

A Logistic Model of Internet Shopping Adoption

* Dr.S.Arulkumar

* Assistant Professor in Business Administration, Directorate of Distance Education.
Annamalai University, Annamalai Nagar, Tamil Nadu India

Abstract

Internet shopping in India is having rapid growth during the past ten years. This paper applies the variables available in Innovation Diffusion Theory, Technology Acceptance Model and Theory of Reasoned Action to find out the difference in demographic characteristics. After analyzing data collected from 200 respondents, some characteristics of current patterns of internet shopping adoption are concluded. A logit model is adopted to present the correlation with demographic variables in this study. This research offers some insights into the links between e-shopping and consumers' decisions to shop or not shop online. This information can help online marketers and retailers to develop appropriate market strategies, make technological advancements, and make the correct marketing decisions in order to retain current customers and attract new customers.

Keywords: Internet shopping, Adoption, Logistic regression.

Introduction

The growth of Internet and its user base in recent years has been truly phenomenal. Despite the phenomenal growth of the Internet over the past few years, the vast potential of conducting business over the Internet remains largely untapped. This study is designed to provide suggestions to online marketers and e-vendors regarding adoption of online shopping. There are several questions that need to be answered so that internet users do online shopping without any risk, fear or lack of trust. Moreover, demographic characteristics like income and education are also anticipated to influence internet shopping adoption. This study was undertaken to find out the various insights to internet shopping adoption.

Need For the Study

The literature review has revealed that most of the studies have mainly concentrated on trust and risk independently. There have been hardly any studies which take into account the relationship between various variables based on several theories. Whether and how adoption is affected by perceived behavioural control, perceived ease of use, perceived usefulness, perceived risk, website factor, subjective norm, trust, attitude and intent to buy (or vice versa) in the Internet shopping context is still very much an open issue. Moreover, on detailed search the researchers have found only little literature available in the Indian context dealing with adoption and the various components of the same. From literature study it can be seen that all these studies have been done in non-Indian context and particularly focus on a single theory. Through this research, the researchers expect to find the different ways in which shopper adoption, which is understood as a function of culture, in Indian context.

Objectives of the Study

1. To find out the willingness level of males and females with respect to variables related to internet shopping adoption
2. To Know the status of adopters and non adopters with respect the variables related to internet shopping adoption

Review of Literature

Peng et al. (2002) indicate that logistic regression is a commonly used technique for describing and testing hypotheses about relationships between a categorical outcome variable and one or more categorical or continuous predictor variable. The literature on using logistic regression analysis for online shopping behavior is sparse. However, there is

an increasing trend in using logistic regression analysis in economic and behavioral research because of the existence of many discrete variables. The research concludes that “in presenting the assessment of logistic regression results, researchers should include sufficient information to address the following: (1) an overall evaluation of the logistic model; (2) statistical tests of individual predictors; (3) goodness-of-fit statistics; and (4) an assessment of the predicated probabilities”.

Clemes, Gan and Zhang (2010) use a logit analysis to analyze the factors that contribute to bank switching in China. The findings of the authors’ study suggest that price, reputation, service quality, effective advertising, involuntary switching, distance, and switching costs are important factors that have an impact on the Chinese customers’ bank switching behaviour.

Maddala (2001) suggests that logistic regression assumes the existence of an underlying latent variable for which a dichotomous realization is observed. In this study, logistic regression is used to establish associations between the dichotomous dependent variables (consumers adopt/not adopt online shopping) and independent variables identified from the review of the literature and focus group discussions.

Research Methodology

It is to be noted that there are several ways of studying and handling the research objectives. The present study fall under the category of descriptive studies as the nature of problem is to determine the relationship among the different variables. This type of research is also called survey based study. The major strength of survey research has its wide scope and ability to collect the detailed information from a sample of large population.

In the proposed study, the researcher used both primary and secondary data. Primary Data is originated by the researcher for the specific purpose of addressing the problem at hand and secondary data has been collected from various national and international journals, websites, books, working papers, online databases and reports like EBSCO and emerald Library Data was collected from 200 samples using convenient sampling technique. Statistical inferences were drawn from the primary data collected by applying statistical tool like SPSS 20 and statistical analysis like logistic regression.

The Model

In any regression analysis, the key quantity is the mean value of the dependent variable, given the values of the independent variable (Equation 1)

$$E(Y|x) = \beta_0 + \beta_1 x$$

Where Y indicates the dependent variable, x indicates value of the independent variables, and the β_0 and β_1 values denote the model parameters. The estimated quantity is called the conditional mean or the expected value of Y given the value of x. Many distribution functions have been proposed for use in the analysis of a dichotomous dependent variable (Hosmer & Lemeshow, 1989). The distribution function used in the logistic regression model is (Equation 2):

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

Where, to simplify the notation, $\pi(x) = E(Y|x)$. The transformation of the $\pi(x)$ logistic function is known as the logit transformation (Equation 3)

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x$$

Hosmer and Lemeshow (1989) conclude the main features of a regression analysis when the dependent variable is dichotomous are:

