8. J. Kruithof. Calculation of Telephone Traffic. De Ingenieur (E. Elektrotechnik), Vol. 52, 1927, pp. E15-E25.
9. B. Lamond and N.F. Stewart. Bregman's Balancing Method. Transportation Research, Vol. 15B, No. 4, 1981, pp. 239-248.
10. A.W. Drake. Fundamentals of Applied Probability Theory. McGraw-Hill, New York, 1967.
11. H. Theil. Principles of Econometrics. Wiley, New York, 1971.
12. S. McNeil. Quadratic Matrix Entry Estimation Methods. Ph.D. thesis. Department of Civil Engineering, Carnegie-Mellon University, Pittsburgh, Pa., Sept. 1983.
13. C. Hendrickson and S. McNeil. Estimation of Origin/Destination Matrices with Constrained Regression. In Transportation Research Record 976, TRB, National Research Council, Washington, D.C., 1984, pp. 25-32.
14. C. Hendrickson and S. McNeil. Matrix Entry Estimation Errors. Presented at the Ninth International Symposium on Transportation and Traffic Theory, The Hague, 1984.
15. G.G. Judge, W.E. Griffiths, R.C. Hill, and T.C. Lee. The Theory and Practice of Econometrics. Wiley, New York, 1980.
16. J.J. Heckman. The Common Structure of Statistical Models of Truncation Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. Annals of Economic and Social Measurement, May 4, 1976, pp. 475-492.
17. U. Landau, E. Hauer, and I. Geva. Estimation of Cross-Cordon Origin-Destination Flows from Cordon Studies. In Transportation Research Record 891, TRB, National Research Council, Washington, D.C., 1982, pp. 5-10.
18. I. Geva, E. Hauer, and U.Landau. Maximum Likelihood and Bayesian Methods for the Estimation of O-D Flows. In Transportation Research Record 944, TRB, National Research Council, Washington, D.C., 1983, pp. 101-105.
19. Y.M.M. Bishop, S.E. Feinberg, and P.W. Holland. Pseudo-Bayes Estimates of Cell Probabilities. In Discrete Multivariate Analysis, M.I.T. Press, Cambridge, Mass., 1975.
20. H.R. Kirby and M.N. Leese. Trip-Distribution Calculations and Sampling Error: Some Theoretical Aspects. Environment and Planning A, Vol. 10, 1978, pp. 837-851.
21. S. Tsygalnitzky. Simplified Methods for Transportation Planning. M.S. thesis. Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, 1977.
22. J. Simon and P.G. Furth. Generating a Bus Route Origin-Destination Matrix from On-Off Counts. Working Paper. Northeastern University, Boston, Mass., 1983.
23. B. Effron. Nonparametric Standard Errors and Confidence Intervals. Canadian Journal of Statistics, Vol. 9, No. 2, 1981, pp. 139-172.
24. B. Effron. A Leisurely Look at the Bootstrap, the Jackknife and Cross-Validation. American Statistician, Vol. 37, No. 1, 1983, pp. 36-48.
25. B. Effron. Nonparametric Estimates of Standard Error: The Jackknife, the Bootstrap and Other Methods. Biometrika, Vol. 68, No. 3, 1981, pp. 589-599.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.

Route Choice Analyzed with Stated-Preference Approaches

PIET H. L. BOVY and MARK A. BRADLEY

ABSTRACT

An application of scenario-based, or stated-preference, survey and analysis techniques is described in the context of cyclists' route choice. Route choice modeling with observed choice data is hampered by the cost of processing network data and by the difficulty of assessing the alternative routes and the perceived attributes of the routes considered by individual travelers. An alternative approach is to obtain stated evaluations of well-defined hypothetical routes. Such data were collected from commuting cyclists in the city of Delft in the Netherlands and analyzed by using functional measurement to estimate the relative importance placed on such route attributes as time, traffic level, and surface quality. Though the techniques used are well founded in the marketing and psychology literature, the route choice context raises issues that are particularly important for their application in transport analysis. A case study of the application of stated-preference techniques to route choice is discussed and empirical results obtained for urban bicycle trips are presented.

An understanding of the relative influence of time and cost versus qualitative factors on route choice is valuable in several types of transport system analysis. Alternative plans for new roads or cycling facilities may involve changes in travel times that must be weighed against costs or benefits in terms of other choice factors. These trade-offs are important in predicting the use of new facilities and assessing the benefits to the users. Such information may also be useful for large, network-based studies to define the paths that best represent travel alternatives and to assign the predicted flows for those alternatives to the network. Despite these clear needs for a more thorough understanding of the trade-off process in route choice behavior, theory building as well as modeling efforts in this field are still in need of development. There may be many reasons for this. For example, a route in a network is a difficult concept to deal with in quantitative and statistical analyses. Also, the choice situation with routes is relatively complex, being composed of many alternatives, which are not all clearly distinguishable and overlap slightly. Tackling the route choice problem with revealed-preference random-utility modeling approaches, therefore, poses serious difficulties to the researcher. One way to overcome a number of these problems is to collect preference data by offering hypothetical travel options to individuals in survey form, each option defined in terms of the attributes assumed to be most important. This general method is what is termed the "stated-preference" approach as opposed to revealed preferences inferred from choices among real options.

An application of stated-preference survey techniques in modeling route choice of bicyclists in the city of Delft in the Netherlands is described.

The relative importance of factors such as travel time, surface quality, traffic level, and cycling facility type was studied by varying them experimentally across sets of hypothetical routes. A secondary focus of this study was a practical one, that is, to assess the relative performance of various techniques for route description, grouping of alternatives, measuring preferences, and estimating preference functions in a hypothetical route choice context.

