



Research on the Influence of Online Comments on the Sales of Group Buying Fruit Commodity

Zixuan Liu and Yongsheng Zhou^(✉)

Beijing Technology and Business University, Beijing, China
zhouys@amss.ac.cn

Abstract. This paper explored fruit, the largest category of agricultural products, as the research object. Based on text analysis and empirical analysis, using crawl data and constructing regression equation to test the hypothesis, this paper discusses the impact on sales from seven perspectives, which are, the number of comments, the depth of comments, the number of additional comments, the number of figure/video comments, the number of returned customers, discount and store rating. It is concluded that the number of comments, depth of comments, discount and store rating have a significant positive impact on fruit sales. The analysis results can be applied to guide and improve the sales of fruit in e-commerce.

Keywords: Online Comments · Fruit Sales · Influence Factor

1 Introduction

Today's society is developing at an unprecedented speed, and people's growing demand for a better life is also increasing. With the rapid development and transformation of the national economy, the consumption level continues to improve, and under the influence of the epidemic, online shopping has become the choice of more users. According to the survey of China Internet Network Information Center (CNNIC), by June 2021, the number of online shopping users in China had reached 812 million, an increase of 29.65 million compared with December 2020, accounting for 80.3% of the total Internet users. In 2020, China's online retail sales of agricultural products reached 610.7 billion yuan, including 129.569 billion yuan of online retail sales of fresh fruits in 2020, with a year-on-year increase of 43.32%, 5.62 percentage points higher than the growth rate of online retail sales of agricultural products, accounting for 22.88%, more than 1/5. It is the largest category of agricultural products and electric commodities in China. Among them, the e-commerce platform represented by Pinduoduo actively promotes the sales of agricultural products in online channels. At the same time, new business formats such as live e-commerce and group buying enable agricultural products to reach more consumers. Therefore, online retail sales are rising rapidly year by year.

Although more and more people choose to buy fresh fruit online, considering fruit is an experiential commodity, there is still significant uncertainty before the purchase. Therefore, customer's evaluation has become a major factor in the purchase decision.

This paper will be based on the group buying environment, taking fresh fruits as the research object, and fully explore the relationship between online comments and sales volume through the research of online comments one by one.

2 Overview of Related Theories and Hypothesis

Before the purchase, consumers will consider the uncertainty of purchase decision and its arising ramifications, which will affect their decision-making. This process is known as consumers' perceived risk. Compared with search-based products, experience-based products have higher risk perception than search-based products. In order to reduce the purchase risk, consumers will try to obtain product information through a variety of ways, including consulting merchants, referring to the graphic information of product pages, comparing different merchants and products, and referring to the comments of buyers. Dellarocas adds the number of comments to the prediction model. The results of empirical analysis show that the overall number of product comments will have a significant impact on the sales volume of products [1]. The information such as product rating and satisfaction of buyers reflected from online comments can help consumers eliminate some doubts, reduce their risk perception and improve consumers' purchase intention. The more words per comment has on average, the more product or service information it carries, that is, when the comment reaches deep aspect, it will reduce consumers' perceived risk of fresh agricultural products. Chenglin Liao and Chunjiang Cai (2013) proposed that consumers' purchase experience and comment depth positively affect the usefulness of comments [4]. Therefore, hypothesis 1 and hypothesis 2 are put forward:

H1: In the group buying situation, the number of comments has a positive impact on product sales.

H2: In the group buying situation, the depth of comments has a positive impact on product sales.

In all comments information, additional comments can reflect the subsequent quality of the goods in the use process, as well as the seller's commodity quality and service attitude. The comment mechanism of pictures and videos is to enable consumers to share the pictures of genuine purchased products, which can not only enable other consumers to more truly feel the appearance and quality of products, but also serves as a warning to those merchants who upload the pictures of commodity pages that are inconsistent with the genuine ones. Therefore, the number of additional comments and the number of images and videos have a great impact on commodity sales. When observing Pinduoduo platform, it is found that in the comment, a consumer will be marked how many times he or she has purchased and whether it is a returned customer. Consumers who have not purchased the product are more likely to believe the comments of returned customers. Forman et al. found that the comments showing some consumer's information are more likely to be believed by other consumers, so as to ensure the quality of comment information and improve the sales volume of products [3], thus assumptions H3, H4 and H5 can be drawn:

H3: In the group buying situation, the number of additional comments has a positive impact on product sales.

H4: In the group buying situation, the number of pictures and video comments has a positive impact on product sales.

H5: In the group buying situation, the number of returned customers has a positive impact on product sales.

