns and pre-process the data into the required format, and the cleaned data was immediately stored into scalable MongoDB server. To avoid the block of our IP, we implemented a recursive mechanism to control the number of requests for Amazon.com pages. With such control to the number of requests per set, the system could continue the extracting of data until all the jobs were done. Finally, all the data records were generated into a Comma Separated Value (CSV) file for following-up analysis.

As mentioned before, this big data architecture is not set up specifically for this project. Therefore, with minor changes to the algorithms, the architecture can also be used to extract, save and process other types of data from different web sources.

Operationalization of variables

Online product sales performance is measured by sales rank provided by Amazon.com. Amazon.com does not provide specific number of sales, instead, it offers the sales rank for products. Although the way Amazon.com calculates sales rank is still a black box, the rank is calculated considering a seller's recent and historical sales, and can be used to represent the overall selling performance of a product. Low sales rank corresponds to good sales performance (Amblee and Bui, 2011). Therefore, this study uses the sales rank as a proxy of actual sales performance.

Online review volume is measured by the total number of customer reviews. Overall customer review rating is the star rating of a product listed in Amazon.com product page. This rating takes consideration of individual reviews and their other attributes such as the age of review to better reflect the overall evaluation of the product. We captured products in the category of cameras, cell phones and clothing as representative search products; home entertainment products and vehicle related products such as tires are included to represent experience products (Jiménez and Mendoza, 2013; Bei et al., 2004; Franke et al., 2004; Nelson, 1970). Number of answered questions is available on the product information page on Amazon.com. Discount value is measured by the monetary value of savings from price deduction. The number of helpful votes is measured by the number of people who vote to agree the review is helpful, and the number is for the most helpful review.

Analysis and Results

In this study, hierarchical multiple regression analysis is used to test the model. The results of regression analysis indicate a good fit of our model (p=0.000; R square=0.209) when including the moderation effects. Compared to the model without moderation effects, our model is significantly improved with 0.021 increase in R square and 9.598 of F change.

Table 1 Results of Regression Analysis

	Model	Coefficient (β)	t-value	Sig. (p)
1	Review volume	-.170	-8.067	.000
	Overall review rating	-.220	-12.639	.000
	Product category	-.276	-16.172	.000
	Number of answered questions	-.023	-1.273	.203
	Discount value	-.055	-3.248	.001
	Number of helpful votes	-.041	-1.982	.048
2	Review volume	-.450	-5.149	.000
	Overall review rating	-.351	-5.146	.000
	Product category	-.273	-15.269	.000
	Number of answered questions	-.155	-1.782	.075
	Discount value	-.047	-2.724	.006
	Number of helpful votes	-.189	-2.464	.014
	Review volume * Product category	.203	2.308	.021
	Overall review rating * Product category	.231	3.909	.000
	Review volume * Number of answered questions	.110	3.461	.001
	Overall review rating * Number of answered questions	.049	.668	.504
	Review volume * Discount value	-.037	-2.029	.043
	Overall review rating * Discount value	.002	.115	.909
	Review volume * Number of helpful votes	.136	1.827	.068
	Overall review rating * Number of helpful votes	.116	4.184	.000
ΔR^2=0.021 (p<0.001)				

Note: (1) Dependent variable: sales rank; (2) Product category is coded as dummy variable where 0 represents search product and 1 represents experience product.

As shown in Table 1, both review volume and overall review rating have strongly significant influence on sales rank (β=-0.450 and -0.351, $p<0.001$), supporting H1 and H2. The negative coefficients for review volume and overall review rating indicates negative influences on sales rank, which represents positive influence on sales performance.

Product category positively and significantly moderates the relationship between review volume and sales rank (β=0.203, $p<0.05$). Moreover, product category also has positive and significant moderation effect on the relationship between overall review rating and sales rank (β=0.231, $p<0.001$). Such results indicate that the effects of review volume and overall review rating on sales rank are more positive for experience product than for search product. Considering the negative coefficient of review volume on sales rank and the value of interaction coefficient, the results mean that review volume has negative effect on sales rank and thus positive effect on sales performance for both search and experience product, and such effect is stronger for search product than for experience product. Similarly, overall review rating also has negative influence on the sales rank for search product and experience product, which means positive influence on sales performance, and such influence is stronger for search product than for experience product. Therefore, H3a and H3b are both supported by the analysis.

Both number of answered question and discount value moderate the effect of review volume on sales rank significantly, thus supporting H4a and H5a. While number of answered question's moderation effect is positive (β=0.110, $p<0.001$), discount value's moderation effect is negative (β=-0.037, $p<0.05$). The results indicate that the influence of review volume on sales rank becomes more positive when number of answered question increases, while review volume affects sales rank more negatively when discount value is large than small. However, number of answered questions and discount value do not moderate overall review rating and sales rank significantly ($p>.100$). The results fail to support H4b and H5b. This study attempt to explain the findings in the next section.

