INCORPORATING DECISION RULES IN SEGMENTING AND MODELING MODE CHOICE

Dave Henley and K. C. Koutsopoulos[*]

INTRODUCTION

In transportation, the explanation of travel choices, based upon behavioral models, is viewed as a means of providing a firmer theoretical basis for travel demand models. As a result a considerable amount of recent research in travel demand forecasting has been directed toward the development of behavioral models of transportation mode choice [6, 22, 28]. In general, these behavioral models attempt to explain the decisions that trip makers make, through the study of variables which affect travel behavior, and to represent this decision-making process.

One set of models that has been described as behavioral are those built upon axiomatic representations of choice. The most exemplary of these are the logit and probit models which are defined by equations 1 and 2 respectively [18, 28, 29]:

(1) $P_i = 1/(1 + e^{f(x)})$

(2) $P_i = 1/(2\pi)^{1/2} \int_{-\infty}^{f(x)} e^{-t/2} \, dt$

where

P_i = probability of choosing mode i

e = base of the natural logarithm

f(x)= set of independent variables describing characteristics of both the transportation system and the traveler.

Examples exist in the transportation literature of attempts to provide interpretations of the logit and probit mode choice models in terms of individual decision processes which are assumed to underlie them [6, 12]. These, however, are strongly tied to the response function (curve) and not to the factors that influence modal choice and consequently the decision-making process. Thus, the models provide, in many instances, reasonable representations of mode choice behavior, but statements as to why the behavior occurs are weak suggesting a correlative rather than a causal relationship [9].

---

[*]Ph.D. candidate and Assistant Professor, respectively, Department of Geography, University of Iowa.

In hopes of achieving a more causal explanation researchers have examined the utility function incorporated into the logit and probit mode choice models [f(x)]. Emphasis has been placed on: first, identifying those transportation attributes which seem to have the greatest influence on individual mode choice decisions [7, 10]; and second, methods for incorporating the perception of individuals with respect to relevant transportation attributes [19, 21]. Little empirical effort, however, has been directed toward investigating how attributes, which are important in the mode choice decisions, affect the form of the utility function.

Furthermore, the logit and probit mode choice models, as presently formu- lated   assume   that the relevant transportation attributes in the decision process, the perception of these attributes, and the way in which they are combined in forming utility functions are invariant across segments of the study population. A developing literature, however, which has examined the modal choice decision-making process, suggests that these assumptions are untenable [14, 22], and also that individuals can be segmented into groups on the basis of decision-making rules [11, 17].

This research is an attempt to demonstrate how the study of individual decision-making might be used in overcoming these two common assumptive drawbacks in the use of logit and probit mode choice models: first, by defining the form and content of the utility function f(x) instead of assuming a priori a constant factor set and linear relationship; and second, by demonstrating that individuals can be segmented by decision-making rules instead of relying only on socio-economic variables.

To achieve this goal, models of decision-making in a hypothetical mode choice situation will be estimated for a study population. Responses in these hypothetical situations will then be used to segment the population into salient groups of decision makers. Logit and probit mode choice models will then be fit to each segment using two sets of data: first, objective measures of mode attributes with an assumed functional form for the utility function (the traditional approach); and second, measures of the subjective or perceived value of mode attributes with a form derived from responses to the hypothetical situations (a decision-making approach). The results of these analyses and their comparison will then be discussed in terms of their implications for the modeling of travel demand and mode choice.

For this research, objective values are defined as engineering or economic measurements of the transportation system (i.e., cost of parking or bus fare), with subjective values being quantified estimates of a travelers perception of the objective factors.

METHODOLOGY

In order to identify possible decision-making models the functional measurement method was employed. The method, which was developed primarily by Norman Anderson [1], has enjoyed considerable empirical success in recent years in modeling human decision-making processes [4, 13, 16], including travel behavior [11, 22].

The functional measurement approach evolved from the view that individuals are information processors. That is, when an individual forms a judgment related to an entity--which could be anything ranging from the likelihood of using a bus system to the "likability" of a person--this judgment is the outcome of an information processing task in which the individual "combines" cognitively his perceptions of a set of attributes of that entity to arrive at an overall evaluation. In addition, the theory posits that each attribute can be described in terms of its "subjective scale value" (its perceived magnitude), and "weight" (its relative importance in the decision-making task), and also that the various attributes are combined according to an implicit decision rule, which can be described algebraically. For example, for a set of attributes this can be expressed mathematically as:

$$(3) \quad R_i = g_i(w_a s_a)$$

That is, the overall response R_i of an individual is a function of the relative "weights" w_a and the "subjective scale values" s_a ascribed to various attributes $\underline{a}$ of the entity under judgment.

