MULTI-ATTRIBUTE ENVIRONMENTAL VALUATION BY MIXED LOGIT

by Michelle PARUMOG **, Shoshi MIZOKAMI*** and Ryuji KAKIMOTO****

1. Introduction

One way of deriving value of non-marketed resource such as road and roadside environment is through stated choice experiment. Attribute-base or conjoint stated choice experiments are generally analyzed using discrete choice models. Various studies on the valuation of time[1)], non-use goods such as road safety[2)] and cultural heritage resource, and corporate decisions have been done utilizing flexibility of conjoint experiment based on willingness-to-pay (WTP) indicators from random utility models. Recent development in analytical models accounting for random taste heterogeneity offers new powerful analytical method for discrete choice models dominated before by multinomial logit and nested logit models. This development also opens possible applications in environmental valuation.

Environmental amenity, just like any hypothetical alternative attribute in a stated choice question, is a complex concept that varies according to individual. For example, a person with a flexible work schedule may seek a route with better environmental quality than a person which is already running late for work. The difference in how individual perceive importance of alternative's attributes like environmental quality causes parameters of discrete choice structural equations to follow certain distribution. Moreover, complex representation of attributes of the environmental goods presents a challenge in the structural equations of models. For example, valuation of water quality where attributes are disaggregated into toxicity and transparency of waters may cause inherent correlation between the two attributes as some individual may perceive good water transparency to equate with low toxicity. Typical logit models present limitations in addressing taste heterogeneity and attributes correlation in environmental valuation problems.

A wide array of discrete choice analytical models taking into consideration taste heterogeneity, attitudes, and correlation structure among attributes of alternatives have been developed. Among these models are probit, GEV, latent class and the current pervasively used mixed logit models. Mixed logit can estimate any random utility problem[3)]. It is more powerful over standard logit is that it allows for random taste variation, unrestricted substitution patterns, and correlation in unobserved factors over time and, unlike probit, it is not limited to normal distribution[4)]. Features of mixed logit to describe a bounded parameter distributions such as log-normal and capacity to induce flexible correlation pattern in model specifications make it very appropriate for valuation problems.

To study the multi-attribute environmental valuation, a stated choice conjoint experiment is performed and estimated using multinomial logit (MNL) model. To account for taste heterogeneity, various specifications of mixed multinomial logit (MMNL) model are estimated. In so doing, this paper aim to answer the questions: (1) how inherent covariance structure of parameter of environmental attributes vector affect marginal utilities, and (2) how willingness-to-pay estimates is affected by specification of mixed logit model.

2. Empirical application

The data used in this experiment was derived from an online pre-test survey conducted of road and roadside environment in Metro Manila (MM) January 20 to February 5, 2005. Respondents are workers from different business districts within MM. Valuations are done on the framework of binary route choice experiment. Environment attributes investigated are: (1) air pollution, (2) noise pollution, (3) landscape, and (4) road safety. At first, the respondent was asked to imagine that his/her regular working trip takes about is 60 minutes and average transportation cost of 30 pesos. Then, he/she was asked to choose between two route options offering environmental improvements. Each route choice problem contains two alternatives in which attribute levels are drawn randomly from the following set. Questionnaire in HTML was embedded with script randomizing attribute levels. The process was repeated three times.

Table 1: Attribute levels

Attributes	Attribute levels set
Travel time:	30 minutes, 45 minutes, 60 minutes, 75minutes
Travel Cost:	30 PhP, 40 PhP, 50 PhP, 100 PhP
Air Quality Improvement:	20% improvement in air quality, 50% improvement in air quality
Reduction in Noise Pollution:	20% reduction in traffic noise, 50% reduction in traffic noise
Landscape:	with improvement, without improvement
Accidents/year:	20 accidents per year, 50 accidents per year, 100 accidents per year

*Keywords: stated choice method, environmental valuation, mixed logit
**Student Member, Graduate School of Science and Technology, Kumamoto University (2-39-1, Kurokami, Kumamoto City, Kumamoto, Japan 860-8555)
***Member, D. Eng., Professor, Department of Environmental and Civil Engineering, Kumamoto University (2-39-1, Kurokami, Kumamoto City, Kumamoto, Japan 860-8555)
****Member, PhD., Associate Professor, Department of Environmental and Civil Engineering, Kumamoto University (2-39-1, Kurokami, Kumamoto City, Kumamoto, Japan 860-8555)

A total of 65 filled questionnaires were received. After it was expanded according to stated route choices, a total of 176 data were used in the analysis after further screening of responses. Majority of the respondents are employees in Makati and Ortigas CBDs. Around 61% of the respondents are male and the average age is 30 years old.

