Quantifying the Benefits of the Internet and Its Applications[1]

by Masanori KONDO*, Akihiro NAKAMURA**,
Hitoshi MITOMO***

Abstract

The objective of this study was to estimate the benefits accruing from the Internet and its applications. With a flat-rate scheme that is common under the broadband access (FTTH, DSL, CATV, etc.) in Japan, it is difficult to measure surplus for each online application with traditional economic methodology as a usage-sensitive market price does not exist. Internet users can enjoy online applications as much as they want. In this situation, subscribing to and using the Internet means that users buy a bundled commodity at a certain fixed price for the services they want, which are included in the package. In this study, we tried to unbundle the composite commodity (i.e., Internet service as a whole) into the applications available on the Internet and measured the willingness to pay (WTP) for each application.

We applied a stated preference (SP) method for analyzing data and used a random parameter logit (RPL) model for estimating WTP for each of such applications as e-mailing, web browsing, and content downloading. The estimated WTP for availability of e-mail and web browsing delivered over personal computers are ¥2,709 and ¥2,914 on a monthly basis, respectively, while average broadband access service costs approximately ¥4,000 in Japan.

Keywords: broadband, stated preference, random parameter logit model, consumer surplus, willingness to pay

* PhD student, Waseda University, Tokyo, Japan
** Associate Professor, Tezukayama University, Nara, Japan
*** Professor, Waseda University, Tokyo, Japan

1. Introduction

Since commercial Internet access service was introduced in 1994, the penetration rate in terms of Internet users had risen to 68.5% by 2006 in Japan, and today, more than 87million people are viewed as users of the Internet.[2] In the process of increasing the number of Internet users, both network infrastructure and the availability of online applications have been developed very rapidly. Additionally, in the present broadband access environment, people can enjoy a wide variety of online applications over the Internet, ranging from e-mail to video programs.[3] With respect to broadband network services such as FTTH, DSL, CATV, and FWA, the number of subscribers had catapulted to more than 26 million households by March 2007.[4] With the broadband service, it is common to apply a flat-rate for access to the Internet, where users are not required to pay any additional fees for the many online applications and services once they have paid the flat-rate. Those applications and services include e-mail, web browsing, word processing, spreadsheet, etc.

While people use these online applications and services at no additional cost, they do anticipate obtaining certain benefits from these applications and services. In other words, the benefits that users enjoy by accessing the Internet do not come from line access per se, which they pay for, but from online applications that are generally supplied free. At the same time, applications such as e-mail and web browsing are provided as a bundled service, and it is difficult to derive the value of each application from a revealed preference (RP) data[5] of bundled services. If we want to use RP data for our analysis, the only available data in the consumer market is the Internet access line fee. In these circumstances, it is difficult to measure consumer surplus with an ordinary economic approach as there is no usage-sensitive market price for each of the online applications. Thus, few studies have been conducted to measure consumers' surplus for these online applications.

The purpose of this paper is estimate the benefit that Internet users can derive from using online applications. One of approaches to deal with the abovementioned issue of nonexistence of market data could be a plan to introduce opportunity cost, that is, cost of time. In the study by Goolsbee and Klenow (2006), the cost of time for using the Internet was introduced. However, they did not apply it for each application, but rather dealt with Internet usage as a whole. It may be difficult to get data of time consumption for each online application and therefore we adopted a stated preference (SP) method in this study.

We estimated consumers' surplus derived from Internet usage and online applications by applying the SP method. We adopted a conjoint method to deal with the SP in which we systematically varied combinations of levels of each attribute, such as availability of e-mail, web browsing, and downloading content (music and/or video).

2. Methodology

2.1 SP

Considering that there is no market price for many online applications, it is considered appropriate to use an SP method, which can apply even to the measurement of consumer benefits of nonexistent goods.[6] Therefore, SP method has been widely used in the study of consumer's valuation for environment or demand analysis for new products that are not yet in the market since data on market choices are unavailable in these fields. (Green and Srinivasan, 1990; Hensher, 1992; Layton, 2000).

