

Segmentation of Electronic Commerce Delivery Service Based on a Stated Preference Survey

Hong-Bin Youn*, Minyoung Park**, Sangho Choo***

Abstract

This paper presents a case study examining delivery service preference of electronic commerce products. With the increasing of electronic commerce and home shopping market, parcel delivery service market has been growing annually by more than 10% in Korea. However, the quality of delivery service and profitability are gradually decreased due to the price competition. A survey conducted by Korea Chamber of Commerce and Industry revealed that nearly half of the respondents selected delivery accuracy and time as major factors in delivery service, implying that online consumers consider service factors such as delivery time more important rather than delivery cost. In this research, a stated preference survey with 16 scenarios was conducted for online consumers to examine delivery preference depending on the type and value of products that they purchased online. The analysis results show that online consumers have a willingness to pay extra cost for rapid delivery or want lower price although it is much slow. It implies that current uniformed delivery cost of online products can be differentiated with different prices for different consumers and products. In addition, the study shows that different online products have different value of time, 27 won per hour on average.

Key Words: *Electronic Commerce, Parcel Delivery, Stated Preference, Value of Time*

JEL classification : *L81 & L87*

* Graduate Student, Graduate School of Logistics, Inha University, Korea. Tel.: +82-10-7728-7789 / Email: iamconge@hanmail.net

** Corresponding Author: Associate Professor, Graduate School of Logistics, Inha University, Korea. Tel.: +82-10-3836-5070 / Email: mypark@inha.ac.kr

*** Associate Professor, Department of Urban Design and Planning, Hongik University, Korea. Tel.: +82-10-3326-6037 / Email: shchoo@hongik.ac.kr

1. Introduction:

With the increasing of electronic commerce and home shopping market, domestic delivery service market has been growing annually by more than 10% (KCCI, 2010). However the quality of delivery service and profitability are gradually decreased due to the price competition. Deregulation has led to increase in the number of delivery service companies.

According to the survey conducted by the Korea Chamber of Commerce & Industry (KCCI), as the reason of using delivery service, 67.6% of the respondents chose the internet shopping (KCCI, 2011). In the case of online stores, most of customers purchase products through the Internet. So, reliable and high-quality logistics services are the key of success for online stores. Nearly half of the respondents selected delivery accuracy and required time as the major factors to improve the henceforth delivery service. It implies that consumers consider service-related factors such as prompt delivery and kindness more important rather than price. Like this, development of electronic commerce and delivery service industry seems to be collinear, and for sustained growth, paradigm should be changed from a price competition to improving service competitiveness.

This study examines delivery service preference depending on shipping cost and delivery requested time and comprehends the needs on delivery service options in online shopping.

2. Literature Review

There have been a few studies on delivery service in electronic commerce. They mainly emphasized delivery time as a major competition factor.

Stalk et al. (1990) addressed about qualitative necessary on time based competition. Time based competition is the management system which shortens required time in business processes where offers service to the consumers. Concept of this means that promptly respond to the customers demand is main competition resources in corporation by utilizing time as the most important competitive source among the important managerial resources such as quality, cost and time.

Hong (2001) established the decision factor on customer service quality to improve delivery service, and analyzed the intension for reuse and recommendation which is the result variable related to customer satisfaction, and how decision factors affect customer satisfaction. Analysis result presented that 3 factors of delivery service quality factors

which is convenience, economic feasibility and accuracy have an effect on customer satisfaction.

Ray *et al.* (2004) presented demand as the linear decreasing function in accordance with price and guaranteed delivery time. And the study examined capacity of price, required delivery time and order processing as the decision making variables for minimum profits in corporation. The result of this deducted that time sensitive customers demand is shorter than price sensitive customers demand with regard to required delivery time, however they willingly pays for the extra price.

Han *et al.* (2010) approached on imaginary delivery service program with diverse features and levels by contingent choice method and measured mean value of willing to pay with examining the delivery service select decision factors in accordant with socioeconomic characteristics such as gender, age and income by analyzing consumers select behavior systematically.

