The Application of Discrete Choice Models In Marketing Decisions

Ateq Al-ghamedii[1] and Melfi Alrasheedi[2]

[1] Department of Statistics, Faculty of Science, King Abdulaziz University

Jeddah, Saudi Arabia. Email: drateq@gmail.com

[2] Department of Quantitative Methods, School of Business, King Faisal University

Hofof, Saudi Arabia. Email: kfu.me@hotmail.com

ABSTRACT

The paper is devoted to the analysis of logit models and their application in the market. A theoretical basis for logit models is determined. Equations for logit probabilities are derived and methods are applied in order to analyze real market situations. The real data set is analyzed to estimate 2 logit models, as well as a probit model. Obtained results are compared with experimentally calculated logit probabilities.

Keywords: Decisions; discrete choice model; logit and probit models; simulation; statistical modeling

1. INTRODUCTION

Scientific research, as well as ordinary life situations, often involves choices and decisions. There are many directions of application for choice models, such as transport demand (Domencich & McFadden, 1975; McFadden, 1974), market research (Malhotra, 1984; Huang & Rojas, 2010), adoption decisions (Adeogun et al., 2008), and so on. In particular, the analysis of demand for different goods is a popular direction. Statistics and probability theory are of great assistance in this type of situation.

In principle, the decision makers (to be referred to as consumers in this paper) can be people, households or firms and the alternatives might be products or courses of action. It should be emphasized that choosing of one alternative implies not choosing any other alternatives. In addition, the number of alternatives is finite (Train, 2003).

Discrete choice models are usually derived under an assumption of utility-maximizing behaviour (Train, 2003). This means that the decision maker n makes the decision j if utility of choice is the largest: $U_{nj} > U_{ni}, \forall i \neq j$. This utility is composed of two parts (Train, 2003):

$$U_{nj} = V_{nj} + \varepsilon_{nj}, \tag{1}$$

where ε_{nj} contains the factors that affect the utility, but are not included in V_{nj}. The problem with this type of discrete choice model is that ε_{nj} is not seen by the researcher. Decomposition (1) is fully general, since ε_{nj} is defined as a difference between the true utility U_{nj} and the part of utility V_{nj} captured by the researcher (Train, 2003). Each characteristic of ε_{nj}, such as its distribution, depends on researcher specifications. The aim of the researcher is to estimate the parameters of this distribution.

In this paper we are going to examine the application of a logit model for discrete choice analysis. The goal of the paper is to estimate the coefficients of the logistic regression. The chosen estimation method for this purpose is a maximum likelihood estimator (MLE) (Myung, 2002), which is the most common method for a model estimation.

The estimated logit coefficients will allow for calculation of the probability of making a particular decision. Such an obtained probability distribution would play a key role in the analysis of the demand for new goods on the market.

In addition to the standard logit model, which is used for the calculation of first choice probability for a given good among the choice set, the ordered probability can be useful. Such probability of the rankings can be obtained from the exploded or ordered logit model (Johnson et al., 2010; Kumar & Kant, 2007). The most important features of both logit and exploded logit models, as well as the methods of their estimation, are discussed in the next section.

2. ANALYSIS OF LOGIT MODELS

Logit or logistic distribution has a closed form of cumulative distribution function

$$\Lambda(X) = \frac{1}{1+\exp(-\mathrm{X})} \tag{2}$$

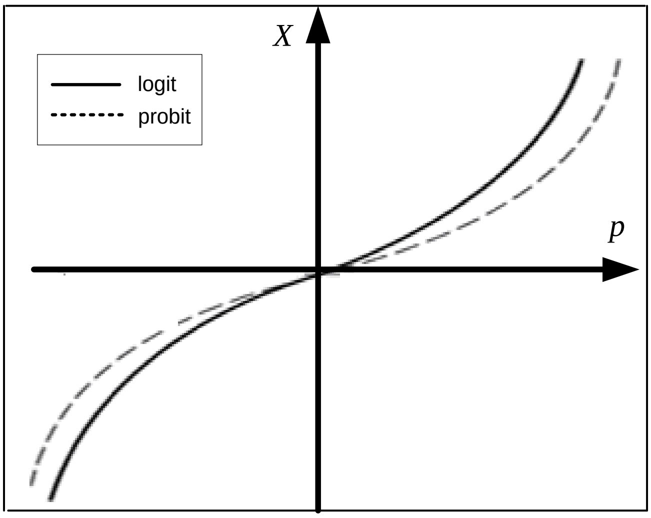

Figure 1: Logit and probit cumulative distributions

One can see from Figure1 that the logit cumulation distribution function (CDF) is similar to the common normal CDF; however, the tails of logit are heavier. The main advantage of the logit distribution is the simplicity of the CDF, while the normal CDF involves an unevaluated integral.

