Research on the impact of Credit, Brand and Service Recovery on Online Sales Based on Web Crawler Technology and Regression Analysis Methods

Mi Luo¹, Jianzong Lin¹*

¹Xiamen University of Technology, Xiamen, Fujian, 361021, China
*e-mail: jianzonglin@xmut.edu.cn

Abstract. In order to explore the relationship between credit, brand, service recovery and the merchandise sales in online transactions, based on web crawler technology and regression analysis methods, the paper empirically analyzes the co-influence of credit, brand and service recovery on online merchandise sales volume by using the data of Baijiu in Tmall platform. The research results show that seller credit score has a significant positive effect on online merchandise sales volume; Merchandise brand has no significant impact on online merchandise sales volume, but brand as a moderator has a significant positive effect on the relationship between seller credit score and merchandise sales volume; The seller’s service recovery has no significant influence on the sales volume of online merchandises.

1. Introduction
The 47th Statistical Report on Internet Development in China shows[1]: As of December 2020, the number of online shopping users in China has reached 782 million, and the online retail sales have reached 11.76 trillion yuan, of which the online retail sales of physical goods were 9.76 trillion yuan, accounting for 24.9% of the total retail sales of social consumer goods.

The rapid development of online shopping has brought great convenience to consumers, but there are great risks in online transactions. It is difficult for consumers to evaluate the credibility of online goods, and to distinguish the quality of service and goods[2]. In order to better manage online transactions, basically all online shopping platforms have built a credit evaluation mechanism, such as eBay, Taobao and so on. Through this credit evaluation mechanism, buyers can score the seller’s credit after the transaction. Credit score is an important basis for other buyers to decide whether to purchase the seller’s goods, which has an impact on the sales volume of goods.

2. Literature review
Many scholars at home and abroad have studied the influencing factors of sales volume in online shopping.

Foreign scholars mainly take the transaction form of auction bidding on the eBay platform as an example to study the influencing factors of sales volume. Ba and Pavlou[3] collected two months of electronic digital product data on the eBay platform for research, they showed that the buyer’s trust in the seller’s credit has an impact on the transaction by generating price premium, and the seller’s credit score can effectively reduce the information asymmetry in the online trading market. Bajari etc.[4] took the American coinage on the eBay platform as the research object, and found that the seller’s total credit value has a significant positive effect on the number of auctions. Livingston[5] collected the auction...
data of eBay for 10 months and found that the impact of reputation on sales volume is nonlinear. There are some problems with the commodities in the form of auction bidding as the research object. For example, the commodity transaction in the form of auction bidding is not very active, the time of collecting data is long, and it is difficult to control the influence of time factors.

Domestic scholars mainly study the influencing factors of sales volume through the example of buy it now transactions on the Taobao platform. Compared with the transaction form of auction bidding, buy it now transaction is more active and the data is better collected. Cui Xiangmei etc.[6] used the transaction data of Kingston SD card in Taobao to conduct research, and found that the price has a significant negative impact on the sales volume, and the number of good, medium and poor ratings in the credit evaluation system have a significant positive impact on the sales volume. Zhou Geng etc.[7] used the transaction data of Nokia mobile phones in Taobao to conduct research, and found that the seller’s credit has a significant positive effect on online purchase behavior, and plays a moderating role in the negative relationship between price and purchase behavior. Li Song et al.[8] used the Tobit regression model to carry out the research, and found that the seller’s credit score, praise rate have a strong correlation with the sales volume. Based on Quantile Regression and threshold regression methods, Xu Qifa et al.[9] constructed a threshold quantile regression model, and found that improving the reputation level of merchants can increase the sales volume of merchants. Zhou Shudong and Miao Chenglin[10] found that credit and price have a significant positive impact on monthly sales, store history has a significant impact on monthly sales, and the independent completion rate of refund has no significant impact on monthly sales.