The main approach used in this study to assess cyclists' trade-offs is called functional measurement. This technique originated in the field of mathematical psychology (1) and has been developed in applications to many choice contexts, some within the transport modeling field (2). Thus, the purpose here is not to argue or extend the theoretical validity of the techniques, although the behavioral assumptions are made clear. Rather, the objective of this paper is to provide empirical evidence of the utility and efficiency of standard, established techniques of the stated-preference approach in route choice analysis and application.

APPROACHES TO ROUTE CHOICE ANALYSIS

In the past, route choice analysis has followed two main approaches. In the motivational-attitudinal approach, travelers are asked to state their reasons or motivations in selecting routes in a network. The results are often in the form of qualitative evaluations of the adequacy and importance of individual route attributes and of overall beliefs about alternative routes (3,4). From such studies it is generally agreed that the single most important influence on a driver's choice of route is travel time but that there are other important factors as well. Such studies provide useful input to other approaches by identifying choice factors, helping to define market segments, providing operational definitions for qualitative factors, and giving insight into the relevant choice sets and choice constraints facing individuals. Although it is possible to relate intended or actual choices to reported perceptions and attitudes (5), it is difficult to apply such relationships in assessing policies that can be characterized only as changes in observable route attributes.

With the revealed-preference approach, observable route characteristics are related directly to observed route choices, often by using an individual random-utility maximization framework, such as that of the logit model (6,7,pp.299-330). This approach requires knowledge or an estimate of the set of alternative routes considered by each person, as well as objective data for all salient attributes for each route in the set. Collection of such data is generally difficult and costly and does not ensure that the route characteristics used in modeling consistently represent the subjective measurements of those attributes made by travelers. This problem is especially relevant to route choice, where travelers perceive many route characteristics continuously, and these perceptions may vary a great deal over the course of travel. A variation on the revealed-preference approach requires travelers to state the choice alternatives considered for a familiar choice context and to provide their perceived values along a given set of attributes for each alternative. With this information, it is possible to estimate the relative importance and interactions of perceived route attributes in behavior.

STATED-PREFERENCE STRATEGY

The stated-preference approach is similar to this latter form of revealed-preference analysis, but with the set of choice alternatives and their attributes given in a hypothetical context and behavior measured as a rating, ranking, or stated choice among the alternatives rather than as an actual choice. Although the responses are not subject to the perception processes and choice constraints of actual choice contexts, Louvière et al. (8) found that the two approaches, carried out on the same sample, resulted in similar trade-offs among important mode-choice factors. In a larger mode-choice study (9), which included bicycle commuting, models from scenario-based data validated well against the sample's actual choice among their self-described mode alternatives.

Though stated-preference methods are still evolving rapidly, applications in the transportation field are already numerous. Theoretical background as well as practical aspects of application may be found in numerous reports (10-12). The relative advantage of the stated-preference approach, in most cases, is the controlled nature of the choice scenarios. This feature allows greater freedom in defining choice contexts, alternatives, and attributes as well as direct comparison with the responses across individuals. The ability to obtain multiple responses from each individual reduces sample size requirements and also enables the estimation of truly individual models. With these advantages comes the liability that the success of the approach depends largely on the consistency of the hypothetical alternatives and the corresponding sets of attributes with their perception in actual choice situations.

The study by Morisugi et al. (13) is the only application of a stated-preference analysis to route choice known to the authors. The study uses hypothetical route attributes to estimate values of time for qualitative route factors such as reliability, comfort, and safety by trading off between two factors at a time. In contrast, the authors designed a study by using the full-profile approach in which each hypothetical route was defined completely in terms of a selected set of variables.

Because of the peculiarities of a route, a simple transfer of experiences from other travel-related choices does not appear justified: spatial perception and visual impressions play an important role in the identification by the traveler. Unlike modes, routes cannot be readily labeled or classified into easily understandable categories and because a route is in fact a chain of different links, it is intrinsically heterogeneous. These aspects of routes require special consideration in a stated-preference study.

To examine the usefulness of stated-preference techniques in route choice analysis, a study was designed to serve both substantive and methodological aims: to analyze trade-offs of cyclists for route characteristics of their regular home-to-work trip (this trip purpose generates a large proportion of urban bicycle travel in the Netherlands) and to study the importance of various experimental design features and presentation techniques in the performance of the stated-preference approach.

CONTEXT OF THE SURVEY

In order to keep comprehension problems to a minimum, the study was directed toward frequent and experienced bicycle commuters. Hypothetical route choice situations were arranged and presented to them in questionnaire form. In the questionnaire, options for stimulus presented (the descriptions of choice contexts and alternative routes) and response measurement (the way of expressing preferences toward route alternatives) were varied methodically. In addition, the questionnaire asked subjects to rate the importance of the route choice factors directly and evaluate the survey in terms of ease of response and similarity to actual choice situations. The sample selected for this study consisted of 134 employees of Delft University, who lived in Delft and commuted by bicycle at least twice a week. As a result of the selection, most members of the sample faced similar traffic conditions and comparable commuting distances and used similar cycling facilities for at least part of their journey.

An attempt was made, however, to achieve a wide cross section across characteristics such as age, sex, profession, other modes available, and cycling frequency (more than twice a week) for sake of segmentation analysis and to explain variations in individuals' preferences.

