

# Analysing Website Choice and Consumer Loyalty:

## The case of Book and CD Markets

**Asmaa KHARIJI<sup>1</sup>**

31<sup>st</sup> May, 2006

### **Abstract**

This paper estimates the relationship between the choice of website and the characteristics of the individual making this choice. What explain consumer website choice and is this consumer loyal to the same website? This paper approaches this question by modelling individual's website choice for buying book and CD online among a set of alternatives within a period of six months (from July 1 to December 31 2002). I find that individuals respond strongly to well-known, branded online retailers and that some individual demographics explain the website choice and the loyalty behaviour. However, the results point to a not switching effect online and rules out there is a consumer loyalty.

---

<sup>1</sup> I would like to offer a special thanks to Philippe Barbet. I am also grateful to Samia Benallah, Yosef Bonaparte, William Brennan, Abe Dunn, Yingyao Hu, Karim Kefi, Christophe Le Guehennec, David Sibley and especially to Eric Delatre and Randal Watson for advices and encouragements. I also thank the CEPN and Benjamin Coriat for financial support.

## 1 Introduction

According to the Pew Internet & American Life Project, 67% of Internet users in the US report buying a product between May and June 2005. Furthermore, according to their February-April 2006 survey, 73% of American adults use the Internet. That currently represents approximately 147 million people. Though Internet is still at an early stage of development, these statistics suggest that it may change dramatically in the coming years. Yet, the Internet is still perceived as a « black box » in which little is provided about individual-level online behaviour. Therefore, a key task for electronic commerce is to find out who the actual and potential customers are and to clarify their choice to buy a media product online. This paper aims to estimate website choice and consumer loyalty online.

My data consists of panel data on a nationally representative sample of internet users. The aim is to understand the behavioural process that leads to the individual's website choice and loyalty online. I address these questions through panel data gathered from *ComScore* in the market for book and CD. *ComScore* data can be associated to clickstream data. Clickstream data provides researchers with an opportunity to observe individual choice behaviour as individuals evaluate the listed alternatives and click on a product offer on a particular website (Goldfarb A., 2002a). This data set provides detailed information on the consumers' internet information and demographics, and book and CD purchases. Individuals then choose to buy at least one product from a website. I model their choices using discrete-choice, multinomial logit and probit regression model. These two models are estimated using a maximum likelihood function.

The main empirical finding is that there is no switching effect across websites and that there is a consumer loyalty for homogeneous goods such as book and CD online. The results also suggest that some individual demographics explain the website choice and the loyalty behaviour. Developing an understanding of online behaviour on many aspects of life today is currently of interest to economists and practitioners as the internet has changed the way the individuals purchase media products. Measures of individual purchases and loyalty can serve as an indicator for site performance.

This approach to analyzing individual's behaviour online complements recent empirical studies. There is a small but growing literature on consumer behaviour on the Internet. Johnson, Lohse, & Bellman (2000), Moe & Fader (2004a, 2004b), and Johnson et al. (2004)

look at search behaviour at various types of e-commerce sites. Bucklin & Sismeiro (2001, 2003) explore the individual online decision between continuing to browse or to exit the site. Brynjolfsson and Smith (2001) find that brand is an important determinant of consumer choice. The identification of switching effect is notoriously difficult in any analysis of individual behaviour. Some previous empirical studies contribute to this related analysis by seeking to identify the extent of switching costs online (Goldfarb 2006). Online switching behaviour was explored by Chen & Hitt (2000) at Internet brokers and by Telang, Mukhopadhyay, & Wilcox (1999) at search engines. Nevertheless, there has been considerable discussion about loyalty online. Shapiro & Varian (1999) emphasize that the competition is just one click away. Gandal (2001b, p. 1105) claims that “there are little (if any) consumer switching costs” at Internet portals while Goldfarb (2002b, 2006) analysis offers evidence of the existence of switching costs in the Internet portal market, and users differ in these costs.

The remainder of this paper is organized as follows. The source of data is covered in Section 2 while section 3 presents the empirical models we use to analyze our data. Section 4 reports results and conclusions.

## **2. Data**

The main data source used in this paper is ComScore from July 1<sup>st</sup> to December 31<sup>st</sup> 2002. Comscore Web Behavior (panelist-level) database captures detailed browsing and buying behavior by 100 thousand Internet users across the United States. The disaggregate dataset is based on a massive random sample from a cross-section of more than 1.5 million global Internet users who have given *ComScore* explicit permission to confidentially capture their Web-wide activity. The original 2002 sample is very large (about 100 million observations, last updated May 2003). In this case, The 2002 *ComScore* dataset contains the result of a survey that is a sample from the set of books and CD purchased in 6 months of behavioural activity.

For each member of their panel, ComScore records every website visited, the date of arrival, the name of the product purchased, the price of the product, the basket total price, the user machine ID, the user session ID, the website name, the site category ID, the individual education, the individual income, the individual age, the household size, the child presence, the racial background, the connection speed and the country of origin. Unfortunately,

ComScore data does not provide detailed information on each website characteristics. These data are used to determine how consumers respond to the presence of a set of websites alternatives. One of the most important aspects of this study is the individual ability to explore different choices and choose the website that is better tailored and provides the highest utility.

## 2.1 Variables

The dependent variables used in this paper are the website choice and the consumer loyalty. I assume that the individual always purchase at least one product and that he/she chooses the website that provides the greatest utility. The original data book and CD samples are very large: respectively 27, 288 observations with 43 different websites and 9,900 observations with 29 websites for the multinomial logit model. Concerning the probit model the book and CD data sets contain respectively 14,029 and 4,299 observations.

I adopt a strategy of utilizing as many observations as possible by focusing on the book basket total price instead of the book product price. I delete observations for each individual by basket total price and by date. Then I delete observations with missing values for any of the consumer attributes I used (e.g., the individual education), and am left with about 11,791 and 5,548 observations for the multinomial logit model and 6,086 and 2,254 observations for the probit regression analysis.

I turn next to a description of the variables of the model. The dependant variables are multivariate variables for the multinomial logit model. Concerning the book market  $Website_i$  is equal to 1 if individual  $i$  chooses Amazon.com over the other websites, is equal to 2 if individual  $i$  chooses Barnesandnoble.com over the other websites, is equal to 3 if individual  $i$  chooses Ebay.com over the other websites. In the case of CD market:  $Website_i$  is equal to 1 if individual  $i$  chooses Amazon.com over the other websites, is equal to 2 if individual  $i$  chooses Cdnnow.com over the other websites, is equal to 3 if individual  $i$  chooses Columbiahouse.com over the other websites, and is equal to 4 if individual  $i$  chooses Ebay.com over the other websites. Individual  $i$ 's characteristics are described by six dummy binary variables:  $CH_i$  for child presence,  $INT_i$  for the Internet connection speed,  $CO_i$  for the country of origin,  $RAC_i$  for the racial background,  $REG_i$  for the census region,  $SEA_i$  for the seasonal period. Basically, child presence is a binary variable that is equal to 1 if individual  $i$  has a child, connection speed is equal to 1 if individual  $i$  has a broadband access, country of origin is equal to 1 if

individual  $i$  is from an Hispanic country, racial background is equal to 1 if individual  $i$  is white, census region of residence is equal to 1 if individual  $i$  lives in the South, seasonal period is equal to 1 if individual  $i$  buys book or CD online during the winter period (from October 1<sup>st</sup> to December 31<sup>st</sup> 2002). The rest of explanatory variables have been used as continuous variables: the individual age  $AGE_i$ , income  $INC_i$ , and the household size  $SIZ_i$ .

The products in the markets are: book and CD. The price of each product is denoted  $PRICE_{ij}$ . The price chosen is the one that includes the transaction price of the number of book or CD purchased and the delivery.

In addition to a multinomial logit model I also use an alternative specification. The multinomial logit model assumes that an individual  $i$  is presented with a set of  $J$  alternatives. However, we can imagine some situations where an individual is facing only two alternatives. Therefore, I also estimate website choice and consumer loyalty with a probit regression analysis as well. The dependant variables are binary variables for the probit model. In both book and CD markets,  $Website_i$  is equal to 1 if individual  $i$  chooses Amazon.com over Barnesandnoble.com, Cdnnow.com, or Ebay.com. Again, I incorporate the same explanatory variables but I add some online specific characteristics as: the time, i.e., the time at which an individual buys a product offer from a particular website; the duration, i.e., how long it takes for an individual to make a decision before buying a product offer; and the pages viewed, i.e., the number of pages visited on a website when an individual decides to make a product (s) purchase.  $TIM_i$  is a binary variable that is equal to 1 if individual  $i$  makes a purchase decision on a night-time (from 8pm to 6am). I use  $PAG_i$  and  $DUR_i$  as continuous variables. However, due to a lack of data the  $TIM_i$ ,  $PAG_i$  and  $DUR_i$  explanatory variables are only used in the last probit regression, i.e., in a situation where the individual chooses Amazon.com over eBay.com.