All the above studies show that the seller’s credit has an impact on sales volume, and consumers can judge the quality of their goods through the seller’s credit score, thus affecting their purchasing behavior. Brand is also an important factor affecting consumers’ purchasing behavior[11]. Wang Ying etc.[12] applied the rice data of Jingdong platform to study, and found that brand strength, the number of comments, the comment valence and the comment timeliness have significant effects on rice sales.

Existing studies separately examine the impact of seller’s credit score and brand on sales volume, but how credit and brand jointly affect sales volume has not been fully studied. This paper analyzes and study this. In traditional offline transactions, both practice and research have verified that service recovery significantly affects consumers’ willingness to continue to consume. However, in online transactions, the impact of service recovery on sales volume has not been fully studied, and this paper will also analyze and study it.

3. Variables and models

3.1. Explained variables selection

The seller’s sales volume within 30 days is selected as the explained variable. Since the cumulative sales time of products on the Tmall platform is different, in order to avoid the problem of difference in sales due to different sales time, the cumulative sales within 30 days is taken as the explained variable.

3.2. Explanatory variables selection

3.2.1. Credit score

The seller’s credit score can provide consumers with a basis for trust. It is known that the credit calculation method of Taobao platform is the total number of people evaluated multiplied by the difference between the favorable comment rate and the poor comment rate. However, the credit score of Tmall platform has no specific numerical value, so the credit calculation method of Taobao platform can be applied to the credit calculation of Tmall platform. The higher the credit score, the more consistent the seller’s product with the description, the more the buyer’s trust can be obtained, thus promoting the purchase behavior. Therefore, hypothesis 1 is proposed: the seller’s credit score has a positive impact on the sales volume of goods.
3.2.2. Brand
Brand is a kind of name, term, mark, symbol or design, or a combination of them, which is used to identify a company’s products or services and distinguish them from those of competitors[13]. Some scholars have proved the influence of brand on purchase intention or sales volume. Bansah et al.[14] found that brands have a positive impact on commodity sales, and consumers are loyal to strong brands. Therefore, the brand can enhance the loyalty of consumers, promote the consumption of consumers, and thus increase the sales of goods. Therefore, hypothesis 2 is put forward: brand has a positive impact on the sales volume of goods.

According to the accessibility-diagnostic model, the possibility of information clues used in consumer cognitive judgment depends on the accessibility and diagnostics of the clues[15], and the accessibility of information is the prerequisite for diagnostic[16]. Brand is the consumer’s existing impression of the product, and it is more accessible to consumers’ cognition and judgment of the product. The seller’s credit score is more specific, reflecting the situation of the product and the seller, and thus is more diagnostic. So when the product brand is not strong, even if the seller’s credit score is high, consumers will not buy the product. Therefore, hypothesis 3 is proposed: brand has a positive moderating effect on the relationship between seller’s credit score and sales volume of goods.

3.2.3. Service recovery
Service recovery is the action taken by service providers against service failures[17], in order to recover consumers who have suffered service failures. Jian Zhaoquan et al.[18] pointed out that compared with psychological remedies such as apology, substantive remedies such as refund can bring higher consumer secondary satisfaction and loyalty. Service recovery commonly used in online transactions include refund, exchange, discount, error correction and apology[19]. So the self-completion rate of refund is chosen to measure the service recovery situation of the merchants, which is the ratio of the number of self-completion refunds by the merchant to the total number of refunds completed by stores. The higher the self-completion rate of refund, the better the quality of service recovery, and the more attractive it will be to consumers. Therefore, Hypothesis 4 is proposed: the seller’s service recovery has a positive impact on the sales volume of goods.

3.3. Control variables

3.3.1. Price
Price is related to the vital interests of consumers. And according to the law of demand, under other conditions unchanged, the quantity of goods demanded is inversely proportional to the price. Therefore, Hypothesis 5 is put forward: price has a negative impact on the sales volume of goods.