The basic definition of the term "route" had to match the experience of the respondents yet be simple enough to represent with a small set of attributes. To this end a verbal description together with a pictorial form using a map were applied. A route was defined that consisted of a trip from home to work. Routes were further defined along a single set of attributes, each of which was assumed to be homogeneous along each alternative. People were asked to assess various routes as if they were single links. Conceptually, if the utility of a route is assumed to be a linear sum of link utilities, then the relative preference between two routes can be modeled as a function of the important differences in noncommon links. This conceptual solution, however, does not preclude difficulties in perceiving routes as homogeneous. Various presentational approaches were used to confront this problem.

ROUTE ATTRIBUTES AND LEVELS OF THE EXPERIMENTS

The set of route attributes was chosen after the results of previous research had been assessed (14,2). On the basis of such work and previous research in Delft, travel time, surface quality, traffic level, and cycle facility type were selected. Descriptions of the factors and levels for the survey are given in Table 1. Three levels were chosen for each attribute to allow the estimation of nonlinear (quadratic) effects. For the quantitative variable, time, a base value of 12 min was chosen, roughly the median of the respondents' reported one-way travel times. The high and low levels were defined as 15 and 9 min, a range equal to half the base level and encompassing the majority of self-reported times. This range was considered large enough to be perceivable in actual choice situations but not so great as to overshadow the influence of changes in the qualitative attributes.

The levels for facility type and surface quality were defined as commonly encountered types of cycle network construction. Though the definitions were made as mutually independent as possible, there is bound to be some correlation in the perception of these two attributes (i.e., separate bicycle paths are most likely to have an adequate surface). The traffic-level factor was the most difficult to define, as can be seen from Table 1 (a translation from the Dutch survey).

Two variations were included in the survey to test the influence of factor and level presentation on the perception of the qualitative variables and on the ease of comparing hypothetical routes. One-half of the sample was given photographs portraying each of the levels for these factors in addition to the normal verbal descriptions. An overlapping half of the sample was asked to classify their own home-to-work route according to the factor levels that best characterized its major portions, using the given verbal and (in some cases) pictorial descriptions.

PRIMARY EXPERIMENTAL DESIGN

In contrast to the trade-off matrices approach used by Morisugi et al. (13), which presents combinations of pairs of attributes, all others held constant, full profiles were used here; that is, choice sets were presented with alternatives varying across all attributes. The full-profile approach has proven more understandable in practice and more stable if there are significant interactions between variables (15).

To allow the estimation of the independent effect of each attribute, the factors and levels in the route-choice scenarios were arranged in an orthogonal design. To estimate all main effects and all interactions, evaluations of all 3^4 = 81 possible route configurations (treatments) would be required. To limit the size and difficulty of the experiment, the analysis of the trade-offs between route factors was performed at two levels of aggregation. At the individual level a simple piecewise linear main-effects model was assumed where the unobserved error term has the same distribution across all routes:

$$U_i(x) = \sum_{k=1}^{K} \sum_{j=1}^{m_k} (\alpha_{ikj} \cdot x_{kj}) + \varepsilon \qquad (1)$$

where

$U_i(x)$ = overall utility or preference measure given to an alternative by individual i,

TABLE 1 Descriptions of Factors and Levels Given

FACTOR	LEVEL	GIVEN VERBAL DESCRIPTION
Facility Type	Physically Separated	This portion of the roadway is meant only for bicycles and mopeds, and is totally separated from other traffic by curbing or plantings. Pedestrians are also provided with a separate walkway.
	Reserved On-Street	This is a full lane of the roadway, reserved for bicyclists and marked with a white stripe on the surface. Now and then there are autos parked on this lane.
	Non-Existent	There is no separate space for bicyclists on the street. They must ride in the same lanes as other traffic.
Surface Quality	Smooth	The surface over the whole route is good asphalt with no large cracks.
	Moderate	In this case, about half of the route is asphalt and the other half brick. There are occasional bumps and cracks.
	Rough	This route has a brick surface from beginning to end with bumps and cracks which one must try to avoid.
Traffic Level	Light	A car or bicycle comes along this route from time to time. There are few crossing pedestrians. Cyclists can easily ride next to each other. You never have to stop for others.
	Moderate	Quite a few cars use this route, but this is not a hinderance to cyclists. You can still ride along with other cyclists when it does not present a problem.
	Heavy	Many cyclists and autos ride along this route. It is very busy here. It is often difficult to enter or cross. You must also wait for other traffic and crossing pedestrians. It is not possible to ride next to other cyclists.
Travel Time		This refers to the total time that you are travelling, including all of the delays encountered on the route. The average travel times used in this survey are:
	Short	- 9 minutes,
	Medium	-12 minutes, or
	Long	-15 minutes.

K = number of attributes of the alternative,
m_k = number of levels of attribute k,
α_{ikj} = the partworth contribution of level j of attribute k to individual i,
x_{kj} = presence or absence of level j of attribute k, and
ε = error term.

With this type of model only a one-ninth fractional factorial design (nine orthogonal alternatives) must be included in each survey to estimate the main effects for each individual. Because individual-level estimates were desired mainly for market segmentation rather than for strict tests of functional form, the simple design was deemed adequate.

Because certain interactions were thought to be potentially important for the qualitative factors, a more extensive block design was used to allow their estimation. To this end, three blocks of nine routes were designed, each block being internally orthogonal. The blocks were distributed evenly across the sample and across other survey variations. Together, the three main-effects designs form a one-third fractional design [see Master Plan 8 by Kocur (10)], allowing aggregate estimation of the two-way interaction between each pair of qualitative variables (although certain interactions may not be separable from other two-way and higher-order terms).

