

Analysis on C2C Online Trading Platform Under the Regression Analysis and Game Theory —— Taking Taobao as an Example

Jiatong Ding

*RDFZ Xishan School
13522650511@163.com*

ABSTRACT

Taobao is one of the biggest C2C platform in the world, because of the information asymmetry and the lack of communication between consumers and sellers, consumers may buy counterfeit or inferior goods in Taobao. The paper will focus on one of measure of the communication, comments. It will analyze the Taobao C2C transaction with or without communication by using game theory and regression analysis. The paper will use Python to crawl enough data from Taobao website and use Stata to analyze the data. The analysis shows that positive comments are correlated with the sales number of products, which will prove that comments from the buyer will have an impact on sales. Besides that, the paper will build a game model to analyze the Taobao C2C transaction without communication, which could prove that the best strategy for sellers are being honest whether consumers return or not return the product. This paper has provided solutions for C2C platform improvement.

Keywords: *C2C, game theory, regression analysis, communication, Taobao*

1. INTRODUCTION

In recent years, with the continuous development of China's online shopping market, more and more people choose to shop online. According to CNNIC (China internet network information Center) , In 2019, China's online retail sales reached 10.63 trillion yuan, including 8.52 trillion yuan of physical goods, accounting for 20.7 percent of the total retail sales of consumer goods. Although the online shopping makes people's life more convenient, it still has some drawbacks that could be critical to consumers. The consumer may face the problem of buying inferior goods or fake goods in online. Moreover, the consumer will also face the problem that the description of goods and the actual goods are mismatched. Under this background, the paper will introduce the method of game theory and focus on the game between buyers and the sellers. The comments and evaluation from buyers are important measures for communication. The paper will take Taobao for example, and it will analyze the online transaction that with or without comments and evaluation for sellers. The meaning of this paper is to provide some specific solutions to Online transaction website; thus sellers will have high quality of transaction and consumers will have better online shopping environment. Moreover, this paper will also lead the sellers to pay more attention on their trust building.

1.1. Analysis of Taobao C2C Online Transaction with Comments

Taobao, one of the biggest C2C website in China, nearly 600 millions of people use it. All kinds of goods are selling in this platform. Taobao have the function that which could show the comments of people who have already bought this good. People could check the buyers' comments and decide to buy or not to buy this item, which depend on the trust and the comments of the buyers for most commodities. Moreover, people have already bought this product could do a review about the experience of after use it for several days, so that the consumers will have a better understand about the durability of the good and does it worth to buy it. The part wants to prove that if this seller has more comments, people will more tend to buy same item in the sellers have more positive items.

This part selects the data of sneakers sellers from Taobao, and the selected features of data are product sales, price, comments number, review number, and picture review, among which the sales number are Y. The software stata was used to analysis the data and Python to attain the data. Firstly, start with the basic descriptive statistics. as figure 1 shows

Secondly, do regression test about the data, as figure 2 shows.

Because $p > 0.05$, so that review and picture review these two variables didn't pass the regression test, price and sales pass the test, so the paper rejects review and picture review these two variables and doing regression test again.

According to figure 3, price and comments both pass the regression test. the paper will use BP test and White test in Stata to test whether there is heteroscedasticity.

As figure 4 shows, the p-value for both two tests are below 0.05, which means that the paper rejects the null hypothesis under 95% confident level, therefore, the heteroscedasticity exist. After that, the paper needs to check the Multicollinearity of data.

As figure 5 shows, because Mean VIF < 10, so the data do not exist Multicollinearity.

From figure 6, the p-value belows 0.05, so it have pass the test; from the regression coefficient, there's a negative correlation between sales and price when price increases 1 yuan, the sales number will decrease 0.2465. There is a positive correlation between sales and comments, when increase each additional comment, the sales number will increase 0.0995. After that, to make sure whether the price or comments have more impact on sales, the paper will use Stata to find the Standard regression coefficient of comment and price.

From the Standard regression coefficient (Beta) in figure 7, price is -0.1695, comment is 0.4035, then the absolute value of these two Beta values were compared in figure 7.

Thus, comment have more impact on sales.