## 2.2 Descriptive Statistics

For the multinomial logit model Table 1, Table 2 and Table 3 compare the distribution of household characteristics in Amazon.com sample to those in Barnesandnoble.com, eBay.com and the other websites. Table 1 provides the fraction of the respective samples in each of six different income groups. Table 2 and Table 3 compare the individual attributes

other than income (age, connection speed). Apparently, the different samples have on average the same shares in percent in different income, age range and type of connection speed. Table 4 provides the mean individual characteristics by type of website chosen. Several relationships between household characteristics and website choice stand out from these tables; some more expected than other. Among the expected, individuals with high income purchase book online and age seems to be so important a determinant of website choice. Somewhat more surprising is that on average individuals do not have any child though average size of household is 3. Interestingly, the Barnesandnoble.com sample has somewhat more individuals living in north-central, and the eBay.com sample has relatively younger individuals purchasing book online.

According to Table 5, Table 6, 7 and Table 8 and compare to the previous book sample, the CD sample has larger fractions of younger individuals. Moreover, the Columbiahouse.com sample has somewhat more individuals having a child. And one alternative, Cdnw.com, has more individuals living in the north-central like in the book dataset.

Concerning the probit model for the book market Table 9, Table 10 and Table 11 compare the distribution of household characteristics (income, age, connection speed) in Amazon.com sample to those in eBay.com. Apparently, the different samples have different percent shares within the group incomes, e.g., Amazon.com sample has a larger fraction of individuals with higher income while eBay.com sample has a larger fraction of individuals with lower income. Table 12 provides the mean individual characteristics by website chosen while Table 13 displays the individual characteristics by type of website. The eBay.com sample has somewhat more individuals living in the north-central while the Amazon.com sample has more individuals living in the south. Moreover, the eBay.com sample has relatively a larger fraction of individuals having a broadband access. Maybe the most striking difference between the two samples is that the eBay.com sample spends significantly more time online to make a purchase decision which can also explain the higher number of pages viewed on eBay.com. Besides, the fraction of individuals purchasing book online is higher on a night-time (from 8pm to 6am) than on a day-time (from 6am to 8pm) as displayed in Table 14. Likewise during the winter season, the fraction of individuals making a purchase decision is slightly higher (Table 15) compare to the summer season.

Finally, for the probit model on the CD market Table 16, Table 17 and Table 18 compare the distribution of household characteristics (income, age, connection speed) in Amazon.com sample to those in eBay.com. Interestingly, Amazon.com still has a larger share of individuals with a higher income. Less difference appears within the range age in both the Amazon.com and the eBay.com samples. Table 19 provides the mean individual characteristics by website chosen while Table 20 displays the individual characteristics by type of website. The mean characteristics between the two alternative samples are similar except that the mean income for the eBay.com sample is relatively lower. Interestingly, the eBay.com sample still spends more time online to make a purchase decision for CD. Nevertheless, the fraction of individuals purchasing CD online is significantly higher on a night-time (Table 21) compare to the book sample. But similarly, the fraction of individuals purchasing CD online is higher during the winter season (Table 22) compare to the book sample.

### 3. The Model and Method of Estimation

#### 3.1 Random Utility Models

Random utility theory (Manski, 1977) enables the probabilistic nature of choice to be incorporated into choice models where individuals' choices are intransitive and inconsistent. An individual, labelled  $i$ , faces a choice among  $J$  alternatives. As I stated previously, the individual would obtain a certain level of utility (or net benefit) from each alternative. The utility that the individual  $i$  obtains from alternative  $j$  is  $U_{ij}$ ,  $j = 1, \dots, J$ . This utility is known to the individual but not by us. Random utility theory assumes that the individual  $i$  will choose the website  $j$  with the highest utility  $U_{ij}$  which consists of two components: a deterministic component  $V_j$  which is specified as a function of the attributes of the alternatives and a random component  $\varepsilon_j$ . The random component represents the inability of the modeller to accurately include all factors affecting preferences. I will assume that the observable and unobservable characteristics are constant over time, and captured by an extreme value error  $\varepsilon_{ij}$ . Thus, individual  $i$ 's utility for website  $j$  can be written as:

$$U_{ij} = V_{ij} + \varepsilon_{ij} \quad (1)$$

Where  $U_{ij}$  is the utility of website  $j$  to individual  $i$  and the deterministic component  $V_{ij}$  is assumed to have an additively separable linear form:

$$V_{ij} = \mathbf{x}_{ij}\boldsymbol{\beta} + \mathbf{z}_i \alpha_i + CB_i \quad (2)$$

Where  $\mathbf{x}_{ij}$  is a vector of variables that vary with  $i$  and  $j$ , for example price,  $\mathbf{z}_i$  is a vector of variables that vary with  $i$ , for example age or income, and  $C$  is a scalar taking the value 1, the dummy constant. Finally, vectors  $\beta$ ,  $\alpha_i$ , and scalar  $B_i$  represent the corresponding parameters.

Discrete choice models are often derived from the principle of maximum random utility (Thurstone, (1927a), Marschak, (1960). It is assumed that an unobserved utility  $U_k$  is associated with the  $k$ th alternative, and the response function  $Y$  is determined by:

$$Y = k \Leftrightarrow U_k = \max \{U_l, 1 \leq l \leq m\} \quad (3)$$

Discrete choice models are now used in a wide variety of situations in applied econometrics (McFadden, 1981) and can be broadly classified as binary or multinomial. Binary models examine an individual's choice to purchase or not to purchase a particular product, while multinomial models examine a consumer's choice among a larger set of possible products. I first model individual's website choice for buying book and CD online by using a multinomial logit model then I use a probit model with a set of two alternatives that includes specific individual characteristics online. I also complete this analysis with a probit regression of switching effect on individual demographics in order to see if online loyalty can be a function of individual demographics. I first sketch the two different models and then turn to the estimation procedure.

### 3.2 Multinomial Logit Model

By far the model specification which is used most often is the multinomial logit model (McFadden, 1973; Train, 2003). The multinomial logit (MNL) model is based on the assumption that  $U_1, \dots, U_m$  are independently distributed and each follows an extreme maxima value distribution (Hoffman and Duncan, 1988). Yet it is widely known there exists a potentially drawback of the MNL model, termed the independence from irrelevant alternatives property (IIA). Debreu (1960) was among the first economists to discuss the implausibility of the independence from irrelevant alternatives assumption. Simply stated, the IIA property holds that the ratio of the choice probabilities of any two alternatives (in response categories) for a particular observation is not influenced systematically by any other alternatives. In this case, I alleviate the IIA problem by modelling the choice of any alternative over the other websites left in our dataset. Indeed, the MNL model requires the specification of an alternative option reflecting the choice of none of the websites considered



(the “inside” choices). Then consider an individual choosing among  $n$  alternatives in a choice set. Let  $\Pi_{ik}$  denote the probability that an individual  $i$  chooses alternative  $k$ , let  $X_i$  represent the characteristics of the individual  $i$ . The MNL model focuses on the individual as the unit of analysis and uses individual characteristics as explanatory variables. The explanatory variables, being characteristics of an individual, are constant over the alternatives. The probability that individual  $i$  chooses alternative  $k$  is (McFadden, 1981):

$$\Pi_{ik} = \frac{\exp(\beta_k X_i)}{\sum_{l=1}^m \exp(\beta_l X_i)} \quad (4)$$

where  $\beta_1, \dots, \beta_m$  are  $m$  vectors of unknown parameters (each of which is different, even though  $X_i$  is constant across alternatives). By setting the last set of coefficients to the null (that is,  $\beta_m = 0$ ), the coefficients  $\beta_k$  represent the effects of the  $X$  variables on the probability of choosing the  $k$ th alternative over the last alternative. In fitting such a model, I estimate  $m - 1$  coefficients.

### 3.3 Probit Model

The probit model is a popular device for explaining binary choice decisions in econometrics (Nelson and Aldrich, 1984)<sup>2</sup>. “Probit” is an abbreviation of the term “probability unit” (the term is attributed to C. I. Bliss, 1935) and was the first such model developed and studied (Finney, 1971).  $Y^*$  is often an observation on the behaviour of an individual in a situation where the individual is faced with a choice of selecting between two alternatives. The rational choice approach (Luce and Suppes, 1965) asserts that the individual has preferences over these two alternatives, and that the individual will choose the most preferred alternative. Here, our example models website choice between Amazon.com and eBay.com. This model is consistent with an econometric approach (Goldberger, 1964; Maddala, 1983) that assumes an underlying response variable  $Y_i^*$  defined by the regression relationship:

$$Y_i^* = \sum_{k=1}^K \beta_{ik} x_{ik} + \varepsilon \quad (5)$$

Where  $Y_i^*$  represents the latent (unobservable) preference for an alternative,  $\beta_{ik}$  measures the effect of a change in  $x_{ik}$  on the unobserved variable  $Y_i^*$ , and  $\varepsilon$  is symmetrically distributed with zero mean and has its cumulative distribution function defined as  $F(\varepsilon)$ . What we do observe is a dummy variable  $Y$ , a realization of a binomial process, defined by:

---

<sup>2</sup> Many examples can be found in Amemiya (1981) and Maddala (1983).

$$Y = 1 \text{ if } Y_i^* \geq 0 \quad (6)$$

or  $Y = 0$  otherwise.