A tool common in most applications of the method is that of the factorial design, in which subjects are presented with all possible combinations of attributes in an experiment. The factorial design enables the researcher to specifically test for the appropriate "combination rule" through an analysis of variance of the response data. In the general case the "subjective scale values" correspond to the "marginal means" of the analysis, and weights may be derived either directly through examination of main effects (if the combination rule is a linear one) or through numerical estimation (if it is non-linear).

It should be noted that the choice situations presented in a controlled experiment may produce models that are unrelated to real world behavior. This question of validation is of basic importance for the further development of travel demand models from an experimental standpoint, and yet it remains largely unanswered. Although sparse, attempts to confirm the correspondence between experimentally derived models and actual behavior in both mode choice (20) and destination choice (23) indicate a strong relationship may exist; much more research is needed.

Experimental Design

A set of relevant factors describing both mode and destination characteristics were incorporated in the factorial design (3[5]). The factors and their levels were: cost of parking (CP) at levels of \$156, \$104 and \$52, bus fare (BF) at levels of 45¢, 30¢ and 15¢, frequency of service (FS) at levels of 45 minutes, 30 minutes and 15 minutes, distance to the bus stop (DB) at levels of 1 block, 3 blocks and 5 blocks, and travel time (TT) at levels of 5 minutes, 15 minutes and 25 minutes. Subjects were instructed to respond to a number of possible combinations of the various levels of these factors on a scale that was anchored on one end by "never take bus" and on the other by "always take bus." Responses were provided by a slash mark on a 150 mm line scale. By recording the length of the mark from the end of the line, to the nearest 5 mm, the responses were

converted into a measurement that served as the dependent variable for the
decision models.   End anchors and fillers were used to minimize bias in the
experimental design [2, 5].

Since the data were obtained from an experimental situation, other variables
which might affect individual travel decisions can be controlled.   For example,
to hold the effect of trip purpose constant, subjects were instructed to consider
only trips to the university.   The sample consisted of faculty and staff members
at Florida State University who used a car or bus as a means of travel to the
university.   They were selected at random from the faculty/staff directory and
received the experimental design through the mail.   The return rate was 40
percent.

<div align="center">ANALYSIS</div>

In an effort to derive models of decision-making for each of the individuals
in the study, the first step was to subject the response data to an analysis of
variance.   The results indicate that the decision-making processes of a majority
of the respondents can be predicted by a simple additive model.   However, a
fundamental difference exists among individuals, whose decision-making process
is captured by this additive model, namely the factors which are considered
important in making the decision.   As a result, individuals were differentiated
into two groups based on the factors incorporated into their decision-making
models.   The first group consisted of individuals who consider only travel costs
and travel time as important to their mode choice decision, and can be rep-
resented by the following model form:

(4) $R_i = W_1 CP_j + W_2 BF_k + W_5 TT_n$

where

R_i = the response of individual $\underline{i}$ on the line scale

CP_j, BF_k, TT_n = independent variables (defined previously)

W_1, W_2, W_5 = weights or importance of factors

The individuals of the second group, while considering cost and time variables to
be important, also deem frequency of service and distance to the bus stop as
significant in the mode choice decision.   This group, then, can be described
by the following model form:

(5) $R_i = W_1 CP_j + W_2 BF_k + W_3 FS_l + W_4 DB_m + W_5 TT_n$

where

R_i = the response on the line scale

CP_j, FB_k, FS_l, DB_m, TT_n = independent variables (defined previously)

W_1, W_2, W_3, W_4, W_5 = weights or importance of factors

For identification ease, the segment of individuals described by (4) will be referred to, hereafter, as group A (n = 142) and those individuals described by (5) as group B (n = 140).

