A SCOBIT-BASED TRAVEL MODE CHOICE MODEL

Junyi Zhang, Hiroshima University, Japan
Harry Timmermans, Eindhoven University of Technology, Netherlands

ABSTRACT

As an additional consumer choice model, this paper suggests using the Scobit model, which includes a skewness parameter, to accommodate a more general structure of marginal effects in the context of binary choice behavior. It is empirically confirmed that the Scobit model outperforms the widely used binary logit model.

INTRODUCTION

Estimating the sensitivity of consumers to changes in product or service attributes is of vital concern in marketing research and in other disciplines, which attempt to model consumer response. How will consumer choice and corresponding market shares change as a function of changed attributes? In many of these disciplines, discrete choice models have found ample applications (e.g., Bucklin et al., 2008, Aribarg and Foutz, 2009; Briesch et al, 2009, Park and Gupta, 2009; van Rosmalen et al., 2010). These models assume that consumers maximize their utility, which is typically expressed as a (linear) function of the attributes of the choice alternatives and an error term. Depending then on the assumptions about the distribution of the error terms, different choice models can be derived.

The most commonly used model is the binomial (BL) or multinomial logit (MNL) model, however, implicitly states that those consumers who are invariant across choice alternatives (50% choice probability in case of two alternatives) are most sensitive to changes in attribute values. This assumption has immediate implications for estimating elasticity and consumer responses; however, it has hardly been tested. Considering that marketers are usually interested in what types of consumers are more sensitive to changes in attributes of products than others, in case that this assumption does not hold, using the BL or MNL model to identify those consumers would be problematic. However, careful literature reviews suggest that no study in marketing has been carried out to test the above assumption about the response sensitivity in the logit model. To test this implicit assumption underlying the logit model, taking binary choice behavior as an example, this paper applied a Scobit model, which has a skewness parameter and includes the BL model as a special case. The context of application is the choice of transport mode (car versus bus).

A BINARY TRAVEL MODE CHOICE WITH SCOBIT STRUCTURE

In this paper, we are dealing with binary choices. Multiple examples of applications in marketing include, to buy or not to buy a product, propensity of shopping, conveying a message to other persons or not, etc. The stipulated problem also applies to multinomial choice, but we leave that for future research.

A Scobit Model

Assume that we have two alternatives in a choice set of travel mode. In this case study, we deal with car and bus. Then, the utilities of the two modes can be defined as follows, where *n* indicates trip maker, u_{n1}, u_{n2} are utility functions of car and bus, v_{n1}, v_{n2} are deterministic terms of u_{n1}, u_{n2}, and e_{n1}, e_{n2} are the corresponding error terms, respectively.

$$\text{Car: } u_{n1} = v_{n1} + e_{n1}, \tag{1}$$

$$\text{Bus: } u_{n2} = v_{n2} + e_{n2}, \tag{2}$$

Then, the probability p_{n1} that trip maker *n* chooses car can be described as,

$$p_{n1} = \Pr(u_{n1} > u_{n2}) = \Pr(e_{n1} - e_{n2} > v_{n2} - v_{n1}). \tag{3}$$

Let us define a new error term $\varepsilon_n = e_{n1} - e_{n2}$ and further assume that it follows a distribution with the distribution function $F(\varepsilon_n)$. Then the probabilities p_{n1} can be expressed as $p_{n1} = 1 - F(-(v_{n1} - v_{n2}))$. Since v_{n1}, v_{n2} are usually assumed to be a

linear function of explanatory variables, let us define them as $v_{n1} - v_{n2} = \sum_k \beta_k (x_{n1k} - x_{n2k})$, where x_{n1k}, x_{n2k} are the kth variables with parameter β_k, respectively. The marginal effect ME_x^p of a change in $x_{n1} - x_{n2}$ can be derived as follows, where $f(\bullet)$ is the probability density function of $F(\varepsilon_n)$.

$$ME_x^p = \partial p_{n1} / \partial (x_{n1} - x_{n2}) = f(-\sum_k \beta_k (x_{n1} - x_{n2}))\beta_k , \quad (4)$$

It is obvious that ME_x^p depends not only on β_k and $x_{n1} - x_{n2}$, but also on the form of $f(\bullet)$. If a normal or Weibul distribution is assumed, then $f(\bullet)$ will have a maximum at $\sum_k \beta_k (x_{n1} - x_{n2}) = 0$. This means that any given variable $x_{n1} - x_{n2}$ will have its greatest effect on those individuals for which $\sum_k \beta_k (x_{n1} - x_{n2})$ is closest to 0, or for which p_{n1} is closest to 0.5. However, if individuals with an initial choice probability other than 0.5 are those most sensitive to the change, then the logit or probit model would result in a misspecification and consequently biased inferences about the marginal effects of changes in any explanatory variable. In this sense, it is necessary to adopt a more general distribution which allows the highest sensitivity to changes in variables at any initial probability. To meet the above requirement, this study applies the Scobit (or skewed logit) model (Nagler, 1994), which to the best of our knowledge is not well known in marketing research. This model can be derived by assuming the following $F(\varepsilon_n)$, in which α is a parameter used to measure the skewness of Burr-10 distribution. This is, in fact, a Burr-10 distribution (Burr, 1942), which is one of the 12 distributions defined by Burr.