The choice perspective says that an individual  $i$  chooses alternative one over alternative two if  $Y_i^* \geq 0$ . But, by equation (7), this means that alternative two is chosen when  $Y^* < 0$ . In our case, alternative one refers to Amazon.com while alternative two refers to eBay.com. Then we are led naturally to a probabilistic statement:

$$P(Y=1) = P(Y_i^* \geq 0) \quad (7)$$

The model is then given by:

$$P(Y = 1|X) = \Phi (\Sigma \beta_{ik} x_{ik}) \quad (8)$$

Where  $P$  is the probability that  $Y$  equals one,  $\Phi(\cdot)$  is the standard normal cumulative distribution function, and  $X$  denotes the set of  $K$  independent variables.

### 3.4 Maximum Likelihood Estimation

Probit and logit parameters are typically estimated by a method called Maximum Likelihood Estimation (MLE). Since our objective is to estimate the “ $\beta$ ” coefficients, we proceed to make this dependence on  $\beta$  explicit by defining the likelihood function,  $L(Y|X, \beta) \equiv P(Y|X)$ . The principle of MLE is to choose as an estimate of  $\beta$  that set of  $K$  numbers which would make the likelihood of having observed this particular  $Y$  as large as possible. In particular, each “value” of  $\beta$  will yield a value of  $L(Y|X, \beta)$ . We take as the MLE estimate  $\tilde{\beta}$  that particular value for  $\beta$  which yields the largest value,

$$L(Y|X, \tilde{\beta}) = \max_{\beta} L(Y|X, \beta) \quad (9)$$

In MLE we proceed to find  $\beta$  so as to maximize the probit likelihood,

$$L(Y|X, \beta) = \prod_{i=1}^N [\Phi(\Sigma \beta_{ik} x_{ik})]^{Y_i} [1 - \Phi(\Sigma \beta_{ik} x_{ik})]^{1-Y_i} \quad (10)$$

with a random sample denoted by  $i, i=1, \dots, N$ .

Or the logit likelihood,

$$L(Y|X, \beta) = \prod_{i=1}^N \left[ \frac{\exp(\Sigma \beta_k X_i)}{1 + \exp(\Sigma \beta_k X_i)} \right]^{Y_i} \left[ \frac{1}{1 + \exp(\Sigma \beta_k X_i)} \right]^{1-Y_i} \quad (11)$$

## 4. Results

### 4.1 Parameter Estimates

Before turning to the estimates for the parameters of the probit model, it will be worth considering a multinomial logit estimation. Tables 23a), b) and c) display the multinomial logit estimates for the parameters of the model for the online book market. On the whole, the coefficients correspond closely to what we would have expected though the connection speed, the country of origin, the household size, the racial background, and the child presence coefficients miss significance. There are three intercepts coefficients and three slope coefficients for each dummy that we already have for some binary variables (child presence, connection speed, country of origin) or that we have created for our estimation (seasonal period, racial background, census region of residence). The rest of explanatory variables have been used as continuous variables (age, income, household size). With the other websites as the reference mode, the intercept for Amazon.com, which is positive, may reflect a higher utility of the individuals for choosing Amazon.com. Likewise the positive intercepts for eBay.com reveal the same behaviour of choosing eBay.com over the other websites. The coefficients in the utility of Amazon.com are consistent with its reputation as having a higher national recognition. Indeed, Amazon.com is a company that has been very successful in building a strong reputation on the Internet (Rindova and Kotha, 1999). According to the *Economist* (1997a: 9), “Companies around the world are studying it [Amazon.com] as perhaps the best model for tomorrow’s successes in electronic commerce.” In this respect Amazon.com is an example of a firm that can be viewed as a “revelatory case” (Yin, 1994). Amazon.com is effectively recognized for using a “recommender system” to help advise customers on their purchases (Brynjolfsson, Smith and Hu, 2003). But not only Amazon.com, also Barnesandnoble.com and eBay.com are now household names. The first one is Amazon.com’s direct competitor in online book retailing—Barnesandnoble.com. Yet Barnesandnoble.com, unlike Amazon.com, is the online division of Barnes & Noble, the US industry leader in “traditional” book retailing. The second one has been successful as an auction site—eBay.com. Among the significant coefficients age has a negative impact, i.e., older individuals do not choose Amazon.com and eBay.com over the other websites. Though of higher magnitudes for Amazon.com and Barnesandnoble.com, the income coefficients reveal that wealthier individuals choose Amazon.com, Barnesandnoble.com and eBay.com over the other websites. Moreover, the seasonal period coefficients reveal that on a winter period the individuals do not choose eBay.com over the other websites. Besides, living in the south has a negative effect of choosing Barnesandnoble.com over the other websites. The coefficients of the individual expenses show almost a null effect of choosing Amazon.com

and Barnesandnoble.com, while the same coefficient reveals that a higher amount of expenses has a negative effect of choosing eBay.com relatively to the other websites.

The above multinomial logit model shows that online individuals respond strongly to well-known, heavily branded online book retailers. The question as to what explain individual website choice can be addressed through probit regression analysis as well. This allows us to distinguish between the two empirical analyses. With Barnesandnoble.com or eBay.com as the reference mode, the intercept for Amazon.com is significant and reflects a higher utility of the individuals for choosing Amazon.com (Tables 25a) and b)). More specifically, the income coefficient reveals that wealthier individuals choose Amazon.com relatively to Barnesandnoble.com or eBay.com. Additionally, the seasonal period has also a significant effect of choosing Amazon.com over eBay.com while the same coefficient reveals a negative effect of choosing Amazon.com relatively to Barnesandnoble.com. And surprisingly, the racial background and the country of origin have a positive influence on choosing Amazon.com over Barnesandnoble.com. Yet, the number of pages viewed coefficient reveals almost a null effect of choosing Amazon.com over eBay.com. Likewise age reveals almost a null effect of choosing Amazon.com over Barnesandnoble.com as well. And interestingly, a higher amount of expenses has a positive effect of choosing Amazon.com over eBay.com while the same coefficient reveals almost a null effect of choosing Amazon.com over Barnesandnoble.com.

Tables 24a), b), c), d) display the multinomial logit estimates for the parameters of the model for the online CD market. With the other websites as the reference mode, the intercepts for Amazon.com and eBay.com reflect a higher utility of the individuals for choosing Amazon.com and eBay.com. This presents an interesting point that needs to be addressed here. The intercept coefficient is, effectively, still significantly higher for choosing eBay.com over the other websites compare to the book sample. And though the intercept for Cdnw.com is not significant the intercept for Columbiahouse.com reflects a higher utility for choosing Columbiahouse.com relatively to the other websites. Concerning its reputation, Cdnw.com was launched before Amazon.com but has a lower level of national reputation. Unlike Amazon.com, Cdnw.com and eBay.com, Columbiahouse.com is rather club based. First launched in 1996, the users have to join Columbiahouse.com by participating in an initial offering that includes a fixed number of products. Of the socio-demographic variables that were included in our model, the seasonal period has a significant but negative effect on the

website choice (except for choosing Amazon.com). Moreover, income has a significant and positive effect of choosing Amazon.com, but it has also a negative effect of choosing Cdnw.com, Columbiahouse.com and eBay.com. Likewise the connection speed and the country of origin have a negative effect on the website choice. Interestingly, the individuals with a white racial background do not choose Columbiahouse.com. An opposite relationship is observed for the census region, as living in the south has a significant and positive effect of choosing Cdnw.com. And though the child presence has a negative effect of choosing eBay.com, surprisingly the household size has a positive effect of choosing eBay.com and Columbiahouse.com. The coefficient of the individual expenses still reveals that a higher amount of expenses has a negative effect of choosing eBay.com like in the previous parameters results for the online book market. However the same coefficients also show almost a null effect of choosing Cdnw.com and Columbiahouse.com.

In the probit model for the online CD market, Tables 26a) and b) show that the intercept for Amazon.com still reflects a higher utility of the individuals for choosing Amazon.com over eBay.com. Additionally, the seasonal period still has a positive effect of choosing Amazon.com over eBay.com or Cdnw.com. Yet, age and the household size have a negative effect of choosing Amazon.com over eBay.com while income still has a positive effect of choosing Amazon.com over eBay.com or Cdnw.com. Surprisingly, the racial background still has a positive influence on choosing Amazon.com over Cdnw.com. Moreover, the amount of individual expenses and the duration coefficients reveal almost a null effect of choosing Amazon.com over eBay.com or Cdnw.com. Nevertheless, the number of pages viewed coefficient reveals a negative though almost null effect of choosing Amazon.com relatively to eBay.com.

#### **4.2 Consumer Loyalty?**

The above empirical evidence, together with the theoretical model developed here, suggest us that a switching effect across websites can play a role in website choice. The identification of switching effect is notoriously difficult in any analysis of individual behaviour. In this paper a positive switching effect is identified if the current website choice is independent from the previous website choice. According to Tables 27, 28, 29, 30, 32, 33, 34, and 35 our evidence points to a not significant switching effect online and rules out there is a consumer loyalty.