3. Methodology

3.1 Stated Preference (SP)

Stated Preference (SP) has been approached since mid 1980's on urban, environment and transportation and usefulness has been verified. Especially it has been used on new means of transportation, demand estimate of products and it was used for estimating values of non-commodity that is not actually transacted in reality.

SP research which refers to the stated preference, can be defined as a sequence of finding individual preference by providing imaginary scenario constructed by statistical experiment program (Fowkes and Wardman, 1988). That is, this is the method to examine individual's immanent preference, consciousness and intention in the imaginary condition.

When new delivery service is introduced to the current market, the existing model is not able to estimate its demand. Therefore the new model needs to be estimated using a SP method.

3.2 Logistic Regression Model

A probabilistic choice model is based on the idea of selecting the most desirable and attractive alternatives among the options, and each individual selects the alternative that

maximizes his/her utility.

According to the random utility theory, the total utility of alternatives is composed of deterministic utility and random utility. The former is the observable utility factor, the latter is the unobserved one. The total utility of alternatives can be expressed as Equation (1).

$$U_{in} = V_{in} + \epsilon_{in} \quad (1)$$

where, U_{in} = total utility of alternative i for individual n

V_{in} = deterministic utility factor of alternative i for individual n

ϵ_{in} = random utility factor of alternative i for individual n

Then, the probability of an individual n chooses alternative i is identical with the probability of alternative i 's usefulness for individual n exceeds every other options and the probability for individual n selecting alternative i , $P_n(i)$ is like as Equation (2).

$$\begin{aligned} P_n(i) &= \text{Prob}(U_{in} \geq U_{jn}, \forall j \in C_n) \\ &= \text{Prob}(V_{in} + \epsilon_{in} \geq V_{jn} + \epsilon_{jn}, \forall j \in C_n) \\ &= \text{Prob}(V_{in} - V_{jn} \geq \epsilon_{jn} - \epsilon_{in}, \forall j \in C_n) \end{aligned} \quad (2)$$

where, $P_n(i)$ = probability of individual n choosing alternative i

C_n = combinations of alternatives that individual n can choose

For model estimation, our dependent variable Y is whether the respondent chooses the provided delivery service, or not. Thus, Y is a binary variable (if the respondent chooses the service, $Y = 1$, otherwise $Y = 0$). If P is the probability of choosing the provided delivery service, the logistic regression model can be formulated as:

$$Y = \log \left[\frac{P}{1-P} \right] = \beta' X,$$

where X is the vector of explanatory variables which influence P , β is the vector of parameters of the explanatory variables. If the formula is transformed into a probability model, it is the same as a binary logit model:

$$P = \frac{\exp(\beta' X)}{1 + \exp(\beta' X)}$$

Thus, this study conducted binary logistic regression analysis since it was examined about delivery use depending on shipping cost and required time. Key independent variables are delivery service characteristics, such as delivery time and shipping cost, and socio-economic variables such as family income and occupation.

The logistic regression is estimated by the maximum likelihood method. In addition, the elasticity of P can be calculated by e^{β} . This study used statistical modeling software, SPSS version 20, to estimate logistic regression models.

In addition, marginal willingness-to-pay can be calculated from parameters of shipping cost and delivery time in the models. This is called value of time (VOT) of online delivery services. It can be figured as:

$$VOT = (dY/dX_1) / (dY/dX_2) = \beta_1 / \beta_2$$

where X_1 = delivery time, X_2 = shipping cost, β_1 = parameter of delivery time, and β_2 = parameter of shipping cost.

4. Data Collection and Analysis

4.1 SP Design

Two variables, delivery time and shipping cost were chosen to the key attributes of each alternative that affects delivery service, in composing of questionnaires for SP research. Moreover, age, gender, occupation, personal income and family income are added as property variable and intend to identify the effects of these on select behavior.