Conjoint analysis, which we used in this study, is an SP method and has been widely applied in many research areas, such as travel, transportation, marketing, psychology, and environment economics (Louviere, 1988; Green and Srinivasan, 1990; Gan and Luzar, 1993; Hensher, 1994; Roe et al., 1996; Layton, 2000; Hensher, 2001; Huber and Train, 2001; Alvarez-Farizo and Hanley, 2002). In conjoint analysis, researchers first make hypothetical bundles of several attributes, describing the characteristics of a commodity or a service, and then ask respondents to state their preferences for some hypothetical alternatives selected from the full set of combinations of alternatives. A statistical technique is applied to the responses to analyze the relative importance of the attributes.

2.2 Random parameter logit model (RPL[7])

To forecast demand for new products or goods/services that have not been launched in the market, the most popular discrete choice econometric models are logit and nested logit (McFadden, 1974, 1978; Ben-Akiva and Lerman, 1985; Ida et al., 2008). These specifications have advantages such as simplicity of estimation. However, they require some restrictions in specifications as well. One is that the coefficients of variables in the model are assumed to be fixed across all users. This assumption means that different users with the same observed characteristics have the same preferences for each factor in the model. The other is that logit and nested logit models require the independence from irrelevant alternatives (IIA) property, over all alternatives and over alternatives within each nest, respectively. Since IIA property means that a change in the attributes of one alternative changes the probabilities of the other alternatives proportionately, it is not always a realistic assumption for many goods and services (McFadden, 1974; Train, 1986; Allenby and Rossi, 1999; McFadden and Train, 2000). The RPL model allows the coefficients of observed variables to vary across users and does not impose the IIA property. Because of these advantages, RPL models have been adopted in recent years as the speed of computers advances (Brownstone and Train, 1999; Revelt and Train, 1996; Train, 1998; Calfee et al., 2001; Kim 2005). Additionally, less

restrictive choice model specifications can illustrate more accurate estimation in certain contexts (Train, 1998; Layton, 2000; Hensher, 2001).

3. Model

3.1 Model specification

We applied the discrete choice model that is based on the random utility theory. According to the random utility framework proposed by McFadden (1974),[8] utility U is composed of the deterministic component and the random component. That is,

$$U = V + \varepsilon$$

where V denotes the deterministic core and ε denotes the random component.

This study applied the RPL model as it is natural to suppose that an evaluation for each online application is not the same among individuals. The RPL model captures the variations of preferences by introducing stochastic terms into the coefficients, which are created by deviations from mean preferences. The RPL model allows these coefficients to be correlated with each other across the attributes.

The RPL specifies the issue for individual n who faces j alternatives to maximize utility functionU_{nj},

$$U_{nj} = \beta'_n x_{nj} + \varepsilon_{nj}$$

where x_{nj} is a vector of variables as a set of attributes when individual n choose alternative j

As the distribution of β_n is unknown, it is common to estimate parameters of the distribution function by assuming parametric probability distribution function. In this study we allowed the coefficients vector β_i, whose elements are $\beta_{i,x}$, to be distributed normally across the population with mean vector b and variance covariance matrix W.

Individual n will choose alternative j if and only if $U_{nj} > U_{nk}$ for any $j \neq k$. With the RPL model, we assume ε_{nj} distributed as independent and identical extreme value distribution and the probability of individual n choosing alternative j who is facing β_n can be described as

$$L_{nj}(\beta_n) = \frac{e^{\beta'_n x_{nj}}}{\sum_k e^{\beta'_n x_{nk}}}$$

As β_n is unknown and has a probability distribution, if we let the distribution function of β_n be $f(\beta)$, then the probability of individual n choosing alternative j out of J alternatives is

$$P_{nj} = \int \frac{\exp(U(\beta_{nj}, X_j))}{\sum_{j=1}^{J} \exp(U(\beta_{nj}, X_j))} f(\beta) d\beta$$