The same figures display the sum of multiples (i.e., the number of times the same individual went online) and the fraction of switches (i.e., the number of switches per individual divided by the number of times the same individual went online) for the probit model. In the online book market in a situation where the individuals choose Amazon.com over Barnesandnoble.com, 90.54% of the sample do not switch (i.e., 6,324 individuals) However, one should also consider the subpopulation that went online once within the same period, i.e., 3,526 individuals. Yet, only 3.58% of the data sample switches half of the time. Concerning the situation where the individuals choose Amazon.com over eBay.com, almost 94% of the individuals do not switch (i.e., 5,719 individuals) while approximately 2% of the data sample switches half of the time. Interestingly, 2,929 out of 5,719 individuals went online once from July 1<sup>st</sup> to December 31<sup>st</sup> 2002.

A similarly relationship is observed when the individuals choose Amazon.com over Cdnnow.com or eBay.com, as less than 2% of the CD sample switches half of the time. More precisely, in a situation where the individuals choose Amazon.com over eBay.com, 1,095 out of 2,151 individuals who do not switch went online once. Likewise, in a situation where the individuals choose Amazon.com over Cdnnow.com, 1,079 out of 1,781 individuals who do not switch went online once. These main results involve that a switching effect is not such a significant aspect for homogeneous goods such as book or CD.

Nonetheless, Tables 37, 38, 40, and 41 from the multinomial logit model data samples suggest some different interpretations. In the online book market, approximately 25% of the population switch across websites with almost 9% out of the 25% switching half of the time. Yet, approximately half of the population who do not switch went online once. Conversely, in the online CD market more than 82% of the population do not switch, i.e. 4,578 individuals. Among individuals who do not switch 2,207 went online once.

Another important aspect of this study will be to explore if some individual characteristics can influence the loyalty behaviour and if the amount of expenses has any influence on the switching effect. A switching effect can be a function of individual characteristics. Tables 31a), 31b), 36a), 36b), 39 and 42 show the results of regressing switching effect on individual demographics in the book market. The child presence, the census region, the individual income and the country of origin were found not to influence switching effect in most cases in the book market. Likewise in the CD market, in addition to

the previous individual demographics (except the individual income) the seasonal period, the individual racial background, and the household size were found not to influence switching effect. Some other variables in the regression like the connection speed and age have a positive influence on switching effect in both book and CD markets.

Yet, when the individuals choose Amazon.com over eBay.com, the mean expenses of individuals who went online once is higher than the mean expenses of those who went online more than once. However, in a situation where the individuals choose Amazon.com over Barnesandnoble.com, I observe some opposite results. Additionally when the switching effect occurs, the mean amount of expenses falls compare to the mean amount of the all sample in a situation where the individuals choose Amazon.com over eBay.com. Conversely, the same amount remains higher compare to the amount of the all sample in a situation where the individuals choose Amazon.com over Barnesandnoble.com.

Now looking at the different CD mean amount of expenses from the probit model data sample, the results show that the mean basket expenses falls drastically when the switching effect occurred when the individuals choose Amazon.com over eBay.com or Cdnw.com. Considering the subpopulation of individuals choosing Amazon.com over eBay.com and going online more than once, the mean expense is still lower than the mean expense of individuals going online only once.

Surprisingly, when looking at the book and CD mean expenses in the MNL model data sample the results show that when the switching effect occurs the mean amount of expenses is higher compare to the mean sample in the CD market. And the mean amount of expenses is slightly lower than the mean sample in the book market. Nevertheless, the mean expenses of individuals who went online once is still higher compare to the mean expenses of individuals who went online more than once.

## **Conclusion**

This paper presents results from two different methods for estimating the website choice for book and CD online and sees if there is any switching effect from one website to another from a set of alternatives. The results suggest that there is no switching effect across websites and that some individual characteristics explain the website choice. Using a discrete choice model allows to estimate households' responses to website choices online. The choices I estimate

incorporate the interdependence of the choice of website with the individual characteristics. Furthermore, the estimation procedure is fairly simple. Our analysis provides substantive information on how households respond to website choice for buying book and CD online. However in terms of the limitations of this study, the data analysed is based on the year 2002 and for this reason there are updated information not included in the dataset which I think would be interested to analyse. For this reason, and bearing in mind the lack of research in this field, I think it would be very useful to complement this study with a recent updated dataset. For this reason, I consider that another interesting line of research would be to contrast the validity of the proposed behavioural model online with samples of consumers from recent years and compare the results obtained.

## References

Aldrich, J. H., Nelson, F. D. (1984), *Linear Probability, Logit, and Probit Models*, Sage University Papers Series, Quantitative Applications in the Social Sciences; No. 07-045, California Sage Publications.

Amemiya, T. (1981), "Qualitative response models: A survey", *Journal of Economic Literature*, 19, 1483-1536.

Bellman, S., Lohse, G.L., Johnson, E.J., (1999), "Predictors of Online Buying Behavior", *Communications of the ACM*, 42(12), 32-38.

Bliss, C. I. (1935), "The calculation of the dosage-mortality curve", *Annals of Applied Biology*, 22, 134-167.

Brynjolfsson, E., Smith, M. D. (2001), "Consumer Choice Behaviour at Internet Shopbots", MIT Sloan School of Management, Cambridge.

Brynjolfsson, E., Smith, M. D. and Hu Yu J. (2003), "Consumer Surplus in the Digital Economy : Estimating the Value of Increased Product Variety at Online Booksellers", MIT Sloan School of Management, Cambridge.



Bucklin, R. E., Sismeiro, C. (2001), "A Model of Web Site Browsing Behavior Estimated on Clickstream Data".

————— (2003), "A Model of Web Site Browsing Behavior Estimated on Clickstream Data", *Journal of Marketing Research*, 40, 306-323.

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

Debreu, G. (1960), "Review of R.D. Luce individual choice behaviour", *American Economic Review*, 50, 186-188.

*The Economist* 1997a, *A Survey of Electronic Commerce*, May 10, 1-18.

Finney, D. J. (1971), *Probit Analysis*. Cambridge: Cambridge University Press. (Originally published 1947).

Gandal, N. (2001), "The dynamics of Competition in the Internet Search Engine Market," *International Journal of Industrial Organization*, 19, 1103-1117.

Goldberger, A. S. (1964), *Econometric Theory*. New York: John Wiley.

Goldfarb A. (2002a), "Understanding Internet Clickstream Data." *Advances in Applied Microeconomics v. 11: The Economics of the Internet and E-commerce*. Ed. Michael Baye, Elsevier Science Ltd. 209-230.

————— (2002b), "State Dependence at Internet Portals." Working Paper.

————— (2006), "State Dependence at Internet Portals." *Journal of Economic and Management Strategy* 15(2), 317-352.

Hoffman, Saul D. and Greg J. Duncan, (1988), "A Comparison of Choice-Based Multinomial and Nested Logit Models: The Family Structure and Welfare Use Decisions of Divorced or Separate Women." *Journal of Human Resources*, 23(4), 550-562.

Johnson E., Moe W., Fader P. S., Bellman S., Lohse G. (2004), "On the Depth and Dynamics of Online Search Behavior", *Management Science*, 50(3), 299-308.

Luce, R. D. and P. Suppes (1965), "Preference, utility, and subjective probability," in R. D. Luce, R. Bush, and E. Galanter (eds.) *Handbook of Mathematical Psychology*, Vol. 3. New York: John Wiley.

Maddala, G. S., (1983), *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

McFadden, D. (1981), "Econometric Models of Probabilistic Choice," in *Structural Analysis of Discrete Data*, ed. By C. Manski and D. McFadden. Cambridge: MIT Press.

————— (1973), "Conditional Logit Analysis of Quantitative Choice Behavior," in *Frontiers in Econometrics*, ed. By P. Zarembka. New York: Academic Press.

Manski , C. (1977), "The Structure of Random Utility Models." *Theory and Decision* 8, 229-254.

Marschak, J. (1960), "Binary-Choice Constraints and Random Utility Indicators." In K. Arrow, S. Karlin, and P. Suppes (eds), *Mathematical Methods in the Social Sciences*. Stanford University Press, 312-329.

Moe W., & Fader P.S. (2004a), "Dynamic Conversion Behavior at E-commerce sites", *Management Science*, 50(3), 326-335.

—————(2004b), "Capturing Evolving Visit Behavior in Clickstream Data", *Journal of Interactive Marketing*, 18(1), 5-19.

Rindova, V. P., and Kotha, S. (1999), "Building Reputation on the Internet Lessons from Amazon.com and its Competitors."

Shapiro, C., and H. R. Varian (1999), *Information Rules: A Strategic Guide to the Network Economy*, Boston: Harvard Business School Press.

Telang, R., Mukhopadhyay, T., Wilcox, R. T., (2001), “An Empirical Analysis of Internet Search Engine Choice”, Dean School of Business Working Paper No. 03-05.

Thurstone, L. L. (1927), “Psychological Analysis.” *American Journal of Psychology* 38, 368-389.