In order to derive the choice probabilities for the logit model, we analyze N consumers and J possible choices. The utility of each choice is known only for the customer, but not for the researcher. This utility U_{nj} can be presented as in (1). The first term in (1), V_{nj}, is usually decomposed into two parts $x'_{nj}\beta$ and k_j. Term $x'_{nj}\beta$ is vector of variables that relate to alternative j as faced by decision maker j, and k_j is a constant that is specific to alternative j. The researcher treats ε_{nj} as random. The joint density of the random vector $\varepsilon_n = \langle \varepsilon_{n1} ... \varepsilon_{nJ} \rangle$ is denoted as $f(\varepsilon_n)$. In this notation the probability that the consumer n chooses alternative j is

$$P_{nj} = P(U_{nj} > U_{ni} \forall i \neq j) = P(V_{nj} + \varepsilon_{nj} > U_{ni} + \varepsilon_{ni} \forall i \neq j) = P(\varepsilon_{ni} - \varepsilon_{nj} < V_{nj} - V_{ni}, \forall i \neq j) \tag{3}$$

This probability is a cumulative distribution; namely, the probability that each random term $\varepsilon_{ni} - \varepsilon_{nj}$ is below the observed quantity $V_{nj} - V_{ni}$. Using the density $f(\varepsilon_n)$, this cumulative probability can be rewritten as

$$P_{nj} = \int I(\varepsilon_{ni} - \varepsilon_{nj} < V_{nj} - V_{ni} \forall i \neq j) f(\varepsilon_n) d\varepsilon_n \tag{4}$$

Here $I(\cdot)$ is indicator function, which is equal 1 if equation in the brackets is true. One should emphasize here that the choice probability P_{nj} depends only on the differences $\varepsilon_{ni} - \varepsilon_{nj}$ and $V_{nj} - V_{ni}$.

The logit model is obtained by assuming that each ε_{nj} is an independent and identically distributed extreme value to each additional ε_{nj} value. Such specific distributions arise in the case of modeling of rare events. The generalized extreme value distribution (GEV) PDF is defined as

$$f(x) = \begin{cases} \frac{1}{\sigma} \exp\left(-(1+kz)^{-1/k}\right)(1+kz)^{-1-1/k} & k \neq 0 \\ \frac{1}{\sigma} \exp(-z - \exp(-z)) & k = 0 \end{cases}, z = (x-\mu)/\sigma \tag{5}$$

where k, σ, μ are shape, scale and location parameters. Various values of shape parameter k yield to extreme value type I, II, III distributions. Specifically, the three cases $k = 0, k > 0, k < 0$ correspond to Gumbel, Frechet and reversed Weibull distributions, respectively. GEV distributions are used in weather predictions, extreme floods and snowfalls, market crashes, and so on (Train, 2003). The main peculiarity of such distributions is that in the case where one generates N data sets from the same distribution, and creates a new data set that includes the maximum values from these N data sets, the resulting data set can only be described by one of the three above-mentioned distributions (Fisher & Tippett, 1928).

In the case of the logit model, the density of unobserved utility ε_{nj} is Gumbel or type I extreme value (Persson & Rydén, 2007)

$$f(\varepsilon_{nj}) = \exp(-\varepsilon_{nj}) \exp(-\exp(-\varepsilon_{nj})) \tag{6}$$

and CDF

$$F(\varepsilon_{nj}) = \exp(-\exp(-\varepsilon_{nj})) \tag{7}$$

The differences between two terms are distributed logistic differences. That is $\varepsilon_{nji} = \varepsilon_{nj} - \varepsilon_{ni}$ follows the logistic distribution

$$F(\varepsilon_{nji}) = \frac{\exp(-\varepsilon_{nji})}{1+\exp(-\varepsilon_{nji})} \tag{8}$$

At this stage, our goal is to derive the choice probabilities. We use (4) for the probability of choice:

$$P_{nj} = P\left(\varepsilon_{ni} - \varepsilon_{nj} < V_{nj} - V_{ni} \forall i \neq j\right) = P\left(\varepsilon_{ni} < \varepsilon_{nj} + V_{nj} - V_{ni} \forall i \neq j\right) \tag{9}$$