3.2 Specification of the utility function

We specify the utility function as below and assume the distribution of β_n to be normal distribution. We also assume that β_{price} is assumed as fixed.

$$U_{nj}(\beta_{nj}, X_j) = \beta_{n,sp} SPEED_j + \beta_{n,ml} MAIL_j + \beta_{n,web} WEB_j + \beta_{n,dlm} DLM_j + \beta_{n,dlv} DLV_j + \beta_{price} PRICE_j + \varepsilon_{nj}$$

where
$U_{nj}(\beta_{nj}, X_j)$ is the utility of individual n in the case of choosing alternative j
SPEED : Access line speed of alternative j
MAIL : Dummy variable for the availability of e-mail
WEB : Dummy variable for the availability of web browsing
DLM : Dummy variable for the availability of downloading music
DLV : Dummy variable for the availability of downloading video
PRICE : Monthly charge for alternative j

4. Data

As it is recommended that the number of attributes should be restricted to as few as possible, preferably less than seven, to avoid confusing the respondents (Huber and Klein, 1991),[9] we selected the six attributes listed in Table 1 and that we considered to be the most influential in the consumer's choice of Internet usage.

In order to measure willingness to pay (WTP) for each attribute such as availability of e-mail and web browsing, we asked respondents to chose his or her preferences out of packaged services alternatives. The packaged services consisted of factors which are chosen for each attribute with a specific level. For example, a packaged service would be the broad band access with e-mailing, web browsing, music downloading but without video downloading and its monthly charge is ¥4,000. An alternative of not to subscribe to internet is also included in the choice. Details of the attributes and their levels are shown in Table 1 below:

Table 1: Design of the conjoint analysis

		Levels			
Attributes	Monthly charges	¥1,000	¥2,000	¥3,000	¥4,000
	Access speeds	Narrow Band	Broad Band		
	e-mailing	Available	Not available		
	Web browsing	Available	Not available		
	Music downloading	Available	Not available		
	Video downloading	Available	Not available		

A) Monthly charges

This attribute ranges from ¥1,000 to ¥4,000, which we chose on the basis of our pre-test survey. Considering the fact that many applications can be used free of charge once a subscription is made for the Internet, the monthly charge includes the access fee plus (anticipated) applications fee, which are available under his/her option choice. Be mindful of the fact that the charge of the contents downloaded is not included in this monthly charge and respondents are informed that additional payment may be necessary.

B) Access speeds

This attribute outlines the transmission speed of access lines that subscribers can enjoy when they access the Internet.

C) E-mailing

This represents the availability of e-mailing services. While the number of online applications available now is increasing, including social networking service (SNS) and voice over internet protocol (VoIP), we chose e-mailing, web browsing, music downloading, and video downloading as the key applications since it is thought that the usage rates of those applications are higher among Internet users than those of emerging online applications.

D) Web browsing

This also represents the availability of the web browsing service, which includes searching, posting, and exchanging information.

E) Music downloading

This attribute is related to the availability of service for the downloading of music that is provided through the Internet. Streaming service is included in this category. Respondents are informed that additional payment would be necessary for this attribute.

F) Video downloading

This attribute is related to the availability of service for the downloading of video, which is rapidly becoming popular. YouTube is the typical application users can enjoy.

As it is impossible to ask the preference for all possible combinations of the attributes, we have extracted minimum sets of the attributes (i.e., the profiles) in accordance with the orthogonal design method. Seven questions were asked to each respondent, to each of which he/she was requested to make his/her best choice from among three service package profiles arbitrarily selected from the profiles and one profile with no usage.

We conducted an online survey from December 25 to 27, 2007. The questionnaire was composed of 11 basic questions such as age and gender and 7 multiple choice questions. The number of respondents was 1,000, who had registered in advance as members of a survey panel for the Internet. As we used the survey panel, there could be a bias in terms of age and/or sex. Therefore, we collected the same number of data samples from each category as shown below and then calculated the weight we should use according to the distribution of Internet users in terms of age and sex in Japan.[10] In addition, there could be differences in preference toward Internet applications between Internet users and non-users. Those differences, if any, may give upward bias in our estimation, since we used samples collected from the respondents who actually use the internet.