Fitting the Logit and Probit Models

The utility function f(x) of the logit and probit mode choice models can be defined in numerous ways. Traditionally the form of the utility function, which has demonstrated the greatest empirical success, is comprised of travel costs and travel times of alternative transportation modes expressed in a difference formulation (3, 6). In symbolic terms it can be written as:

(6) $f(x) = b_0 + b_1(C_i - C_j) + b_2(T_i - T_j)$

where

C_i = cost associated with mode i

C_j = cost associated with mode j

T_i = travel time associated with mode i

T_j = travel time associated with mode j

b_1, b_2 = estimated coefficients

The independent variables in (6) can be defined by incorporating the factors used in the hypothetical mode choice experiment (described earlier). In symbolic terms:

(7) $f(x) = b_0 + b_1(CP_j - BF_k) + b_2(TT_{auto} - TT_{bus})$

where

CP_j = cost of parking

BF_k = bus fare

TT_{auto} = travel time of auto

TT_{bus} = travel time of bus

b_1, b_2 = estimated coefficients

The input values for the independent variables in (7) can be measured in two dimensions. The objective values, the first dimension, are the levels of the factors presented in the hypothetical mode choice situation and can be used, as it has been done traditionally in logit and probit models. However, since the measures of perception in modeling mode choice are important and need to be emphasized, the subjective values, the second dimension, can also be utilized. Subjective values are the perceived values of the factors, derived from the analysis of the experiments' results (ANOVA's marginal means) [15].

Furthermore, in order to incorporate decision rules in the definition of f(x) and thereby overcome the serious drawback of a priori definition of the

function, the models developed in the mode choice experiments (Equations 4 and 5) can be used. In addition, since each of these models represent a segment of individuals who use a similar decision strategy in making a model choice, the assumption made by logit and probit models, that the study population possesses a homogeneous utility function, is also satisfied. Therefore, four models are available and will be used in defining $f(x)$: first, the traditional difference formulation defined in terms of objective values; second, the same formulation defined in terms of subjective values; third, the decision-making rule defined by Equation 4; and fourth, the decision-making rule defined by Equation 5.

Estimation

In estimating the logit and probit models a cross validation approach was used. That is, each group, A and B, was divided into two equal segments, one for estimating the mode choice models, the second for making mode choice predictions. For each group, $f(x)$ of the logit and probit models was defined by both the decision model, unique to that group, and the travel cost and travel time difference models (with objective and subjective values) represented by (7). For example, consider group A in which individuals employed the additive deicision model (CP + BF + TT). This additive relationship and the travel cost and travel time differences models defined in both the objective and subjective dimensions were used to define $f(x)$ of the logit and probit mode choice models. In each case maximum likelihood procedures were used to estimate the models [26]. These models, then, were used to predict mode choices for the second data segment [27].

The models were evaluated in two ways: first, in terms of their ability to correctly predict mode choices, measured by the number of mode choice mis-classifications and second, in terms of their statistical significance as measured by the correlation ratio, which is accompanied by an F-statistic corrected for sample size.

Stopher [25] reports on the strengths and weaknesses of various statistical goodness-of-fit measures and concludes that the correlation ratio is more useful, than other measures, for assessing non-linear travel demand models, since it is not based on a linearity assumption. However, the correlation ratio and associated F-statistic should only be used as a criterion for evaluating which one of a set of models best fits the data, that is, as a comparative and not as an absolute measure.

Because of the possibility that subjects were using different parts of the response scale in the mode choice experiment, a transformation was used to convert the likelihood of mode choice to a binary variable. This transformation involved determining the mean of an individual's response scale and allowing it to allocate the response into either a 0 or 1 category, with 0 representing a car choice and 1 a bus choice. This transformation assumes that the responses are monotonically related to a car or bus choice, that is a 0 or 1 transformation. However, the binary responses allow the derived models to be compared with previous research

dealing with mode choice because the logit and probit models are discrete choice models.[1]

The results for the models tested are shown in Table 1. For group A, the best fit was provided by the logit and probit models incorporating the decision rule (CP + BF + TT). The logit and probit models defined by the functional relationship of travel cost and travel time differences, with factors measured at the objective level, were also significant. In terms of misclassifications, only a small absolute number separated these latter models from the models defined by the decision rules. However, the logit and probit models incorporating subjectively measured travel cost and travel time differences were unable to accurately fit the data when compared to the other models tested.

For group B, the best fit was obtained when the decision rule (CP + BF + FS + DB + TT) was used to define the logit and probit models. The logit and probit models defined by travel cost and travel time differences in either the objective or subjective dimension provided poor fits to the data.

DISCUSSION AND CONCLUSIONS

The research objectives were given empirical support. For both groups the incorporation of decision rules produced predictive models that perform better than traditional methods of defining logit and probit mode choice models. The use of decision rules, however, provided, in addition, a means of overcoming the assumptive drawbacks of the traditionally formulated logit and probit mode choice models. Tripmakers who used similar decision rules in making their modal choice were defined and isolated, while the form of the utility function was explicitly investigated and not assumed a priori to be of a specific nature.