The scenario design of the attributes was constituted with 16 combinations (4 levels times 4 levels) of 3,000 won (1 dollar = 1,030 won), 4,000 won, 5,000 won, and 6,000 won for shipping cost, and 6 hours, 12 hours, 24 hours, and 48 hours for delivery time. Average delivery price was probed as 2,506 won in the basis of 2012. The price has been decreased for 40% compared to 4,000 won in 1997 and that the unit cost per box is only a quarter compared to that of the U.S. The reason for this is pointed out as excessive competition and master-servant relationship between client and delivery service companies. Shipping cost has been decreasing every year while on-line shopping mall and home shopping has been increasing since 2000. The standard (base) price is set as 3,000 won, considering of price realization of delivery corporations initiated by Hyundai Logistics which had increased 500

won per box. The distinction between each level is set as 1,000 won. Required delivery time is set with the gap of day, or half-day to be easily understood by the respondents.

4.2 Data Collection

An questionnaire was designed for figuring out tendency of on-line shopping mall delivery service users to conduct SP survey in this study. Subject target had to be everyone who used the delivery service, however it was restricted due to cost and time constraints. The survey area is Seoul Metropolitan Area, the capital region in Korea, to facilitate collecting the questionnaires. The respondents were limited to persons over 20 years old who have the purchasing power.

We employed a stratified random sampling method, and for survey instrument face-to-face interview as well as email methods were adopted. The survey period was from March 12th to March 26th, two weeks. One thousand questionnaires were distributed, and among them 285 respondents completed the questionnaire. Of these, 248 questionnaires were used in the analysis after excluding invalid 37 questionnaires.

4.3 Analysis of Data

4.3.1 Sample Characteristics

Of the respondents, 128 (51.6%) were male, 120 (48.4%) were female. In case of the age, 20's were 100 (40.3%), 30's were 93 (37.5%), 40's were 36 (14.5%), over 50 were 19 (7.7%). This corresponded to sampling standard as stated above.

Table 1.
Sample Characteristics

Variable		Frequency	Percentage (%)
Gender	Male	128	51.6
	Female	120	48.4
Age	20s	100	40.3
	30s	93	37.5
	40s	36	14.5
	Over 50	19	7.7

Variable		Frequency	Percentage (%)
Occupation	Student	99	39.9
	Office worker	149	60.1
Monthly personal income	Under 1 million won	100	40.3
	1~1.99 million won	26	10.5
	2~2.99 million won	58	23.4
	3~3.99 million won	26	10.5
	4~4.99 million won	12	4.8
	Over 5 million won	26	10.5
Monthly family income	Under 1 million won	3	1.2
	1~1.99 million won	5	2.0
	2~2.99 million won	41	16.5
	3~3.99 million won	46	18.5
	4~4.99 million won	44	17.7
	Over 5 million won	109	44.0
Residence Area	Seoul	77	31.0
	Incheon	104	41.9
	Gyeonggi-do	67	27.1

Table 2.

The Average Number of Purchases of the Respondents over the Past Month

Variable	Min	Max	Average	Variance
Number of purchases	1	30	3.14	3.090

The average number of purchases using on-line shopping mall in a recent month was 3.14 times of the past month, as shown in Table 2. In the student group, the average is 2.85 times and 3.43 times in the office worker group. It indicates that office workers are more frequently using online stores.

Table 3.

The Proper Shipping Cost and Delivery Time that Respondents Think

Item	Min	Max	Average	Sample Variance
Proper shipping cost (won)	1,000	10,000	2,965.7	1,016.7
Proper delivery time (hour)	6	72	36.5	13.9

The proper shipping cost and delivery time by the respondents is presented in Table 3. The proper shipping cost that respondents think is 2,965.73 won on average. It is much higher than now (2,500 won). Also, the proper delivery time on average is 36.5 hours. It is similar with the current delivery service, within 2 days.

4.3.2 Model Estimation

The logistic regression model for all item was first estimated, and then that for each item category was also estimated to explore the differences of the explanatory variables between item categories.

The estimated result of the logistic regression model for all items is shown in Table 4, McFadden $p^2(=1-L(\beta)/L(0))$ which refers to conformity shows 0.40, relatively higher in this context, and explanatory variables were statistically significant at the 95% level of significance. In addition, the estimated model by item category appears in Table 5, the value of p^2 ranges from 0.3245 to 0.3955, which is high enough value of conformity. The chi-squared test also showed all models are statistically better than the market share model, which contains only the constant term.