Table 2: Number of samples collected

	age					
	<30	<40	<50	<60	60>=	total
male	100	100	100	100	100	500
female	100	100	100	100	100	500

Percentage of Internet-user distribution in Japan

	Age					
	<30	<40	<50	<60	60>=	total
male	16.1	10.3	10.2	9.6	6.3	52.54
female	15.6	9.9	9.5	8.3	4.2	47.46

5. Estimation results

The maximum likelihood simulation method was used to estimate this model. Estimation results of the RPL model are shown below (Tables 3 and 4). Table 4 shows that the RPL model is an appropriate model for this analysis as most of the standard errors of the random parameters are statistically significant in deviating from 0.

From Table 3, we can see that price (the monthly charge) has a negative effect on the choice probability as anticipated. The coefficients with stochastic distribution are evaluated at the mean value. The coefficients of e-mailing, web browsing, and video downloading are statistically significant at the 1.0% level, while that of music downloading is at the 5.0% level. These results are considered to be natural because many current Internet users actually use these applications.[11]

Table 3: Estimation results (random parameter logit)

Variables	Coefficient	S.E.	p-value	
Random parameters in utility functions				
SPEED	0.2202	0.0158	0.0000	***
MAIL	2.0798	0.0932	0.0000	***
WEB	2.2370	0.0993	0.0000	***
DLM	0.1512	0.0624	0.0154	**
DLV	0.3387	0.0567	0.0000	***
Nonrandom parameters in utility functions				
PRICE	-0.0008	0.0000	0.0000	***

p<0.10*, p<0.05**, p<0.01***

Table 4: Standard deviations of parameter distributions

Variables	Coefficient	S.E.	p-value	
Derived standard deviations of parameter distributions				
SPEED	0.2200	0.0190	0.0000	***
MAIL	1.5620	0.1043	0.0000	***
WEB	1.8227	0.1098	0.0000	***
DLM	0.8123	0.0963	0.0000	***
DLV	0.8193	0.1179	0.0000	***

p<0.10*, p<0.05**, p<0.01***

Table 5: Estimation results (conditional logit)

Variables	Coefficient	S.E.	p-value	
SPEED	0.1223	0.0084	0.0000	***
MAIL	1.1810	0.0400	0.0000	***
WEB	1.3374	0.0422	0.0000	***
DLM	0.0475	0.0380	0.2106	
DLV	0.1251	0.0345	0.0003	***
PRICE	-0.0005	0.0000	0.0000	***

p<0.10*, p<0.05**, p<0.01***

In this study, price factor as a monthly charge for packaged services has been included in the estimation equation. As the random utility function is described in a linear equation, the WTP for each service can be calculated by dividing the coefficients of services by the coefficient of the price.[12] This can be illustrated through the following simple formula:

$$\partial PRICE / \partial z = -\partial U / \partial z * \partial PRICE / \partial U = -\beta_z / \beta_{price}$$

Where z denotes: SPEED, MAIL, WEB, DLM, DLV

Table 5 shows the WTPs for the applications available on the Internet. It should be noted that these figures are estimated in terms of average user. The WTP for e-mail service is about ¥2,700 and that of web browsing is about ¥2,900, and the amount adding them up exceeds the average cost of monthly broadband access charges in Japan.[13] This means that average internet user can get consumer surplus with the current monthly fee for Internet access as long as he/she can use both e-mail service and web browsing. In other words, the results illustrate that these two applications can be considered as fundamental applications (“killer contents”) for Internet users since they account for a large proportion of the benefit users derive. (There may be additional factors that have a positive impact on Internet users’ benefit. However, as we can see from the estimated WTPs for DLM or DLV, new online services do not provide large benefit yet and we believe that it is rational in current internet user environment to understand e-mailing and web browsing as “”killer contents””.)