In conclusion, from the analysis of the sellers in sneaker area, this part could prove that comments are important in C2C business, which is an important part of communication. Without communication, the sales and reputation will both have some impact on it. For Buyers, they could compare the same product that from different sellers depending on their positive comments and reviewing. For sellers, they could increase the quality of their product and take out new sales strategy like providing discounts to attach more consumers.

```
. summarize sale price comment review picture_review
```

Variable	Obs	Mean	Std. Dev.	Min	Max
sale	566	362.1413	457.4516	13	3932
price	566	477.9263	314.4652	88	2299
comment	566	1166.528	1854.694	0	9926
review	566	171.2032	296.3215	0	1995
picture_re-w	566	40.11307	68.98843	0	495

Figure 1 Descriptive analysis of the data

```
. regress sale price comment review picture_review
```

Source	SS	df	MS	Number of obs =	566
Model	26691104.7	4	6672776.18	F(4, 561)	= 40.89
Residual	91541930	561	163176.346	Prob > F	= 0.0000
				R-squared	= 0.2257
				Adj R-squared	= 0.2202
Total	118233035	565	209262.008	Root MSE	= 403.95

sale	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
price	-.2658533	.0563874	-4.71	0.000	-.3766095 -.1550971
comment	.0608483	.0244583	2.49	0.013	.0128073 .1088893
review	.2429977	.1665851	1.46	0.145	-.084209 .5702045
picture_review	.0858389	.7217403	0.12	0.905	-1.331805 1.503482
_cons	373.1732	35.49987	10.51	0.000	303.4443 442.9021

Figure 2 Regression test of the data

```
. regress sale price comment
```

Source	SS	df	MS	Number of obs =	566
Model	26119132.9	2	13059566.5	F(2, 563)	= 79.82
Residual	92113901.8	563	163612.614	Prob > F	= 0.0000
				R-squared	= 0.2209
				Adj R-squared	= 0.2181
Total	118233035	565	209262.008	Root MSE	= 404.49

sale	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
price	-.2465614	.0554057	-4.45	0.000	-.3553886 -.1377343
comment	.0995247	.0093941	10.59	0.000	.081073 .1179765
_cons	363.8811	35.14132	10.35	0.000	294.857 432.9052

Figure 3 Regression test of the price and comment

```
. estat hettest ,rhs iid
```

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: price comment

chi2(2) = 23.10
Prob > chi2 = 0.0000

```
. estat imtest,white
```

White's test for Ho: homoskedasticity
against Ha: unrestricted heteroskedasticity

chi2(5) = 39.58
Prob > chi2 = 0.0000

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	39.58	5	0.0000
Skewness	14.30	2	0.0008
Kurtosis	3.93	1	0.0473
Total	57.82	8	0.0000

Figure 4 Heteroscedasticity test

```
. estat vif
```

Variable	VIF	1/VIF
comment	1.05	0.953925
price	1.05	0.953925
Mean VIF	1.05	

Figure 5 Multicollinearity test

```
. regress sale price comment,robust
```

Linear regression		Number of obs	=	566
		F(2, 563)	=	51.31
		Prob > F	=	0.0000
		R-squared	=	0.2209
		Root MSE	=	404.49

sale	Robust				
	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
price	-.2465614	.0383446	-6.43	0.000	-.3218774 - .1712454
comment	.0995247	.0128334	7.76	0.000	.0743176 .1247319
_cons	363.8811	28.48681	12.77	0.000	307.9277 419.8345

Figure 6 Linear regression model

```
. regress sale price comment ,beta
```

Source	SS	df	MS	Number of obs	=	566
Model	26119132.9	2	13059566.5	F(2, 563)	=	79.82
Residual	92113901.8	563	163612.614	Prob > F	=	0.0000
				R-squared	=	0.2209
				Adj R-squared	=	0.2181
Total	118233035	565	209262.008	Root MSE	=	404.49

sale	Coef.	Std. Err.	t	P> t	Beta
price	-.2465614	.0554057	-4.45	0.000	-.1694933
comment	.0995247	.0093941	10.59	0.000	.4035136
_cons	363.8811	35.14132	10.35	0.000	.

Figure 7 Standard regression analysis

1.2. Analysis on Taobao C2C Online Transaction without Communication

If buyers do not communicate with the sellers and do not read the comments, the buyers may face the problem of information asymmetric, because the buyers do not know whether the sellers are trustworthy[2]. Therefore, it could lead the result of buying fakes or the sellers won't ship the goods. In this situation, we had built game model to show the problem.

Taking taobao as the example, following assumptions for the model will be present for convenience.

- (1) The insurance company will bear the cost of shipping the goods back if the buyers don't want it; and the freight are same for each time[3].
- (2) The price of the good is P .
- (3) The cost of authentic good is C_1 , the cost of the fake good is C_2 .
- (4) The cost of the seller's products being returned is R , the technical cost of the seller's releasing false information is T , the loss cost of the seller's increasing return rate is a , and the probability of the seller's fraud being found is q .
- (5) The buyer's perceived value when purchasing the trustworthy merchant's product is $V(v > p)$, and the buyer's psychological loss when purchasing the fake inferior product is L .
- (6) The psychological loss of the buyer who returns the good faith product is L' .
- (7) The insurance fee is M .