If ε_{nj} is considered given, then (9) is the cumulative distribution for each ε_{ni}, which is defined as

$$\exp(-\exp(\varepsilon_{nj} + V_{nj} - V_{ni}))$$

As far as we can use the assumption of ε independence, the CDF over all $\forall i \neq j$ is the product of individual CDFs:

$$P_{nj} \mid \varepsilon_{nj} = \prod_{i \neq j} \exp(-\exp(\varepsilon_{nj} + V_{nj} - V_{ni})) \tag{10}$$

ε_{nj} are not given; therefore, the choice probability is the integral of (9) over all values of ε_{nj}, weighted by its density

$$P_{nj} \mid \varepsilon_{nj} = \int \prod_{i \neq j} \exp(-\exp(\varepsilon_{nj} + V_{nj} - V_{ni})) \exp(-\varepsilon_{nj}) \exp((-\exp(-\varepsilon_{nj}))) d\varepsilon_{nj} \tag{11}$$

After algebraic manipulations, one can obtain the closed form for the needed choice probability

$$P_{nj} = \frac{\exp(V_{nj})}{\sum_{i \neq j} \exp(V_{ni})} \tag{12}$$

It is evident that the choice probability is between 0 and 1 and that the normalization condition is fulfilled. This model will be estimated with MLE in the next section.

Here, β_j and σ_j is the average random coefficient and its standard deviation. These two parameters are important for the simulation of logit model and their influence will be tested in the next section.

3. SIMULATION AND DISCUSSION

In this section, the analysis of a real dataset is accomplished. A market research firm collected this data as a part of research study to evaluate the demand for a new good. The consumers were asked to respond to several choice experiments. The experiment required that each consumer completely rank the goods that are preferred over outside goods. There were 10 price levels and 8 brands in the dataset. The price index “0” and brand index “0” corresponds to a hypothetical outside good.

In each choice experiment the customer ranks only the goods that are preferred over outside goods. For instance, if for the given brand j with price level p_j the utility $U_j > U_0$, then the rank for this brand is assigned. Several choice experiments are performed with 170 independent customers. Therefore, in cases where when the current brand is ranked

$$U_{nj} > U_{n0}\,, \tag{13}$$

where U_{n0} is utility of the outside good for the n^{th} customer. In order to analyze the input dataset, it would be useful to calculate some quantitative dependencies. The histogram of prices frequencies is shown in Figure 2. It illustrates the number of experiments in which the particular price value was used.

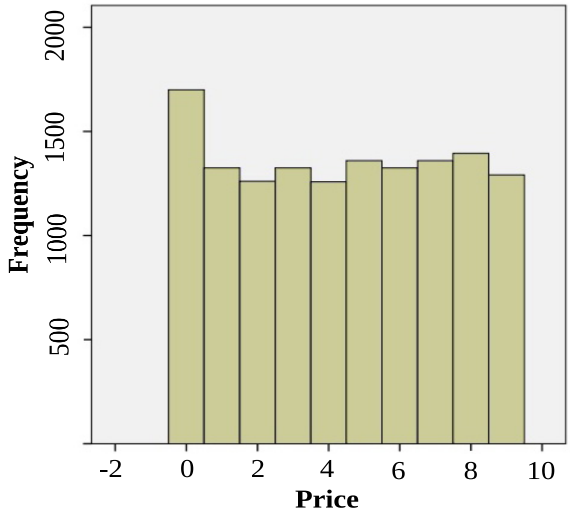

Figure 2: Price frequencies

One can see from Figure 2 that the histogram is relatively flat, except for the number of experiments with the outside good as an input. Figure 3 illustrates the frequency of different rank assignments.

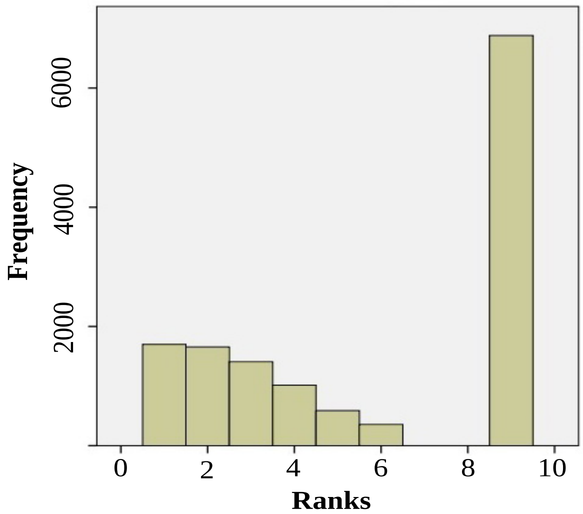

Figure 3: Frequency of assigned ranks

If we recall to the initial conditions of the discrete choice experiment, one can make a conclusion that the outside good was often preferred to the examined brands. In other words, condition (16) was often not held.

