International Journal of Psychosocial Rehabilitation, Vol. 25, Issue 01, 2021
ISSN: 1475-7192

A psychological approach to Model Iranian Demand for Internet

Mohammad Amin Kouhbor*, Ebrahim Ghasemi Varnamkhasti, Homayoun Yousefi

Abstract--- *This article is aimed to identify the factors affecting demand for internet use among Iranian people with a special consideration of its psychological foundation. For this reason, we used Iran's Income expenditure survey data in 2019, Households internet expenditure as dependent, and some Economic, Geographic and demographic variables as control variables. Based on Double hurdle estimations, the expenditure on home internet is governed by two decisions, the decision to participate in internet connection and the decision on the amount of data to be used. Also, these two decisions, could determine by two separate sets of variables. Some household characteristics such as age, number of under 8 years members in a household, education and the economic variables found to be significant in both decisions, while the geographic ones, were only significant in the participation equation.*

Keywords--- *Psychology, Consumption behavior, Internet Demand, Limited dependent variable models*.

I. INTRODUCTION

Technological innovations has reorganized economic and business affairs and household budget undergone a significant alterations via the introduction of personal computers and internet to the homes. The proportion of people participated in PC and internet consumption are increasing day by day, for example computer ownership rose from 34 in 2011 to 67.9 in 2018. Also, percentage of active internet users which was 21.8 in 2011, is now reached to almost 82% of the whole population (www.iriu.ir/matma).

Internet usage is now unavoidable, as far as people are frequently use it for the different purposes such as to connect to the other people, sharing information, entertainment, getting latest news, education and research. Moreover, due to the corona pandemic, telecommuting jobs are highly accelerating. According to Income Expenditure survey, published by Iran Bureau of statistic (2019), most popular use of internet among Iranian people are social media and electronic mail (37.5%), news and weather application (21%), education and research (11%), music, film and applications download (10%), call (8.5%), shopping (less than 3%). However, among the total internet users, almost two thirds use mobile data to access internet, more than one third have local area connection and less than 5 percent of households reported to access via the dial up connection. Also, more than half of the population use internet in a daily basis and many users which has not any home access, no matter how much, have used it in the public areas such as universities or airports. For this reason small minority of people are now has never used internet at all (www.mis.iran.ir).

Based on Income-Expenditure data (2019), more than 10 percent of households reported zero minutes for the internet usage during the last month. The reasoning behind this, and also the econometric issues arises from zero consumption will be of great importance in our study.

Mohammad Amin Kouhbor, Ebrahim Ghasemi Varnamkhasti, Assistant Professor at Khorramshahr University of Marine Sciences and Technology, Department of Economics, Faculty of Economics and Managements, Iran.*Email: Aminkuhbor@yahoo.com*

Homayoun Yousefi, Associate Professor at Khorramshahr University of Marine Sciences and Technology, Department of Marine trasport, Faculty of Economics and Managements, Iran.

II. Literature review

As said before, this study has examined consumer behavior in internet traffic usage. There is evidence enclosing that the demand for internet traffic has been growing (Sabagh kermani-Najafi, 2007; Salem, A, and Morovat; 2019) Ngugi & Nyoro, 2007; Weinberger, Pasquini, Kasambula, & Abukutsa-Onyango, 2011). Considering the household level of the data, two factors should be regarded to conduct an appropriate statistical modelling. The first one is to consider zero inflated nature of the dependent variable and the second one is called infrequent consumption.

Time as a valuable and productive resource plays an important role in consumer utility, preferences and also their choices. The more one person has a valuable time, the less utility will have time consuming services for him or her. For example a stuck man will pay less time for the cinema or expense to watch a football match. This story would be highly applicable in the case of internet usage and could explain a big part of time differences devoted for this service among different classes of people. For example a surgeon or a professor is expected to waste less time in the social media or entertainment websites.