At the aggregate level, therefore, models were specified consisting of main effects and selected two-way interaction terms:

$$U(x) = \sum_{k=1}^{K} \sum_{j=1}^{m_k} [(\alpha_{kj} \cdot x_{kj}) + \sum_{h=1}^{K-1} \sum_{l=1}^{m_k} (\beta_{h,kj} \cdot x_{kj} \cdot x_{hl})] + \varepsilon \tag{2}$$

for specified kj and hl.

In order to minimize the loss of variation through aggregation, sample segments analyzed with aggregate models were kept as homogeneous as possible in terms of preferences. The individual models were used to guide segmentation, particularly where the aggregate model included only the same main effects. With respect to the statistical specification of the aggregate model, it should be noted that in contrast to the individual model, the error terms presumably will not be identically and independently distributed across all observations. Especially because of the repeated-measurement type of observations, the error terms will tend to be more highly correlated for repeated observations within individuals than for observations across individuals. Assuming independence within individuals will not bias the estimates but may lead to an underestimate of their standard errors, as discussed later.

RESPONSE MEASUREMENT SCALES

In the survey, three response scales were attempted:

- A verbally defined seven-point scale assigning a value to the strength of preference (e.g., "always choose option A," "slightly prefer option B," "no preference"),
- An extension of this scale to a continuous one on which the percentage probability of choosing each option can be indicated, and
- Ranking of the options in order of preference.

For the presentation of the rating scale and percentage score methods, the set of alternatives was transformed into sets of route pairs with every alternative placed against a common base route. This base route was chosen to have the middle level of all four factors. Use of this pairwise format rests on an assumption that responses from partial choice sets are transitive across the remaining choice pairs.

Each respondent was asked to complete each of the three response tasks, which, for sake of comparison, were offered in a fixed order. The first nine routes to be rated were presented one at a time, all against a common base route. The second set of routes, presented on separate cards, had to be ranked. This method uses the full choice set of nine alternatives. The final set was presented in the same pairwise manner as the first set, but now the respondent had to indicate the percentage of times he would choose each alternative route.

To ensure that the three blocks of routes would be well distributed across all three response methods and that subjects would not simply transfer preferences from one scoring method to another, each subject was given all three blocks of nine routes, randomly assigned to one of the six possible permutations of the three sets.

GRAPHICAL ANALYSIS OF AVERAGE SCORES

An overall picture of the relative importance of the choice factors and factor levels and of nonlinearities and interactions can be gained from a graphical plot of the average responses. In Figure 1, the average (utility) scores for the rating scale are plotted according to the attributes of the routes they represent. Because the chosen design accounted for interactions between facility type and the other factors, three separate plots are presented for each of these combinations of two factors against the average scores. If these interactions were not present, each plot would show parallel lines for the three facility types. The relative importance of each of the factors is shown by the slopes of the lines connecting the factor levels or can be approximated from the range of the average scores of the extreme levels. Changes in travel time in the chosen range appear to have the greatest linear effects, directly followed by surface quality, which appears slightly less important. Changes in traffic level and facility type clearly have less influence on the average score values. All the main effects have the expected direction (sign) and are more or less linear, with the exception of traffic level. Certain interaction terms appear to be present: a smooth surface has the least effect when a separate bicycle path is concerned and a rough surface has the smallest effect when no cycling facility is present. These results suggest that the facility type itself has implications regarding surface quality. The same type of interaction is suggested by the reduced influence of light traffic when combined with physically separated bikeways. It is probable that these are not true interactions in the behavioral sense but are due to an interaction in the perception of these attributes from the survey presentation. At this level of average scores, roughly the same relationships were discovered with both of the other response methods (ranks and percentages).

ESTIMATION OF PREFERENCE FUNCTIONS

Estimation Approaches

The metrically scaled data were analyzed by using ordinary (dummy) least-squares regression. To estimate the partworth utilities in Model 1 only $\sum_{k=1}^{K}(m_k - 1)$ linearly independent variables are needed to completely specify the preference model. Therefore, each attribute with m_k-levels is converted into $(m_k - 1)$ dummy variables, where the omitted level serves as a reference (16). In this case, the base route in the paired comparisons, which combines the middle levels of all four factors, was taken for reference. The original Model 1 was then estimated as follows:

$$U(x) = \beta_0 + \sum_{k=1}^{K} \sum_{j=1}^{m_k-1} (\beta_{kj} \cdot x_{kj}) + \varepsilon \quad (3)$$

where

β_0 = the utility for the choice alternative, which has been coded zero for all attributes (base route);
β_{kj} = the differential partworth utility of level j of attribute k, which is the difference in utility between each attribute and the reference; and
x_{kj} = 1 if level j of the kth attribute is present in a choice alternative; 0 otherwise.

At the individual level the main-effects model could be estimated algebraically because there are nine orthogonal equations (observations) to determine the same number of unknown parameters.

At the aggregate level ordinary least-squares regression (OLS) was applied. For efficient regression estimates, the error terms ε must be assumed to be independent and identically distributed across route alternatives and individuals. Because all alternatives represent the same generic choice (a route), this common assumption was considered reasonable for the application-oriented nature of the study. Yet it must be recognized that the error terms will tend to be more highly correlated for repeated observations within individuals than for observations across individuals and that each individual's taste for certain types of routes may vary along factors not controlled for in the design. As mentioned previously, this will cause the standard errors given by the OLS procedure to be biased downward. A conservative adjustment is to assume that the error terms are completely correlated within individuals, adjust the standard errors, and check for significance. If doubt remains, one can use more complex generalized linear regression (GLS) methods.