$$F(\varepsilon_n) = 1/(1 + \exp(-\varepsilon_n))^{\alpha} \quad (5)$$

Having defined the above distribution function $F(\varepsilon_n)$, the probability of choosing car can be derived as,

$$p_{n1} = 1 - F(-(v_{n1} - v_{n2})) = 1 - 1/(1 + \exp(v_{n1} - v_{n2}))^{\alpha} , \quad (6)$$

When α is equal to 1, equation (6) returns to the logit model. Thus, the popular logit model is nested within the Scobit model. The Scobit model is also called the skewed logit model because it allows a skewed response curve, which is different from the symmetric curve (symmetric about zero) derived from the logit model.

Representing Heterogeneous Skewness Parameter

It is expected that the skewness may be different across individuals, i.e., some individuals may show the highest sensitivity to change at $p_{n1} = 0.5$, some at $p_{n1} < 0.5$, and some at $p_{n1} > 0.5$. However, it is difficult for analysts to figure this out in advance. To accommodate such heterogeneity, this study therefore defines α as a function of some individual attributes (z_{nq}), where θ_q is the parameter of the qth variable z_{nq}, and π is a constant term.

$$\alpha_n = \exp(\pi + \sum_q \theta_q z_{nq}) , \quad (7)$$

Note that the exponential function is adopted to meet the requirement that $\alpha_n > 0$. In the empirical analysis shown later, the Scobit model with heterogeneous α will be compared with the model with homogeneous α.

AN EMPIRICAL ANALYSIS

To assess the effectiveness of the Scobit model in representing binary travel mode choice behavior, this study adopted a panel data collected in Hiroshima City, Japan in 1987, 1990, 1993 and 1994. This panel data includes only car and bus as the alternative modes for commuting (at the time of the survey, these were the only available modes). The valid sample size for each wave is 226 respondents. Looking at respondents' attributes in the 1st wave, 69% of the respondents are male, most of which are aged between 30 and 60 years old. All the female respondents are younger than 50 years old. And, 97% of the respondents are employed, and 72% belong to a household with 2 or more members. The shares of bus ranged between 40%~45% across the four waves. In this case study, we estimated the Scobit and logit models for each wave, respectively, based on the standard maximum likelihood estimation method. To evaluate whether the skewness parameter shows heterogeneity across consumers (here, trip makers), two types of models were estimated: one assumes the skewness

parameter homogeneous across trip makers, and the other assumes the heterogeneous skewness parameter, where the latter defines the skewness parameter as an exponential function of socio-demographic characteristics of trip makers.

Usually, two types of explanatory variables are introduced into a choice mode: one are the alternative-specific attributes and the other are the alternative-generic attributes. In this case study, we adopted travel time and cost for the car and the bus as the alternative-specific attributes and used socio-demographic characteristics of trip makers as the alternative-generic attributes. We assumed that the utility of the car is a function of the travel time and cost differences between these two transport modes. We further assumed in the model with heterogeneous skewness parameter that the skewness is influenced by socio-demographic characteristics of trip makers. More specifically, gender, age, employment and household size were used.

Results of Model Estimation

We first estimated the Scobit model with homogeneous skewness as well as the relevant logit model for the four waves. Due to the space limitation, the results of the estimation are not shown here. Adjusted McFadden's Rho-squared values range from 0.2 to 0.5, suggesting that both Scobit and logit models are effective to represent the car and bus choice behavior. These results indicate that the goodness-of-fit of both models is good, given standard rules of thumb in the relevant literature. The two models estimate that all the parameters of the travel time and cost are negative except for the travel time parameters in 1987. To test whether the Scobit model outperforms the logit model or not, the Chi-square test is conducted for all the waves. It is found that in 1990 and 1994, the accuracy of the Scobit model is higher than that of the logit model at 90% and 95% levels, respectively. Looking at the skewness parameter, they are all significantly different from 0 at the 95% level, but not all of them are different from 1. At the 90% significance level, the skewness parameters in 1990 and 1994 are different from 1, where the parameter in 1990 is significantly different from 1 even at the 95% level. These test results suggest that the Scobit model is superior to the logit model in 1990 and 1994. This implies that the assumptions underlying the binary logit model regarding consumer sensitivity to changes in attributes are not supported by the data in 1990 and 1994. The values of travel time in 1990 and 1994 estimated by the Scobit model are about 20% higher than those by the logit model, respectively, and the value of travel time in 1993 by the Scobit model is almost the same as that by the logit model.