3. Fixed Coefficient Model

To determine robustness of models incorporating taste heterogeneity, we first estimated binary logit model with linear utility function for each *j* alternative specified linearly as:

$$V_j = \beta_1 TT_j + \beta_2 TC_j + f(\beta_{ATT}, ATT_j) + \varepsilon_j, \quad \text{j=1,2} \tag{1}$$

where *TT* stands for travel time, *TC* stands for transportation cost and $f(\cdot)$ stands for the function describing environmental attributes effects. For investigation purposes, a linear specification of the environmental attributes vector below is described in this study.

$$f(\beta_{ATT}, ATT) = \beta_3 AIR + \beta_4 NOISE + \beta_5 LA + \beta_6 ACC + \varepsilon_{ATT} \tag{2}$$

In this function, *AIR* stands for air pollution level (1-% reduction), *NOISE* stands for noise pollution level (1-% reduction), *LA* is a dummy representing improvement (with or without) and *ACC* is number of accidents.

Assuming effects of attributes are captured in deterministic part (i.e. $\varepsilon_{ATT} = 0$), willingness to pay (WTP) indicator in this choice problem can be described by the subjective elasticity of attribute coefficient, say air pollution reduction (*SVAIR*), with respect to cost computed for each observation *i*. In a simple linear specification of utility, the value of a unit of air pollution reduction (*VAIR*) can then be computed from the ratio of β_3 and β_2 which are the air quality and cost coefficient respectively.

$$SVAIR_i = \frac{\partial V / \partial AIR}{\partial V / \partial TC} = \frac{\beta_3}{\beta_2} \tag{3}$$

For nonlinear specification, to compute for the value of a unit improvement in air quality, *SVAIR* should be obtained for each individual and averaged over the number of observations *i*.

4. Random Coefficient Model

(1) Definition and estimation

Since tastes and perception varies per individual, it is not likely for the estimated coefficients to be fixed or common across observations. To consider this, we estimated an MMNL model[3] where not only stochastic part of the indirect utility, but also alternative attribute coefficients, varies randomly. Assuming for instance that β follows a continuous normal distribution the choice probability is specified as:

$$P = \int \left(\frac{e^{V_j}}{1 + e^{V_j}} \right) \phi(\beta \mid b, \Omega) d\beta \tag{4}$$

where $\phi(\beta \mid b, \Omega)$ is a normal density with mean *b* and covariance Ω. This can be estimated by maximum simulated likelihood where *P* is estimated by drawing values of β from assumed density, then calculating average to compute the simulated probability $\widetilde{P}_{nj}$ as follows:

$$SSL = \sum_{n=1}^{N} \sum_{j=1}^{J} \theta_{nj} \ln \widetilde{P}_{nj} \tag{5}$$

where θ_{ni} is the choice dummy. Mixed logit allows for flexible specification as it permits heterogeneity, correlation and taste variations in data.

(2) Issues in valuation

In attribute base discrete choice analysis with linear specifications, the willingness to pay indicator for changes in attribute is equal to the coefficient of the attribute over the coefficient of the cost vector. This mean that estimates are based on point estimates or average value in sample. Several issues on the computation of willingness to pay may arise in the use of mixed logit. One is the possibility of positive attribute or price coefficient particularly when distribution is unbounded (e.g. normal distribution). There are two ways of looking at this issue. Lack of explanatory power of positive coefficients may lead it to be categorized as misspecification. On the other hand, share of positive value coefficients derived from unbounded distributions can be explained by data impurities or observations not following counter intuitive behavior. Unbounded distributions (i.e. lognormal distribution) offer better alternative to normal distribution as it can restrict sign of parameter. While some studies have found that lognormal distribution performed well than unbounded distribution, other find difficulty in making it converge [2]. The possible reason for this is the long tail on the unbounded side. Better bounded distributions are found to be easier to use than the log normal (e.g. triangular distribution).

Another issue related to multi-attribute valuation is the attribute correlation. The capability of mixed logit to incorporate covariance structure among alternatives addresses this issue.

(3) Model specification

Based on the issues discussed above, the following mixed logit specifications are used in this paper. Standard mixed logit with normally distributed parameter of price, time, and environmental attributes vectors (MMNL) is first estimated. Then, to investigate costs effects, mixed logit with fixed parameter for cost vector and normally distributed time and environmental attributes vector (MMNLFC) is ran. To account for the inherent covariance structure of the environmental attributes, mixed logit with normally distributed parameter of price and time vectors, and environmental attributes with variance-covariance structure for environmental attributes vector (MMNLCOV) is estimated. Finally, mixed logit with log-normally distributed time and environmental attributes vector to describe effect of bounded distribution like lognormal are implemented. The models are evaluated based on the robustness of models and estimates (-LN).