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

Yin, R. K. (1994), *Case Study Research: Design and Methods*. Applied Social Research Methods Series, 5, second edition, Thousand Oaks: Sage Publications.

5 Tables

**Table 1: Comparison of Individual Samples by Income Group**

<b>Income Range</b>	<b>% in Amazon.com</b>	<b>% in B&amp;N.com</b>	<b>% in eBay.com</b>	<b>% in Outside Good</b>
< 15	4.08%	4.68%	5.31%	4.40%
15-24,999	8.25%	7.93%	10.35%	10.19%
25-34,999	12.61%	12.78%	15.81%	15.14%
35-49,999	18.61%	18.95%	20.49%	23.12%
50-74,999	27.48%	28.88%	23.71%	26.63%
75-99,999	13.41%	13.07%	12.17%	11.70%
100+	15.54%	13.70%	12.2%	8.81%
<b>Total</b>	100	100	100	100

*Note:* Income in thousands dollars.

**Table 2: Comparison of Individual Samples by Age Group**

<b>Age Range</b>	<b>% in Amazon.com</b>	<b>% in B&amp;N.com</b>	<b>% in eBay.com</b>	<b>% in Outside Good</b>
18-20	2.73%	2.85%	2.62%	1.72%
21-24	4.57%	3.94%	4.86%	3.44%
25-29	6.50%	5.94%	7.84%	5.37%
30-34	8.01%	7.76%	9.24%	9.64%
35-39	8.01%	7.76%	7.61%	9.08%
40-44	9.15%	9.25%	9.33%	7.91%
45-49	17.56%	15.12%	19.42%	16.86%
50-54	17.96%	19.00%	16.97%	17.62%
55-59	9.55%	9.13%	7.46%	9.91%
60-64	7.87%	10.96%	9.33%	7.23%
65-74	8.08%	8.28%	5.31%	11.22%
<b>Total</b>	100	100	100	100

**Table 3: Comparison of Individual Samples by Type of Connection Speed**

<b>Connection Speed</b>	<b>% in Amazon.com</b>	<b>% in B&amp;N.com</b>	<b>% in eBay.com</b>	<b>% in Outside Good</b>
Narrowband	55.15%	53.42%	52.55%	55.40%
Broadband	44.85%	46.58%	47.45%	44.60%
<b>Total</b>	100	100	100	100

**Table 4: Comparison of Individual Samples by Demographics**

<b>Variable</b>	<b>Amazon.com Mean</b>	<b>B&amp;N.com Mean</b>	<b>eBay.com Mean</b>	<b>Outside Good Mean</b>
Internet	Narrowband	Narrowband	Narrowband	Narrowband
Region	South	North-Central	South	South
Size	3	3	3	3
Age	50-54	50-54	45-49	50-54
Income	50k-74.999k	50k-74.999k	50k-74.999k	50k-74.999k
Child	No	No	No	No
Race	White	White	White	White
Country of Origin	Non-Hispanic	Non-Hispanic	Non-Hispanic	Non-Hispanic

**Table 5: Comparison of Individual Samples by Income Group**

<b>Income Range</b>	<b>% in Amazon</b>	<b>% in Cdnw</b>	<b>% inColumbiahouse</b>	<b>% in eBay</b>	<b>% in Outside Good</b>
< 15	4.40%	10.26%	6.88%	4.15%	5.51%
15-24,999	7.95%	11.73%	12.39%	12.30%	9.91%
25-34,999	12.43%	14.08%	16.28%	16.74%	17.40%
35-49,999	20.51%	20.53%	23.63%	23.12%	16.52%
50-74,999	24.52%	24.93%	24.77%	26.61%	25.11%
75-99,999	13.41%	10.85%	9.63%	10.01%	11.89%
100+	16.76%	7.62%	6.42%	6.87%	13.66%
<b>Total</b>	100	100	100	100	100

Note: Income in thousands dollars.

**Table 6: Comparison of Individual Samples by Age Group**

<b>Age Range</b>	<b>% in Amazon</b>	<b>% in Cdnw</b>	<b>% inColumbiahouse</b>	<b>% in eBay</b>	<b>% in Outside Good</b>
18-20	1.97%	3.23%	1.38%	1.82%	1.10%
21-24	5.26%	6.74%	3.90%	3.72%	4.40%
25-29	6.84%	4.11%	5.73%	7.44%	7.71%
30-34	10.12%	8.21%	16.97%	10.16%	12.33%
35-39	8.88%	7.33%	10.09%	8.08%	8.37%
40-44	8.55%	9.09%	12.15%	11.84%	9.47%
45-49	18.80%	21.70%	21.79%	20.53%	22.70%
50-54	17.03%	19.94%	16.05%	16.09%	17.18%
55-59	9.93%	5.57%	2.98%	6.58%	6.17%
60-64	6.11%	5.28%	4.59%	8.30%	4.63%
65-74	6.51%	8.80%	4.36%	5.44%	5.95%
<b>Total</b>	100	100	100	100	100

**Table 7: Comparison of Individual Samples by Type of Connection Speed**

<b>Connection Speed</b>	<b>% in Amazon</b>	<b>% in Cdnw</b>	<b>% inColumbiahouse</b>	<b>% in eBay</b>	<b>% in Outside Good</b>
Narrowband	56.15%	55.72%	59.86%	55.11%	47.80%
Broadband	43.85%	44.28%	40.14%	44.89%	52.20%
<b>Total</b>	100	100	100	100	100

**Table 8: Comparison of Individual Samples by Demographics**

<b>Variable</b>	<b>Amazon Mean</b>	<b>Cdnw Mean</b>	<b>Columbiahouse Mean</b>	<b>eBay Mean</b>	<b>Outside Good Mean</b>
Internet	Narrowband	Narrowband	Narrowband	Narrowband	Broadband
Region	South	North-Central	South	South	South
Size	3	3	3	3	3
Age	45-49	45-49	45-49	45-49	45-49
Income	50k-74.999k	50k-74.999k	50k-74.999k	50k-74.999k	50k-74.999k
Child	No	No	Yes	No	No
Race	White	White	White	White	White
Country of Origin	Non-Hispanic	Non-Hispanic	Non-Hispanic	Non-Hispanic	Non-Hispanic



**Table 9: Comparison of Individual Samples by Income Group**

<b>Income Range</b>	<b>% in Amazon.com</b>	<b>% in eBay.com</b>
< 15	4.15%	4.61%
15-24,999	8.24%	10.52%
25-34,999	12.67%	20.04%
35-49,999	18.53%	21.14%
50-74,999	27.70%	25.85%
75-99,999	13.32%	8.22%
100+	15.40%	9.62%
<b>Total</b>	100	100

Note: Income in thousands dollars.

**Table 10: Comparison of Individual Samples by Age Group**

<b>Age Range</b>	<b>% in Amazon.com</b>	<b>% in eBay.com</b>
18-20	2.71%	1.90%
21-24	4.64%	6.51%
25-29	6.51%	5.01%
30-34	7.97%	9.42%
35-39	7.95%	8.82%
40-44	9.28%	7.82%
45-49	17.47%	18.34%
50-54	17.88%	16.43%
55-59	9.72%	9.02%
60-64	8.00%	10.42%
65-74	7.87%	6.31%
<b>Total</b>	100	100

**Table 11: Comparison of Individual Samples by Type of Connection Speed**

<b>Connection Speed</b>	<b>% in Amazon.com</b>	<b>% in eBay.com</b>
Narrowband	55.43%	51.4%
Broadband	44.57%	48.6%
<b>Total</b>	100	100

**Table 12: Comparison of Individual Samples by Demographics**

<b>Variable</b>	<b>Amazon.com Mean</b>	<b>eBay.com Mean</b>
Internet	Narrowband	Narrowband
Region	South	North-Central
Size	3	3
Age	50-54	45-49
Income	50k-74.999k	50k-74.999k
Child	No	No
Race	White	White
Country of Origin	Non-Hispanic	Non-Hispanic

**Table 13: Individual Characteristics by Type of Website**

<b>Variable</b>	<b>Amazon.com Mean</b>	<b>eBay.com Mean</b>	<b>All Sample</b>
Duration	24-43 min	43-456 min	24-43 min
Page viewed	28-48 pages	48-456 pages	48-544 pages

**Table 14: Dummy Variable for Time (All sample)**

<b>Time</b>	<b>Frequency</b>	<b>Percent</b>	<b>Cumulative Percent</b>
Day-time	2,563	42.11	42.11
Night-time	3,523	57.89	100.00

**Table 15: Dummy Variable for Seasonal Period (All sample)**

<b>Seasonal Period</b>	<b>Frequency</b>	<b>Percent</b>	<b>Cumulative Percent</b>
Summer	2,964	48.70	48.70
Winter	3,122	51.30	100.00

**Table 16: Comparison of Individual Samples by Income Group**

<b>Income Range</b>	<b>% in Amazon.com</b>	<b>% in eBay.com</b>
< 15	4.52%	4.28%
15-24,999	8.00%	15.03%
25-34,999	12.32%	12.97%
35-49,999	20.39%	26.34%
50-74,999	24.39%	26.07%
75-99,999	13.51%	8.00%
100+	16.79%	7.31%
<b>Total</b>	100	100

*Note:* Income in thousands dollars.