Next, we estimated the Scobit model with heterogeneous skewness, which is defined as an exponential function of socio-demographic attributes including sex, age, employment, and number of household members. For the purpose of comparison, the same set of socio-demographic attributes is introduced into the logit model as explanatory variables together with travel time and cost variables. Estimation results are shown in Table 1. Looking at the model accuracy, it is demonstrated that introducing socio-demographic attributes of trip makers remarkably improved the model accuracy in the sense that the Adjusted McFadden's Rho-squared values in Table 1 are about 20%~50% higher than the corresponding values of the model with homogeneity. It is found that most of the socio-demographic attributes in the two models have statistically significant parameters at the 95% level. The average values of the skewness parameters across samples are about 0.9~1.5 with the standard deviations 0.5~1.3. These values are substantially higher than those estimated in the Scobit model with homogeneous skewness parameter. Different from the model with homogeneity, the Scobit model in Table 1 estimates a smaller value of travel time than that by the logit model (-6% ~ -15%). Except the model in 1990, the accuracy of the Scobit model is about 2~8% higher than that of the logit model. To further understand the difference of the two models, the choice probabilities for the car and the bus from these two models are calculated. In the first wave (i.e., 1987), larger differences between the two models are observed in the sides of smaller (about 0.1~0.2) and larger (about 0.7~0.9) choice probabilities for both car and bus. In the second wave (1990), the Scobit model estimates a larger choice probability of the bus at the probability space 0.0~0.7 than the logit model; in contrast it calculates a smaller choice probability of the car at the probability space 0.3~0.9. Relatively large differences between the two models are observed across the whole choice probability space in the third wave (1993). In the last wave, it is observed that the two models show larges differences below the probability of 0.5 for the bus, but over the probability of 0.5 for the car.

Elasticity Analysis

Here, we calculated the elasticity of travel time and travel cost differences with respect to the car choice probability. The elasticity formula adopted here is shown below.

$$E_{x_{ntk}}^{P_t(y_t=1)} = \alpha_n \exp(v_t)\beta_k x_{ntk} P_{nt}(y_{nt}=0)/(P_{nt}(y_{nt}=1)(1+\exp(v_t))) \quad (8)$$

In the first wave (1987), larger discrepancies in the elasticity between the two models are observed when the travel time/cost of car are shorter/smaller than those of bus. More specifically, when the level of service provided by the car is better than that by the bus, the logit model overestimates the elasticity in most cases. In contrast, the logit model underestimates the elasticity in the second wave (1990) when the car service outperforms the bus service. The two models estimate similar elasticity for both car and bus in the third wave (1993) and it is also true to the travel time difference in 1994. For the travel cost difference in 1994, the logit model overestimates the elasticity when the car service is better than the bus service. However, the above-observed differences are not supported by the T-test. This means that in this case study, the Scobit and logit models generate the indifferent elasticity of both travel time and cost.

CONCLUSIONS

Most choice models applied in marketing research and in other disciplines implicitly assume that consumer sensitivity is highest when choice probabilities are equal across alternatives in a choice set. This property is derived from the assumed symmetric response curve, but is hardly tested. In this paper, we therefore tested this assumption by comparing the performance of a binary logit model against a binary Scobit model, which adds a skewness parameter in the context of transport mode choice decisions. Note that the logit model is a special case of the Scobit model when the skewness parameter is equal to 1. Thus, the Scobit model is more general than the logit model in representing consumer choice behavior.

We conducted the first study in marketing to examine whether consumer choice behavior show the highest sensitivity when choice probabilities are equal across alternatives in a choice set. This is done by estimating a Scobit model. Based on an analysis using a 4-wave panel data of commuting travel mode choice (car and bus), it is confirmed that in case that the skewness parameter is homogeneous across trip makers, in two out of the four waves, the accuracies of the Scobit model are higher than those of the logit model and the skewness parameters are statistically different from 1. In case that the skewness parameter is heterogeneous across trip makers, in three out of the four waves, the accuracies of the Scobit model are slightly higher than those of the logit model. Statistical test results suggest that on average the Scobit model and the logit model show equal performance in the estimated choice probabilities and the elasticity of travel time and cost. Nevertheless, the resulting skewness parameters and elasticity show large discrepancies between the two models. Thus, it is still too earlier to make a conclusion that the two models are not different with each other at least at the disaggregate level. Therefore we can conclude that at least for the data used in the present analysis, the Scobit model bears comparison with the binary logit model, and especially at the disaggregate level, it is superior to the binary logit model. In this study, we examined the performance of the Scobit model in the context of commuting travel mode choice, which is routinely performed on a daily basis.