5. Results and Discussion

Models MNL, MMNL, MMNLFC, MMNLCOV and MMNLFC-LN were estimated using a non-commercial estimation package BIOGEME 1.2. Pseudo random numbers were used to simulate normal distribution of the coefficients. Table 2 below shows parameter estimates of the different models, t-statistics are shown in the parentheses. In all the models, 100 draws to approximate probability are done. Model MMNL-LN was likewise implemented but is not presented because of poor estimates.

Table 2: Parameter estimates of MNL, MMNL, MMNLFC, MMNLCOV, MMNLFC-LN

	MNL		MMNL $\beta \sim N(\mu,\sigma)$		MMNLFC $\beta \sim N(\mu,\sigma)$		MMNLCOV $\beta \sim N(\mu,\sigma)$		MMNLFC-LN $\beta \sim LN(\mu,\sigma)$	
α_2	-0.302	(-1.47)	-2.590	(-0.85)	-1.935	(-0.79)	-11.108	(-2.15)	1.358	(0.44)
β_1 (Travel time)	-0.048	(-5.35)	-0.514	(-1.05)	-1.052	(-1.13)	-1.879	(-2.64)	-3.690	(0.37)
σ_1			0.149	(0.95)	0.808	(-1.13)	-0.270	(-1.50)	1.542	-(1.00)
β_2 (Travel cost)	-0.023	(-3.88)	-0.297	(-0.97)	-0.393	(-1.13)	-0.912	(-2.37)	-4.079	(0.32)
σ_2			-0.365	(-1.00)			-0.399	(-2.17)		
β_3 (Air pollution)	-0.032	(-3.02)	-0.343	(-0.95)	-0.525	(-1.08)	-0.759	(-2.09)	-14.527	(0.00)
σ_3			0.451	(1.01)	-0.23	(-1.03)	-0.319	(-1.34)	20.085	(1.00)
β_4 (Noise pollution)	-0.009	(-0.93)	-0.075	(-0.74)	-0.277	(-1.07)	-0.590	(-2.60)	-5.353	(0.00)
σ_4			0.512	(0.95)	-0.198	(-1.08)	-0.541	(-2.37)	-12.988	(0.00)
β_5 (Landscape)	-0.345	(-1.17)	-4.127	(-0.99)	-0.344	(-0.12)	4.846	(1.00)	2.892	(2.11)
σ_5			1.570	(0.57)	0.444	-0.16	-5.127	(-0.92)	-35.585	(14.65)
β_6 (Road safety)	-0.027	(-5.18)	-0.246	(-0.99)	-0.919	-1.13	-2.594	(-2.45)	-5.370	(0.77)
σ_6			0.027	(0.58)	-1.044	(-1.13)	0.041	(0.47)	0.520	(0.94)
$\Omega(\beta_4,\beta_3)$							1.352	(2.51)		
$\Omega(\beta_5,\beta_3)$							0.000	(0.00)		
$\Omega(\beta_5,\beta_4)$							-1.782	-(0.38)		
$\Omega(\beta_6,\beta_3)$							0.564	(2.23)		
$\Omega(\beta_6,\beta_4)$							3.307	(2.49)		
$\Omega(\beta_6,\beta_5)$							-0.260	-(1.63)		
Parameters	**7**		**13**		**12**		**19**		**12**	
N	**176**		**176**		**176**		**176**		**176**	
LR	**83.375**		**100.48**		**106.255**		**109.73**		**95.8682**	
Adjusted ρ^2	**0.284**		**0.305**		**0.337**		**0.294**		**0.491**	

Based on goodness of fit indicator adjusted ρ^2, except for the MMNLFC-LN, random coefficient models provide more robust than the fixed parameter model. MMNL model shows that cost parameter β_2 did not significantly vary across sample which is also reflected in better MMNLFC model where it is held fixed. The explanatory power of the landscape and noise parameters appears to be marginal across all models. On the other hand air pollution and accident parameters provide more robust estimators. The covariance value estimates in the MMNLCOV provide robust models. The model supports strong covariance among the attributes *AIR*, *NOISE* and *ACC*. Relationship between air and noise betas, for instance show very significant relationship. In practical terms, this is very realistic as air and noise are environmental attributes which may be hard to differentiate from each other. The lognormal mixed logit models is the most difficult to converge among the models. Moreover, it sometimes fails to provide reliable estimators as can be seen in MMNLFC-LN above.

Figure 2 shows the cumulative distribution function (CDF) of the coefficient estimate of the MMNL model. From the figure, the tendency of some of the parameters (i.e. *TC, AIR, NOISE*) to have positive values in some part of the distribution can be seen. The wide variance of the coefficient of variable *LA* can be also observed indicating perception difficulty for the variable.