1. The conditional mean of the regression equation must be formulated to be bounded between 0 and 1
2. The binomial, not the normal, distribution upon which the analysis is based;
3. The principles that guide an analysis using linear regression will also apply for logistic analysis

If Y is coded as 0 or 1 (binary variable), the expression $\pi(x)$ given in equation provides the conditional probability that Y is equal to 1, given x, denoted as $P(Y=1|x)$. It follows that the quantity $1 - \pi(x)$ offers conditional probability that Y is equal to 0, given x, $P(Y=0|x)$. For those pairs (x_i, y_i) where $y_i = 0$, the contribution to the likelihood function is $1 - \pi(x_i)$, where the quantity $\pi(x_i)$ denotes the values of $\pi(x)$ computed at x_i (Hosmer and Lemeshow, 1989).

A convenient way to state the contribution to the likelihood function for the pair (x_i, y_i) is (Equation 4):

$$\zeta(x_i) = \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i}$$

Since x values are assumed to be independent, the product for the terms given in Equation (4) yields the likelihood function (Equation 5):

$$I(\beta) = \prod \zeta(x_i)$$

The log of equation (5) gives the following log likelihood expression (Equation 6):

$$L(\beta) = \ln[I(\beta)] = \sum \{y_i \ln[\pi(x_i)] + (1 - y_i) \ln[1 - \pi(x_i)]\}$$

Maximizing equation (6) with respect to β and setting the resulting expressions equal to

0 will produce the following values of β (Equation 7 and 8).

$$\sum [y_i - \pi(x_i)] = 0$$

$$\sum x_i [y_i - \pi(x_i)] = 0$$

These expressions are called likelihood equations. An interesting consequence of equation (7) is:

$$\sum y_i = \sum \bar{\pi}(x_i)$$

That is, the sum of the observed values of y is equal to the sum of the expected values.

After estimating the coefficients, the significance of the variables in the model is assessed. That is, the observed values obtained from model with and without the variables in the equation. In logistic regression the comparison is based on the log likelihood function defined in equation (6). The likelihood ratio is given as follows (Equation 9)

$$D = -2 \sum \left\{ y_i \ln \left(\frac{\bar{\pi}_i}{y_i} \right) + (1 - y_i) \ln \left(\frac{1 - \bar{\pi}_i}{1 - y_i} \right) \right\}$$

Where $\bar{\pi}_i = \bar{\pi}(x_i)$.

The dependent variable in this study, adoption of online shopping, is dichotomous.

Therefore, the logit model is:

P (adopters) = Equation 10

$$\pi(x) = \frac{e^{g(x)}}{1 + e^{g(x)}}$$

P (non adopters) = Equation 11

$$1 - \pi(x) = \frac{1}{1 + e^{g(x)}}$$

where $g(x)$ represents the independent variables: perceived behavioural control, perceived ease of use, perceived usefulness, perceived risk, website factor, subjective norm, trust, attitude and intent to buy, adoption, and demographic characteristics (male and female). Due to the fact that the online shopping adoption status is a binary variable, the logit model is used in this research. This model is tested by SPSS software.

Results

In the online context there is a behavioral difference between males, females, adopter and non-adopters. To find out the likelihood of the respondents in making online purchases, respondents were asked to indicate their intention to buy over the internet on a five point scale (0= definitely will not buy and 1 = definitely will buy). Based on a median split approach, the researchers could then classify the respondents into two groups namely adopters (willing to buy) and non-adopters (not willing to buy).

Logistic regression was conducted to assess whether the eleven predictor variables significantly predict the criterion variable. The criterion variable is dichotomous (willing to buy or not willing to buy) and the other predictor variables in the study are gender (Male Female), Perceived behavioural control, Perceived Ease of Use, Perceived Usefulness, Perceived Risk, Attitude, Trust, Website Factors, Subjective Norm, Intention to Buy and Adoption. Respondents have higher score indicating greater levels of influence on particular variable.

The result of logistic regression indicates that the eleven predictors model provides a statistically significant improvement over the constant model (Chi-Square =27.62, $p < .001$). The Nagelkerke Pseudo R^2 indicates that the model accounted for 71 percent of total variance. This suggests that the set of predictors discriminates between those who are willing to buy and not willing to buy. Prediction success for the cases used in the model was relatively high with an overall predication success rate of 78.7 percent and the correct prediction rate of 99.8 percent for willing to buy and 5.4 percent for those respondents not willing to buy. Table 4.40 presents the regression coefficient (B), the Wald statistics, sig level odds ratio (Exp (B)) and the 95 percent confidence interval for odd ratios for each predictor. The WALD test reports that except Perceived risk, other predictors are not significant.