The psychologist's point of view suggest that zero internet consumption may arise from two different sources. The first one is in the situations that internet is not in consumers preferences. In other word, consumer doesn't get satisfaction from allocating a time for the internet. And the other may be due to low value of internet for the consumer. Hence, it is not reasonable for a consumer to allocate a time or money to use internet.

Another important reason for the zero reported consumption would be infrequent purchase which might be more probable for some specific goods and services, especially durable ones. if you buy an automobile once in ten years, but use it actively during this period, it is normal to report zero monthly expenditure while being an active user if the data collected more than a month after the car is purchased. This will be applicable for many of internet user who buy traffic package data for more than a month or other durable goods and services. In other words, zero inflated data will be more likely in cases that the time period between two purchases are larger than the period for which the data is reported. Finally, Corner solution is another source for zero inflated data and occurs when a specific good or service is not a priority of the consumer.

However, following Tobin (1958) applying a workhorse OLS on this kind of data may seems reasonable at a first glance, but is apparent that we are ignoring an important information. Deeper probing, will reveal more serious issues. Disregarding a considerable amount of data, when it is available, not only affect the degree of freedom, but also it will makes the estimations to be biased. To overcome the mentioned problems he proposed some amendments on the probit likelihood function to handle a variable with a large probability of having zero value and the remaining probability spread over a positive interval. The Tobin's version of Probit which called Tobit are now well known among researcher and has been widely used in consumer behavior studies.

The Tobit model facilitated the emergence of the next generate of microeconometric modelling which developed by Hekman(1976). His correction was a statistical method to correct the bias arises from non-randomly selected samples. This is achieved by modelling the individual sampling probability of each observation (selection equation) together with the conditional expectation of the dependent variable (outcome equation). Heckman also developed a two-step procedure to estimate this model which avoids the inflexibility of having to estimate both equations in a same equation, albeit at the cost of inefficiency. So, he first suggested a logit to estimate the probability of non-zero consumption and then, a truncated classical OLS for a part of sample who reported a positive consumption. The inverse Mills ratio must be generated from the estimation of a former model to correct for the selection bias. The difference between the heckit and the Tobit is that the latter observes the process in a two- step procedure and then it allows the use of different sets of explanatory variables in both stages of estimations where as the tobit uses a one-step procedure and assumes that the explanatory variables

affecting both decisions to participate and consume are the same. For this reason, the heckit is viewed as a 'generalized version of the Tobit model.

Cragg(1971) introduced a more flexible model that is a modification of the Tobit model and the Heckman. The difference between heckit model and Craggs double hurdle model is that heckits assumes that in the second stage, there will be no zero observations once the first stage is passed, whereas the double hurdle still considers that there might be a possibility of a zero observation which may arise from the individual's choice (the case of corner solution) or random circumstances.

This research is to examine the effect of factors affecting both participation and the amount of expenditure each household expensed for the internet in the separate regression equations.

Many of authors applied Heckit selection model in their empirical research (Cheng and Caps; 1998, Heien & Wessels; 1990, Heien & Durham; 1991). But following the abovementioned reasons, Double Hurdle model is more pervasively used by researcher to analyze consumer behavior. Some of them are in social sciences and the study of the time parents allocated for taking care about their child (Kalenkoski et all; 2005, Kimmel and Connelly; 2004)

Data

The data is used in this study is Household Income-Expenditure data collected from more than 34000 Iranian urban and rural households which collected by Iran's bureau of statistics in 2019. Regarding the aim of the article, the data for the suspected factors is drawn out for each household. To do this, a pivot reporting table is extracted in the access data base. Finally some socio economic data containing demographic traits such as household age, household head's sex, the number of men in a house hold, education level, Hosehold size and geographical location and also economic ones such as income and price is provided in an excel file consisting 18210 urban and 15941 rural households.

According to data, less than 10 percent of households had home LAN connection, and the rest of the sample households, mostly used individual data internet provided by sim card. Average cost of internet was IRR 15712 for the home sample and IRR 22431 for the internet consumer sample. Nearly 13% of users live in south, and almost on third live in the north of Iran. The proportion of internet users live in the west and east were 22 and 28 percent respectively. Surprisingly, 70% of them live in the urban area, while less than 30% of them were rural ones and finally, almost 90 percent of internet user families have educated head.

