경영학 석사학위논문

An Analysis of Hit List Effects: Field Experiments in an Electronic Marketplace

“인기 상품 목록”의 효과: 전자상거래에서 현장 실험을 통한 분석

2014년 2월

서울대학교 대학원

경영학과 경영학 전공

한 동 교

Abstract

An Analysis of Hit List Effects:

Field Experiments in an Electronic Marketplace

Tongyo Han

College of Business Administration

The Graduate school

Seoul National University

Electronic marketplace has recorded tremendous growth, and apparel sales are rising rapidly. Two sequential field experiments in an apparel online shopping mall on *Taobao*, the most dominant marketplace in China, are carried out to analyze the effects of online promotion on hit list. This study conduct one pilot test for 17 products and second experiment for 290 products and record the daily sales for each by posting selected products on hit list. Difference-in-Differences method and Propensity score matching method are used to analyze the data. D-in-D estimation is effective both in identifying causality and precisely measuring its impact. In the settings of e-commerce web sites, the method identifies the difference brought by the treatment between two groups.

This study finds that once the products are displayed on hit list, product sales increase by 1.8 units per day. Utilizing Propensity score matching method, the sales promotion effect was measured as approximately 0.99 to 1.05 units depending on the matching algorithm. Furthermore, our analysis shows that the treatment is more effective when the price of a product is more than 200 and less than 300 Chinese RMB and the cardigan and T-shirt categorical groups show the larger sales promotion.

This study empirically proved the effect of hit list in online shopping mall and estimated the effect of it. This study can also provide extremely practical implications and give better insights into the understanding of the impact of online promotion on hit list.

Keywords: Online marketplace, Hit list, Difference-in-differences, propensity score matching

Student Number: 2012-20533

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURE

CHAPTER 1 INTRODUCTION

The electronic marketplace is growing at such a pace as to replace traditional marketplaces. One of the unique characteristics of e-commerce is the separation of buyers and sellers, and consequently consumers are exposed to a high degree of uncertainty (Lucking-Reiley, 2000; Ba et al., 2003; Pavlou et al., 2007; Sun, 2006). The purpose of this study is to identify the effects of 'Hit list' on sales using the Difference-in-Difference method and propensity score matching method based on a micro-transactional data set of an online shopping mall in Taobao's Tmall.

Taobao, a Chinese electronic marketplace, enables retailers and individuals to sell products to buyers on the Internet. As the B2C (Business to Consumer) and C2C (Consumer to Consumer) markets grow rapidly, total sales on Taobao for 2012 exceeded 1 trillion yuan, equivalent to USD160 billion (Forbes, 2011; ChinaInternetWatch, 2012). Total online sales in China are expected to eclipse those in the United States, rising to $356.1 billion in 2016 (Mozur, 2013). Taobao, China's biggest e-commerce site, has expanded from C2C to include B2C in order to meet the need for increasingly high quality services by establishing sales channel Tmall, in which thousands of companies are located, and thus provide one-stop quality e-services.

The Chinese fashion industry continues to expand at an exponential rate and has undergone extraordinary change within the past decade. Fashion sales in China are expected to reach US$27 billion by 2015 which is almost one fifth of the

world's total, up from US$10 billion in 2009, to become the largest fashion industry by country within the next five years (The World of Chinese, 2013). Even only considering the mainland China ranked sixth in the world for spending on luxury goods according to study of Bain & Company in 2011, and in 2010, it was a US$17.7 billion market. With China's fashion sector possessing such great potential along with the emergence of e-commerce in China, the online fashion industry is also showing tremendous promise, as Chinese women in particular begin to catch up fashion-wise to their female counterparts elsewhere in the world.

Several specific issues regarding the fashion online business exist, and inventory issues haves been investigated through a number of studies. Lemieux et al. (2012) assert that fashion consumers are demanding a greater variety of designs, better quality and service, including both faster delivery and reliability, and thus, achieving new modes of evaluation and innovative processes are important concerns to ensure customer satisfaction.

A significant amount of research has been undertaken to understand E-services from various IS perspectives such as website design (Palmer, 2002; Schlosser et al., 2006) and customer service (Ba and Johansson, 2008). Delivery time and service quality in both offline and online transactions have been studied for some time (Li, 1992; Kumar and Kalwani, 1997; Chen and Hitt, 2002; Ho and Zheng, 2004; Ba and Johansson, 2008; Cenfetelli et al., 2008; Rao et al., 2011; Luo et al., 2012).

In e-marketplace, consumers are exposed to a higher degree of uncertainty (Lucking-Reiley, 2000; Ba et al., 2003; Pavlou et al., 2007; Sun, 2006) than in the

traditional marketplace. Online promotion is one of the critical issues of uncertainty in the area of e-services (Ba and Johansson, 2008; Cenfetelli et al., 2008). However, it is not easy to find literature on the impacts of hit list in the empirical context of exogenous causality.

The purpose of this study is to identify the effects of hit list on sales using the Difference-in-Difference method and propensity score matching based on a micro-transactional data set of an online shopping mall in Taobao's Tmall. This study sets up two experiments to identify the causal effects of distinction between two product groups, while avoiding the issue of omitted variables by using the Difference-in-Differences method. To make another precise estimation, and verify the accuracy of results from Difference-in-Differences method, this study also implemented propensity score matching which allows comparison of treatment group and selected control group which are selected in the criteria of similar characteristics of the treatment group.

CHAPTER 2 THEORETICAL BACKGROUND

Information from the perspective of online shopping services, several papers on hit list stand out. Ba and Johansson (2008), and Cenfetelli et al. (2008), present the online promotion as a critical feature of e-commerce websites. Herd behavior resulting from the online promotion is particularly conspicuous in the economics and IS field. Users of computer frequently adopt popular software products subsequently making them even more prevailing (Brynjolfsson and Kemerer 1996). As an illustration, a book's appearing on the New York Times bestseller list results in modest growth for book sales as shown in Sorensen's (2007) empirical analysis. An experiment with a bestseller list on a menu in a Chinese restaurant (Cai et al. 2009) shows that the demand for a popular dish increases when the dish's popularity ranking is revealed to customers. Duan et al. (2009) empirically examined the impact of ranking information in the context of software adoption, while Ghose et al. (2009) found empirically that the monetary value of a click by user is not uniform throughout all positions in a search result. Furthermore, Grahl et al (2013) analyze the causal effect of social recommendations on online shopping through a randomized field experiment. Grahl et al(2013) investigated how social recommendations influence shopping behavior using a field experiment, and found that social recommendations and "Likes" caused increase of revenue about 12.97%

There are gender differences in online shopping behavior in relation to

uncertainty. Women perceive a higher level of risk which is related to the probability of negative outcomes by loss of privacy, credit card misuse, product failure, and shipping problems (Garbarino, 2004). In addition, Kim and Kim (2004) show that female consumers show greater online purchase intentions for clothes, jewelries and accessories than do male consumers. These studies imply that women tend to be more sensitive to e-commerce service regarding the risk of uncertainty including the characteristics of products purchased and also delivery or other services.