This study has also demonstrated that the traditional method of defining logit and probit mode choice is inadequate, since the travel cost and travel time difference formulation ignores, in many instances, mode choice variables which are important in the decision-making process of the population segments. This can be seen in group B by the inability of travel cost and travel time differences to accurately predict modal choice and the predictive success of the decision rule model. Thus, although the travel cost and travel time difference relationship can be adequate, in certain cases, when variables other than cost and time become relevant in the decision process, a better rule exists for defining the logit and probit mode choice models.

---

[1]An alternative approach is to standardize the responses, thus preserving the metric information. However when this standardization approach was employed in a second model fitting process (not reported here), no significant differences were apparent.

TABLE 1: Empirical Results

	Logit						Probit					
	Group A			Group B			Group A			Group B		
	η	F	Misclass.	η	F	Misclass.	η	F	Misclass.	η	F	Misclass.
f(x)=travel cost/ travel time diff. (objective)	.34	7.36*	13/72	.11	1.53	24/70	.34	7.36*	13/72	.11	1.37	24/70
f(x)=travel cost/ travel time diff. (subjective)	.11	1.49	21/72	.09	1.12	26/70	.11	1.32	22/72	.05	0.98	27/70
f(x)=CP+BF+TT	.38	9.42*	12/72				.35	8.35*	12/72			
f(x)=CP+BF+FS+ DB+TT				.25	3.76*	18/70				.27	4.85*	18/70

*significant at .01 level

η = correlation ratio

On the other hand, results from group A indicate that although the logit and probit models employing the decision rule provide better predictions than the other models tested, the traditional method of using cost and time differences cannot be discarded since these models also demonstrate good predictive ability. This suggests that if the correct decision rule is not known a priori, given that the relationship is linear or can be described by a linear process and the relevant variables are included, different methods of defining f(x) can yield similar results. This may be useful for predictive purposes, especially if data collection is difficult. Yet if the effects of system changes or the explaining of future travel demand are important, the confidence placed in the decision based mode choice model would be much greater. The parameters in the model could be given a more valid interpretation with respect to mode choice and evaluation of transportation system attributes since it represents a causal system.

Although the research emphasized the importance of utilizing measures of perception in modeling mode choice, the incorporation of subjective values in the defining relationship of travel cost and travel time differences in both groups provided insignificant results. This was expected for group B, since the models incorporating the objective values were also unable to accurately predict the choices being made. One should not expect the transformation of the incorrect rule into the subjective dimension to improve the model significantly. In terms of group A, a good data fitting model was produced by the travel cost-travel time difference relationship defined in the objective dimension, thus the incorporation of the subjective dimension was expected to do as well. A possible explanation, however, for this inconsistency might be that the scale values (ANOVA's marginal means) provide reliable subjective estimates only for the original factors used in the hypothetical mode choice situation and not for the travel cost-travel time difference relationship which represents new factors.

In conclusion this research indicated that the use of decision rules to define mode choice models can increase the researcher's ability to explain the mode choice process while still retaining good predictive models. More importantly, it has been demonstrated that individuals may be segmented on the basis of decision functions instead of the traditional socio-economic and demographic variables which have proved less than fruitful in developing a general theory of travel behavior. By investigating traveler decision-making processes, rules which relate, for example, modal choice to transportation system attributes, are established, thereby suggesting directions that should be pursued in developing theoretical statements of travel behavior. In addition, by utilizing perceived values of relevant transportation system variables researchers can effectively evaluate the factors involved in travel decisions. Although the research dealt with a small sample and examined only one type of work trip, it clearly indicated that the reliance upon decision rules will allow a better fundamental understanding of mode choice.

Even though the variable set used was select, other factors could have been included in the experimental design, thereby creating a more robust set of models for testing. It was felt, however, that such additions would not change dramatically the conclusions reached. Finally, no contention is made that the ultimate mode choice model has been developed, but presently used mode choice methodologies are inadequate and hopefully, this research points to new directions for overcoming these shortcomings.

REFERENCES

1. Anderson, N. H. "Functional Measurement and Psychological Judgment," Psychological Review, 77 (1970), 153-170.

2. Anderson, N. H. "Information Integration Theory: A Brief Survey," in Krantz, Atkinson, Luce, and Suppes (eds.), Contemporary Developments in Mathematical Psychology, Volume 2, San Francisco: W. H. Freeman, 1974.

3. Ben-Akiva, M. "Disaggregate Multimodal Model for Work Trips in the Netherlands, Transportation Research Record, 569 (1976), 107-123.