It is expected that the performance of the Scobit model might differ from context to context in representing consumer behavior. Comparative studies are required to cover different contexts. Needless to say, such comparative analyses should be extended to multinomial choice cases.

REFERENCES

Aribarg, A. & Foutz, N.Z. (2009). Category-based screening in choice of complementary products. *Journal of Marketing Research*, 46(August), 518–530

Briesch, R.A., Chintagunta, P.K., & Fox, E.J. (2009). How does assortment affect grocery store choice. *Journal of Marketing Research*, *46*(April), 176–189

Bucklin, R.E., Siddarth, S., & Silva-Risso, J.M. (2008). Distribution intensity and new car choice. *Journal of Marketing Research, 45*(August), 473–486

Burr, I.W. (1942). Cumulative frequency functions. *Annals of Mathematical Statistics*, *13*, 215-232.

Nagler, J. (1994). Scobit: An alternative estimator to logit and probit. *American Journal of Political Science*, *38*, 230-255.

Park, S. & Gupta, S. (2009). Simulated maximum likelihood estimator for the random coefficient logit model using aggregate data. *Journal of Marketing Research, 46*(August), 531–542.

van Rosmalen, J., van Herk, H., & Groenen, P.J.F. (2010). Identifying response styles: A latent-class bilinear multinomial logit model. *Journal of Marketing Research*, *47*, 157-172.

Table 1. Estimation Results of Scobit and Logit Models

Explanatory variable	Logit Model		Scobit Model	
	Parameter	t-statistic	Parameter	t-statistic
(1) Model for the year of 1987				
Travel time difference (car-bus) (minute)	0.0428	1.858	0.0525	2.110
Travel cost difference (car-bus) (yen)	-0.0064	-6.790	-0.0070	-5.563
Value of travel time (yen/hour)	-403		-452	
			(Skewness parameter)	
Sex (1: Male; 0: Female)	1.2821	2.224	0.9606	2.330
Age	-0.1025	-3.939	-0.0758	-3.965
Employment (1: employed; 0: unemployed)	4.2383	4.048	3.3121	4.351
Number of household members	-0.3965	-1.661	-0.3824	-2.421
Converged log-likelihood	-84.05		-80.35	
McFadden's Rho-squared	0.4635		0.4871	
Adjusted McFadden's Rho-squared	0.4488		0.4731	
(2) Model for the year of 1990				
Travel time difference (car-bus) (minute)	-0.0210	-1.543	-0.0166	-1.176
Travel cost difference (car-bus) (yen)	-0.0035	-6.090	-0.0032	-5.321
Value of travel time (yen/hour)	358		311	
			(Skewness parameter)	
Sex (1: Male; 0: Female)	1.4874	2.850	0.9602	2.616
Age	-0.0630	-2.710	-0.0404	-2.726
Employment (1: employed; 0: unemployed)	3.1081	2.914	1.5661	2.563
Number of household members	-0.6557	-2.877	-0.2718	-2.013
Converged log-likelihood	-104.16		-107.51	
McFadden's Rho-squared	0.3351		0.3137	
Adjusted McFadden's Rho-squared	0.3169		0.2950	
(3) Model for the year of 1993				
Travel time difference (car-bus) (minute)	-0.1265	-3.378	-0.1189	-3.751
Travel cost difference (car-bus) (yen)	-0.0048	-6.031	-0.0048	-7.120
Value of travel time (yen/hour)	1592		1499	
			(Skewness parameter)	
Sex (1: Male; 0: Female)	2.9801	3.377	1.8141	4.179
Age	-0.1255	-3.932	-0.0723	-4.493
Employment (1: employed; 0: unemployed)	2.1011	1.437	1.1865	1.120
Number of household members	0.8661	2.508	0.5146	2.397
Converged log-likelihood	-65.69		-64.23	
McFadden's Rho-squared	0.5807		0.5900	
Adjusted McFadden's Rho-squared	0.5692		0.5788	
(4) Model for the year of 1994				
Travel time difference (car-bus) (minute)	-0.0478	-2.372	-0.0472	-2.113
Travel cost difference (car-bus) (yen)	-0.0037	-6.135	-0.0043	-6.116
Value of travel time (yen/hour)	781		664	
			(Skewness parameter)	
Sex (1: Male; 0: Female)	2.7388	4.384	2.3982	5.040
Age	-0.0738	-3.437	-0.0566	-3.521
Employment (1: employed; 0: unemployed)	0.5824	0.676	0.0602	0.091
Number of household members	0.3625	1.437	0.3776	1.921
Converged log-likelihood	-86.97		-81.68	
McFadden's Rho-squared	0.4448		0.4786	
Adjusted McFadden's Rho-squared	0.4297		0.4644	
Sample size (persons): Same across 4 waves	226			
Initial log-likelihood: Same across 4 waves	-156.65			