**Table 17: Comparison of Individual Samples by Age Group**

<b>Age Range</b>	<b>% in Amazon.com</b>	<b>% in eBay.com</b>
18-20	1.97%	0.97%
21-24	5.25%	3.72%
25-29	6.89%	4.97%
30-34	10.3%	11.59%
35-39	8.85%	9.24%
40-44	8.59%	11.03%
45-49	18.62%	18.76%
50-54	16.92%	17.66%
55-59	10.16%	7.03%
60-64	5.97%	7.17%
65-74	6.49%	7.86%
<b>Total</b>	100	100

**Table 18: Comparison of Individual Samples by Type of Connection Speed**

<b>Connection Speed</b>	<b>% in Amazon.com</b>	<b>% in eBay.com</b>
Narrowband	56.39%	56.41%
Broadband	43.61%	43.59%
<b>Total</b>	100	100

**Table 19: Comparison of Individual Samples by Demographics**

<b>Variable</b>	<b>Amazon.com Mean</b>	<b>eBay.com Mean</b>
Internet	Narrowband	Narrowband
Region	South	North-Central
Size	3	3
Age	45-49	45-49
Income	50k-74.999k	35k-49.999k
Child	No	No
Race	White	White
Country of Origin	Non-Hispanic	Non-Hispanic

**Table 20: Individual Characteristics by Type of Website**

<b>Variable</b>	<b>Amazon.com Mean</b>	<b>eBay.com Mean</b>	<b>All Sample</b>
Duration	13-25 min	46-454 min	25-46 min
Page viewed	30-54 pages	54-575 pages	30-54 pages

**Table 21: Dummy Variable for Time (All sample)**

<b>Time</b>	<b>Frequency</b>	<b>Percent</b>	<b>Cumulative Percent</b>
Day-time	910	40.37	40.37
Night-time	1,344	59.63	100.00

**Table 22: Dummy Variable for Seasonal Period (All sample)**

<b>Seasonal Period</b>	<b>Frequency</b>	<b>Percent</b>	<b>Cumulative Percent</b>
Summer	992	44.01	44.01
Winter	1,262	55.99	100.00

**Table 23a): Parameter Estimates (Multinomial Logit Model: the case of Book)**

<b>Amazon.com</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
Intercept	1.4747 (0.1704)	8.65	0.000	1.140642	1.808834
Age	-0.0870 (0.0022)	-3.79	0.000	-.0131998	-.0042088
Income	0.8129 (0.1029)	7.90	0.000	.6111297	1.014671
Expenses	-0.0014 (0.0005)	-2.58	0.010	-.0024263	-.0003317
Likelihood Ratio: -13,343					
LR chi2(30): 2812.29					
Pseudo R2: 0.0947					
Observations: 11,791					

*Notes:* standard errors in parentheses.

**Table 23b): Parameter Estimates (Multinomial Logit Model: the case of Book)**

<b>Barnesandnoble.com</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
South	-0.1827 (0.07660)	-2.38	0.017	-.332851	-.0325565
Income	0.6324 (0.1218)	5.19	0.000	.3935619	.8713787
Expenses	-0.0028 (0.0007)	-3.91	0.000	-.0043007	-.00143
Likelihood Ratio: -13,343					
LR chi2(30): 2812.29					
Pseudo R2: 0.0947					
Observations: 11,791					

*Notes:* standard errors in parentheses.

**Table 23c): Parameter Estimates (Multinomial Logit Model: the case of Book)**

eBay.com	Coef.	z	P> z	[95% Conf. Interval]	
Intercept	3.2625 (0.1915)	17.03	0.000	2.887074	3.638024
Winter	-0.3227 (0.0665)	-4.85	0.000	-.4532876	-.1922421
Age	-0.1661 (0.0025)	-6.52	0.000	-.0216164	-.0116189
Income	0.4845 (0.1148)	4.22	0.000	.2593681	.7096325
Expenses	-0.0662 (0.0018)	-35.97	0.000	-.0699119	-.0626875
Likelihood Ratio: -13,343					
LR chi2(30): 2812.29					
Pseudo R2: 0.0947					
Observations: 11,791					

Notes: standard errors in parentheses.



**Table 24a): Parameter Estimates (Multinomial Logit Model: the case of CD)**

<b>Amazon.com</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
Intercept	0.7325 (0.2955)	2.48	0.013	.1533856	1.311766
Income	0.4955 (0.1843)	2.69	0.007	.134159	.8568875
Internet	-0.3541 (0.1084)	-3.27	0.001	-.5666234	-.1416965
Country	-0.5686 (0.1926)	-2.95	0.003	-.946209	-.1910114
Likelihood Ratio: -6419.1991					
LR chi2(40): 1323.16					
Pseudo R2: 0.0934					
Observations: 5,548					

*Notes* : standard errors in parentheses.

**Table 24b): Parameter Estimates (Multinomial Logit Model: the case of CD)**

<b>Cdnnow.com</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
Winter	-0.4350 (0.1457)	-2.98	0.003	-.7206816	-.1493769
South	0.3290 (0.1550)	2.12	0.034	.0251857	.6329911
Income	-0.7837 (0.2588)	-3.03	0.002	-1.29113	-.2763363
Internet	-0.2970 (0.1458)	-2.04	0.042	-.5829667	-.0110531
Country	-0.8138 (0.2808)	-2.90	0.004	-1.364301	-.2633315
Expenses	-0.0090 (0.0021)	-4.24	0.000	-.0132265	-.004865
Likelihood Ratio: -6419.1991					
LR chi2(40): 1323.16					
Pseudo R2: 0.0934					
Observations: 5,548					

*Notes* : standard errors in parentheses.

**Table 24c): Parameter Estimates (Multinomial Logit Model: the case of CD)**

<b>Columbiahouse.com</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
Intercept	1.2940 (0.3585)	3.61	0.000	.5913751	1.996785
Winter	-0.4071 (0.1365)	-2.98	0.003	-.6747323	-.1395021
Size	0.1788 (0.0656)	2.72	0.006	.0501432	.3075707
Income	-0.8651 (0.2420)	-3.57	0.000	-1.339548	-.3907288
White	-0.5796 (0.1922)	-3.02	0.003	-.9563953	-.2029086
Internet	-0.4912 (0.1374)	-3.57	0.000	-.7606567	-.221914
Country	-0.8820 (0.2540)	-3.47	0.001	-1.379917	-.3842047
Expenses	-0.0040 (0.0016)	-2.41	0.016	-.0072907	-.0007535
Likelihood Ratio: -6419.1991					
LR chi2(40): 1323.16					
Pseudo R2: 0.0934					
Observations: 5,548					

Notes : standard errors in parentheses.

**Table 24d): Parameter Estimates (Multinomial Logit Model: the case of CD)**

<b>eBay.com</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
Intercept	2.7205 (0.2897)	9.39	0.000	2.15269	3.288501
Winter	-0.3608 (0.1065)	-3.39	0.001	-.5696938	-.1520842
Size	0.1686 (0.0524)	3.22	0.001	.0659105	.2713166
Income	-0.5468 (0.1831)	-2.99	0.003	-.9057621	-.1878382
Child	-0.2975 (0.1381)	-2.15	0.031	-.5683784	-.0267441
Internet	-0.2619 (0.1058)	-2.47	0.013	-.4694694	-.0543636
Country	-0.4443 (0.1857)	-2.39	0.017	-.8083804	-.0803455
Expenses	-0.0445 (0.0020)	-21.77	0.000	-.0485353	-.0405174
Likelihood Ratio: -6419.1991					
LR chi2(40): 1323.16					
Pseudo R2: 0.0934					
Observations: 5,548					

Notes : standard errors in parentheses.

**Table 25a): Parameter Estimates (Probit Model: the case of Book)**

<b>Amazon.com relatively to B&amp;N.com</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
Intercept	0.6997 (0.0910)	7.69	0.000	.521312	.8781501
Winter	-0.0816 (0.0329)	-2.48	0.013	-.1461983	-.0171662
Age	-0.0033 (0.0012)	-2.64	0.008	-.0057575	-.0008534
Income	0.1106 (0.0548)	2.02	0.044	.0032094	.218082
White	0.1361 (0.0525)	2.59	0.010	.0331347	.2390794
Country	0.1378 (0.0692)	1.99	0.047	.0020617	.2736391
Expenses	0.0008 (0.0003)	2.35	0.019	.0001396	.0015358
Likelihood Ratio: -3,915.3825					
LR chi2(12): 37.67					
Pseudo R2: 0.0048					
Observations: 6,985					

Notes: standard errors in parentheses.

**Table 25b): Parameter Estimates (Probit Model: the case of Book)**

<b>Amazon.com relatively to eBay.com</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
Intercept	0.2866 (0.1219)	2.35	0.019	.0476989	.5256933
Winter	0.2439 (0.0427)	5.71	0.000	.1602153	.3276101
Income	0.4303 (0.0730)	5.89	0.000	.2871887	.5735732
Pages	-0.0074 (0.0009)	-7.83	0.000	-.0092631	-.0055523
Expenses	0.0249 (0.0011)	22.44	0.000	.0227616	.0271176
Likelihood Ratio: -2,228.0421					
LR chi2(12): 975.25					
Pseudo R2: 0.1796					
Observations: 6,086					

Notes: standard errors in parentheses.