In addition women tend to buy more apparel in online than men do. Given that this research experiment has been conducted in an apparel e-marketplace, where most of the customers are women, the results can extend the findings of former studies. Since the gender of the customer of online fashion shopping mall which this study have analyzed is usually female, this research can expect that the study will make a further understanding of online shopping behavior of women.

This study is closely related to the Difference-in-Difference method introduced by Card and Krueger (1994), which analyzes the effects of minimum wage on employment in two U.S. statges. The method is extremely useful in identifying the causal effects of treatment between two separated groups, while avoiding the issue of omitted variables (Angrist and Pischke, 2009). Pischke (2007) investigates the effects of school term length on academic performance in regard to a sudden policy change in Germany. Angrist and Evans (1999) examine the effects of changes in state abortion laws on teen pregnancy with the method. Reinstein and Snyder (2005)

make use of a Difference-in-Difference approach to identify the effects of movie critics on sales.

In the field of IS, Rishika et al. (2013), regarding the intensity of the relationship between the firm and its customers, examined the effect of customers' participation in a firm's social media efforts and quantified the impact of social media participation on profitability of customer.

CHAPTER 3 METHODOLOGY

3.1 Research Design

This study uses two datasets taken from two experiments conducted with a Korean costume seller in *Taobao*'s Tmall where vendors are located in Korea. Normally products are displayed by the registration time from the top. Meanwhile, products added to the *Tenli* list (which is consisted with Hit items and labeled as "Hit list") are listed outstandingly in the separate section in order to draw attentions on the trendy products. This study assumes that, if the products on the *Tenli* list(Hit list) are selected randomly, both the treated and controlled groups have the same characteristics. This model of treatment corresponds to the case in which one group of restaurants in New Jersey adopted a minimum wage system while another group in Pennsylvania did not (Card and Krueger 1994). For the robustness checks, this study implements the propensity score matching methods as well.

3.2 Difference-in-Differences Model

Let S_i be the sales revenue for products $i = 1, \ldots,$ with i measured over the time period D_i. Consumers are likely to seek determinants of delivery service quality prior to purchase by checking whether the products are on the Hit list, denoted as T_i. The price of product i is denoted as P_i. Assume these variables influence consumer demand, and thus product sales, according to the following equation (Angrist and Pischke, 2009):

$$S_i = \beta_0 + \beta_1 Dummy_i * T_i + \beta_2 Dummy_i + \beta_3 T_i + \delta P_i + \varepsilon_i \quad (1)$$ [1]

where β_0, β_1, β_2, β_3 and δ are coefficients and ε_i is an error term. The variable $Dummy_i$ is a dummy equal to zero if the product i is not on the Hit list T_i before the end of the period D_i, and equal to one if T_i is on the list.

Simply, this study implements the Difference-in-Difference method as an ordinary least square estimator. This study has an outcome of interest S_i, an indicator whether an observation belongs to the treatment or control group: $T_i = 1$ treated, $T_i = 0$ control, and (at least) two time periods denoted by $D_i = 0$ and $D_i = 1$. Where the treatment, changes take place between the two periods, the Difference-in-Difference estimator is:

D-in-D estimator

$$= [E(S|D=1, T=1)- E(S|D=1, T=0)] - [E(S|D=0, T=1)- E(S|D=0, T=0)]$$

$$= [(\beta_0 + \beta_1 + \beta_2 + \beta_3) - (\beta_0 + \beta_2)] - [(\beta_0 + \beta_3) - \beta_0]$$

$$= \beta_1 \quad (2)$$ [2]

[1] Angrist, J.D., Pischke, J.S., 2009. "Mostly Harmless Econometrics: An Empiricist's Companion," Oxford: Princeton University Press

The estimator of β_1 is the difference between the effects of being not on the list of Tenli("Hit list") before the end of period D_i and being on the list after D_i.

3.3 Propensity score matching

Propensity score matching is one of the most popular approaches to estimate causal treatment effects (Heckman et al., 1997; Greene, 2003). It can be applied in diverse social situation such as evaluating labor market policies. This study estimates not only the treatment effect between two groups but also estimates their standard errors by implementing propensity score matching.

Clearly, it is not possible or available to observe both outcomes for the same individual or product at the same time. In the case of this research, there cannot be a record of same product with both case; one with listed in the 'hit list' and the other with not listed in 'hit list. Taking the mean outcome of non-treatment group as an approximation is not advisable, since the treatment group and control group (non-treatment group) usually differ even in the absence of treatment.

A simple example can be driven from this case, where popular products, which this popularity is due to its own characteristics, have a higher probability of selling more than those which have less popular characteristics. The matching

---

[2] Angrist, J.D., Pischke, J.S., 2009. "Mostly Harmless Econometrics: An Empiricist's Companion," Oxford: Princeton University Press

approach is one of the possible solutions to this problem –"Selection Bias". Its fundamental idea is to select in a large group of control groups (non-treatment groups) those individuals or products who or which are similar to those of treatment group in all relevant pre-treatment characteristics. By executing this process, differences in outcomes of this well selected and thus adequate control group and of participants can be attributed to the treatment.