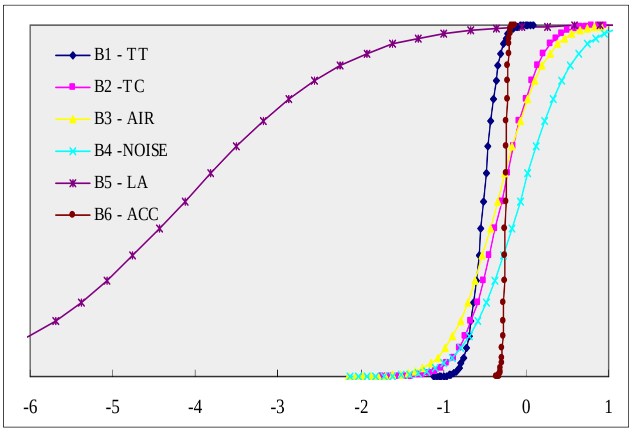

Figure 2: CDF of MMNL Coefficient Estimates

Estimates of WTP for the different road environment attributes based on equation 3 and its standard error in parentheses are shown in Table 3. The approximation was done by computing variance of two estimators and using Taylor series where second order equation was ignored. From the values below, it can be seen that random coefficient models present better estimates. MNL, on the other hand, unlike random coefficient models present no room for negative values. The goodness of estimated value within the mixed logit models depends greatly on how the models are specified. From the models with normally distributed random coefficient, MMNLCOV shows the most reliable estimate, though it failed to give positive value estimate for landscape. Moreover, MMNLFC provide slightly better estimates than MMNL. This could probably be attributed to the noise in estimates distribution of the cost coefficient is giving in MMNL. Though the values derived from the MMNLFC-LN are significantly different from all the values, the estimates reliability can be superior to the others. The difference in the mean WTP estimate of the normal and lognormal distributed random parameters may be explained by the tendency of the log normal distribution to misrepresent true mean because of the distribution shape.

Table 3: Estimated WTP for time and environmental attributes for different models and its standard deviation in ( ).

	MNL	MMNL	MMNLFC	MMNLCOV	MMNLFC-LN
β_1 / β_2 (PhP/minute)	2.09 (0.581)	1.73 (0.347)	2.68 (0.371)	2.06 (0.216)	0.90 (0.109)
β_3 / β_2 (PhP/1%improvement in air)	1.39 (0.513)	1.16 (0.356)	1.34 (0.267)	0.83 (0.168)	3.56 (0.282)
β_4 / β_2 (PhP/1%improvement in noise)	0.39 (0.427)	0.25 (0.336)	0.70 (0.212)	0.65 (0.136)	1.31 (0.104)
β_5 / β_2 (PhP/landscape improvement)	15.00 (13.153)	13.91 (7.866)	0.88 (7.241)	-5.31 (5.088)	-0.71 (0.507)
β_6 / β_2 (PhP/accident)	1.17 (0.330)	0.83 (0.145)	2.34 (0.273)	2.84 (0.214)	1.32 (0.217)

6. Conclusion

Results of estimation of pre-test data done in this study shows that random coefficient models and estimate provided more robust estimate than fixed model. It should be noted however that run time of random coefficient model is significantly longer than that of fixed parameter. In terms of multi-attribute analysis of environmental change, better interpretation of the model can be done using random coefficient model on the context of: (1) variation of attribute parameter across individuals, (2) inherent correlation structure depicting how respondent perceived multi-attribute choice. Random coefficient models are therefore better estimated in project evaluation involving multi-attribute environmental change

It should also be understood that reliability of estimates of random parameter models greatly depends on the specification of the models. Estimate from the different models shows that WTP values from random parameter model have tighter confidence interval than that of the fixed parameter models. In terms of the parameter of the cost vector, estimates can be tighter if cost parameter is held fix. With regards to the correlation structure of the environmental attributes vectors, including covariance structure in the attributes improve both model and estimates.

Based on the investigation of mixed logit in multi-attribute valuation done in this study, the authors recommend: (1) use of mixed logit with covariance structure for the attribute set to better explain goods in question; and (2) use of fixed cost coefficient for more centered estimate.

From the richness of mixed logit model and the robustness of the different WTP values estimated, various environmental policies can be made and evaluated. Extension of the models to include socio-economic variables, attitude data and environmental level perception is recommended to be done in future studies.

References

1) Hess, S., Bierlaire, M. and Polak, J.W. (2005) Estimation of value of travel-time savings using Mixed Logit models, Transportation Research A, Volume 39, Issue 3, Pages 221-236
2) Iraguen P., and Ortuzar J.D. (2004) Willingness-to-pay for reducing fatal accident risk in urban areas: an Internet-based Web page stated preference survey, Accident Analysis & Prevention, Volume 36, Issue 4, 1 July 2004, Pages 513-524
3) McFadden, D. and K. Train (2000) “Mixed MNL Models for Discrete Response”, Journal of Applied Econometrics 15(5), 447-470.
4) Train K. (2003). Discrete Choice Methods with Simulation, Cambridge University Press