Table 5: Willingness to pay (WTP)

SPEED	¥286.84	***
MAIL	¥2,708.77	***
WEB	¥2,913.55	***
DLM	¥196.93	**
DLV	¥441.16	***

5. Conclusion

In this study we showed that the total WTP for both e-mail and web browsing applications exceeds the cost of having broadband Internet access.

We conducted this study to analyze present consumers' behavior and estimate their surplus derived from use of the Internet and its applications. The findings we show in this study will change as new Internet applications and technologies develop. For instance, no one had anticipated the use of PCs as music player, television, or telephone. As new applications such as net shopping, net auctioning, online gaming, web logging, and SNS are being further developed and disseminated, the composition of applications that Internet users valuate highly will change. Here we have merely calculated the WTP of online applications employing the conjoint method and there may be other approaches to measuring consumers' evaluation for online applications such as introducing usage time (opportunity cost) in utility analysis.

NOTES

1. This research was partially supported by a grant from the Study Program on Information and Communication Policies, 2007.
2. White Paper (2007), Ministry of Internal Affairs and Communications, Japan
3. The volume of traffic data of broadband users has also expanded rapidly from 269.4 Gigabits per second (Gbps) in September 2004 to 636.6 Gbps in November 2006 (data from the six main ISPs in Japan).
4. The number of broadband subscribers was 3.9 million in March 2002, 9.4 million in 2003, 15.0 million in 2004, 19.6 million in 2005, 23.3 million in 2006, and 26.4 million in 2007 (Ministry of Internal Affairs and Communications, 2007).
5. In many cases, the only available data as RP is Internet access line fee.
6. Common criticism of the SP method is the hypothetical nature of the questions and respondents' choice.
7. RPL is also called "mixed logit" (McFadden and Train, 1997; Train 1997) and "random coefficient logit" (Kim, 2005).
8. McFadden (1974) pioneered the concept of the model that values the effects of the explanatory variables from consumers' choice data
9. Huber et al. (1991) recommend that the number of attributes should be restricted to about six.
10. "Survey of household's usage on information and communication services," Ministry of Internal Affairs and Communications (2006).
11. New applications have been continuously introduced in the market, such as social network society (SNS) and consumer generated media (CGM). In this sense, e-mail, web browsing, downloading music, and downloading video are not the only applications available on the Internet. In our pre-interview survey, however, many interviewees said that e-mail and web browsing were their main reasons for using the Internet, and we chose the abovementioned four applications because of the technical limitations of analysis.
12. As the random parameter logit model allows the coefficients to be distributed, the estimated level of willingness to pay illustrates the one for the average individual.
13. The price depends on various consumer package services. The actual flat-rates for broadband access range from approximately ¥2,000 to ¥4,000 for ADSL (1-50 Mbps).

REFERENCES

Allenby, G. and Rossi, P., (1999) “Marketing models of consumer heterogeneity.” *Journal of Econometrics* 89, 57-78.

Alvarez-Farizo, B. and Hanley, N. (2002) “Using conjoint analysis to quantify public preferences over the environmental impacts of wind farms: an example from Spain.” *Energy Policy* 30, 107-116.

Ben-Akiva M. and Lerman S. (1985) “Discrete Choice Analysis,” MIT Press, Cambridge, MA.

Brownstone, D. and Train, K. (1999) “Forecasting new product penetration with flexible substitution patterns.” *Journal of Econometrics* 89, 109-129.

Calfee, J., Winston, C. and Stempski, R. (2001) “Econometric issues in estimating consumer preferences from stated preference data: a case study of the value of automobile travel time.” *The Review of Economics and Statistics* 83 (4), 699-707.

Chiang, J., Chib, S. and Narasimhan, C. (1999) “Markov chain Monte Carlo and models of consideration set and parameter heterogeneity.” *Journal of Econometrics* 89, 223-248.

Chintagunta,P., Jain D., and Vilcassim N. (1991), "Investigating heterogeneity in brand preference in logit models for panel data", *Journal of Marketing Research*. Vol.28.