Table 1- Descriptive Statistic

Participated Sample			Whole Sample			Variable
Maximum	Minimum	Mean	Maximum	Minimum	Mean	
85000.1	52314.5	22431.7	0	42316.7	15712.1	Internet expenditure
143961	17412172	8261233	152341	15432180	7522091	Income
0	1	0.12	0	1	0.13	South
0	1	0.38	0	1	0.34	North
0	1	0.23	0	1	0.22	West
0	1	0.22	0	1	0.27	East
0	1	0.70	0	1	0.65	Urban
1	12	3.9	1	19	4.2	HH Size
1	28	11.1	0	28	9.2	Education
0	10	1.9	0	10	1.8	Men Number

Source: Iranian Households Income-Expenditure published data for 2019

III. RESEARCH METHODOLOGY

Following the abovementioned reasons, Double Hurdle model is used to estimate the coefficients of internet demand. As said before, following a psychological principle, two separated hurdle should be passed to observe a positive consumption. First, there must be at least one internet user in each household, namely that household should be potential internet user. The second condition, implies that the latent variable (which is equilibrium consumption) should be positive. This can be defined as follows:

$$\begin{aligned} &\textit{Participation Decision}: \quad y_1^* = x_1\alpha + u_i \\ &\textit{Consumption Decision}: \quad y_2^* = x_2\beta + v_i \end{aligned} \tag{1}$$

So, to hold latent variable positive

$$y_i = \begin{cases} y_{2i}^* & \textit{if}: y_{2i}^* \succ 0 \ \ \& \ \ y_{1i}^* \succ 0 \\ 0 & \textit{Otherwise} \end{cases} \tag{2}$$

This implies disturbance terms to follow a bivariate joint distribution. By the assumption of normality we have:

$$\begin{aligned} &\left(\varepsilon_1, \varepsilon_2\right) \square\ N(0, \Omega) \quad , \\ &\Omega = \begin{bmatrix} 1 & \rho \\ \rho & \sigma^2 \end{bmatrix} \end{aligned} \tag{3}$$

So, the likelihood function of the double hurdle model can be rewritten as:

$$\begin{aligned} L = &\prod_0 \left[1 - \Phi\left(X_1\beta_1, \frac{X_2\beta_2}{\sigma}, \rho\right)\right] * \\ &\prod_+ \left[\Phi\left(\frac{X_1\beta_1 + \frac{\rho}{\sigma}(y - X - 2\beta_2)}{\sqrt{1-\rho^2}}\right)\frac{1}{\sigma}\phi\left(\frac{y - X_2\beta_2}{\sigma}\right)\right] \end{aligned} \tag{4}$$

Rho will be zero if a household makes two decisions separately, and a non-zero value shows that both decision are made simultaneously. In the latter state, double hurdle would be reduced to a Hackman model (Mc Dowell; 2003, Aristei & Pieroni.; 2012).

Another point to be considered is among the wide range of possible specifications, what variable should be included in each equation. Unlike Tobit, purchasing decision in the double hurdle and heckit models will be made in two separate steps and hence, we should decide about variables to include in each decision. Although there is no clear rule (Aristei; 2012, Woodajo; 2013, Newman et all; 2003) but some psychologists believe that participation is greatly under the influence of demographical factors, while the amount of expenditure is more affected by economic parameters (Jones; 1992, Yen and Jensen; 1996).

Estimation

The estimated parameters for the double hurdle model is presented in table2. The dependent variable is stated in equation1 is the probability of participation in internet consumption in a logit model, while this is amount of expenditure in the second equation.