Table 4.
Estimated Coefficients of the Model for All Items

Variable	Coefficient	P-value	95% CI		Exp(β)
			Lower	Upper	
Shipping cost	-.002	.000	-0.002	-0.002	.998
delivery time	-.054	.000	-0.060	-0.048	.947
Male*	.373	.000	0.197	0.549	1.453
Age	.406	.000	0.283	0.529	1.500
Student*	-.261	.025	-0.490	-0.032	.770
Family income	.305	.000	0.236	0.374	1.356
Number of purchases	.063	.000	0.036	0.090	1.065
Constant	5.734	.000	5.117	6.351	309.29
Sample size(n)	3,968				
L(0)	5430.9				
L(β)	3285.4				
p^2	0.40				

Note: * means dummy variable

Table 4 represents coefficients, p-value, and elasticity ($\exp()$) of each explanatory variable. The significant variables are shipping cost, delivery time, gender, age, occupation, family income and number of purchases. Only family income variable was included due to multicollinearity of personal income and family income. In here, the coefficients of delivery time and shipping cost have negative signs, and this means that higher values of delivery time and shipping cost induces lower possibility of the service use. In contrast, higher age, family income and number of purchases mean higher possibility of the service use.

As the dummy variable, the male variable has a positive sign which refers to higher use of males rather than females, and the student variable has a negative sign which refers to office workers use the service more than students.

Odd ratio is the conversion value of individual variable x 's regression coefficient β to $\exp(\beta)$, and is the increase multiple in chance of winning acquired by increasing x to value 1. In here, when the coefficient has a negative sign, its odd ratio is less than 1, therefore selection probability decreases. For example, in case of the delivery time, its odd ratio is 0.947, which means when delivery time increase by one minute, the selection chance increases to 0.947 times. The most effective variable is age, which means 1.5 times higher selection probability in case of one year old increase.

Table 5.
Estimated Coefficients of the Models by Item Category

Variable	Clothing/ Fashion	Household Supplies	Electronics/ computer	Food/ beverages	Sports/ Leisure	Book/ magazine
Shipping cost	-.002	-.002	-.001	-.001	-.002	-.002
delivery time	-.053	-.048	-.055	-.075	-.047	-.043
Male*	.345	.498	.519	.206	.336	.361
Age	.279	.279	.266	.223	.273	.171
Student*	-.454	-.382	-.280	-.066	-.101	.173
Family income	.267	.282	.252	.191	.246	.154
Number of purchases	.063	.049	.071	.049	.085	.085
Constant	6.190	5.938	5.128	5.131	5.762	6.236
Sample size(n)	3968	3968	3968	3968	3968	3968
$L(0)$	5392.531	5373.810	5500.715	5446.434	5415.143	5268.208
$L(\beta)$	3259.517	3256.215	3712.435	3698.742	3380.535	3239.817
p^2	0.3955	0.3940	0.3250	0.3245	0.3757	0.3850

The estimated results of the logistic regression models for six item categories are shown in Table 5. Similar to the model for all items, the models presents negative utilities of shipping cost and delivery time. All dummy variables have negative signs except the category of book/magazine/record, which can be interpreted as higher selection probabilities in the office worker group compared to the student one.

When all else are equal, the required delivery time for the service users is more sensitive than the shipping cost. This indicates varied preference forms in accordance with the required delivery time. Males have a higher selection probability than female. Additionally, higher age or higher family income is also likely to show a higher selection probability.

In detail, the shipping cost coefficients in both electronic/computer products and food/beverages models were -0.001 which is relatively low than other items. It implies that these items are less sensitive to shipping cost. On the other hand, the coefficients of the required delivery time in the models present 0.055 and 0.075, respectively, which is relatively high. It is considered that consumers concern delivery time importantly since value of food can be degraded depending on time. Also electronic/computer products are relatively expensive, therefore consumers can be sensitive on the required delivery time. Thus it can be considered that consumers will pay an extra price to shorten required delivery time in case of these two categories. Based on the model results, corporation can increase the venue by differentiating service such as providing price policy that differs depending on the items.