3.3.2. The duration of store establishment
The duration of store establishment refers to the difference between the year when the seller created the store and the year when the data was collected. The longer a store is established, the more transaction records it has, so as to provide more reference for consumers. Therefore, hypothesis 6 is put forward: the duration of store establishment has a positive impact on the sales volume of goods.

The variable description is shown in Table 1.

| Variable name | Variable meaning |
|---------------|------------------|
| Explained variable | Sales volume | The number of transactions completed by the seller within 30 days |
| Explanatory variable | Brand | The symbol of quality of goods and services |
| | Credit score | The total number of evaluators multiplied by the difference between the favorable rate and the poor rate |
| | Service recovery | Actions taken by service providers for service failures |
### 3.4. Model establishment

Taking the logarithm can eliminate the possible heteroscedasticity in the regression, and at the same time convert the potential nonlinear relationship into linear relationship, so logarithm of each explanatory variable is taken. In addition, since the brand value may be 0, it is added to 1 to avoid invalid values. Therefore, according to the establishment process of multiple regression model, the model is established as in equation (1):

$$\ln Y = C + \alpha_1 \ln X_1 + \alpha_2 \ln (X_2 + 1) + \alpha_3 \ln X_3 + \alpha_4 \ln X_4 + \alpha_5 \ln X_5 + \alpha_6 \ln X_1 \ln (X_2 + 1) + \varepsilon \tag{1}$$

Where, $Y$ is the sales volume, $X_1$ is the seller’s credit score, $X_2$ is the brand, $X_3$ is the service recovery, $X_4$ is the price of the product, $X_5$ is the duration of store establishment, and $\varepsilon$ is the random disturbance item, that is, the influence of other factors other than brand, seller’s credit score, service recovery, price and the duration of store establishment on the sales volume.

The conceptual model of the study is shown in Figure 1.

![Conceptual Model](image)

Figure 1. Conceptual Model

### 4. Empirical research

#### 4.1. Data sources

Since Baijiu has obvious brand division and the trading activity of Baijiu in Tmall platform is high, the data of 500 ml of 52-degree Baijiu was selected for research. Considering the lack of representativeness of sellers with zero sales volume and low credit score, the transaction data with zero sales volume within 30 days and sellers with credit score less than 10 were excluded, so a total of 382 valid data were collected. Since the brand can’t be measured specifically, NGA etc.[20] distinguished the brand into strong brands and weak brands through the brand ranking with credibility. This study follows this practice. According to the Competitiveness Index Report of Chinese Baijiu Enterprise (2019), jointly issued by the China Liquor Circulation Association and China Liquor News Network, the top 10 brands in the report are regarded as strong brands and coded as 1, and other brands are regarded as weak brands and coded as 0. The descriptive statistical results of data are shown in Table 2.

| Variable               | N  | Minimum value | Maximum value | Mean value | Standard deviation |
|------------------------|----|---------------|---------------|------------|-------------------|
| Y sales volume         | 382| 1.0           | 1060.0        | 53.992     | 117.0528          |
| X₁ credit score        | 382| 29.0          | 6721447.0     | 227627.704 | 1175786.3182      |
| X₂ brand               | 382| 0.0           | 1.0           | 0.581      | 0.4934            |
| X₃ service recovery    | 382| 0.8378        | 1             | 0.988501   | 0.016011          |

Table 2. Descriptive statistical results
4.2. Empirical test

Eviews8.0 econometric analysis software was used to conduct LS analysis on the sample data, and the relevant data of five variables were imported to get the analysis results as shown in Table 3. The results showed that brand X2 is not significant at the 5% significance level, which indicates that brand has no significant effect on the sales volume of goods.