Gan, Ch. And Luzar, E. (1993), “A conjoint analysis in waterfowl hunting in Louisiana.” *Journal of Agricultural and Applied Economics* 25 (2), 36-45.

Goolsbee, A., and Klenow, P. J., (2006) ”Valuing Consumer Products by the Time Spent Using Them: An Application to the Internet.”, *The American Economic Review*, 96; 2, 108-113

Green, P.E. and Srinivasan, V. (1990), “Conjoint analysis in marketing research: new developments and directions.” *Journal of Marketing* 54 (4), 3-19.

Hensher, D. (1992), “Stated-preference analysis of travel choices: The state of the practice.” *Transportation* 21, 107-133.

Hensher, D.A., (1994) “Stated preference analysis of travel choices: the state of the practice.” *Transportation* 21, 107-133.

Hensher, D., (2001) “The valuation of commuter travel time savings for car drivers: evaluating alternative model specifications.” *Transportation* 28, 101-118.

Hensher, D.A., J.M. Rose, and W. H. Greene (2005), Applied Choice Analysis: A Primer, Cambridge University Press

Huber, J., and Klein, N.M., (1991), "Adapting Cutoffs to the Choice Environment: The Effects of Attribute Correlation and Reliability", *The Journal of Consumer Research*, 18;3, 346-357

Huber, J. and Train, K., (2001). "On the similarity of Classical and Bayesian estimates of individual mean partworths." *Marketing Letters* 12, 257-267.

Ida T., Kinoshita S. and Sato M., (2008). "Conjoint analysis of demand for IP telephony: the case of Japan." *Applied Economics* 40, 1279-1287

Kim Y., (2005). "Estimation of consumer preferences on new telecommunications services: IMT-2000 service in Korea", *Information Economics and Policy* 17; 73-84

Kuehl, R.O. (1999), Design of Experiments: Statistical Principles of Research Design and Analysis: Second Edition, Duxbury Press.

Layton, D.F., (2000). "Random coefficient models for stated preference surveys." *Journal of Environmental Economics and Management* 40, 21-36.

Louviere, JJ., (1988). "Analyzing Decision Making: Metric Conjoint Analysis." *Sage University Paper* No. 67, Newbury Park, Beverly Hills.

Louviere, J.J, D.A. Hensher, and J.D. Swait (2001), Stated Choice Methods: Analysis and Applications, Cambridge University Press

Luce, R.D. and J.W. Tukey (1964) "Simultaneous conjoint measurement: A new type of fundamental measurement", *Journal of Mathematical Psychology,*

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zaremlka (ed.), Frontiers in Econometrics, Academic Press, New York.

McFadden, D., (1978). Modeling the choice of residential location. In Karquist, A. et al. (Eds.), Spatial Interaction Theory and Planning Models. North-Holland, Amsterdam.

McFadden, D., and Train, K., (1997). "Mixed MNL models of discrete choice." Working paper, Department of Economics, University of California, Berkeley.

Revelt, D., and Train, K., (1996), Incentives for appliance efficiency: random parameters logit models for households; choices, Department of Economics, University of California, Berkeley.

Roe B., Boyle K.J. and Teisl M.F., (1996). "Using Conjoint Analysis to Derive Estimates of Compensating Variations." *Journal of environmental Economics and Management* 31, 145-159.

Rosen, S. (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*, 82, 34-55.

The Ministry of Internal Affairs and Communications, (2006), Information and Communications Statistics Database.

The Ministry of Internal Affairs and Communications. (2007). White Paper

Train K., (1986), Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand, MIT Press, Cambridge, MA.

Train, K., (1997), Mixed logit models for recreation demand. Valuing the Environment Using Recreation Demand Models. New York: Elgar Press.

Train, K., (1998). "Recreation demand models with taste differences over people." *Land Economics*, 74(2) 230-239.

Train, K. (2003), Discrete Choice Methods with Simulation, Cambridge University Press, Cambridge.