The ranked data were exploited for estimation by using a so-called "exploded" logit analysis (17). This is a procedure for exploiting the information of ranked choice sets to estimate the parameters of

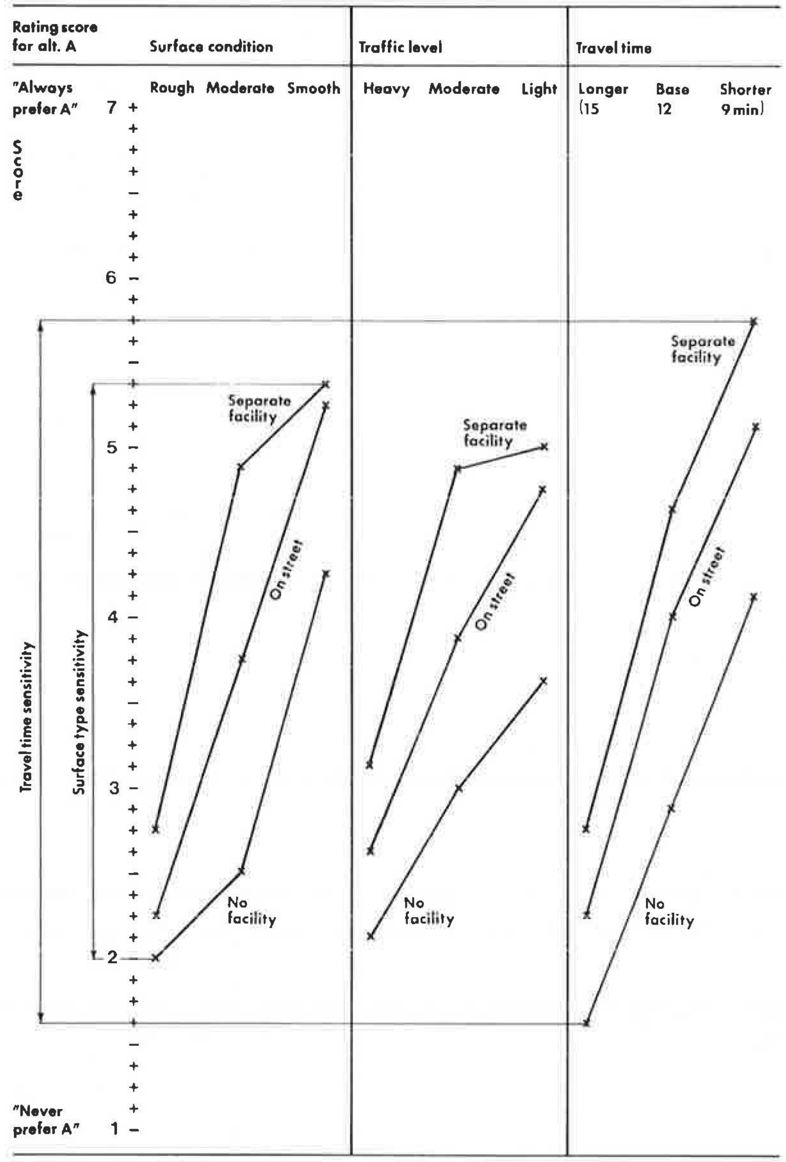

FIGURE 1 Graphical analysis of average rating scores.

the utility function in the multinominal logit model. The "explosion" means the decomposition of a single ranked choice set into a series of unranked and statistically independent choice sets. Each separate ranking can be treated as being chosen over all alternatives that rank equally below it. The choice model to be estimated thus has the following form:

$$P_i\{n \mid n \in N \subset M\} = \exp\{U_i^n(x)\} \Big/ \sum_{n=1}^{N} \exp\{U_i^n(x)\} \qquad (4)$$

where $P_i(n)$ is the probability that individual i ranks alternative $n \in N$ highest in subset M of the choice set L and U_i^n are the utility functions as, for example, in Equation 3.

The parameters are estimated by using maximum-likelihood techniques. This exploded logit approach requires that two important assumptions be made: first, the validity of the I.I.A. property, which means that the utility of an alternative is not influenced by whatever other alternatives are in the choice set, and second, as with the metrically scaled data regression analysis, the independence and identical distribution across all alternatives and individuals of the error terms. The logit approach was used as a comparison to the OLS approach, though only as a first step toward a more appropriate but more complicated procedure, including tests of the I.I.A. property.

Estimation Results with Different Approaches

Table 2 contains the results of the most aggregate main-effects model estimated for the entire sample by using each type of response data and the maximum number of observations possible (only the top six ranks were chosen for the logit model, however). To facilitate comparison of the parameter values of the different models as well as for better assessment of the trade-offs, all parameters are also expressed relative to the travel-time values as minute equivalents. Overall, the estimation results look plausible and consistent. For the regression models, the constant terms (grand mean) were not significant and were very small, and they are not reported. All

TABLE 2 Aggregate Models for Each Response Data Type

response method	scores (1-7)		percent (0-100)		ranked (1-6)	
individuals	119		114		121	
observations	1071		1026		741	
R-squared (RHO-squared)	.56		.56		(.23)	
	OLS coefficients		OLS coefficients		LOGIT coefficients	
	raw (T-stat)[b]	minutes[a] (range)	raw (T-stat)[b]	minutes[a] (range)	raw (T-stat)[b]	minutes[a] (range)
1. no	-0.96	-2.0	12.1	-1.5	-0.65	-1.8
facility	(7.9)		(6.2)		(6.0)	
		(.51)		(.35)		(.52)
2. separate	0.48		5.0		0.53	
path	(4.0)	1.0	(2.5)	0.6	(5.2)	1.4
3. rough	-1.44	-3.1	-21.4	-2.7	-1.22	-3.3
surface	(11.8)		(10.7)		(10.5)	
		(.92)		(.87)		(.90)
4. smooth	1.15		20.6		0.80	
surface	(9.5)	2.4	(10.3)	2.6	(7.9)	2.2
5. heavy	-1.23	-2.6	-16.1	-2.0	-0.95	-2.6
traffic	(10.1)		(7.9)		(8.2)	
		(.62)		(.49)		(.51)
6. light	0.52		7.8		0.20	
traffic	(4.2)	1.1	(3.8)	1.0	(1.9)	0.6
7. longer	-1.66	---	-25.7	---	-1.42	---
time	(13.4)	6.0	(12.5)	6.0	(11.8)	6.0
		(1.0)		(1.0)		(1.0)
8. shorter	1.15		22.6		0.82	
time	(9.3)	---	(11.1)	---	(8.0)	---