Figure 4 shows the distribution of how many times the customers did not choose any products. One can see that if we remove the first column, the distribution will be close to normal.

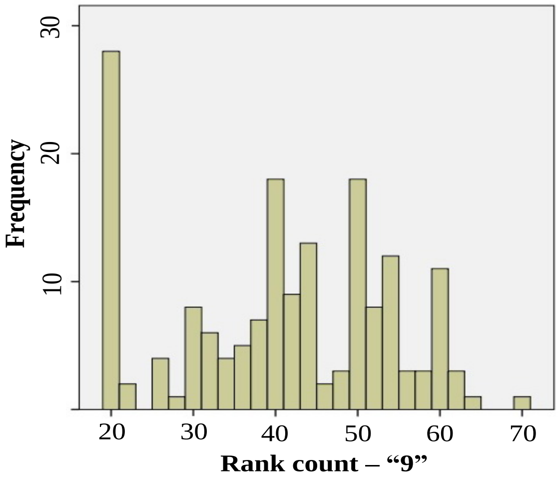

Figure 4: Frequency of assigned ranks

One can calculate the experimental probability for the given brand to be ranked first for each price. This probability Pr is defined as follows:

$$\Pr(p) = \frac{N_B^1}{N_B^\Sigma} \tag{14}$$

where N_B^1 is the number of times when the given brand with price was 1st ranked, N_B^Σ - is the total number of experiments with given brand from price p.

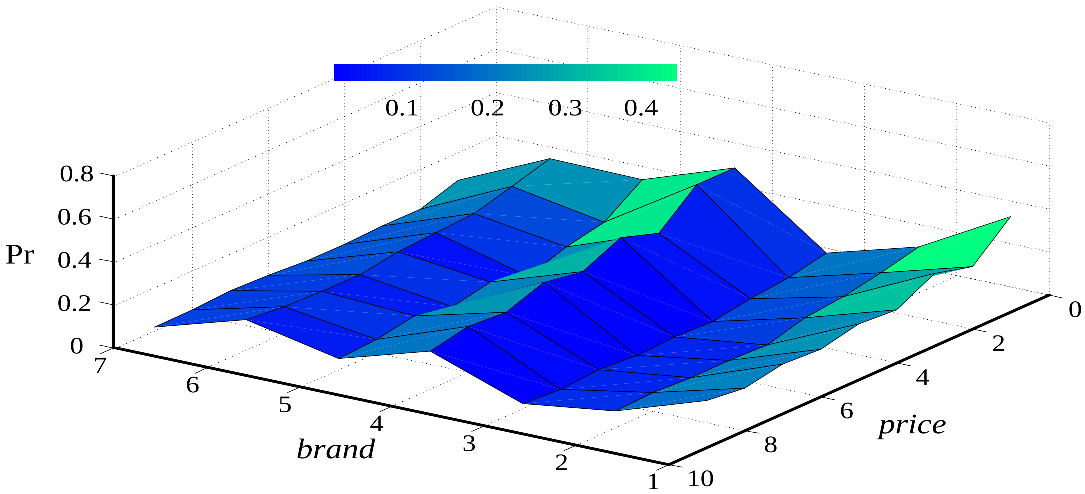

Figure 5: Experimental probabilities to be 1st ranked

Figure 5 illustrates the calculated probabilities for all brands. One can observe the highest probabilities for brands with indexes 1 and 4. 2D plots of $\Pr(p)$ for these brands are shown in Figure 6.