Table 2. Estimation Results for Double Hurdle and Heckit

Variable	Heckman		Double Hurdle	
	Expenditure	Participation	Expenditure	Participation
Economic Variables				
Computer Ownership	1.64**	1.23	0.35	1.22
Income	0.43	0.23	0.23	0.21
Geographical Variables				
East	-.0.58	-0.40	-0.26	-0.41
West	-0.21*	-0.23	-0.001*	-0.23
North	0.37*	0.26	0.07	0.26
South	0.33	-0.11**	0.46	-0.11
Urban	0.57	0.31	0.26	0.31
Demographic Variables				
HH Size	0.04*	0.09	0.04*	0.08
Under 8 aged	0.02*	0.03	0.01*	0.03
Above 40 aged	0.15	0.14	0.15	0.14
Educated Number	-0.10*	-0.11	-0.14*	-0.10
Men number	0.04*	0.10	0.01*	0.10
If HH is Educated	0.003	0.002	0.001	0.002
Intercept	7.00	-0.65	7.24	-0.56
Likelihood Ratio	-10268.467		-8736.835	

** Significant in 95% * insignificant

Before interpretation, Voung test is used to select the best model, as we can see in table 3:

Table 3: Specification Test

Models	d.f	Test stat	Best Specification
Heckit vs Double Hurdle	14	229.20406	Double Hurdle

Results suggest that the difference between two specifications are not considerable in the participation equation. However there is a bit difference between them in the latter stage. Moreover both models are the same in term of parameter's sign. Finally, indebted to high degree of freedom, almost all of the variables found to be significant in 95% confidence in both specification models. Following the specification results, conditional and unconditional elasticities are computed as presented in table 4:

Table 4- Elasticities of the internet demand model

Variable Name	Double Hurdle Model		
	Unconditional Elasticity	Participation Elasticity	Conditional Elasticity
Economic Variables			
Computer Ownership	1.63	1.37	1.26
Income	5.47	0.83	4.64
Geographical Variables			

East	-16.14	-0.11	-16.036
West	4339.59	-0.808	4339.67*
North	77.68	0.12	77.57*
South	0.86	-0.019	0.88
Urban	30.10	0.21	29.89
Demographic Variables			
HH Size	-1235.16	0.41	-1235.58*
Under 8 aged	865.19	-0.18	865.21*
Above 40 aged	1134.54	0.21	1134.34*
Educated Number	-2889.63	-0.36	-2889.61*
Men number	-1889.10	0.03	-1889.08*
If HH is Educated	431.19	0.006	431.18

IV. CONCLUSION

Follwing the results obtained from Voung test, Double hurdle model is selected as the best specification to analyze consumer behavior in demand for internet. So, its estimated parameters which is presented in table 2, were used to calculate conditional and unconditional elasticities for demand with respect to different factors. Our findings suggest that each decision did not affect by the same set of repressors. For example, most of the demographic variables such as education, average age of the household, the number of men in a household, number of children and the number adolescents per household, just affect probability of participation in the internet consumption and doesn't influence amount of expenditure. Another interesting result is the contradictory effect of some geographical variables on participation and consumption decisions. For instance, southern inhabitants while have a less participation, expense more to access the internet. Finally, Computer ownership and Income level are two economic factors which have a positive impact on both decisions.