4. Bettman, J. R., N. Capon and R. J. Lutz. "Cognitive Algebra in Multiattribute Models," Journal of Marketing Research, 12 (1974), 151-164.

5. Corry, F. A. A Psychological Analysis of the Mode Choice Problem, unpublished Masters thesis, University of Iowa, 1975.

6. Domencich, T. A. and D. McFadden. Urban Travel Demand, North Holland/ American Elsevier, 1976.

7. Golob, T. F., E. T. Canty and R. L. Gustafson. "An Analysis of Consumer Preferences for a Public Transportation System," Transportation Research, 6 (1972), 13-39.

8. Hartgen, D. and G. H. Tanner. "Investigations of the Effect of Traveler Attitude in a Model of Mode-Choice Behavior," Highway Research Record, 369 (1971), 1-14.

9. Hartgen, D. and M. Wachs. "Disaggregate Travel Demand Models for Special Context Planning: A Dissenting View," Transportation Research Board, Special Report 149 (1974), 154-163.

10. Hille, S. T. and T. K. Martin. "Consumer Preference in Transportation," Highway Research Record, 197 (1967), 36-41.

11. Koutsopoulos, K. C. and R. Meyer. "Mass Transit Decision Making Market Segmentation: A Pilot Study," Technical Report 69, Institute of Urban and Regional Research, University of Iowa, 1976.

12. Lave, C. "A Behavioral Approach to Modal Split Forecasting," Transportation Research, 3 (1969), 463-480.

13. Leon, M., G. C. Oden and N. H. Anderson. "Functional Measurement of Social Values," Journal of Personality and Social Psychology, 27 (1973), 301-310.

14. Levin, I. "Information Integration in Transportation," in Kaplan and Schwartz, (eds.), Human Judgment and Decision Processes: Applications in Problem Settings, New York: Academic Press, 1976.

15. Louviere, J. J. "Psychological Measurement of Traveler Attributes," in Hensher and Dalir, (eds.), Identification and Valuation of the Determinants of Travel Choice, London: McGraw-Hill, 1977.

16. Louviere, J. J. and R. Meyer. "A Model of Residential Impression Formation," Geographical Analysis, 9 (1976), 479-486.

17. Louviere, J. J., et al. "Travel Demand Segmentation: Some Theoretical Considerations Related to Behavioral Modeling," Transportation Research Board, Special Report 165, 1976.

18. Luce, R. D. Individual Choice Behavior, New York: John Wiley and Sons, 1959.

19. McFadden, D. "Conditional Logit Analysis of Qualitative Choice Behavior," in Zarembka, (ed.), Frontiers in Econometrics, New York: Academic Press, 1973.

20. Meyer, R. and I. Levin. "Functional Determinants of Mode Choice," Technical Report 88, Institute of Urban and Regional Research, University of Iowa, 1977.

21. Nicholaidis, G. C. "Quantification of the Comfort Variable for Use in Disaggregate Mode-Choice Models," Transportation Research, 9 (1976), 55-61.

22. Norman, K. L. and J. J. Louviere. "Integration of Attributes in Public Bus Transportation: Two Modeling Approaches," The Journal of Applied Psychology, 59 (1974), 753-758.

23. Piccolo, M. and J. Louviere. "Information Integration Theory Applied to Real-World Choice Behavior: Validation Experiments Involving Shopping and Residential Choice," Great Plains Rocky Mountain Geographical Journal, 6 (1977) 49-63.

24. Stopher, P. R. "A Probability Model of Travel Mode Choice for the Journey to Work," Highway Research Record, 283 (1969), 57-65.

25. Stopher, P. R. "Goodness-of-fit Measures for Probabilistic Travel Demand Models," Transportation, 4 (1975), 67-83.

26. Stopher, P. R. PROLO: Maximum Likelihood Program for Binary Logit and Probit Mode Choice Models, unpublished mimeo, Cornell University, 1974.

27. Stopher, P. R. STATP: Program for Cross-Validating Binary Logit and Probit Models, unpublished mimeo, Northwestern University, 1974.

28. Stopher, P. R. and J. O. Lavender. "Disaggregate Behavioral Travel Demand Models: Empirical Tests of Three Hypothesis," Transportation Research Forum, Vol. 13 (1972), 321-336.

29. Talvities, A. "Comparison of Probabilistic Modal-Choice Models: Estimation Methods and System Inputs," Highway Research Record, 392 (1972), 111-120.