Table 1 The game model of seller and buyer

			Seller	
			Honest	Dishonest
Buyer	Buy	Return	$U_{i1} = V - L' - M, U_{s1} = P - C_1 - a - R$	$U_{i2} = -L - M, U_{s2} = q(-a - T - R) + (1 - q)(P - C_2 - a - T - R)$
		Not return	$U_{i3} = V - P - M, U_{s3} = P - C_1$	$U_{i4} = -P - M, U_{s4} = q(-T) + (1 - q)(P - C_2 - T)$
	Not buy		$U_{i5} = 0, U_{s5} = 0$	$U_{i6} = 0, U_{s6} = -T$

As table1 shows, If the buyer chooses to buy and return the good, and seller wants to be honest, the profit for seller is $U_{s1} = P - C_1 - a - R$, the profit for buyer is $U_{i1} = V - L' - M$.

If the buyer chooses to buy and return the good, seller chooses to be dishonest, the profit for seller is $U_{s2} = q(-a - T - R) + (1 - q)(P - C_2 - a - T - R)$, and the profit for buyer is $U_{i2} = -L - M$.

If the buyer chooses to buy but not return the good, and seller wants to be honest, the profit for seller is $U_{s3} = P - C_1$, the profit for buyer is $U_{i3} = V - P - M$.

If the buyer chooses to buy but not return the good, and seller wants to be dishonest, the profit for seller is $U_{s4} = q(-T) + (1 - q)(P - C_2 - T)$, the profit for buyer is $U_{i4} = -P - M$.

If the buyer chooses not to buy and seller choose to be

honest, the profit for seller is $U_{s5} = 0$, the profit for buyer is $U_{i5} = 0$.

If the buyer chooses to not buy but the seller choose to be dishonest, the profit for seller is $U_{s6} = -T$.

The profit for buyer is $U_{i6} = 0$.

According to the analyze of the game model, the three situations will be present.

If $U_{s1} > U_{s2}$, the profit that seller choose to be honest will higher than the profit that seller choose to be dishonest. Therefore, $U_{s3} > U_{s4}$, $U_{s5} > U_{s6}$. And the profit that the buyer chooses to buy is greater than the profit that does not buy, so that the buyer chooses to return or not return the goods depends on the profit $U_{i1} = V - L' - M$ and $U_{i3} = V - P - M$. which is the psychological loss of the buyer's returning the authentic product L' and the good price P . If buyer don't satisfy from the authentic good and want to return, so $L' < P$, $U_{i1} < U_{i3}$, the Nash equilibrium is (buy and return, honest), If buyer are satisfy about the authentic good, so $L' > P$, the Nash equilibrium is (buy and not return, honest).

If $U_{s1} < U_{s2}$, which means the profit that seller choose to be honest will lower than the profit that seller choose to be dishonest. Therefore, $U_{s3} < U_{s4}$, $U_{s5} < U_{s6}$. If the seller chooses to be dishonest, the buyer will not buy, and only if the seller chooses to be honest, the seller will buy. This is a mixed strategic problem, and the buyer and the seller cannot achieve a win-win situation at the same time. There is no unique Nash equilibrium solution. Therefore, to achieve the long run profit and have good reputation, choose dishonest will Damage the seller's reputation, so the best strategy for seller is to be honest. Thus, the Nash equilibrium in this situation is (buy, honest).

If $U_{s1} = U_{s2}$, which means the profit that seller choose to be honest are same with the profit that seller choose to be dishonest. Because the goals for seller is to attach more consumers and have a good reputation, which all the effort is for the long run profit. Also the profit for consumer to buy this good is higher than the profit that don't buy this good, therefore, the only Nash equilibrium is (honest, buy).

In conclusion, the best strategy for seller is to be honest in all the situation, whether consumer return or not return the products, they could still have good reputation and do not have the risk been caught that the merchant is selling fake and inferior products.

1.3. Discussion

The paper has already analyzed the communication in C2C online transaction, it shows that without communications, consumers may buy fake and inferior goods and the rights and interests of consumers are infringed. The government and companies could increase and develop new penalty measures to reduce the fraud and dishonest happened to the consumers. As for Taobao, Taobao have the most non official seller in the online shopping website, Taobao could Further divide the evaluation system, replace the level evaluation such as "good, moderate, bad" to percentage level, so that consumer could present a more specific

comment about the product[6]. Moreover, Taobao could introduce specific evaluation system in different categories of goods. For example, Taobao could introduce such evaluation indexes as whitening degree, water supplementation and mildness in the evaluation of cosmetics[1]. Therefore, the consumers will find the cosmetics that fit for them and won't waste money and time for the cosmetic that not fit for them.