a) Coefficients per minute travel time are normalized using the design range of 6 minutes between the time levels, and the estimated time coefficients. The range is the difference between the coefficients for the extreme levels of each factor, normalized to the range of the travel time factor.

b) t-values based on independence across all observations. Very conservative estimates result by dividing with $\sqrt{9}$, 9 being the number of designed responses per subjects (see also (18)).

other estimates appear statistically significant; longer travel time shows the most precise and separate facility the least precise coefficients. [Even when complete dependence between observations within subjects is assumed, by dividing t-values with $(9)^{1/2}$, most parameter estimates remain significant.] All models indicate, by the range between the normalized effects of the extreme levels, that travel time is most important (in the given time ranges), directly followed by surface type. From the models it can be observed, for example, that for an average trip length of 9 min, an improvement from "no facility" to "separate path" will compensate for a travel time loss of about 3 min. An improvement from a rough to a middle-quality surface also can compensate for a travel time detour of 3 min. These trade-off values suggest a fairly high sensitivity of bicycle travelers toward changes in qualitative route factors.

The three data types yield similar models in terms of the explained variation and the relative importance of the factors. Models that included the interactions identified in Figure 1 showed them to be marginally significant [for details see report by Bovy et al. (19)].

Estimation Results with Various Segmentations

Various segmentations were attempted to improve the explanatory and predictive power of the models. First, the respondents were grouped according to their individual coefficients for time, as estimated from their scores on the rating scale. Table 3 provides evidence that this type of clustering is effective, because the rho-squared values for internal segmentation increase noticeably. The coefficients for both time levels are significantly higher for the time-sensitive group. The other group shows significantly higher disincentives from rough surface, heavy traffic, and no facility. This appears to be the comfort-sensitive segment. Whereas the time-sensitive bicyclists are willing to spend only 1.5 min of extra time to use a route that has a separate bicycle path instead of no facility, the comfort-sensitive segment appears willing to accept a detour in such a case of more than 6 min (on an average trip length of 9 min).

A second segmentation was done compositionally, according to age. The travel time coefficients for the respondents under 40 are higher than those of the over-40 segment, whereas the coefficients of the other variables are not significantly different between both segments.

The corresponding time-valued figures show that the older cyclists are willing to sacrifice much more travel time for better route quality than the younger riders. Although this particular external segmentation is less effective in improving the model, groupings of this type are more useful in application, and the composition of the internal segments can help to define more effective groupings.

TABLE 3 Estimated Coefficients and Validation Results for Segmentations (Logit Model, Top Three Ranks)

TYPE OF SEGMENTATION	NONE		INTERNAL: HIGH WEIGHT ON TIME		INTERNAL: LOW WEIGHT ON TIME		EXTERNAL: UNDER 40 YEARS		EXTERNAL: OVER 40 YEARS	
OBSERVATIONS	374		208		166		179		195	
RHO-SQUARED	.29		.40		.34		.34		.27	
	RAW (T-STAT)[b]	NORM[a] *(RANGE)*	RAW (T-STAT)[b]	NORM[a] *(RANGE)*	RAW (T-STAT)[b]	NORM[a] *(RANGE)*	RAW (T-STAT)[b]	NORM[a] *(RANGE)*	RAW (T-STAT)[b]	NORM[a] *(RANGE)*
1. NO FACILITY	-0.64 (4.0)	*-1.27* *(0.49)*	-0.36 (1.5)	*-0.44* *(0.24)*	-1.07 (4.2)	*-3.07* *(1.09)*	-0.52 (2.1)	*-0.77* *(0.32)*	-0.79 (3.6)	*-1.87* *(0.66)*
2. SEPARATE PATH	0.83 (5.6)	*1.64*	0.82 (3.2)	*1.00*	1.21 (5.7)	*3.48*	0.79 (3.4)	*1.17*	0.89 (4.6)	*2.11*
3. ROUGH SURFACE	-1.72 (6.6)	*-3.41* *(0.85)*	-1.02 (2.2)	*-1.25* *(0.52)*	-2.31 (6.0)	*-6.64* *(1.71)*	-1.67 (3.5)	*-2.48* *(0.66)*	-1.80 (5.5)	*-4.27* *(1.08)*
4. SMOOTH SURFACE	0.94 (6.6)	*1.86*	1.54 (4.6)	*1.89*	1.27 (5.7)	*3.65*	0.99 (3.3)	*1.47*	0.93 (4.6)	*2.20*
5. HEAVY TRAFFIC	-1.15 (6.1)	*-2.28* *(0.55)*	-0.87 (2.6)	*-1.07* *(0.20)*	-1.57 (5.8)	*-4.51* *(1.11)*	-1.23 (3.8)	*-1.83* *(0.43)*	-1.13 (4.7)	*-2.68* *(0.67)*
6. LIGHT TRAFFIC	0.52 (2.9)	*1.03*	0.12 (0.4)	*0.15*	0.76 (3.2)	*2.18*	0.49 (1.6)	*0.73*	0.56 (2.5)	*1.33*
7. LONGER TIME	-1.30 (6.0)	- 6.0 *(1.0)*	-2.01 (4.6)	- 6.0 *(1.0)*	-1.25 (4.6)	- 6.0 *(1.0)*	-1.96 (4.4)	- 6.0 *(1.0)*	-1.04 (4.0)	- 6.0 *(1.0)*
8. SHORTER TIME	1.73 (9.8)	-	2.87 (9.0)	-	0.84 (3.5)	-	2.08 (6.8)	-	1.49 (6.6)	-