When studying consumers' risk perception in online shopping [2], some scholars found that the price has a significant positive impact on consumers' perceived risk in several dimensions, which indicates that the increase of price will lead to the increase of consumers' social perceived risk, which is not conducive to consumers' purchase intention. For consumers participating in shopping, the emergence of price discount allows consumers to buy at a more favorable price and reduces consumers' perception of financial risk thus promotes individual consumption. In addition, on the Pinduoduo mall platform, choosing a discount means buying immediately. Whether the store is brand and whether the score is high will also affect consumers' judgment of purchase. Junyun Liao took hotels as the research object to study the impact of online comments on product sales, and brand trust played a regulatory role [5]. Based on this, putting forward assumptions H6 and H7:

H6: In the group buying situation, the discount has a positive impact on product sales.

H7: In the group buying situation, store scores have a positive impact on product sales.

3 Data Source and Model Construction

In view of the above research problems, this study takes Pinduoduo as an example, using Python to write a crawl program for data collection, taking "fresh fruit" as the keyword and sales volume as the sorting method, and collecting 16 days of data at one time, a total of 60 commodity samples and 9523 pieces of data. After data processing, 56 commodities with a total of 9470 data are obtained, and panel data are generated for analysis.

The regression equation is constructed according to the above theoretical basis, research assumptions and variable selection. The standardized expression of the regression equation is as follows:

$$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 \ln X_3 + \beta_4 \ln X_4 + \beta_5 \ln X_5 + \beta_6 X_6 + \beta_7 X_7 + C \quad (1)$$

The meaning of variables in the formula is shown in Table 1.

Considering the large divergence among different values, it will have a great impact on the standardized final equation coefficients and the error of regression results. Therefore, in this study, we first take the natural logarithm of the number of additional comments, the number of figure/video comments and the number of returned customers for the independent variables, then bring the obtained values into the multiple linear regression equation for empirical analysis, so as to reduce the error affecting the conclusion.

Table 1. Variable meaning

Variable	Variable meaning
Y	Sales volume (Sales ranking)
X ₁	Number of comments in 16 days
X ₂	Comment depth (average length)
X ₃	Additional comment quantity
X ₄	Figure/video comment quantity
X ₅	Returned customers quantity
X ₆	Discount
X ₇	Store rating

Table 2. Descriptive statistics

	N	Minimum	Maximum	Average	Standard Deviation
X ₁	56	22	490	163.11	106.11
X ₂	56	20.65	54.11	42.050	6.88175
X ₃	56	305	30000	5372.89	5810.318
X ₄	56	2673	100000	32962.82	31369.837
X ₅	56	821	100000	19211.98	23502.381
X ₆	56	0.127	0.9653	0.73519	0.18687
X ₇	56	4.0	5.0	4.616	0.3692
Effective N	56				

4 Results and Analysis

4.1 Results

This paper uses SPSS 26 carry out descriptive statistical analysis on the sorted data, and the specific analysis results are shown in Table 2.

Before the empirical analysis, first we can make descriptive statistics on the structured and unstructured data of online comments, so as to have a general understanding of the overall data. It can be seen from the table that there are great divergences among structured and unstructured data of online comments of 56 commodities in the number of additional comments, figure/video comments and returned customers, indicating that the sample is well representative and can be used for subsequent analysis. And consumers generally have good comments for fresh fruit stores, with an average score of 4.616 (Table 2).

Through the regression analysis of the sample data, first we can judge whether there is a correlation between the independent variables and the dependent variables. If the

Table 3. Correlations

		Y
Y	Pearson Correlation	1
	Sig. (2-tailed)	
X ₁	Pearson Correlation	.382**
	Sig. (2-tailed)	.004
X ₂	Pearson Correlation	.486**
	Sig. (2-tailed)	.000
X ₃	Pearson Correlation	.285*
	Sig. (2-tailed)	.033
X ₄	Pearson Correlation	.318*
	Sig. (2-tailed)	.017
X ₅	Pearson Correlation	.277*
	Sig. (2-tailed)	.039
X ₆	Pearson Correlation	.304*
	Sig. (2-tailed)	.023
X ₇	Pearson Correlation	.491**
	Sig. (2-tailed)	.000

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 4. Model Summary

Model	R	R Square	Adjusted R Square	Std Error of the Estimate
1	.491 ^a	.242	.227	14.335
2	.600 ^b	.360	.336	13.292
3	.659 ^c	.434	.401	12.620
4	.692 ^d	.478	.437	12.234

value corresponding to the regression coefficient is positive, the relationship between the independent variable and the dependent variable is positive. If the value corresponding to the regression coefficient is negative, the relationship between the independent variable and the dependent variable is negative. From the results of correlation analysis alone, the number of comments in 16 days, store rating and comment depth were significantly positively correlated with sales at the level of 0.01; The number of figure/video comments, additional comments, returned customers and discount are significantly positively correlated with sales at the level of 0.05.