| Variable       | Coefficient | Std. Error  | t-Statistic | Prob. |
|----------------|-------------|-------------|-------------|-------|
| C              | 4.913663    | 0.435433    | 11.28454    | 0.000 |
| LNX1           | 0.303292    | 0.046897    | 6.467257    | 0.000 |
| LNX2           | 0.532855    | 0.606979    | 0.877881    | 0.380 |
| LNX3           | 6.296129    | 3.110509    | 2.024147    | 0.043 |
| LNX4           | -0.781754   | 0.048059    | -16.26671   | 0.000 |
| LNX5           | 0.113835    | 0.049963    | 2.278386    | 0.023 |
| LNX1X2         | 0.070754    | 0.077091    | 0.917799    | 0.359 |

Therefore, the variable of brand X2 was excluded and the regression analysis was carried out, and the analysis results were shown in Table 4.

| Variable       | Coefficient | Std. Error  | t-Statistic | Prob. |
|----------------|-------------|-------------|-------------|-------|
| C              | 5.179117    | 0.313222    | 16.53498    | 0.000 |
| LNX1           | 0.269559    | 0.026876    | 10.02961    | 0.000 |
| LNX3           | 6.155120    | 3.105412    | 1.982062    | 0.048 |
| LNX4           | -0.782701   | 0.048032    | -16.29549   | 0.000 |
| LNX5           | 0.119599    | 0.049514    | 2.415448    | 0.016 |
| LNX1X2         | 0.136377    | 0.018843    | 7.237489    | 0.000 |

White test was used for heteroscedasticity test, and the analysis results were shown in Table 5. The results showed that the p values were less than 0.05 (at the 5% significance level), indicating the existence of heteroscedasticity.

| Heteroscedasticity Test: White          |          |
|-----------------------------------------|----------|
| F-statistic                             | 2.374728 |
| Prob. F(19,362)                         | 0.0011   |
| Obs*R-squared                          | 42.33588 |
| Prob. Chi-Square(19)                    | 0.0016   |
| Scaled explained SS                    | 47.56086 |
| Prob. Chi-Square(19)                    | 0.0003   |

The weighted least square method was used to modify the model, and the general weight was set to 1 / LNX1. The results were shown in Table 5. The results showed that the duration of store establishment X5 is not significant at the 5% significance level, which indicates that the impact of the duration of store establishment on the sales volume is not significant.
Table 6. Heteroscedasticity correction results

| Variable | Coefficient | Std. Error | t-Statistic | Prob.  |
|----------|-------------|------------|-------------|--------|
| C        | 4.254952    | 0.311380   | 13.66484    | 0.0000 |
| LNX₁     | 0.352971    | 0.029512   | 11.96007    | 0.0000 |
| LNX₃     | 5.746432    | 2.876992   | 1.997375    | 0.0465 |
| LNX₄     | -0.724310   | 0.046586   | -15.54779   | 0.0000 |
| LNX₅     | 0.046210    | 0.044240   | 1.044528    | 0.2969 |
| LNX₁X₂   | 0.137275    | 0.019256   | 7.128989    | 0.0000 |

Weighted Statistics

- R-squared: 0.528948
- Adjusted R-squared: 0.522684

Excluding the variable of the duration of store establishment X₅ for regression analysis and heteroscedasticity test, the analysis results were shown in Table 7 and Table 8. In the stepwise regression process, it is found that the variable of service recovery X₃ is not significant at the 5% significance level, while the seller's credit score X₁, price X₄ as the explanatory variable and brand X₂ as the moderating variable are all significant at the 5% significance level, and there is no heteroscedasticity. This shows that service recovery has no significant effect on the sales volume, the seller’s credit score and price have a significant impact on the sales volume, and the moderating effect of brand on the relationship between the seller’s credit score and the sales volume is significant.