Guo and Fraser(2010) introduces this idea with more precise modification. 'Average Treatment effect on the Treated' (ATT) is the parameter which can be interpreted as the "Treatment Effect" without selection bias and this can be defined as:

$$\tau ATT = E(\tau|D = 1) = E[Y(1)|D = 1] - E[Y(0)|D = 1] \quad (3)$$ [3]

Considering the counterfactual mean for those being treated - $E[Y(0)|D = 1]$ – cannot be observed, the proper substitute has to be chosen in order to estimate ATT. The true parameter τATT can be identifed, under the condition of:

$$E[Y(0)|D = 1] - E[Y(0)|D = 0] = 0: \quad (4)$$ [4]

[3] Guo, S. and Fraser , M.W. (2010), The propensity Score Analysis : Statistical Methods and Application. SAGE Publication Inc

[4] Guo, S. and Fraser , M.W. (2010), The propensity Score Analysis : Statistical Methods and Application. SAGE Publication Inc

In most of the social experiments where assignment to treatment is usually cannot be random, this can cause the major problem. Regarding the fact that it is not possible to measure the treatment effect separately, because the treatment group and control group can be different in the first place, considering the fact that one group is treated by “treatment” and other group is not treated by “treatment”, propensity score matching can be appropriate solution by selecting the ones in the control group which is similar with those of in the treatment group and make an assumption for equation (4).

In conclusion, the propensity score matching estimator is the mean difference in outcomes over the common support. This study implements three algorithms among “greedy matching”; the Nearest Neighbor matching, Kernel matching and Stratification matching. The price variable is used as pre-treatment characteristics to estimate propensity score.

CHAPTER 4 DATA AND EXPERIMENTS

In the first pilot experiment, 17 products from 3 vendors were selected. Nine products were assigned to the Hit list, while 8 products were observed as a control group between August 22 and September 23, 2012. For each product, the number of daily sales was recorded. In the second experiment, 290 products were selected from 8 vendors. Forty-five products were chosen for the treatment group, while 245 products were taken as a control group from October 9 to December 7, 2012. Products were randomly assigned to two different experimental conditions which remained consistent throughout the study.

Table 1. Descriptive Statistics			
		Treatment Group	Control Group
First Experiment			
Total Number of Products		9	8
Frequency		52.94	47.06
Cumulative Number of Sales	Mean	64.3333	25
	Standard Deviation	28.1203	12.9173
	Minimum	34	12
	Maximum	122	51
Price (Chinese RMB)	Mean	187.768	179.865
	Standard Deviation	66.5811	55.4433
	Minimum	83.6867	94.8449
	Maximum	277.84	256.081
Category	Dress	2	2
	Shirt	5	3
	T-shirt	2	3
Second Experiment			
Total Number of Products		45	245
Frequency		15.52%	84.48%
Cumulative Number of Sales	Mean	28.41333	8.261633
	Standard Deviation	22.5636	6.5098
	Minimum	3	0
	Maximum	151	37
Price (Chinese RMB)	Mean	265.4049	258.9182
	Standard Deviation	149.4593	122.374
	Minimum	99.68	79.05
	Maximum	893.41	793.59
Category	Cardigan	5	22
	Knit	5	25
	Shirt	5	20
	T-shirt	12	42
	Others	18	36

Table 1 illustrates data descriptions of 17 and 290 products, respectively, in each experiment. Using the first experiment as a pilot, the marketing team manager picks 17 products which are supplied by three “friendly” vendors. Figure 1 shows the average cumulative sales before and after the first experiment which applies that it should make the approach of Difference-in-Difference in the second experiment.

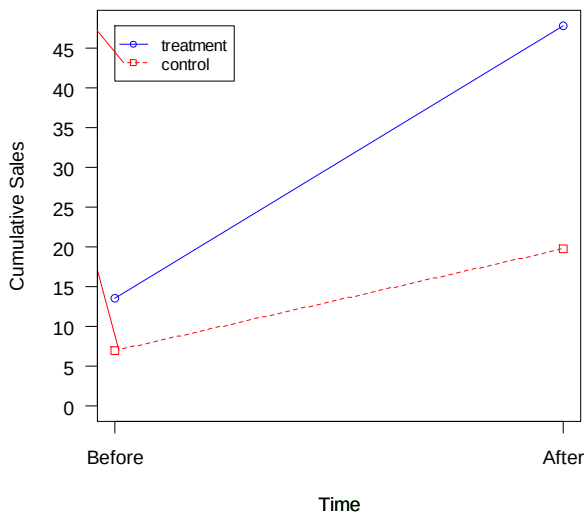


Figure 1. Treatment Effects in the First Experiment

Later the manager was asked to randomly assign the 45 products to the experiment group in order to eliminate any potential selection bias in the second experiment. The cumulative sales figures in the Table 1 let us assume that the Hit list treatment may increase the total number of sales eventually, though the price levels of two groups are not significantly different. This study basically compare

the cumulative sales for each 5 days before and after the treatment, as some of the products in the treatment groups get sold out after 5 days. Appendix 1 presents the detailed descriptions of the variables of the panel data set used in the analysis in the second experiment.

For the matching, this study used the same data from the prior analysis; the data of 2nd experiment. There are 45 products in the treated group and 245 products in non-treated group which means the former 45 products were displayed on the 'Hit list' and others haven't. The propensity score was calculated with the records and characteristics of these products and matching algorithms were used to estimate the effect of "Hit list" and this study become able to compare the results of the prior analysis which used Difference-in-Differences method.

CHAPTER 5 RESULTS

5.1 Difference-in-Differences Model

This study presents two sets of results for the first and second experiments: an OLS without D-in-D estimate and a pooled OLS specification. The results of the first experiment and estimations are in Table 2. This study compared the cumulative sales for each 5 days before and after the treatment, as some of the products in the treatment groups get sold out after 5 days. Week, brand, and category have been controlled for all estimations.

Table 2. Regression Results for the First Experiment

	Traditional Way (without D-in-D estimate)	Pooled OLS
Treatment	11.39*** (1.006)	7.601*** (1.356)
Period	0.0448 (1.641)	-5.341** (2.077)
D-in-D estimator		7.498*** (1.893)
Price(Chines RMB)	-0.121*** (0.0185)	-0.122*** (0.0177)
Week Control	Yes	Yes
Brand Control	Yes	Yes
Category Control	Yes	Yes
Constant	29.19*** (5.708)	33.04*** (5.546)
Observations	170	170
R-squared	0.798	0.816

The dependent variable represents the number of cumulative sales. The standard errors are reported in parentheses.

* Significant at 10% level. ** Significant at 5% level. *** Significant at 1% level.

The result of the second experiment is as following table 3. The D-in-D estimator shows approximately 6.325 more units were sold for 5 days which is 1.1 less than the result of the first experiment which showed 7.498 more units for 5 days.