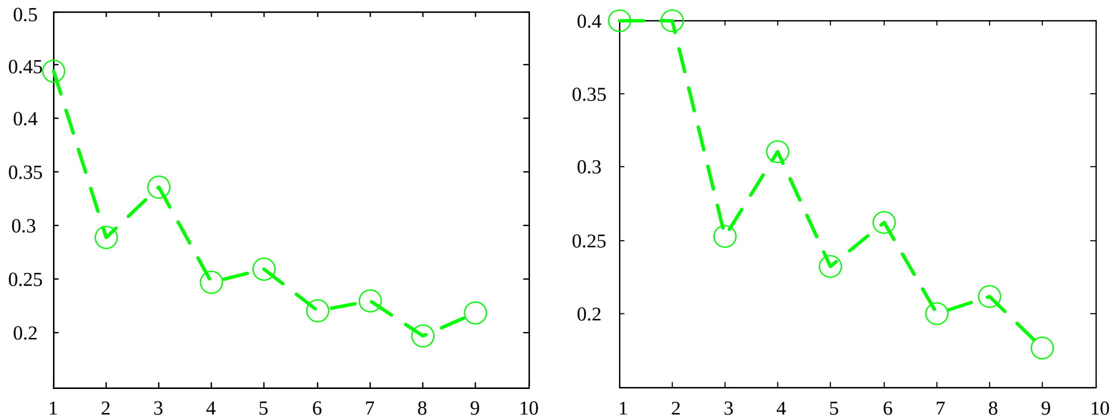

Figure 6: Experimental probabilities for brands 1 and 4 (x – price index, y - probability)

In order to examine the demand for the outside good, the logit model estimation should be performed. One of the aims of this procedure is to compare the fixed effect and random coefficients logit models. The main difference between these two models is related to individual consumer taste for brands. In fixed effect or simple logit model, individual taste is accounted for in the logistic regression procedure. In contrast to this model, the random coefficients logit model is accomplished with consideration of all choice experiments, without any consideration of individual tastes.

As a result of logit model estimation, price and brand coefficients can be obtained. Hence, the probability of given brand to be first ranked at given price can be calculated as follows:

$$P_{jp} = \frac{e^{\beta_j b - \alpha_p p}}{\sum_i e^{\beta_i b - \alpha_p p}}, \tag{15}$$

where α_p, β_j are price and taste coefficients for the current brand. Here we have used the fact that the particular choice of consumer depends on the brand type and price.

After the logistic regression, one can obtain the list of estimated coefficients. The coefficients are usually presented in two ways. The first option is to estimate the coefficients in log-odds units. This means that

$$\log\left(\frac{p}{1-p}\right) = \sum_{j=1}^{Nbrands} \beta_j b_j + \sum_{p=1}^{Nprices} \alpha_p p_p \tag{16}$$

These estimates describe the relationship between the independent variables (brand and price) and the dependent variable rank, where the dependent variable is on the logit scale. These estimates tell us the amount of increase in the log odds of rank that would be predicted by a 1-unit increase in that predictor, holding all other predictors constant. Because these coefficients are in log-odds units, they can be difficult to interpret, so they are often converted into odds ratios. This conversion is performed via a simple exponentiation. The results of the estimation procedure for the logit and probit regressions with obtained odds ratios are illustrated in Table 1.

Table 1: Estimated odds ratios for two logit models and probit model

Coefficientss	**Odds ratio (fixed - effect)**	**Odds ratio (simple logit)**	**Odds ratio (ordered probit)**
α_1	25.675	0.069	0.353
α_2	19.672	0.094	0.569
α_3	18.909	0.100	1.324
α_4	16.992	0.113	0.343
α_5	14.173	0.124	0.609
α_6	13.102	0.132	0.545
α_7	13.570	0.127	0.612
α_8	12.508	0.127	0.784
α_9	12.324	0.130	0.834
β_1	3.044	0.402	0.960
β_2	1.839	0.670	1.071
β_3	0.868	1.197	1.146
β_4	2.972	0.423	1.098
β_5	1.729	0.726	1.140
β_6	1.600	0.637	1.140

From Table 1, one can observe that odds obtained via a fixed-effect logit indicates higher probabilities than simple logit and probit models, suggesting that the fixed effect model best fits the solutions.

After fitting data to the model, an odds ratio cross-tab can be generated in order to illustrate the odds product to be ranked first for particular scenarios, based on data. The crosstab is shown on Table 2.