REFERENCES

[1] Akinbode, S, O.,& Dipeolu, A, O. (2012).Double-hurdle model of fresh fish consumption among urban households in south-west Nigeria, Current Research Journal of Socialsciences,4(6): 431-439, 2012.
[2] Aristei, D., &Pieroni, L.(2012). A double hurdle approach to modelling tobacco consumption in Italy, Applied economics, 40:19, 2463:2476.
[3] Burton, M., R. Dorsett, and T. Young (1996) "Changing Preferences for Meat: Evidence from UK Household Data, 1973-1993," European Review of Agricultural Economics, 23 (357-370).
[4] Chern,W.,K,Ishibashi,& K,Tokoyama.(2003).Analysis of Food Consumption of Japańs Household .FAO Economic and Social Development. 68:152-190.
[5] Carlin, P. and L. Flood (1997) "Do Children Affect the Labor Supply of Swedish Men? Time Diary vs. Survey Data," Labour Economics, 4 (167-183).
[6] Cheng, H. and O. Capps, Jr. (1988) "Demand Analysis of Fresh and Frozen Finfish and Shellfish in the united States," American Journal of Agricultural Economics, 70:3 (533-542).
[7] Flood, L. and U. Gråsjö (1998) Regression Analysis and Time Use Data A comparison of Microeconometric Approaches with Data from the Swedish Time Use Survey (HUS), Göteborg University, School of Economics and Commercial Law, Working Papers in Economics, No 5.
[8] Flood, L. and U. Gråsjö (2001) "A Monte Carlo Simulation Study of Tobit Models," Applied Economics Letters, 8 (581-584).
[9] Cragg, J.(1971). Some Statistical models for limited dependent variables with application to the demand for the durable goods, Econometrica, 23, 59-80.
[10] Deaton, A. and M. Irish (1984) "Statistical Models for Zero Expenditures in Household Budgets," Journal of Public Economics, 23:1-2 (59-80).
[11] Dow, W. and E. Norton (2003) "Choosing Between and Interpreting the Heckit and Two-Part Models for Corner Solutions," Health Services & Outcome Research Methodology, 4 (5-18).

[12] Heckman, J. (1979). Sample selection bias as a specification error, Econometrica, 47, 153-61.
[13] Heien, D. and C. Durham (1991) "A Test of the Habit Formation Hypothesis Using Household Data," The Review of Economics and Statistics, 73:2 (189-199).
[14] Heien, D. and R. Wessells (1990) "Demand Systems Estimation with Microdata: A Censored Regression Approach," Journal of Business and Economic Statistics, 8:3 (365-371).
[15] Jones, A, M. (1992). A note on computation of the double hurdle model with dependence with an application of Tobacco expenditures, bulletin of economic research, 44, 67-74.
[16] Jones, A. M. (1989). A Double-Hurdle Model of Cigarette Consumption, Journal of Applied Econometrics, 4, 23-39.
[17] Jones, A, M., Yen, S, T.(2000). A box-cox double hurdle model, Manchester school, 68, 203-221.
[18] Jones, A, M., & Yen, S. (1996). Individual cigarette consumption and addiction: a flexible limited dependent approach, Health economic, 5, 105-117.
[19] Kalenkoski, C., D. Ribar, and L. Stratton (2005) "Parental Child Care in Single-Parent, Cohabiting, and Married-Couple Families: Time-Diary Evidence from the United Kingdom," American Economic Review, 95:2 (194-198)
[20] Kimmel, J. and R. Connelly (2007) "Determinants of Mothers' Time Choices in the United States: Caregiving, Leisure, Home Production, and Paid Work," Journal of Human Resources, 42:3 (643-681)
[21] McDowell, A. (2003) "From the help desk: hurdle models," The Stata Journal, 3:2 (178- 184).
[22] Moffatt, P. (2005) "Hurdle Models of Loan Default," Journal of the Operational Research Society, 56 (1063-1071).
[23] Saha, A,, O. Capps, Jr. and P. Byrne (1997) " Calculating Marginal Effects in Models for Zero Expenditures in Household Budgets Using a Heckman-Type Correction" Applied Economics, 29:10 (1311-116).
[24] Sabaghkermani, M. and Najafi, N(2005), "Estimating the Internet Demand Function: Case Study of Iran. " Iranian Journal of Trade Studies, 36: 9(14-31)
[25] Salem, A, and Morovat, H.,(2018), " Estimation of Internet Demand by Iranian Urban Households" Quarterly Journal of New Economy and Trade, 13:4 (109-130)
[26] Sayer, L., S. Bianchi and J. Robinson (2004) "Are Parents Investing Less in Children? Trends in Mothers' and Fathers' Time with Children," American Journal of Sociology, 110:1 (1-43).
[27] Tobin, J. (1958). Estimation of relationship for limited dependent variables, Econometrica, 26, 24-36.
[28] Voung, Q, H. (1989). Likelihood ratio test for model selection and non-nested hypothesis, Econometrica, 57, 303-333.