Table 3. Cumulative Number of Sales for the Second Experiment

	Traditional Way (without D-in-D estimate)	Pooled OLS
Treatment	17.81*** (0.854)	14.66*** (0.899)
Period	4.521 (3.159)	1.101 (3.279)
D-in-D estimator		6.325*** (1.549)
Price (Chinese RMB)	-0.00627*** (0.00156)	-0.00624*** (0.00155)
Week Control	Yes	Yes
Vendor Control	Yes	Yes
Category Control	Yes	Yes
Constant	-7.179*** (1)	-4.017*** (1.301)
Observations	2,900	2,900
R-squared	0.598	0.605

The dependent variable represents the number of cumulative sales. The standard errors are reported in parentheses.

* Significant at 10% level. ** Significant at 5% level. *** Significant at 1% level.

In pooled OLS the D-in-D estimator is statistically significant in both experiments; this study can also examine how the effect of the hit list has been alleviated with D-in-D estimator compared to those resulted from traditional OLS method. As D-in-D estimator measured, this study can make a conclusion that products on the hit list is applicable are sold at a rate of approximately 1.8 more units over a day than are goods for which listing on the Hit list is not applicable.

The effectiveness of hit list varied among the characteristics of products. This study implement the same D-in-D estimation among the groups for similar characteristics in “price” and “categories”. (Table 4, Table 5.) This study reaches the conclusion that the effectiveness of hit list is most apparent when the price of a product is more than 200 and less than 300 Chinese RMB.

Table 4 Treatment Effects by Price

	Price <200	200<=Price<300	300<=Price<500	500<=Price
Period	-2.52 (1.629)	-6.178** (2.588)	0.848* (0.47)	0.366 (0.319)
D-in-D estimator	6.334*** (1.813)	10.69*** (2.955)	4.480*** (1.313)	1.839 (1.09E+00)
Observations	1,040	1,060	630	170
R-squared	0.394	0.509	0.406	0.397
Number of products	104	106	63	17

The dependent variable is the number of cumulative sales. The standard errors are reported in parentheses.

* Significant at 10% level. ** Significant at 5% level. *** Significant at 1% level.

In the categorical group, the cardigan group shows the largest sales promotion for the online promotion of hit list with the coefficient of 10.26. The results show that Cardigan and T-shirt have the stronger promotion effects than Shirt and Knit when this study compares the statistically significant coefficients of D-in-D estimators.

Table 5. Treatment Effects by Category

	Cardigan	Knit	Shirt	T-shirt
Period	-2.086* (1.062)	-0.218 (0.583)	0.1 (0.347)	-5.637*** (1.821)
D-in-D estimator	10.26*** (1.701)	4.368*** (1.242)	3.990*** (1.309)	6.275** (2.421)
Observations	270	300	250	540
R-squared	0.683	0.404	0.516	0.462
Number of products	27	30	25	54

The dependent variable is the number of cumulative sales. The standard errors are reported in parentheses.

* Significant at 10% level. ** Significant at 5% level. *** Significant at 1% level.

5.2 Propensity score matching

The propensity score must be calculated before the matching because it has to be used in the matching algorithm. To calculate the propensity score this study included the variables for control which are unaffected and unrelated by Hit-list such as price of the products. This study used price of the products as a matching estimator to calculate propensity score.

Table 6. Propensity Score

Inferior of block of pscore	Non-hit list	Hit list	Total
0.0301	203	15	218
0.2	16	11	27
0.4	5	6	11
0.6	3	4	7
0.8	1	9	10
Total	228	45	273

In the prior analysis using D-in-D estimator, this study had a result of 6.325 more product sales during 5 days which means the product in “Hit-list” sell 1.26 more units per day than other products which are no listed in “Hit-list”.

In this analysis, even though there are slight differences of ATT (Average Treatment effect on the Treated) depending on which algorithm this paper uses, the

results are fairly consistent to each other and also with that of Difference in Differences estimation. If Nearest Neighbor matching algorithm is used, 29 products were chosen as a reasonable control group to be compared with the treatment group, and the result showed treatment effect is 5.289 which is similar with 6.325 but little less. The reason for this difference will be the regulation effect that propensity score matching have. In difference-in-differences method, this study used the whole 290 products as a subject of analysis but in propensity score matching the model chooses the proper control group based on the propensity score. This makes the analysis of propensity score matching more sophisticated eliminating exogenous variable problem.

This lessens the problem occurred due to the other differences between the treatment group and control group except for the difference this research want to measure. As you can check on the table 7, the result of Kernel matching and Stratification matching is also similar with those of other estimator.

Table 7. Propensity Score matching results

Algorithm	Number of treatment	Number of Control	ATT	Std.Err	t	p
Nearest Neighbor Matching	45	29	5.289***	1.834	2.884	0.004
Kernel Matching	45	228	4.990	2.553	1.954	0.052
Stratification Matching	45	228	5.040*	2.344	2.150	0.0324

ATT*: Average Treatment effect on the Treated

CHAPTER 6 CONCLUSION

This study examines the effects of “Hit list” information on sales by a Korean costume market in Taobao. Experiments were set up to identify the causal effects of treatments between two separated groups in order to avoid the issue of omitted variables. Using the actual transactional data set of a store in Tmall, this research makes use of the Difference-in-Differences method and propensity score matching to measure the impact of “Hit list”. Specifically, products for which are listed in “Hit list” sold at a rate of 6.4 more units over 5 days after the treatment than normal goods to which such were not listed in “Hit list”. Using propensity score matching, the effect of “Hit list” on sales estimated the result for 4.9 to 5.3 more units over 5 days. Consumers are likely to recognize the information of “Hit list” as a positive signal regarding the purchase.

This study provides extremely practical implications for business. Through this research, it verified the model of Difference-in-Difference to identify the effects of “Hit list” on sales and also implemented propensity score matching to verify if the result can be strengthened. The first pilot experiment convinced the company to undertake another follow-up experiment to verify whether “Tenli” treatment which means displaying the products on the “Hit list” is indeed effective. This study reported the initial findings from both experiments to the company, which is now applying the “Tenli” treatment to a greater number of products, as they witnessed a revenue increase following the two experiments. At the end of the day, the

company has become able to expect the “Hit list” has a promotion effect on sales with actual estimated number of sale units.

The results of this study show that “Hit list”, which provides the information that the products in the list are popular and selling well, promotes the customer to purchase the products more likely. This raises the interesting question of how other characteristics in e-commerce websites may influence transactions.