Table 2: Calculated odds ratios for particular scenarios (fixed-effect logit model)

price / *brand*	1	2	3	4	5	6	7
1	78.01	47.06	22.14	76.49	44.31	41.10	25.68
2	59.77	36.05	16.96	58.60	33.95	31.49	19.67
3	57.45	34.66	16.31	56.33	32.63	30.27	18.90
4	51.63	31.14	14.65	50.62	29.32	27.20	16.99
5	43.06	25.97	12.22	42.22	24.46	22.69	14.17
6	39.81	24.01	11.30	39.03	22.61	20.97	13.10
7	41.23	24.87	11.70	40.43	23.42	21.72	13.57
8	38.00	22.92	10.78	37.26	21.588	20.02	12.50
9	37.44	22.58	10.63	36.71	21.27	19.73	12.32

One can observe from Table 2 that the highest odds ratios correspond to the brands 1 and 4. This means that according to the estimated model, the probabilities of these 2 brands being ranked first are the largest. It is important to emphasize that this result is correlated with the experimental probabilities obtained in this section.

4. CONCLUSIONS

The equations for logit model probabilities were derived and analyzed. The logit regression procedure was applied in order to estimate the demand for different goods in the market. The results have shown that the fixed effect logit model is more suitable than alternative models for analysis of real data. This is because this type of model accounts for individual tastes of consumers. The obtained odds ratios for the fixed-effect logit model show the correspondence with experimentally obtained probabilities. Obtained results can be used for price management in real markets, in order to increase the demand of a given product.

APPENDIX

DERIVATION OF LOGIT PROBABILITIES: This appendix shows how the logit probabilities (12) can be derived from (11).

$$P_{ni} = \int_{s=-\infty}^{\infty} \left(\prod_{j \neq i} e^{-e^{-(s+V_{ni}-V_{nj})}} \right) e^{-s} e^{-e^{-s}} ds, \tag{1}$$

where $s = \varepsilon_{ni}$. One can rewrite the multiplication of the exponents as a summation in the one resulting exponential term

$$P_{ni} = \int_{s=-\infty}^{\infty} \left(\prod_{j \neq i} e^{-e^{-(s+V_{ni}-V_{nj})}} \right) e^{-s} ds = \int_{s=-\infty}^{\infty} \exp\left(\sum_{j} e^{-(s+V_{ni}-V_{nj})} \right) e^{-s} ds \tag{2}$$

One can make a substitution $t = \exp(-s), -\exp(-s)ds = dt$. After such substitution (1) is easily evaluated

$$P_{ni} = \int_{\infty}^{0} \exp\left(-t \sum_{j} e^{-(V_{ni}-V_{nj})} \right)(-dt) = \int_{0}^{\infty} \exp\left(-t \sum_{j} e^{-(V_{ni}-V_{nj})} \right) dt =$$
$$= \frac{\exp\left(-t \sum_{j} e^{-(V_{ni}-V_{nj})} \right)}{-\sum_{j} e^{-(V_{ni}-V_{nj})}} \Big|_0^{\infty} = \frac{\exp(V_{ni})}{\sum_{j \neq i} \exp(V_{nj})} \tag{3}$$

Bibliography

Adeogun, O. A., Ajana, A. M., Ayinla, O. A., Yarhere, M. T., & Adeogun, M. O. (2008). Application of logit model in adoption decision: A study of hybrid clarias in Lagos State, Nigeria. *American-Eurasian Journal of Agriculture and Environmental Sciences*, *4*(4), 468-472.

Domencich, T. A., & McFadden, D. (1975). *Urban Travel Demand - A Behavioral Analysis* (No. Monograph). Oxford: North-Holland Publishing Company, 1975.

Fisher, R. A., & Tippett, L. H. C. (1928, April). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society* (Vol. 24, No. 02, pp. 180-190). Cambridge University Press.

Huang, D., Rojas, C. (2010). Eliminating the Outside Good. Bias in Logit Models of Demand with Aggregate Data. Amherst: University of Massachusetts at Amherst.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, *3*(4), 303-328.

Johnson, J., Bruce, A., & Yu, J. (2010). The ordinal efficiency of betting markets: an exploded logit approach. *Applied Economics*, *42*(29), 3703-3709.

Kumar, S., & Kant, S. (2007). Exploded logit modeling of stakeholders' preferences for multiple forest values. *Forest Policy and Economics*, *9*(5), 516-526.

Malhotra, N. K. (1984). The use of linear logit models in marketing research. *Journal of Marketing Research*, 20-31.

Myung, I. J. (2002). Maximum likelihood estimation. *Manuscript submitted for publication.*

Persson, K., & Rydén, J. (2007). *Exponentiated Gumbel distribution for estimation of return levels of significant wave height*. Department of Mathematics, Uppsala University.

Train, K. (2003). *Discrete choice methods with simulation*. Cambridge: Cambridge University Press.