For the problem of fake and inferior products, Taobao should strengthen the supervision mechanism, so they could strengthen the audit of sellers publishing false information, so that it will increase the cost of being dishonest, and make the sellers tend to choose integrity strategy[4]. buyers could through learning and identify some commonly used skills of fraud (such as phishing site, etc.), increase the difficulty of the dishonest seller cheat customer. For the C2C transaction website, they could introduce a third-party platform to conduct the identification of fake and inferior products[5]. If the counterfeit products are found, sellers can apply to the C2C platform for compensation, thus protecting their own rights and interests.

In the future, because technology have developed more advanced, the C2C platform will have more cutting-edge solutions to protect the consumer's right and Interests. Electric business platform could use big data analysis, purchase, sampling and more measures to locate the online fake products, but at the same time also need linkage including industry and commerce bureau, Police department, Quality and Technology Supervision Bureau, State Intellectual Property Office (SIPO), customers, and merchants. So that it could achieve a long-term mechanism for coordinated anti-counterfeiting from online tracking and identification to offline investigation and law enforcement, and all departments will work together to eliminate the sources of counterfeiting and infringement[7]. In the same time of using big data method to crack down the Internet counterfeit products, it also needs to apply stricter identification and off-limits system to promote a law-abiding business atmosphere. In addition, the violation by merchants shall be recorded and Shared to a certain extent, so as to make counterfeiters pay due costs and force manufacturers and sellers to pay attention to the basic rules of honest and law-abiding operation.

2. CONCLUSION

Through the analysis of Taobao C2C online transaction that with or without communication in different method, which could present following conclusion. Communication is an indispensable part of online shopping; consumer will tend to buy the products in the seller that have more positive comments and more reviews. Consumer will give priority for the evaluation about the products, but not the price when they are choosing items to buy. Therefore, positive comments are correlate with the sales number of products. For Sellers, According to the analysis of the game model, whether buyers return or not return the products, they best strategy for them is to be honest, because they have the risk been found by the consumer and C2C website, Sellers

would also have great loss in selling fake and inferior goods, because if they were caught, the cost of releasing fraud information and the reputation of the sellers will both get strong impact on it. Being honest is the basic rule for a merchant. However, the paper still has some drawbacks, As the analysis of transaction with communication for example, the data analysis only focuses on only one category of the commodity, the sneaker market. Taobao have various of category, In the future research, the author could do more research in different field of products to make the research become more preciseness and get different solutions for different area. Besides that, although the research have consider some important index for sellers, it still need to have more specific consideration, like when did merchant begin to sell on the C2C platform, the promises of sellers, etc. Thus, in the future research could consider introducing these variables in the research range, so that the research could have more specific understanding of the factors that would influence the sales.

ACKNOWLEDGMENT

I would like to express my gratitude to all those who helped me during the writing of this thesis. I gratefully acknowledge the help of Professor Putterman, who has offered me valuable suggestions in the academic studies. I also owe a special debt of gratitude to my friend Evan, who has given me some useful instructions on collecting data. Finally, I should like to express my gratitude to my beloved parents who have always been helping me out of difficulties and supporting without a word of complaint.

REFERENCES

- [1] Ou Jianli. Game Theory Analysis of integrity in Cosmetics C2C online shopping [D]. Shanghai University,2013.
- [2] Li Zhiting, Yu Huiling, Gu Yuan. Research on information Asymmetry in Online Shopping -- A Case study of Taobao.com [J]. Shang,2015(28):216+208.
- [3] Zhao Jinlu, Zheng Tingyu. Game model analysis of two parties in taobao purchase and transaction under C2C mode [J]. Journal of management,2014,27(03):58-61.
- [4] Lou Yingjun. Game Analysis and Countermeasures on the Integrity of C2C E-commerce Transactions under the New Situation [J]. Productivity Research, 2018(03):18-22+28.
- [5] Yang Liqing. Research on spatial Differentiation of Taobao store service quality in China and its influencing factors [D]. Northeast Normal University,2018.
- [6] Pang Xin. Research on the Influence of Dynamic scoring on sellers' sales on Taobao based on Evolutionary Game [D]. Shandong University,2018.
- [7] Wang Yifan. Research on existing Problems of Taobao Platform based on Game Theory [J]. Market Modernization,2020(02):16-17.