This study’s results are consistent with previous research related to this study. An experiment with a bestseller list on a menu in a Chinese restaurant showed that the demand for a popular dish increases when the dish’s popularity ranking is revealed to customers. (Cai et al. 2009) The study by Grahl et al (2013) analyzed the causal effect of social recommendations on online shopping through a randomized field experiment and the result is also consistent with this study. Furthermore, these results have implications for business opportunities as improvements in China’s logistics infrastructure increase e-commerce transactions.

It should also be emphasized that D-in-D estimation is very useful both to identify the causality and measure its impact precisely. In the settings of e-commerce web sites, any treatment between two groups can be analyzed with this method as long as the operator agrees on the execution of the experiments. This study has pursued the methodology to mitigate the methods to identify the true estimator in the long run.

The robustness check with propensity score matching methods also fortified the conclusion of this study. Propensity score matching method can make more

rigorous estimation. In the settings of e-commerce web sites, any treatment between two groups can be analyzed with this method as long as the operator agrees on the execution of the experiments.

Certain limitations regarding this study still remain. First of all, the products in the “Hit list” were serviced with 24 hour delivery service, which can also make a promotion effect on sales. At the moment, the results by category come from the treated and controlled groups which included a relatively small number of products for each category. Furthermore, the products in this study have a unique feature of seasonality, and thus it is necessary to apply the method to other types of products in order to secure the ability to generalize by increasing the number of products in both groups in future studies. Finally, there are additional research opportunities to identify the causal factors of hit list by implementing the models with moderating variables.

For further research, given the fact that the experiments were conducted in China, it will be interesting to see if other cultures or countries show results that are consistent with those reported here. This study contributes to the landscape of IS studies by extending the applications of the Difference-in-Difference and propensity score matching methods in the e-commerce domain.

REFERENCES

Angrist, J.D., Evans, W.N., 1998. “Children and Their Parents’ Labor Supply: Evidence from Exogenous Variation in Family Size”, *American Economic Review*, 88, pp. 450–477.

Angrist, J.D., Pischke, J.S., 2009. “Mostly Harmless Econometrics: An Empiricist’s Companion,” Oxford: Princeton University Press

Ba, S., Johansson, W.C., 2008. “An Exploratory Study of the Impact of e-Service Process on Online Customer Satisfaction,” Production and Operations Management (17:1), January-February, pp.107-119.

Ba, S. Whinston, A.B., Zhang, H. 2003. “Building trust in online auction markets through an economic incentive mechanism,” Decision Support Systems (35:3), pp. 273–286.

Bain & Company, 2011, “2011 China Luxury Market Study”, *Bain Point of View* , December ,http://www.bain.com/Images/2011%20Bain%20China%20Luxury%20Market%20Study.pdf /(accessed May 2013)

Brynjolfsson, E. and Kemerer, C. F. Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Market. Management Science. 42(12), 1996.

Cai, H., Chen, Y. and Fang, H. Observational Learning: Evidence from a Randomized Natural Field Experiment. American Economic Review. 99(3),

2009, pp. 864-82.

Card, D., Krueger, A., 1994. “Minimum Wages and Employment: A Case Study of the FastFood Industry in New Jersey and Pennsylvania,” American Economic Review (90:5), December, pp.1397-1420

Cenfetelli, R.T., Benbasat, I., Al-Natour, S., 2008. “Addressing the What and How of Online Services: Positioning Supporting-Services Functionality and Service Quality for Business-to-Consumer Success,” Information Systems Research (19:2), June, 161-181.

Chen, P.Y., and Hitt, L., 2002. ‘‘Measuring Switching Costs and the Determinants of Customers Retention in Internet-Enabled Businesses: A Study of the Online Brokerage Industry,’’ Information Systems Research (13:3), pp.255–274.

ChinaDaily, 2013, Alibaba sales 'to exceed entire US e-commerce, http://europe.chinadaily.com.cn/business/2013-09/14/content_16969703.htm

China Internet Network Information Center, 2013. “Statistical Report on Internet Development in China”, http://www1.cnnic.cn/IDR/ReportDownloads/201302/P020130221391269963814.pdf(accessed on Dec 2013).

Duan, W., Gu, B. and Whinston, A.B. Informational Cascades and Software Adoption on the Internet: An Empirical Investigation. MIS Quarterly. 33(1), 2009, pp. 23-48.

Forbes, 2011, “Pulling Back the Curtain on Taobao and Yahoo!'s Hidden Value,”

Jackson, E. http://www.forbes.com/sites/ericjackson/2011/05/03/pulling-back-the-curtain-on-taobao-and-yahoos-hidden-value/ (accessed on March 2013).

Garbarino, Ellen., Michal Strahilevitz., 2004. "Gender differences in the perceived risk of buying online and the effects of receiving a site recommendation." Journal of Business Research (57:7), pp.768-775.

Ghose, A. and Yang, S. An empirical analysis of search engine advertising: Sponsored Search in Electronic Markets. Management Science. 55 (10), 2009, pp. 1605-1622

Grahl, J., Rothlauf, F. and Hinz, O., 2013 “How do social recommendations influence shopping behavior? A Field experiment.” Working Paper.

Greene, W. H., 2003, “Econometric Analysis,” New York University, New York.

Guo, S. and Fraser , M.W. (2010), The propensity Score Analysis : Statistical Methods and Application. SAGE Publication Inc

Heckman, J., H. Ichimura, and P. Todd, 1997. “Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme,"Review of Economic Studies, 64

Ho, T., Y. Zheng. 2004. “Setting customer expectation in service delivery: An integrated marketing-operations perspective,” Management Science (50:4), pp.479–488.

Kim, Eun Young, and Youn-Kyung Kim., 2004. "Predicting online purchase

intentions for clothing products." European journal of Marketing (38:7), pp.883-897.

Kumar, P., Kalwani, M. U., Dada, M. 1997. “The impact of waiting time guarantees on customers’ waiting experiences,” Marketing Science (16:4), pp.295–314.

Lemieux, Andrée–Anne. 2012 "A new analysis framework for agility in the fashion industry." International Journal of Agile Systems and Management 5.2, pp.175-197.

Li, L. 1992. “The role of inventory in delivery-time competition,” Management Science (38:2), pp.182–197.

Lucking-Reiley, D., 2000. “Auctions on the Internet: What’s Being Auctioned, and How?,” The Journal of Industrial Economics (48:3), pp.227-252.

Luo, J., Ba, S., Zhang,H., 2012, “The Effectiveness of online shopping characteristics and well-designed websites on satisfaction,” MIS Quarterly (36:4), pp.1131-1144.