VALIDATION RESULTS	NONE	INTERNAL	EXTERNAL
ORIGINAL SAMPLE : 1107 RANKINGS (123 SUBJECTS)			
SPEARMAN RANK CORRELATION :	.745	.778	.750
HOLD BACK SAMPLE: 99 RANKINGS (11 SUBJECTS)			
SPEARMAN RANK CORRELATION :	.746	.768	.753

The predictive improvement from the segmentations can be checked by examining the correlation between the actual and predicted rankings. Table 3 shows a small increase in the Spearman correlation from the clustering segmentation but only a marginal improvement from segmentation by age. Most of this improvement is likely in the top few rankings, on which the models are based and which are most important for forecasting.

Validity and Reliability Tests

The true validity of these models in explaining behavior can be judged only with respect to independent data on observed choices. There were, however, several steps taken to increase confidence in the results. In the preceding sections the robustness of the results to differences in composition of choice sets, response method, and type of estimation has been shown. On the other hand, the modeling results appeared to be clearly sensitive to segmentation of the subjects.

The internal validity of the models was tested by using the unsegmented and segmented models to predict rankings for a small hold-back sample of 11 individuals. Even with this limited sample, the recovered rankings are at a level almost identical to that for the original sample (bottom of Table 3). The external validity of the model can only be tested with an independent source of data. Unfortunately, limited data exist for bicycle route choice.

External validity can be judged somewhat through comparison with the results of similar studies. In a concurrent stated-preference analysis in Wisconsin, Axhausen (2) identified similar trade-offs among the same four attributes (by using distance instead of time). He also found slope, land use, and cycling experience to be influential. These factors, however, exhibit much less variation in the Netherlands and were assumed not to affect the route choice of most cyclists. In another stated-preference experiment (9), also in Wisconsin but this time for automobile versus bicycle choice, time, surface, traffic, and facility were shown to have roughly the same relative influence as that reported in the preceding discussion.

A further test dealt with the method of presenting the qualitative factor levels, an aspect that appears of crucial importance in the context of route choice. Four different methods of presentation were included in the experimental design and were randomly distributed among the subjects: a purely verbal description of each factor level, the same verbal description illustrated with photographs, the verbal description including an exercise to use this for a categorization of the actual home-to-work route, and all three elements combined. Table 4 shows the relative factor utilities normalized with respect to time.

The breakdown indicates that because the qualitative factors were portrayed more clearly, they were generally given more importance relative to travel time and that the inclusion of photographs appears to have helped clarify these factors to a greater extent than asking subjects to categorize their actual route.

The results also indicate that using all presentation approaches in combination was generally no more effective than using either one separately. If the "all methods" category can be assumed to contain the most informed responses, it appears that the use of photographs (column 2) without some relation to the actual routes (columns 3 and 4) may make qualitative factors appear more homogeneous and more important than in actual choice contexts.

Finally, after the scenario comparisons, each respondent was asked to assess the importance of the four attributes in his daily route choice, the clarity of these attributes in the survey presentation, and the overall difficulty of relating the scenarios to his own choice situations. A preliminary analysis of these evaluations showed that the choice factors were generally well understood and considered important in choosing a route. This information also supported the relative importance of the factors found in the analysis.

Over three-fourths of the subjects reported little or no difficulty in comparing hypothetical choices. Of the individual factors, only traffic level appeared to present some difficulty in comprehension. This is the factor that encompasses the widest variety of physical attributes and is likely to be quite variable over an actual route. Interestingly, those who were asked to classify their own route in these terms reported a less clear understanding of this factor. Perhaps the task brought out the inconsistency between simplified and actual routes. Those who were given photographs, on the other hand, reported less trouble understanding the attributes.

As for the different scoring methods, the respondents found the ranking of routes with cards to be easiest and the verbal scale slightly more difficult. The percentage scale was reported as decidedly the most difficult.

SUMMARY AND CONCLUSIONS

By itself, the stated-preference approach appears to give stable estimates of the trade-offs among specific choice factors in a specific context at a cost usually much lower than that of alternative methods. The information on trade-offs within different market segments is useful in ranking alternative policies and in identifying advantaged and disadvantaged groups of users affected by those policies. The differential weights placed by bicyclists on various route factors can help planners in designing bicycle facilities. The outcomes also suggest that extending current minimum-time traffic assignment models with other route factors should be seriously considered for applications in bicycle traffic.

The magnitude of the estimated coefficients suggests a fairly high sensitivity of bicyclists toward the chosen route factors, a result that might be partly due to the specific nature of a stated-preference survey. By showing that the differences in attributes between alternatives vary explicitly, the subject's responses presumably are far more differential than when he is confronted with changes in real alternatives. For more detailed policy analyses it is suggested therefore that the stated-preference models be validated (and probably scaled down) by calibrating them with actual choice data as far as possible. The findings also indicate that the method of presenting the choice context and attributes can have a significant effect on estimates of individuals' trade-offs. This difficulty is essentially what distinguishes this analysis from most other contexts in which scenario-based analyses have been performed successfully. The use of photographs and maps appears to be a useful aid in understanding the experiment, but further research into methods of defining and presenting qualitative route factors should have priority in extending these techniques to a wider range of route choice contexts.