Table – 1: Omnibus Tests of Model Coefficients

	Chi-square	Df	Sig.	Nagelkerke R Square
Step 1 step	27.623	11	.004**	.071
Block	27.623	11	.004**	
Model	27.623	11	.004**	

Source: Primary data, ** Significant at five percentage level

Table – 2: Internet Shopping Adopters and Non-Adopters Classification

Observed	Predicted		
	Likeliness		Percentage Correct
	Non Adopters	Adopters	
Non Adopters	7	122	5.4
Adopters	1	448	99.8
Overall percentage			78.7

a. The cut value is .500 Source: Primary data

Table – 3: Variables in the Equation

Variables	B	S.E.	Wald	df	Sig.	Exp(B)
Perceived Behavioural Control	.137	.163	.705	1	.401	1.147
Perceived Ease of Use	-.370	.197	3.540	1	.060	.690
Perceived Usefulness	.117	.206	.323	1	.570	1.125
Perceived Risk	.209	.100	4.323	1	.038	1.232
Trust	.152	.164	.863	1	.353	1.164
Subjective Norms	-.132	.160	.683	1	.409	.876
Website Factors	.096	.213	.204	1	.651	1.101
Attitude	.894	.990	.817	1	.366	2.446
Intention	-.357	1.012	.125	1	.724	.699
Adoption	-.134	.260	.268	1	.605	.874
Gender	.062	.234	.071	1	.791	1.064
Constant	-2.005	1.011	3.935	1	.047	.135

Source: Primary data

The influence of gender is strong; males are .80 times (CI=.673 and 1.683) more likely to purchase than females adjusting for perceived behavioral control score. For each point increase in perceived behavioral control score, there is 0.40 times greater likelihood of purchase, controlling gender: For perceived risk 0.69, perceived usefulness 0.57, perceived risk 0.04, trust 0.35, subjective norm 0.41, website factors 0.65, attitude 0.37, intention to buy 0.72 and adoption 0.61 increase in likelihood of purchase in future.

Discussion

The logistic regression analysis on the set of variables analysed has highlighted the fact that age, frequency of Internet shopping are the variables which best predict adoption

level. Gender and frequency of Internet use have not turned out to be determinant factors in the purchase decision. The possible differences due to gender in internet shopping (and all virtual environments in general) tend to disappear as a consequence of changes in social habits and in the greater level of introduction and development of the new technologies [Modahl 2000; Siegel 2003]. Thus, the findings of this study confirm the results of other research [Modahl 2000; Rosenberg and Hirschman 1980] which highlights the fact that women are becoming increasingly familiar with virtual environments and that Internet use is becoming more and more widespread [AIMC 2005; Siegel 2003].

Limitations

The present finding should be interpreted with some caution, based on several limitations. However, necessary remedial actions have been taken to minimize the possible effect of these limitations in the results.

Firstly, there is always the issue of generalizability in consumer behavior studies, and the present study is no exception. As this study was conducted on internet shoppers in Tamilnadu, whose purchasing power and patterns were influenced by socio-economic studies and lifestyles, the generalizability of the findings in other parts of India may be limited.

Secondly, the current investigation was executed in the context of internet shopping. Therefore, the findings of the study were not focused on any particular industry.

Finally, the measures of all the research constructs were collected at the same point of time and via the same instrument; so the potential for common method variance may exist. However, there was no indication of lack of discriminate validity among the principal constructs, the usual sign of common method variance. In future controlled experimental manipulation prevents respondents from providing uniform responses across constructs.

Conclusion

Based on 200 responses, this study examined the nature of demographic differences between adopters and non-adopters of internet shopping. This results show that the influence of gender is strong; males are .80 times (CI=.673 and 1.683) more likely to purchase than females adjusting for perceived behavioral control score. For each point increase in perceived behavioral control score, there is 0.40 times greater likelihood of purchase, controlling gender: For perceived risk 0.69, perceived usefulness 0.57, perceived risk 0.04, trust 0.35, subjective norm 0.41, website factors 0.65, attitude 0.37, intention to buy 0.72 and adoption 0.61 increase in likelihood of purchase in future. There is no doubt that the success of this new direct sales channel will be conditioned by both the group of product/service offered and their cost, since the focus groups and previous studies [Net size 2004] show that the products/services most acquired by users of these systems are those which are in fashion and offer the least perceived purchase risk

References

1. Clemes, M. D., Gan, C., & Zhang, D. (2010). Customer switching behaviour in the Chinese retail banking industry. *International Journal of Bank Marketing*, 28(7), 519-546.
2. Hosmer, D. W., & Lemeshow, S. *Applied Logistic Regression* (1989) New York.NY: John Wiley & Sons
3. Maddala, G. S. (2001). *Introduction to econometrics* (3rd ed.). New York: Wiley
3. Modahl, M., *Now or never*. Harper Collins: New York, 2000
4. Netsize, *Developing the multimedia mobile market, The Netsize guide*, pp.1-81, 2004.
4. Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic Regression analysis and reporting. *The Journal of Educational Research*, 96(1), 3-14.
5. Rosenberg, L. And E. Hirschman, "Retailing without stores", *Harvard Business Review*, Vol. 58, No.4:103-112, 1980
6. Siegel, C., *Internet Marketing: Foundations and Applications*. Houghton Mifflin Company: Boston, 2003.