Marco Cliendo, Sabine Kopeinig .,2005 “ Some practical guidance for the implementation of propensity score matching", IZA DP No.1588

Mozur, P., 2013. “Scooters Rule as E-Commerce Grows in China,” The Wall Street Journal, http://online.wsj.com/article/SB10001424127887323644904578269460030807102.html

Oriana, Luquetta., 2013, “The fashion industry and china”, The world of Chinese, March 21, http://www.theworldofchinese.com/2013/03/the-fashion-industry-and-china/ (accessed May 2013)

Palmer, J.W., 2002. “Web Site Usability, Design, and Performance Metrics,” Information Systems Research(13:2), June, pp.151-167.

Pavlou, P.A., Liang, H., Xue, Y., 2007. “Understanding and mitigating uncertainty in online exchange relationships: a principal- agent perspective,” MIS Quarterly (31:1), pp.105-136.

Pischke, J.S., 2007. “The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years,” Economic Journal (117), pp.1216–1242.

Rao, S., Goldsby, T. J., Griffis, S. E., Iyengar, D. 2011, “Electronic Logistics Service Quality e-LSQ): Its Impact on the Customer’s Purchase Satisfaction and Retention,” Journal of Business Logistics (32:2), pp.167-179.

Reinstein, D.A., Snyder, C.M., 2005. “The influence of expert reviews on consumer demand for experience goods: a case study of movie critics,” The Journal of Industry Economics LIII (1), March, pp.27-51.

Rishika R, Kumar A, Janakiraman R, Bezawada R., 2013. The effect of customers? Social media participation on customer visit frequency and profitability: An empirical investigation. Inform. Systems Res. 24(1):108–127

Schlosser, A.E., White, T.B., Lloyd, S.M., 2006. “Converting Web Site Visitors into

Buyers: How Web Site Investment Increases Consumer Trusting Beliefs and Online Purchase Intentions," Journal of Marketing (70:2), April, pp. 133-148

Sorensen, A. T., Bestseller Lists and Product Variety. The Journal of Industrial Econ. 2007.

Sun, Y., Han, Z., Yu, W., Liu, KJR., 2006. "A Trust Evaluation Framework in Distributed Networks: Vulnerability Analysis and Defense Against Attacks," IEEE Infocom.

The Wall Street Journal, 2013. "Scooters Rule as E-Commerce Grows in China,", http://online.wsj.com/article/SB10001424127887323644904578269460030807102.html.

Appendix 1: Treatment Effects by Price with Pooled OLS Model

	Price <200	200<=Price<300	300<=Price<500	500<=Price
Treatment	17.48*** (1.24)	22.18*** (1.297)	9.347*** (0.87)	6.153*** (0.733)
Period	-4.948*** (1.895)	-2.24 (1.788)	4.533* (2.346)	-0.343 (1.082)
D-in-D estimator	6.334*** (1.643)	10.71*** (1.701)	4.559*** (1.156)	1.770* (0.907)
Week Control	Yes	Yes	Yes	Yes
Vendor Control	Yes	Yes	Yes	Yes
Constant	-16.88*** (5.177)	-4.913 (5.089)	14.35** (6.967)	-1.633 (3.118)
Observations	1,040	1,060	630	170
R-squared	0.523	0.558	0.743	0.864

Appendix 2: Treatment Effects by Category with Pooled OLS Model

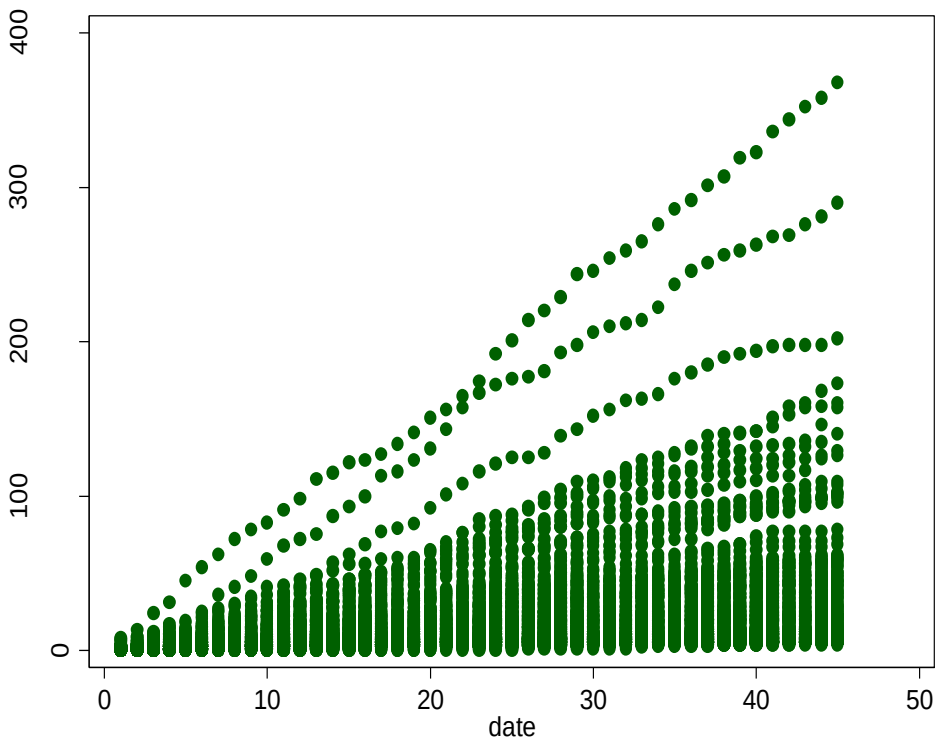
Variable	Cardigan	Knit	Shirt	T-shirt
Treatment	16.36*** (1.833)	9.169*** (1.053)	9.216*** (0.984)	15.48*** (1.291)
Period	-9.757** (4.511)	0.502 (4.322)	0.704 (3.732)	2.169 (2.979)
D-in-D estimator	10.66*** (2.315)	4.343*** (1.295)	3.942*** (1.111)	5.556*** (1.711)
Price	-0.0199** (0.00819)	-0.00501* (0.00283)	0.0123** (0.00508)	-0.0116 (0.00735)
Week Control	Yes	Yes	Yes	Yes
Vendor Control	Yes	Yes	Yes	Yes
Constant	5.768 (4.815)	8.999** (4.443)	2.901** (1.392)	-6.01 (8.105)
Observations	270	300	250	540
R-squared	0.713	0.767	0.788	0.608