**Table 26a): Parameter Estimates (Probit Model: the case of CD)**

<b>Amazon.com relatively to Cdnow.com</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
Winter	0.3160 (0.0699)	4.52	0.000	.1790022	.4531301
Income	0.6445 (0.1180)	5.46	0.000	.4131936	.8758097
White	0.2505 (0.1039)	2.41	0.016	.0468457	.454199
Expenses	0.0059 (0.0012)	4.74	0.000	.0035121	.0084717
Likelihood Ratio: -837.8120					
LR chi2(12): 97.43					
Pseudo R2: 0.0550					
Observations: 1,862					

Notes: standard errors in parentheses.

**Table 26b): Parameter Estimates (Probit Model: the case of CD)**

<b>Amazon.com relatively to eBay.com</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
Winter	0.3650 (0.0587)	6.22	0.000	.2500119	.4801766
Size	-0.0664 (0.0284)	-2.34	0.019	-.122121	-.010708
Age	-0.0057 (0.0023)	-2.47	0.013	-.0102896	-.0011902
Income	0.7692 (0.1030)	7.46	0.000	.5671846	.971313
Pages	-0.0105 (0.0012)	-8.27	0.000	-.0129912	-.0080131
Duration	0.0053 (0.0014)	3.70	0.000	.0025351	.0082611
Expenses	0.0077 (0.0006)	12.41	0.000	.0065056	.0089462
Likelihood Ratio: -1,217.4053					
LR chi2(12): 398.19					
Pseudo R2: 0.1406					
Observations: 2,254					

Notes: standard errors in parentheses.

**Table 27: Sum of Multiples**

Sum Multiple	Frequency	Percent	Cumulative Percent
0	3,526	50.48	50.48
2	1,524	21.68	72.15
3	837	11.98	84.14
4	456	6.53	90.67
5	245	3.51	94.17
6	126	1.80	95.98
7	77	1.10	97.08
8	24	0.34	97.42
9	36	0.52	97.94
10	60	0.86	98.80
11	22	0.31	99.11
16	16	0.23	99.34
20	20	0.29	99.63
26	26	0.37	100.00
<b>Total</b>	<b>6,985</b>	<b>100.00</b>	

*Note:* From the Probit Model data sample (Decision to choose Amazon.com over Barnesandnoble.com).

**Table 28: Fraction of Switches**

Fraction Switch	Frequency	Percent	Cumulative Percent
0	6,324	90.54	90.54
0.1428571	14	0.20	90.74
0.1666667	12	0.17	90.91
0.2	25	0.36	91.27
0.25	56	0.80	92.07
0.3	20	0.29	92.36
0.3333333	162	2.32	94.67
0.4	45	0.64	95.32
0.4285714	7	0.10	95.42
0.5	250	3.58	99.00
0.6	7	0.07	99.07
0.6666667	39	0.56	99.63
0.7142857	7	0.10	99.73
0.75	8	0.11	99.84
0.8	5	0.07	99.91
0.8333333	6	0.09	100.00
<b>Total</b>	<b>6,985</b>	<b>100.00</b>	

*Note:* From the Probit Model data sample (Decision to choose Amazon.com over Barnesandnoble.com).

**Table 29: Sum of Multiples**

Sum Multiple	Frequency	Percent	Cumulative Percent
0	2,929	48.13	48.13
2	1,272	20.90	69.03
3	672	11.04	80.07
4	416	6.84	86.90
5	270	4.44	91.34
6	132	2.17	93.51
7	105	1.73	95.23
8	40	0.66	95.89
9	9	0.15	96.04
10	40	0.66	96.70
11	11	0.18	96.88
12	36	0.59	97.47
14	14	0.23	97.70
15	15	0.25	97.95
17	17	0.28	98.23
18	36	0.59	98.82
21	21	0.35	99.16
25	25	0.41	99.57
26	26	0.43	100.00
<b>Total</b>	<b>6,086</b>	<b>100.00</b>	

*Note:* From the Probit Model data sample (Decision to choose Amazon.com over eBay.com).

**Table 30: Fraction of Switches**

Fraction Switch	Frequency	Percent	Cumulative Percent
0	5,719	93.97	93.97
0.0952381	21	0.35	94.31
0.1428571	7	0.12	94.43
0.2	20	0.33	94.76
0.25	28	0.46	95.22
0.2857143	21	0.35	95.56
0.3333333	51	0.84	96.40
0.4	20	0.33	96.73
0.4166667	12	0.20	96.93
0.5	130	2.14	99.06
0.5714286	7	0.12	99.18
0.5833333	12	0.20	99.38
0.6	5	0.08	99.46
0.6666667	33	0.54	100.00
<b>Total</b>	<b>6,086</b>	<b>100.00</b>	

*Note:* From the Probit Model data sample (Decision to choose Amazon.com over eBay.com).

**Table 31a): Loyalty Probit Regression**

**(Decision to choose Amazon.com over Barnesandnoble.com)**

<b>Non-Loyalty</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>
Intercept	-1.7837 (0.1228)	-14.52	0.000	-2.024504 -1.543048
South	-0.1155 (0.04-4)	-2.49	0.013	-.2065336 -.0246337
Likelihood Ratio: -2168.9841				
LR chi2(12): 36.38				
Pseudo R2: 0.0083				
Observations: 6,985				

*Notes:* standard errors in parentheses.

**Table 31b): Loyalty Probit Regression (Decision to choose Amazon.com over eBay.com)**

<b>Non-Loyalty</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>
Intercept	-1.1547 (0.1451)	-7.96	0.000	-1.439269 -.8702976
Duration	-0.0035 (0.0014)	-2.40	0.016	-.0064349 -.0006546
Child	-0.2001 (0.0728)	-2.75	0.006	-.3428902 -.057321
Pages	0.0034 (0.0010)	3.36	0.001	.0014449 .0054914
Income	-0.2659 (0.0905)	-2.94	0.003	-.4433914 -.0884882
Internet	0.1038 (0.0526)	1.97	0.049	.0006484 .207067
Country	-0.3093 (0.1207)	-2.43	0.015	-.5583576 -.060253
Expenses	-0.0058 (0.0009)	-5.91	0.000	-.0078549 -.0039441
Likelihood Ratio: -1337.34				
LR chi2(12): 98.08				
Pseudo R2: 0.0354				
Observations: 6,086				

*Notes:* standard errors in parentheses.

**Table 32: Sum of Multiples**

<b>Sum Multiple</b>	<b>Frequency</b>	<b>Percent</b>	<b>Cumulative Percent</b>
0	1,079	57.95	57.95
2	390	20.95	78.89
3	192	10.31	89.21
4	52	2.79	92.00
5	55	2.95	94.95
6	30	1.61	96.56
7	35	1.88	98.44
9	9	0.48	98.93
20	20	1.07	100.00
<b>Total</b>	<b>1,862</b>	<b>100.00</b>	

*Note:* From the Probit Model data sample (Decision to choose Amazon.com over Cdnw.com).

**Table 33: Fraction of Switches**

<b>Fraction Switch</b>	<b>Frequency</b>	<b>Percent</b>	<b>Cumulative Percent</b>
0	1,781	95.65	95.65
0.1428571	7	0.38	96.03
0.2	5	0.27	96.29
0.2857143	7	0.38	96.67
0.3333333	24	1.29	97.96
0.4285714	7	0.38	98.34
0.5	26	1.40	99.73
0.6	5	0.27	100.00
<b>Total</b>	<b>1,862</b>	<b>100.00</b>	

*Note:* From the Probit Model data sample (Decision to choose Amazon.com over Cdnw.com).



**Table 34: Sum of Multiples**

Sum Multiple	Frequency	Percent	Cumulative Percent
0	1,095	48.58	48.58
2	420	18.63	67.21
3	252	11.18	78.39
4	96	4.26	82.65
5	75	3.33	85.98
6	78	3.46	89.44
7	42	1.86	91.30
9	18	0.80	92.10
10	10	0.44	92.55
11	22	0.98	93.52
13	26	1.15	94.68
14	14	0.62	95.30
16	16	0.71	96.01
19	19	0.84	96.85
20	20	0.89	97.74
22	22	0.98	98.71
29	29	1.29	100.00
<b>Total</b>	<b>2,254</b>	<b>100.00</b>	

*Note:* From the Probit Model data sample (Decision to choose Amazon.com over eBay.com).

**Table 35: Fraction of Switches**

Fraction Switch	Frequency	Percent	Cumulative Percent
0	2,151	95.43	95.43
0.1538462	13	0.58	96.01
0.25	8	0.35	96.36
0.2857143	7	0.31	96.67
0.3333333	18	0.80	97.47
0.4	5	0.22	97.69
0.5	42	1.86	99.56
0.6666667	6	0.27	99.82
0.75	4	0.18	100.00
<b>Total</b>	<b>2,254</b>	<b>100.00</b>	

*Note:* From the Probit Model data sample (Decision to choose Amazon.com over eBay.com).