Table 7. Final regression analysis results

| Variable  | Coefficient | Std. Error | t-Statistic | Prob.  |
|-----------|-------------|------------|-------------|--------|
| C         | 4.231577    | 0.305628   | 13.84552    | 0.0000 |
| LNX₁      | 0.351321    | 0.028125   | 12.49134    | 0.0000 |
| LNX₄      | -0.720768   | 0.046358   | -15.54781   | 0.0000 |
| LNX₁X₂    | 0.134847    | 0.019193   | 7.025688    | 0.0000 |
| R-squared | 0.523458    |            |             | 0.519676 |

Table 8. Final heteroscedasticity test results

|                 |          |            | Prob. Chi-Square(8) | 0.7721 |
|-----------------|----------|------------|---------------------|--------|
| F-statistic     | 0.601175 | Prob. F(8,373) | 0.7769              |
| Obs*R-squared   | 4.862743 | Prob. Chi-Square(8) | 0.6828 |
| Scaled explained SS | 5.682199 | Prob. Chi-Square(8) | 0.6828 |

In summary, the seller’s credit score has a significant positive impact on the sales volume, which is consistent with Hypothesis 1. The seller’s credit score can measure the quality of the seller’s goods and services. The higher the credit score, the more trust the buyer will have, and the more likely the buyer will have purchasing behavior. Another reason may be that the seller’s credit score is one of the factors that affect the search ranking of Tmall platform. Sellers with high credit score are easier to be searched, and the probability of successful transaction is higher.

The influence of brand on the sales volume is not significant, which is inconsistent with Hypothesis 2. However, the brand has a positive moderating effect on the relationship between the seller’s credit score and the sales volume, which is consistent with Hypothesis 3. It shows that the influence of brand on the sales volume of goods is mainly realized by adjusting the credit score of merchants. The stronger the brand, the easier it is for consumers to generate loyalty, which makes consumers pay attention to positive information and resist negative information in the case of brand preference. Therefore, brand can strengthen the positive impact of high credit score on sales volume, and weaken the negative impact of low credit score on sales volume. Moreover, the stronger the brand, the greater the impact of seller’s credit score on sales volume.

The effect of service recovery on sales volume is not significant, which is inconsistent with Hypothesis 4. The reason may be that the service recovery of sellers in online transactions is generally...
good, and table 2 shows that the average value of the self-completion rate of refund is as high as 0.988501, so consumers don’t pay attention to the small gap in service recovery quality.

Commodity price has a significant negative impact on sales volume, which is consistent with Hypothesis 5. Consumers in online transactions are rational. With the increase of price, the occurrence of consumers’ purchasing behavior decreases.

The impact of the duration of store establishment on sales volume is not significant, which is inconsistent with Hypothesis 6. The reason may be that sellers generally stay in for a long time. Table 2 shows that the average value of the duration of store establishment is 3.9 years, and the maximum value is 12 years. It can be seen that the sellers who stay on Tmall platform have been tested for a long time, so consumers pay less attention to the duration of store establishment.

5. Conclusion and limitation

5.1. conclusions and suggestions

The following conclusions are obtained and recommendations are put forward based on the conclusions.

- The seller’s credit score has a significant positive impact on the sales volume of online goods. The higher the seller’s credit score, the greater the sales volume of online goods. The seller’s credit score is accumulated through the buyer’s score after purchasing goods. It is suggested that sellers should improve the sales volume on the basis of ensuring the quality of goods and services, so as to enter a virtuous circle of credit score improvement and sales volume increase.
- Brand has a positive moderating effect on the relationship between seller’s credit score and sales volume. Brand strengthens the influence of seller’s credit score on online merchandise sales. The stronger the brand, the greater influence of the seller’s credit score on the sales volume. It is recommended that sellers should strengthen their choice of commodity brand while working hard to improve the credit scores.
- Service recovery has no significant effect on the sales of online goods. Sellers’ service recovery in online transactions are generally better. In order to keep up with the pace of the market, sellers still need to pay attention to service recoveries. It is suggested that in case of service failures, sellers should coordinate and communicate with the consumers and take substantive service recoveries, such as refund and return, to eliminate the dissatisfaction of consumers.