Appendix 3: Variable Description

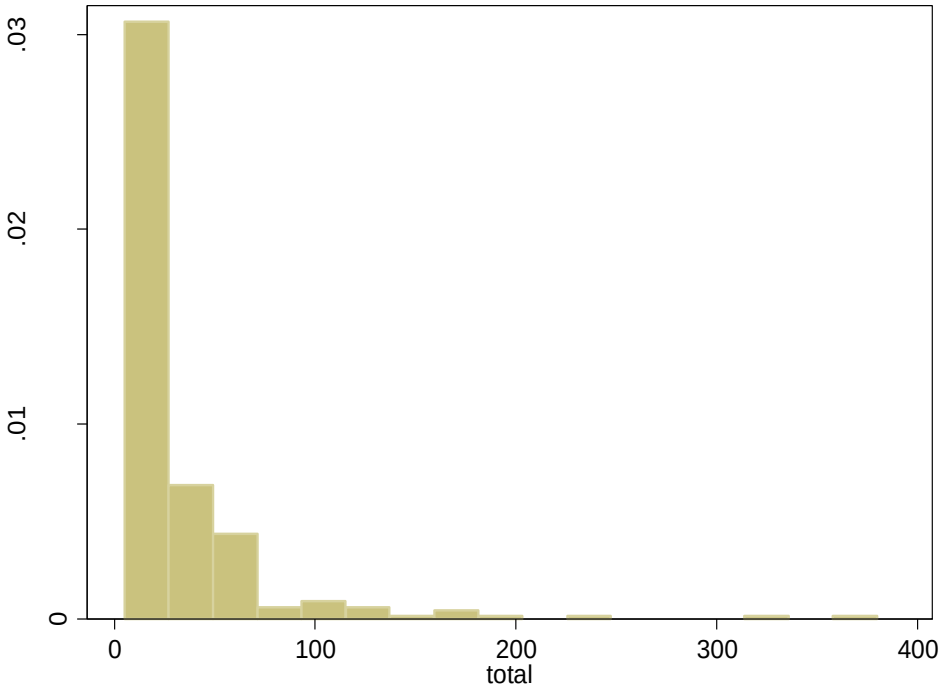
Definitions of Variables and Descriptive Statistics							
Variable	Definition	Unit	Observation	Mean	Standard Deviation	Maximum	Minimum
Number of Sales	Cumulative number of sales per product. The numbers are counted for a period of 10 days, 5 days prior to treatment and 5 days following	Number of products sold	2,900	11.38862	12.95788	151	0
Treatment	Dummy which indicates the product had Tenli treatment(listed on "Hit list" (providing 24-hour delivery service) or not	Dummy	2,900	0.1551724	0.362131	1	0
Period	The dummy for date, for which 1 indicates the product was sold in Tenli((listed on "Hit list"), and 0 indicating that the product was not sold in Tenli(Hit list)	Dummy	2,900	0.4982759	0.500083	1	0

*D-in-D estimator (Treatment*Period)*	Interaction term between treatment and period, which estimates the D-in-D parameter	Dummy	2,900	0.0775862	0.267566	1	0
Price	Product price	Chinese RMB	290	259.9248	126.6889	893.41	79.05
Price Control	Dummy for price range	Dummy	3,360	2.252976	1.087832	4	1
Week	Number of week of sales						
Week Control	Dummy for week	Dummy	2,900	3.481379	0.527281	5	1
Vendor Control	Product seller		8				
Category Control	Product category		19				