Number of helpful votes has positive and significant moderating effect on the relationship between review volume and sales rank (β=0.136, $p<0.1$) as well as on the relationship between overall review rating and sales (β=0.116, $p<0.001$). The results support that the effects of review volume and overall review rating on sales rank are more positive when the number of helpful votes is larger, while the effects are more negative when the number of helpful votes is smaller. Therefore, H6a and H6b are also supported. The findings are summarized in Table 2.

Table 2 Summary of Results

	Description	**Result**
H1	Online review volume influences online sales performance.	Supported
H2	Overall review rating influences online sales performance	Supported
H3a	Product category moderates the effect of online review volume on online sales performance.	Supported
H3b	Product category moderates the effect of online overall review rating on online sales performance.	Supported
H4a	Number of answered questions moderates the effect of online review volume on online sales performance.	Supported
H4b	Number of answered questions moderates the effect of online overall review rating on online sales performance.	Not Supported
H5a	Discount value moderates the effect of online review volume on online sales performance.	Supported
H5b	Discount value moderates the effect of online overall review rating on online sales performance.	Not Supported
H6a	Number of helpful votes moderates the effect of online review volume on online sales performance.	Supported
H6b	Number of helpful votes moderates the effect of online overall review rating on online sales performance.	Supported

Discussion and Conclusion

This paper examines the moderation effects of product category, answered questions, discount and review usefulness on the relationship between customer review and product sales performance in online marketplace. Based on the results, it can be found that product category, answered questions, discount and review usefulness play important roles in understanding how customer review influences sales performance. Consistent with previous findings, large number of reviews and positive overall review rating lead to better sales performance than small volume or negative rating (e.g. Chong et al., 2015). Also consistent with the findings of Cui et al. (2012), overall review rating shows stronger influence on search product's sales performance, suggesting that rating as an evaluative cue has greater importance and persuasive power for more information-demanding products. However, in contrast to Cui et al. (2012)'s finding, we found that review volume also has stronger effect on the sales performance of search product. This is probably because search product is usually associated with high involvement (Cui et al., 2012). For high involvement product, review volume is more than a numeric index of popularity. Larger number of reviews means more information and evaluation available that are necessary for the purchase of high involvement product. Moreover, volume and overall rating can be less important for sales when number of helpful votes increases. Furthermore, review volume can be less important for sales rank when there are more answered questions and less discount. These findings show interesting interactions between different dimensions of factors and provide new perspectives to existing literature.

Number of answered question was found to have no significant moderating effect on the relationship between overall review rating and sales performance. A possible reason is that customers may not always associate the number of answered questions with trust towards existing customers. Answers can be given by the seller as well as other customers (Chong et al., 2015). When not distinguishing the source, answered questions may not reflect the true level of social interactions within customers and thus may not lead to trust towards customers, which makes customer review more important for sales performance. The results also failed to support discount value as a moderator of the relationship between overall review rating and sales performance. Overall review rating's effect does not differentiate across discount values. This is perhaps because that when the discount value is large, the price is often also high, and thus the assessment of product is important in reducing risks; if the discount value is small, the assessment of product is also important. The high price may offset the advantage of large discount value.

This study makes several theoretical and practical contributions. While most existing research focused on examining customer review and many other variables separately as individual factors that influence product sales performance in online marketplaces, this study finds that the interactions between these variables can also influence sales performance. Moreover, while existing research often adopt experiment or self-reported approach to collect data, this study help provides an example where big data architecture helps the testing of theoretical research model.

This study also provides practical implications to managers and practitioners. Existing research has illustrated the importance of customer reviews, and practitioners now usually pay attention to encouraging positive customer reviews. Our study highlights that the effects of customer review can be influenced by other factors. Based on our findings, practitioners should develop their strategies differently according to the types of product. For example, sellers of search products should spend more efforts in encouraging the customers to post positive reviews. Also, sellers, especially those with not many customer reviews, should pay attention to answer customer questions appropriately and timely. In addition, discount promotional strategy should also be tailored to fit the situation, such that setting high discount values can be helpful to magnifies the positive influence of large review volume to improve sales performance.

Moreover, for sellers with confidence to their products but few reviews, it is important to encourage high quality reviews that will potentially get large number of helpful votes. For example, sellers can encourage high quality reviews via sending out free trial products for detailed experience and evaluation report. To sum up, our findings suggest that practitioners can and should manage different factors strategically in online marketplace to improve sales performance.

This study has several limitations and leaves room for future studies. First, future studies could extend the sample to generalize our findings. Despite the fact that we choose representative search and experience products with reference to previous studies, strictly speaking, the findings can only be generalized to those products. Future studies thus could extend the choice of products and test if our findings can be generalized to other products. Future research could also confirm our findings from other websites. A second limitation is that we only use the number of helpful votes from the most helpful review. Although the most helpful review is often seen as of the highest exposure to customers, future studies could ask customers to report the number of reviews that they think influence their purchase decisions, and make more accurate decisions on the number of helpful votes. Another limitation is that, although this study takes an important information source- answered questions- into consideration, we did not distinguish the answers from customers and from the seller. Future studies could differentiate the interactions between customers and the seller with the interactions among customers, and examine the effects of answered questions in more detail.