Table 5. Coefficients^a

Model		Sum of Squares	df	Mean Square	F	Sig.
4	Regression	6996.29	4	1749.07	11.685	.000 ^e
	Residual	7633.70	51	149.68		
	Total	14630.0	55			

Table 6. ANOVA^a

Model		Unstandardized		Coefficients	t	Sig.	Collinearity Statistics	
		<i>B</i>	<i>Std.error</i>	<i>Beta</i>			<i>Tolerance</i>	<i>VIF</i>
4	(constant)	-88.09	21.15		-4.16	.000		
	X ₁	.03	.01	.24	2.37	.021	.93	1.06
	X ₂	.81	.26	.34	3.13	.003	.84	1.18
	X ₆	19.12	9.18	.21	2.08	.042	.92	1.08
	X ₇	13.43	4.87	.30	2.75	.008	.84	1.18

As can be seen from Table 4, Model 4 has the best fit, and the variables used are X₁, X₂, X₆, X₇. The independent variable in model 4 accounts for 43.7% of the variable variation (Table 4).

The significance value is about $\theta < 0.05$, reaching a significant level, which shows that the regression equation is valid.

The final regression equation expression obtained from Table 6.

$$Y = 0.038X_1 + 0.815X_2 + 19.12X_6 + 13.44X_7 - 88.10 \quad (2)$$

4.2 Analysis

- The number of comments in 16 days has a remarkable positive impact on fruit sales, which proves H1. Comments can provide consumers with more information, helping them to make the purchase decisions, and promote sales volume consequently. Comment depth has a significant positive impact on fruit sales, which proves that H2. The longer the comment, the more other consumers know about the product and the higher the recognition of the comment.
- The number of figure/video comments, additional comments and returned customers have no significant impact on the sales of fruits. Therefore, we cannot compare their effects on dependent variables, and H3, H4 and H5 have not been proved. Since we only analyze the number of figure/video comments, additional comments and returned

customers of a commodity, intuitive figures may not stimulate consumers' consumption well. With the development of online shopping, more and more consumers choose to spend time reading the evaluation content.

- The discount has a significant positive impact on the sales of fruits, which proves that H6. When there is a greater divergence between the group price and the non-group price, it will promote the sales of fresh fruits. Store score has a significant positive impact on fruit sales, which proves H7. It shows that most consumers will choose businesses with higher store rating for consumption.

5 Countermeasures and Suggestions

Based on the comment data analysis of Pinduoduo platform, this study makes a text analysis and Empirical Analysis on the impact of online comments on the sales of fresh fruits. The empirical analysis results show that the number of comments, comment depth, discount and store rating will have a great impact on consumers' purchase decision, which is positively correlated with product sales. The number of figure/video comments, additional comments and returned customers did not affect the sales volume of the product. According to the research results, this paper puts forward the following suggestions: it's important for the stores to guide and encourage consumers to comment more on the platform, share the real user experience, and to timely analyze the comment contents of negative tendencies then to adjust the marketing means like increasing discount rate and improving service attitude, so as to find practical solutions and improve the score of stores. For consumers, they can make better use of comments to objectively and truly express their feelings about products, so as to better guide future consumers to make purchase decisions. For the online shopping platform, a better online comment supervision mechanism can be established to avoid malicious comments caused by competition among merchants, so as to reduce the credibility of comments.

References

1. DELLAROCAS C, ZHANG Xiaoquan, AWADNF. Exploring the Value of Online Product Reviews in Forecasting Sales: The Case of Motion Pictures [J]. *Journal of Interactive Marketing*, 2007, 21 (4): 23–45.
2. FENG Jian. An Empirical Study on The Impact of Product Factors on Consumers' Purchase Intention [D]. Xi'an University of Technology, 2011.
3. Forman C, Ghose A, Wiesenfeld B. Examining the Relationship Between Reviews and Sales: The Role Of Reviewer Identity Disclosure in Electronic Markets [J]. *Information Systems Research*, 2008, 19(3):291–313.
4. LIAO Chenglin, CAI Chunjiang, LI Yi. An Empirical Study on Influence Factors of Online Reviews' Helpfulness in E-commerce [J]. *Soft Science*, 2013, (27): 46–50.
5. LIAO Junyun, & HUANG Minxue. (2016). An Empirical Study on Online Product Reviews, Brands and Product Sales Based on Hotel Sales. *Journal of Management*, 13(1), 122–130.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