Appendix 4: Cumulative number of sales by day

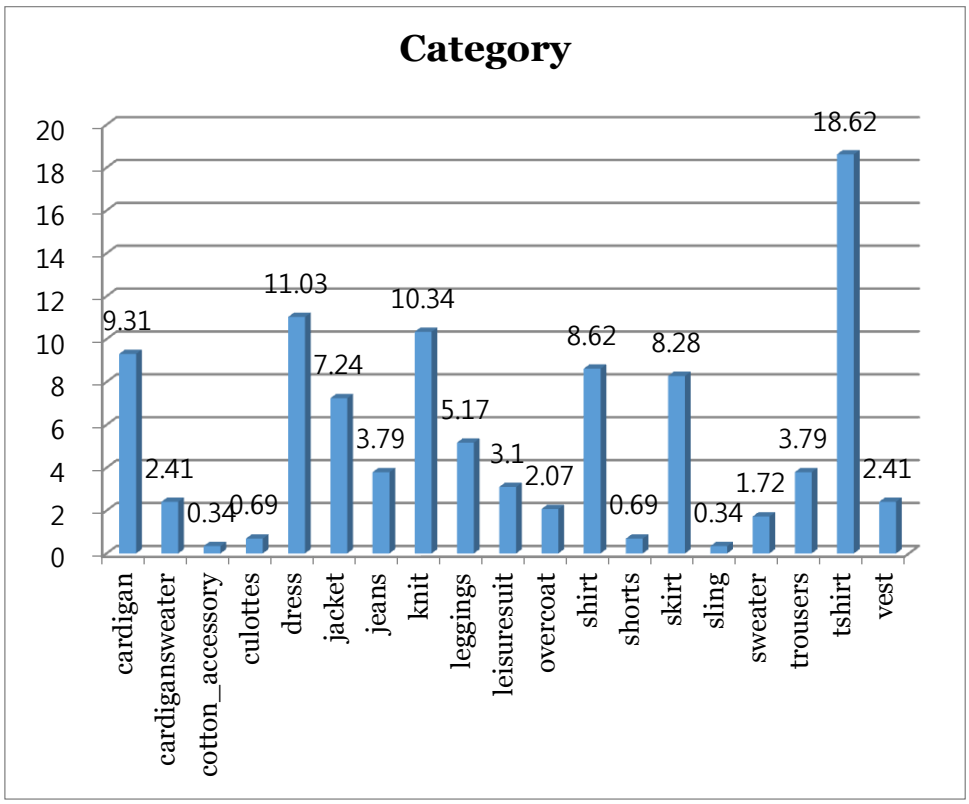


Appendix 5: Distribution of Total number of sales by product



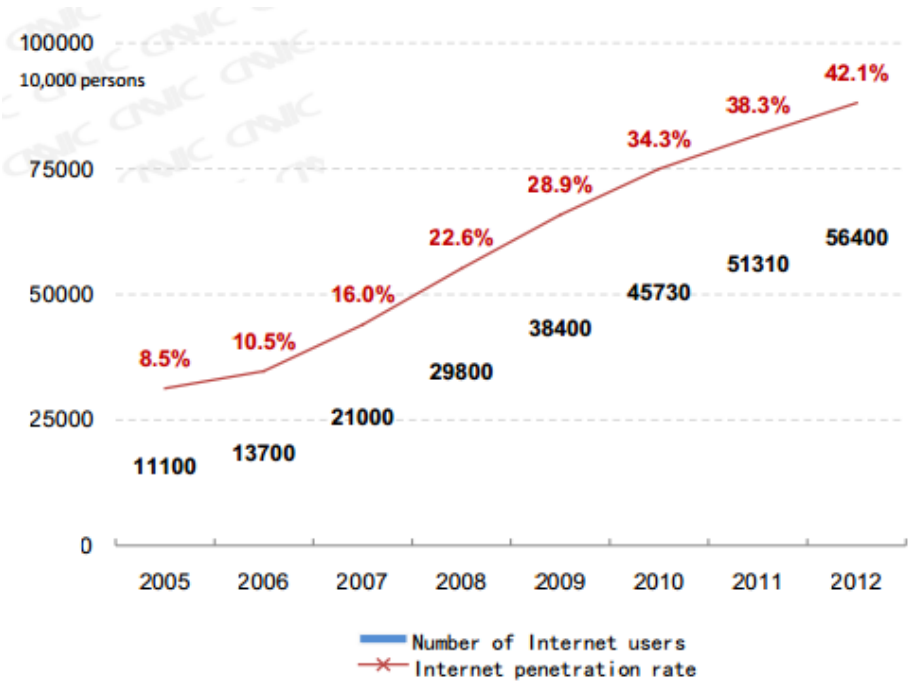
Variable	Obs	Mean	Std. Dev.	Min	Max
Total number of sales	290	30.55862	42.68552	5	380

Appendix 6: Distribution of Product category



Unit of scale: percent(%)

Appendix 7: Size of Chinese Internet users and Internet penetration rate



Source: CNNIC(China Internet Network Information Center) Statistical Survey on Internet Development in China

Appendix 8: Number of Users and Utilization Ration of Online shopping

Source : CNNIC(China Internet Network Information Center) Statistical Survey on Internet Development in China

Appendix 9 :Ranking of Best Sold Items over Internet in China

	Items	Purchased by % of Online Buyers
1	Apparel, accessories and footwear	57.1%
2	Books, CDs and DVDs	46.0%
3	Cosmetics	40.6%
4	Electronics (Cameras, MP3)	35.4%
5	Rechargeable cards for games and cell phones	32.1%
6	Home decorations	29.2%
7	Cell phones and electronic accessories	27.4%
8	Gifts and toys	24.7%
9	Computer, laptops and hardware	23.0%
10	Home appliances	22.1%

Source: iResearch report 2008

국문 초록

“인기 상품 목록”의 효과:

전자상거래에서 현장 실험을 통한 분석

전자상거래 시장은 등장한 이래 빠른 속도로 성장해왔으며, 그 거래 상품에 있어서 의류의 거래도 많은 비중을 차지하고 있다. 본 연구는 중국의 의류 온라인 쇼핑몰에서의 두 번의 현장 실험을 통해 “인기 상품 목록” 이 그 상품의 판매 증진 효과를 가져오는지 검증하고, 그 효과의 크기를 측정하고자 했다.

첫 번째 실험은 17개의 상품을 대상으로 진행되었다. 9개의 상품은 일반적으로 판매되다가 어느 시점에서 “인기 상품 목록” 이라는 정보가 제공되는 카테고리에서 판매되었으며, 8개의 상품은 “인기 상품 목록”에서 판매되지 않고 계속해서 일반적으로 판매되었다. 두 번째 실험은 같은 방법으로 총 290개의 상품을 대상으로 진행되었으며 그 중 총 45개의 상품이 “인기 상품 목록” 에서 판매되었다. 판매 증진 효과를 분석하기 위해서 Difference-in-Differences 방법과 Propensity score matching 방법이 사용하였다. 또한 판매 분석 기간은 “인기 상품 목록” 카테고리에서 판매되는 시점을 기준 전 5일과 후 5일로 한정하여 그 처치 효과를 더 효과적으로 보고자 하였다.

Difference-in-Differences 분석 결과, “인기 상품 목록” 에서 판매되는 처치가 있었던 실험군 상품들이 첫 번째 실험에서는 5일간 7.498개, 두번째 실험에서는 6.325 개 더 팔린 것으로 분석되었으며, Propensity score

matching 분석을 통해서는 매칭 알고리즘에 따라 4.99에서 5.24 개의 판매 증진 효과가 있는 것으로 분석되었다. Propensity score matching 분석을 사용하였을 때 그 판매 증진 효과가 적게 나타나는 것은 Difference-in-Differences 분석이 통제군 전체를 실험군과 비교한 것에 비해 Propensity score matching 은 통제군 전체에서 실험군과 유사한 일부 상품만을 통제군으로 새롭게 구분한 후 비교하였기 때문인 것으로 보인다. 또한 상품의 가격이나 종류에 따라서도 그 효과가 각각 다르게 나타나는 결과도 볼 수 있었다.

본 연구는 전자상거래에서 “인기 상품 목록” 이 판매 증진에 미치는 영향을 현장 실험을 통해 그 효과의 존재를 증명하고 그 크기를 측정하였다는 점에서 그 가치가 있다. 또한 분석 방법에 있어 Difference-in-Differences 분석과 Propensity score matching 분석을 사용하여 단순 회귀 분석의 한계를 극복하고 현장 실험의 이점을 활용하여 보다 효과적으로 “인기 상품 목록”의 효과를 측정하고자 하였다. 기존 연구에서 경험 데이터를 통해 의류 온라인 쇼핑몰에서“인기 상품 목록”의 효과를 직접적으로 측정한 적이 없었던 것에서 본 연구의 학문적 기여가 있으며 또한 “인기 상품 목록”의 정보가 주는 효과는 이와 관련된 의사결정을 하는 실제 비즈니스 주체에게 실용적으로 활용될 수 있을 것이다.

주요어: 전자상거래, 인기 상품 목록, 판매 증진 효과, 온라인 쇼핑몰

학 번: 2012-20533