References:

Amblee, N., & Bui, T. (2011). Harnessing the influence of social proof in online shopping: The effect of electronic word of mouth on sales of digital microproducts. International Journal of Electronic Commerce, 16 (2), 91-114.

Baek, H., Lee, S., Oh, S., & Ahn, J. (2015). Normative Social Influence and Online Review Helpfulness: Polynomial Modeling and Response Surface Analysis. Journal of Electronic Commerce Research, 16 (4), 290.

Bei, L., Chen,. E. Y. I., Widdows, R. (2004) Consumers' online information search behavior and the phenomenon of search vs. experience products. Journal of Family and Economic Issues, 25 (4), pp. 449-467.

Berbegal-Mirabent, J., Mas-Machuca, M., & Marimon, F. (2016). Antecedents of online purchasing behaviour in the tourism sector. Industrial Management & Data Systems, 116(1), 87-102.

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.

Chandon, P., Wansink, B., & Laurent, G. (2000). A benefit congruency framework of sales promotion effectiveness. *Journal of Marketing*, *64*(4), pp. 65-81.

Chen, K., Li, X., & Wang, H. (2015). On the model design of integrated intelligent big data analytics systems. Industrial Management & Data Systems, 115(9), 1666-1682.

Cheung, C. M., Lee, M. K., & Rabjohn, N. (2008). The impact of electronic word-of-mouth: The adoption of online opinions in online customer communities. Internet Research, 18(3), 229-247.

Cheung, C.M.K., and Thadani, D.R. (2012). The impact of electronic word-of-mouth communication: A literature analysis and integrative model, Decision Support Systems, 54 (1), pp. 461–470.

Chevalier, J., and Mayzlin, D. (2006). The effect of word of mouth on sales: online book reviews. Journal of Marketing Research, 43 (3), pp. 345-354.

Ch'ng, E. (2015). The bottom-up formation and maintenance of a Twitter community: Analysis of the #FreeJahar Twitter community. Industrial Management & Data Systems, 115 (4), pp. 612-624.

Chong, A. Y. L., Ch'ng, E., Liu, M. J., & Li, B. (2015). Predicting consumer product demands via Big Data: the roles of online promotional marketing and online reviews. International Journal of Production Research, pp. 1-15.

Chong, A.Y.-L. and Zhou, L. (2014). Demand chain management: Relationships between external antecedents, web-based integration and service innovation performance. International Journal of Production Economics, 154, pp. 48–58.

Clemons, E., Gao, G., and Hitt, L. (2006). When online reviews meet hyperdifferentiation: A Study of the craft beer industry, Journal of Management Information Systems, 23 (2), pp. 149-171.

Cui, G., Lui, H., and Guo, X. (2012). The effect of online consumer reviews on new product sales, International Journal of Electronic Commerce, 17 (1), pp. 39–58.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. The journal of abnormal and social psychology, 51(3), 629.

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), pp. 233-242.

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.

Franke, G. R., Huhmann, B. A., Mothersbaugh, D.L. (2004) Information content and consumer relationship of print ads: a comparison of search and experience products. Journal of the Academy of Marketing Science, 32 (1), pp. 20-31.

Gefen, D., & Straub, D. W. (2004). Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services. Omega, 32(6), pp. 407-424.

Huang, J. H., & Chen, Y. F. (2006). Herding in Online Product Choice, Psychology & Marketing, 23(5), pp. 413-428.

Jiménez, F. R., & Mendoza, N. A. (2013). Too popular to ignore: The influence of online reviews on purchase intentions of search and experience products. Journal of Interactive Marketing, 27(3), pp. 226-235.

Lichtenstein, D. R., Netemeyer, R. G., & Burton, S. (1990). Distinguishing coupon proneness from value consciousness: An acquisition-transaction utility theory perspective. Journal of Marketing, 54(3), pp. 54-67.

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

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

Marshall, R. & Leng, S.B. (2002). Price threshold and discount saturation point in Singapore. *Journal of Product & Brand Management*, 11, pp. 147–159.

McNeill, L. (2013). Sales promotion in Asia: successful strategies for Singapore and Malaysia. Asia Pacific Journal of Marketing and Logistics, 25, pp. 48–69.

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.

Nelson, P. (1970). Information and consumer behavior. Journal of Political Economy, 311-329.

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), pp. 209-230.

Poston, R., and Speier, C. (2005). Effective use of knowledge management systems: a process model of content ratings and credibility indicators. *MIS Quarterly*, 29(2), 221-244.

Richins, M. L., & Bloch, P. H. (1986). After the new wears off: The temporal context of product involvement. Journal of Consumer research, 13 (2), pp. 280-285.

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

Wu, M., Wang, L., Li, M., & Long, H. (2014). An approach of product usability evaluation based on Web mining in feature fatigue analysis. Computers & Industrial Engineering, 75, 230-238.

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.

Zhuang, Y., & Lederer, A. L. (2006). A resource-based view of electronic commerce. Information & Management, 43(2), 251-261.

Highlights

- Confirming the predictive power of product review volume and rating on sales
- Examining product type, answers, discount and information usefulness as moderators
- Using big data architecture to collect data for model testing