4. 4 Estimated Model by Market Segmentation

Since the estimated models assumed identical parameter in the individual samples, it can't be reflected appropriately when differences between each individual features are significant. Therefore subgroup estimation is separately performed in this study by dividing subgroups according to the socioeconomic features of the delivery service users.

Market segmentation has advantages on reducing errors by grouping similar members of marginal utility, considering member's characteristics. In this study, types of the respondents were divided by gender, age, occupation and family income. From this, it is able to know the responses to differentiated shipping costs depending on consumer's features.

Table 6.
The Criteria of Market Segmentation

The criteria of market segmentation		Percentage (%)
Gender	Male	51.6
	Female	48.4
Age	20s	40.3
	30s	37.5
	Over 40	22.2
Occupation	Student	39.9
	Office worker	60.1
Family income	Low(Less than 2 million won)	3.2
	Middle(2~4.99 million won)	52.7
	High(more than 5 million won)	44.1

Table 6 shows the criteria of the segmentation. If the component ratio is differed much each other, it is hard to represent, therefore grouping was conducted by proper ratio.

Table 7.
Estimated Results between Subgroups

Variable	Gender		Age			Occupation		Family Income		
	Male	Female	20s	30s	40s	Student	Office worker	Low income	Middle income	High income
Shipping cost	-.002	-.002	-.002	-.002	-.002	-.002	-.002	-.003	-.002	-.002
delivery time	-.054	-.055	-.043	-.059	-.079	-.047	-.062	-.063	-.046	-.067
Male*			.536	.345	.040	.653	.249	.400	.478	.258
Age	.395	.426				-.128	.416	.321	.366	.499
Student*	-.100	-.431	-.569	-.394	-1.606			1.609	.119	-.823
Family income	.239	.383	.070	.150	.390	.091	.233			
Number of purchases	.077	.024	-.024	.121	.149	.100	.083	-.421	.015	.106
Constant	6.087	5.856	6.951	6.545	8.783	6.617	5.927	12.798	6.457	8.233

Variable	Gender		Age			Occupation		Family Income		
	Male	Female	20s	30s	40s	Student	Office worker	Low income	Middle income	High income
Sample size(<i>n</i>)	2064	1904	1600	736	864	1584	1616	112	2096	1744
L(0)	2851.534	2561.525	2103.269	1014.021	1190.340	2076.566	2240.051	140.659	1416.049	2805.867
L(β)	1766.501	1507.286	1389.912	638.780	630.796	1308.337	1383.800	54.280	876.239	1725.299
p^2	0.380	0.411	0.339	0.370	0.470	0.369	0.382	0.614	0.381	0.385

As the result of the model estimation for the gender subgroups in Table 7, the conformity of the models (p^2) are 0.3805 and 0.4115. Thus the models are comparatively appropriate. Generally, the coefficients between the two models are not much different, therefore it shows that the selection probability depending on gender might not be different.

The model results of the age groups represent that the family income and number of purchases variables in 20's, the male and student variables in 30's, the male variable in 40's were not significant. All model conformity was 0.3391~0.4700, which is close to 0.4, thus it is relatively appropriate. The coefficient of required delivery time's differentiation is increasing in accordance with the age, and this means higher sensitivities to delivery time in older age. It is considered that higher selection probabilities in higher purchasing capacity age groups.

For the results of the model estimation for occupation groups, the age and family income variables in the student group and the male variable in the office worker group are not significant. The model fits with 0.3699~0.3822 of p^2 values show appropriate. Examining the coefficient of the required delivery time shows higher value in the office worker group compared to the student group, which means higher sensitivity to the required delivery time. In case of the student group, male showed 1.9 times higher selection on the delivery service compared to female. In case of the working group, older age showed 1.5 times higher selection on delivery service compared to younger age.

The results of model estimation for income groups showed conformity value of 0.6141 in the low income group which is inappropriate. This is due to the fewer samples of 112, compared to other groups. The middle income and high income groups showed 0.3812~0.3851 for model conformity which is close to 0.4, thus this is appropriate. The coefficient of the required delivery time shows higher differentiation in the high income group, which means higher sensitivity on delivery time. In the middle income group, male showed 1.64 times higher selection on delivery service compared to female. In the high income group, older age showed 1.6 times higher selection on delivery service compared to younger age. That is, higher income and older age induces more selection on prompt

delivery with extra charge, and this corresponds to the hypothesis.