TABLE 4 Relative Linear Utilities of the Qualitative Factors Normalized to Time by Presentation Subgroup[a]

Factor	verbal alone n=22	verbal + photo's n=35	verbal + own route n=37	all methods n=34	average n=128
	1	2	3	4	5
Facility type	2,8	7,1	3,2	3,5	4,5
Surface type	4,0	8,5	6,4	6,8	6,6
Traffic level	3,5	6,4	5,9	3,5	5,0

[a]These utilities are based on averages from the individual-level models using the rating scale data

In terms of measuring preferences, ranking of the routes appears to be easiest for the respondents and most comparable to revealed-preference data. Metric scales, on the other hand, allow simpler regression analysis and provide more information at an individual level. Before such scales are used extensively in mail-out experiments, however, it appears that improvement and simplification of the scale presentation and grouping of alternatives are necessary. The percentage scale is not recommended: it is the most difficult for the subjects and can easily lead to response errors. The estimation of the models appears robust with both regression and discrete choice methods. Apart from these methodological issues, further work should be done to study additional route choice factors. For example, qualitative route factors such as safety, variability, and signposting are important for many policy areas and could be incorporated into a stated-preference analysis.

ACKNOWLEDGMENTS

The research presented in this paper was partially funded by the Dutch Ministry of Transport and Public Works. Thanks are due to Dick den Adel of the Delft Institute for Town Planning Research for doing all of the survey production and administration and to Hugh Gunn and Ben Jansen for their constructive criticisms. The authors wish to thank the anonymous referees for their very helpful comments and criticisms.

REFERENCES

1. N.H. Anderson. Foundations of Information Integration Theory. Academic Press, New York, 1981.
2. K.W. Axhausen. Bicyclists Evaluate Their Environment: Some Results. M.Sc. thesis. Department of Civil and Environmental Engineering, University of Wisconsin, Madison, 1984.
3. S.M. Carpenter. Drivers' Route Choice Project: Pilot Study. Paper 085/WP42. Transport Studies Unit, Oxford University, Oxford, England, 1977.
4. R.D. Huchingson et al. Survey of Motorist Route-Selection Criteria. In Transportation Research Record 643, TRB, National Research Council, Washington, D.C., 1977, pp. 45-48.
5. M. Wachs. Relationships Between Drivers' Attitudes Toward Alternate Routes and Driver and Route Characteristics. In Highway Research Record 197, HRB, National Research Council, Washington, D.C., 1967, pp. 770-787.
6. R. Hamerslag. Investigation into Factors Affecting the Route Choice in Rijnstreek-West with the Aid of a Disaggregate Logit Model. Transportation, Vol. 10, 1981, pp. 373-391.
7. M. Ben Akiva et al. Modelling Interurban Route Choice Behavior. In Proceedings of the Ninth International Symposium on Transportation and Traffic Theory (J. Volmuller and R. Hamerslag, eds.), VNU Science Press, Utrecht, Netherlands, 1984.
8. J.J. Louviere et al. Laboratory-Simulation Versus Revealed-Preference Methods for Estimating Travel Demand Models. In Transportation Research Record 794, TRB, National Research Council, Washington, D.C., 1981, pp. 42-51.
9. Development of Wisconsin Work Trip Models for Forecasting Modal Choice. Division of Planning and Budget, Wisconsin Department of Transportation, Madison, 1981.
10. G. Kocur et al. Guide for Forecasting Travel Demand with Direct Utility Assessment. UMTA, U.S. Department of Transportation, 1982.
11. J. Benjamin and L. Sen. A Manual of Procedures to Analyze Attitudes Towards Transportation. U.S. Department of Transportation, 1982.
12. S.R. Lerman and J.J. Louviere. Using Functional Measurement to Identify the Form of Utility Functions in Travel Demand Models. In Transportation Research Record 673, TRB, National Research Council, Washington, D.C., 1978, pp. 78-86.
13. H. Morisugi et al. Measurement of Road User Benefits by Means of Multiattribute Utility Function. Papers, Regional Science Association, Vol. 46, 1981, pp. 31-43.
14. W.F. Upcott. Cyclists' Route Choice: A Behavioural Analysis of the Journey to School. M.Sc. thesis. Cranfield Institute of Technology, Cranfield, Bedfordshire, England, 1976.
15. M.N. Segal. Reliability of Conjoint Analysis: Contrasting Data Collection Procedures. Journal of Marketing Research, Vol. 19, 1982, pp. 139-144.
16. H. Timmermans. Decomposition Multiattribute Preference Models in Spatial Choice Analysis: A Review of Some Recent Developments. Progress in Human Geography, Vol. 8, No. 2, 1984, pp. 189-221.
17. R.G. Chapman and R. Staelin. Exploiting Rank Ordered Choice Set Data Within the Stochastic Utility Model. Journal of Marketing Research, Vol. 19, 1982, pp. 288-301.
18. J.J. Louviere and G. Woodworth. Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data. Journal of Marketing Research, Vol. 20, Nov. 1983, pp. 350-367.
19. P.H.L. Bovy et al. Toepassing van Stated Preference Methoden Voor de Analyse van Fietsroute-Keuzegedrag (in Dutch). Institute for Town Planning Research, Delft University of Technology, Delft, Netherlands, 1984.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.