5.2. Limitations

The deficiencies of the research are mainly reflected in two aspects: first, only the Baijiu trading data is used to study, and whether the conclusion is applicable to other categories of commodities. Second, the research divides the brands into strong brands and weak brands according to the existing data. Is it necessary to make a more detailed division of the brands? In the future, more accurate data will be used to quantify the brands.

Acknowledgments

Funding: This research was funded by the Graduate Science and Technology Innovation Program of Xiamen University of Technology (YKJCX2019085).

References

[1] China Internet Network Information Center. (2020) The 47th Statistical Report on Internet Development in China. http://www.cac.gov.cn/2021-02/03/c_1613923423079314.htm.
[2] Liu J.S., You Z.X., Le G.X., et al. (2018) Research Progress of Network Trust. Computer Science,45:20-35+43.
[3] Ba S., Pavlou A., Paul A. (2002) Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior. MIS Quarterly,26:243-268.
[4] Bajari P., Hortacsu A. (2003) The Winner’s Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions. Rand Journal of Economics,34:329-355.
[5] Livingston J. (2005) How Valuable Is A Good Reputation? A Sample Selection Model of Internet Auctions. The Review of Economics and Statistics, 87: 453-65.

[6] Cui X.M., Huang J.H. (2010) Empirical Study on the Impact of the Reputation System on Buy-it-Now Online Transaction. Chinese Journal of Management, 7: 50-56.

[7] Zhou G., Wang Q.S., Zhu Z. (2010) Reputation, Third-party Seal and Online Purchasing Behavior: an Empirical Study of Taobao. Economic Management, 32: 143-152.

[8] Li S., Li F.C. (2012) Research on the Relationship Between Reputation and Sales Volume of Online Shopping- An Empirical Analysis Based on Taobao. Modern Management Science, (11): 92-94.

[9] Xu Q.F., Jia J.Y., Jiang C.X., et al. (2016) Analysis of Factors Affecting Online Goods Sales Through Threshold Quantile Regression Approach. Journal of Business Economics, (07): 5-14.

[10] Zhou S.D., Miao C.L. (2019) Analysis on the Impact of E-commerce Sales of Chinese Rural Characteristic Food -- Based on the Data Mining of Taobao ‘Characteristic China Pavilion’. World Agriculture, (5): 65-72.

[11] Erdem T., Swait J., Valenzuela A. (2006) Brands as Signals: A Cross-country Validation Study. Journal of Marketing, 70: 34-49.

[12] Wang Y., Yang D., Wang L., et al. (2018) Impact of online reviews on the sales of rice under the influence of brand effect: An empirical study. Journal of China Agricultural University, 23: 178-187.

[13] Philip K. (1997) Marketing Management: Analysis, Planning, Execution and Control. Shanghai People’s Publishing House, Shanghai.

[14] Bansah P.F., Dabi M., Anita S., et al. (2015) The Effect of Branding on Consumer Buying Behavior Among Textile Ghana Fabric Users in the Ho Municipality of Ghana. Journal of Business Market Management, 7: 117-125.

[15] Feldman J.M., Lynch J.G. (1988) Self-generated Validity and Other Effects of Measurement on Belief, Attitude, Intention, and Behavior. Journal of Applied Psychology, 73: 421-435.

[16] Menon G., Raghubir P., Schwarz N. (1995) Behavioral Frequency Judgments: An Accessibility-diagnosticity Framework. Journal of Consumer Research, 22: 212-228.

[17] Gronroos C. (1988) Service Quality: The Six Criteria of Good Perceived Quality Service. Review of Business, 9: 10-13.

[18] Jian Z.Q., Ke Y. (2017) The Impacts of Service Failure, Service Recovery on Customer Satisfaction and Loyalty in the Context of Online Shopping. Management review, 29: 175-186.

[19] Forbes L.P., Kelley S.W., Hoffman K.D. (2005) Typologies of E-Commerce Retail Failures and Recovery Strategies. The Journal of Services Marketing, 19: 280-292.

[20] Nga N.H., 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: 37-53.