4.5 Value of Time (VOT) in the Models

Firstly, looking at the results of VOT assessment depending on items in Table 8, food/beverages and electronic/computer products had 2~3 times higher than other items. It is considered as in case of purchasing these items, consumers will pay extra price in accordance with the required delivery time. The VOT of book/magazine/record was lowest.

Table 8.

Value of Time by Item Category

Items	Average	Clothing/ Fashion	Household Supplies	Electronic/c omputer	Food/ beverages	Sports/ Leisure	Book/ magazine
Won/hour	27.0	26.5	24	55	75	23.5	21.5

The result of model estimation in market division in Table 9 showed higher measurements in male compared to female and in terms of age, it was increased in order of 20's, 30's, 40's. Also the office worker group's VOT was higher than the student group, and in terms of income, the high income group's VOT was measured larger than that of low income group.

Table 9.

Value of Time by Group

Market segment	Male	Female	20s	30s	Over 40
Won/ hour	27	27.5	21.5	29.5	39.5
Market segment	Student	Office worker	Low income	Middle income	High income
Won/ hour	23.5	31	21	23	33.5

5. Conclusions

This study assumed as there would be diverse demands for delivery service from online customers such as willing to pay extra price for rapid delivery or want lower price although it is much slow. The premise was established as diverse price policy is possible through differentiation of current uniformed shipping cost. SP research was used since this is the analysis of non-actual goods which is not transacted in reality. The SP survey was conducted for consumer's preference by 16 scenarios.

Firstly, looking into the delivery service preference in accordance with the items, electronic/computer products and food/beverages presented lower coefficient in shipping cost which means unsusceptible to shipping cost and presented delivery time coefficient as 0.055, 0.075 each, which is relatively high. Therefore it can be considered that consumers will pay an extra cost to shorten the required delivery time in case of these two categories.

Next, the responses to the differentiation of delivery service depending on features of consumers were examined. It was demonstrated that older age, working group and higher income has higher possibility to select the delivery service with high quality. In other words, higher income and older aged groups have more tendencies to pay extra cost on rapid delivery which corresponds to the hypothesis.

For better understanding of online customers' preference for delivery service, it is worthy of noting that a multiple range of customer's needs and preferences for new services should be primarily recognized. There are still many other types of customers needs for online products delivery that have not explicitly been examined in this research, such as same-day home delivery or geographical differences of deliver. Larger samples of online users would result in better results of this research. Many other possible questions remain topics for further research.

Acknowledgements

This work was supported by National Research Foundation of Korea Grant funded by the Korean Government (MOEHRD) (NRF-2011-413-B00008).

References

- Fowkes, Tony and Mark Wardman. (1988) "The Design of Stated Preference Travel Choice Experiments," *Journal of Transport Economics and Policy*, Vol.22, No.1
- Han *et al.* (2010) "Estimating Economic Values of Parcel Service Attributes," *Korean Society of Transportation*, Vol. 28, Issue 5, pp. 65-75.
- Hong *et al.* (2001) A Study on the Influence of Courier Service Quality on Customer Satisfaction. A Master Thesis, Dongduk Women's University
- Korea Chamber of Commerce and Industry. (2010) Analysis of Consumer Usage Patterns of Internet Shopping Malls in Korea.
- Korea Chamber of Commerce and Industry. (2011) Research of the Usage Pattern of Delivery Service.
- Ray S. and Jewkes E.M. (2004) "Customer Lead Time Management When Both Demand and Price are Lead Time Sensitive," *European Journal of Operational Research*, Vol. 153, Issue 3, pp. 769-781.
- Stalk G. and Hout T.M. (1990) *Competing Against Time: How Time-based Competition is Reshaping Global Markets.* NewYork : *The Free Press.*
- Statics Korea. (2013) 2012 Trends of Electronic Commerce and Cyber Shopping.