**Table 36a): Loyalty Probit Regression****(Decision to choose Amazon.com over Cdnw.com)**

<b>Non-Loyalty</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
Intercept	-2.0861 (0.3143)	-6.64	0.000	-2.702277	-1.470049
Winter	-0.3214 (0.1075)	-2.99	0.003	-.5322201	-.1105888
Size	0.1282 (0.0542)	2.36	0.018	.0219046	.2346381
Child	-0.4353 (0.1462)	-2.98	0.003	-.722067	-.1487314
Age	0.0152 (0.0041)	3.71	0.000	.0072223	.023375
Expenses	-0.0048 (0.0021)	-2.27	0.023	-.0091026	-.0006736
Likelihood Ratio: -311.5469 LR chi2(12): 43.19 Pseudo R2: 0.0648 Observations: 1,862					

Notes: standard errors in parentheses.

**Table 36b): Loyalty Probit Regression (Decision to choose Amazon.com over eBay.com)**

<b>Non-Loyalty</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
Intercept	-3.0486 (0.3099)	-9.84	0.000	-3.656107	-2.441118
Size	0.1948 (0.0454)	4.29	0.000	.105844	.2839055
Child	-4.4663 (0.1258)	-3.71	0.000	-.7130904	-.2196835
Age	0.0182 (0.0039)	4.64	0.000	.0105238	.0258944
Internet	0.4306 (0.1006)	4.28	0.000	.2333678	.6278856
Expenses	-0.0042 (0.0014)	-2.86	0.004	-.0071502	-.0013354
Likelihood Ratio: -378.5718 LR chi2(12): 79.74 Pseudo R2: 0.0953 Observations: 2,254					

Notes: standard errors in parentheses.

**Table 37: Sum of Multiples**

Sum Multiple	Frequency	Percent	Cumulative Percent
0	4,490	38.08	38.08
2	2,386	20.24	58.32
3	1,443	12.24	70.55
4	1,016	8.62	79.17
5	560	4.75	83.92
6	420	3.56	87.48
7	245	2.08	89.56
8	200	1.70	91.26
9	153	1.30	92.55
10	150	1.27	93.83
11	66	0.56	94.39
12	36	0.31	94.69
13	65	0.55	95.24
14	84	0.71	95.95
15	30	0.25	96.21
16	16	0.14	96.34
17	34	0.29	96.63
18	18	0.15	96.79
19	38	0.32	97.11
20	40	0.34	97.45
21	63	0.53	97.98
22	22	0.19	98.17
23	69	0.59	98.75
26	52	0.44	99.19
27	27	0.23	99.42
33	33	0.28	99.70
35	35	0.30	100.00
Total	11,791	100.00	

*Note:* From the MNL Model data sample (Book market).

**Table 38: Fraction of Switches**

Fraction Switch	Frequency	Percent	Cumulative Percent
0	8,918	75.63	75.63
0.0869565	23	0.20	75.83
0.0909091	22	0.19	76.02
0.1176471	17	0.14	76.16
0.1428571	7	0.06	76.22
0.1666667	48	0.41	76.63
0.1739130	23	0.20	76.82
0.1818182	11	0.09	76.91
0.2	85	0.72	77.64
0.2142857	14	0.12	77.75
0.2222222	9	0.08	77.83
0.2380952	21	0.18	78.01
0.25	180	1.53	79.54
0.2631579	19	0.16	79.70
0.2857143	49	0.42	80.11
0.3	20	0.17	80.28
0.3333333	444	3.77	84.05
0.3636364	22	0.19	84.23
0.375	24	0.20	84.44
0.3846154	13	0.11	84.55
0.4	130	1.10	85.65
0.4285714	35	0.30	85.95
0.4444444	18	0.15	86.10
0.4615385	13	0.11	86.21
0.5	1,036	8.79	95.00
0.5151515	33	0.28	95.28
0.55	20	0.17	95.45
0.5555556	27	0.23	95.67
0.5714286	35	0.30	95.97
0.5833333	12	0.10	96.07
0.6	55	0.47	96.54
0.625	24	0.20	96.74
0.6666667	258	2.19	98.93
0.7142857	7	0.06	98.99
0.75	76	0.64	99.64
0.8	25	0.21	99.85
0.8333333	18	0.15	100.00
Total	11,791	100.00	

*Note:* From the MNL Model data sample (Book market).

**Table 39: Loyalty Probit Regression**

Non-Loyalty	Coef.	z	P> z	[95% Conf. Interval]	
Intercept	-0.7440 (0.0708)	-10.50	0.000	-.882921	-.6051904
Winter	-0.0503 (0.0253)	-1.98	0.047	-.1001004	-.0006332
Size	0.0335 (0.0126)	2.66	0.008	.0088856	.0583079
South	-0.0636 (0.0272)	-2.34	0.019	-.1170113	-.0103659
Income	-0.0920 (0.0426)	-2.16	0.031	-.1757011	-.0084387
Age	0.0036 (0.0009)	3.76	0.000	.0017487	.0055559
Child	-0.1743 (0.3469)	-5.03	0.000	-.2423307	-.1063476
Internet	0.0706 (0.0254)	2.77	0.006	.0207131	.120671
Country	-0.2058 (0.0550)	-3.74	0.000	-.3137052	-.0980114
Expenses	-0.0006 (0.0002)	-2.16	0.030	-.0012078	-.0000599
Likelihood Ratio: -6507.4775					
LR chi2(10): 79.25					
Pseudo R2: 0.0061					
Observations: 11,791					

Notes: standard errors in parentheses.

**Table 40: Sum of Multiples**

Sum Multiple	Frequency	Percent	Cumulative Percent
0	2,207	39.78	39.78
2	960	17.30	57.08
3	531	9.57	66.65
4	352	6.34	73.00
5	300	5.41	78.41
6	174	3.14	81.54
7	133	2.40	83.94
8	72	1.30	85.24
9	72	1.30	86.54
10	90	1.62	88.16
11	44	0.79	88.95
12	24	0.43	89.38
13	26	0.47	89.85
14	28	0.50	90.36
15	15	0.27	90.63
16	32	0.58	91.20
17	34	0.61	91.82
18	18	0.32	92.14
19	76	1.37	93.51
20	20	0.36	93.87
21	42	0.76	94.63
30	60	1.08	95.71
31	31	0.56	96.27
32	32	0.58	96.85
35	35	0.63	97.48
41	41	0.74	98.22
49	49	0.88	99.10
50	50	0.90	100.00
Total	5,548	100.00	

*Note:* From the MNL Model data sample (CD market).

**Table 41: Fraction of Switches**

Fraction Switch	Frequency	Percent	Cumulative Percent
0	4,758	82.52	82.52
0.0487805	41	0.74	83.26
0.0625	32	0.58	83.83
0.0666667	30	0.54	84.37
0.0909091	11	0.20	84.57
0.1111111	9	0.16	84.73
0.1176471	17	0.31	85.04
0.1333333	15	0.27	85.31
0.1666667	6	0.11	85.42
0.2	35	0.63	86.05
0.2222222	9	0.16	86.21
0.25	76	1.37	87.58
0.2857143	42	0.76	88.34
0.3	10	0.18	88.52
0.3125	16	0.29	88.81
0.3333333	132	2.38	91.19
0.3571429	14	0.25	91.44
0.4	35	0.63	92.07
0.4285714	14	0.25	92.32
0.4761905	21	0.38	92.70
0.5	306	5.52	98.22
0.6	10	0.18	98.40
0.6666667	63	1.14	99.53
0.7142857	7	0.13	99.66
0.75	8	0.14	99.80
0.8	5	0.09	99.89
0.8333333	6	0.11	100.00
Total	5,548	100.00	

*Note:* From the MNL Model data sample (CD market).

**Table 42: Loyalty Probit Regression**

<b>Non-Loyalty</b>	<b>Coef.</b>	<b>z</b>	<b>P&gt; z </b>	<b>[95% Conf. Interval]</b>	
Intercept	-1.1140 (0.1125)	-9.90	0.000	-1.334627	-.8935289
Winter	-0.1073 (0.4017)	-2.67	0.008	-.186102	-.0286142
Size	-0.0769 (0.0201)	-3.83	0.000	-.1163762	-.0375716
South	-0.1075 (0.4369)	-2.46	0.014	-.1932443	-.0219544
Income	0.1910 (0.0701)	2.72	0.007	.0534248	.3285755
Age	0.0092 (0.0016)	5.70	0.000	.0060416	.0123667
White	-0.1357 (0.0625)	-2.17	0.030	-.2583975	-.0131347
Internet	0.1374 (0.0402)	3.42	0.001	.0586345	.2162421
Country	-0.1914 (0.0851)	-2.25	0.025	-.3584031	-.0245373
Likelihood Ratio: -2524.4001					
LR chi2(10): 93.92					
Pseudo R2: 0.0183					
Observations: 5,548					

*Notes:* standard errors in parentheses.