Transportation Planning and Technology
1978, Vol. 4, p. 219–226

ANALYSIS OF INTENDED BUS USAGE
An Application of the Analysis of Covariance

K. S. KRISHNAN,[†] G. C. NICOLAIDIS[†] and J. N. SHETH[‡]

Transportation and Urban Analysis Department, General Motors Research Laboratories,
Warren, Michigan 48090 (U.S.A.)

(Received November 3, 1976; in final form July 7, 1977)

In order to plan bus operations, it is necessary for transit planners to understand what factors may influence travelers' choice of buses for travels within a city. The proposed method involves various scenarios of a hypothetical bus operation which was rated by a group of individuals.

Analysis of Covariance technique is employed to analyze people's sensitivities to their perceived levels of bus service characteristics. The technique involves:

1) testing for the significant effects of varying levels of service characteristics upon people's intentions to use bus service, and

2) assessing differences among various population segments in their sensitivity patterns towards bus service characteristics.

Results from the application of the technique to attitudinal data collected by the Orange County Transit District indicate that bus service characteristics do influence, independently and jointly, respondents' stated intentions to use buses.

Sensitivity pattern differed across the five homogeneous segments identified in an earlier research based on socioeconomic characteristics.

One segment (an older, predominantly male population segment with higher home ownership level and lower income than the rest of the sample) was relatively insensitive to changes in bus fare and was influenced by changes in headway independent of changes in access distance. Another segment consisting of fewer registered voters with lower education also exhibited similar independent impact of headway and access distance.

The technique is especially useful in reducing a large number of proposed alternative bus systems to a smaller set for further planning considerations by specifying the ranges within which variation of service characteristic would cause substantial changes in the intended usage responses.

INTRODUCTION

In order to plan successful bus operations, it is necessary to know what factors influence people's choices of buses for their travels. In particular, transit planners need to know potential users' travel habits and preferences and their sensitivities to varying levels of perceived service characteristics of a bus system.

If people's past travel behavior is not known, necessary information may be obtained either by direct experimentation or by simulation. Such information could be used for preliminary analysis of people's evaluations of alternative bus systems. Typically, experimentation is very expensive and hence is seldom adopted for preliminary evaluations. Simulation of analytic models of mode-choice has been used extensively in travel demand studies. Reviewing the mode-choice literature over the last decade, Louviere, Beavers, Norman and Stetzer concluded that despite their widespread applications, mode choice models do not adequately replicate individuals' actual choices.[1] Consequently, Louviere *et al.* recommended that future research should focus on simulating selective travel characteristics so as to assess people's sensitivities to the travel characteristics as a first step toward modeling the mode-choice behavior. Following their own recommendations, Louviere *et al.* presented results of a simulation which makes use of Andersons' Integration Theory principles.[2]

The research reported herein is a further step in this recommended direction, and employs simu-

---

[†] Transportation and Urban Analysis Department, General Motors Research Laboratories.

[‡] School of Business Administration, University of Illinois, Urbana—Champaign, Illinois, and Consultant to the General Motors Research Laboratories.

lation to describe how individuals perceive changes in specific service characteristics. The method involves various scenarios of a hypothetical bus system presented to a sample of individuals in a home interview survey. Each scenario is specified by three bus service characteristics: bus fare, headway and access distance. Each survey respondent was requested to evaluate several bus scenarios and project their intentions to use buses for intra-urban travels if such bus systems were presently available. Statistical analyses of these intentions are carried out to identify significant influences of the service characteristics. Although intentions to use buses do not imply individuals' commitment to use buses, they indicate the respondents' sensitivities toward the perceived levels of the three service characteristics.

The scenario approach investigated in this research is similar to the one reported by Louviere *et al.* (1973) with a group of students from the University of Iowa.[1] Louviere *et al.* presented each student with a variety of bus scenarios. A 3 x 3 x 3 factorial design was adopted and the students' responses were analyzed using Analysis of Variance (ANOVA). The ANOVA results led Louviere *et al.* to theorize alternative response functions and postulate the behavioral mechanism.

The present research differs from Louviere *et al's* study in at least two respects. First, this research considers a larger number of scenarios than the Louviere's study. The particular survey design adopted by the OCTD† officials necessitates the use of a covariate analysis rather than the analysis of variance. Secondly, this research is complementary to an earlier work reported by Nicolaidis and Sheth who identified five homogeneous groups of respondents as a part of market segmentation study.[3] The present research is concerned with investigating whether these five groups of respondents do indeed differ in their sensitivities to perceived levels of the three service characteristics. Such a knowledge is essential for designing bus systems tailored to the needs and sensitivities of one or more homogeneous groups.

The research reported herein postulates that people's preference for a bus system is based on their perceived levels of each of several service characteristics such as headway time, bus fare and access distance. This technique differs from the one devised by the authors (Nicolaidis and Krishman) for analyzing people's intentions under different postulate.[4] There the authors postulate that people would consider each bus operatio scenario as a single stimulus and express their preference for the stimulus without explicity evaluating each service characteristic as done in thi research.

Although results obtained from the simulatio can be used for predictions of mode-choices, w regard the procedure merely as a prelude t developing an analytic model of mode-choic behavior. Results from this study could be used t identify attributes which are salient in consumers choices and determine the perceived range fo each attribute within which the respondent is relatively indifferent. When data on people's choice of modes are available, such salient attributes may also be gleaned through revealed preferences However, when planning for a new system, pas data are not available. In the latter case, the simulation of hypothetical scenarios is perhaps the only reasonable recourse available to a transit planner.

DATA

The data used in this study were obtained from a random sample of 1804 households through a home interview survey administered in Orange County, California, by the Orange County Transit District.

The section of the survey which elicited respondents' intentions to use buses under various scenarios of bus operations is the main data source for this study. The OCTD officials adopted the following sampling design for the survey: 54 scenarios were presented to the respondents, each scenario being specified by a combination of bus fare, headway and distance from home to nearest bus station. These 54 combinations corresponded to six levels of bus fare (0, 25, 35, 50, 75 and 100 cents), three levels of headway (15, 30 and 60 minutes) and three levels of distance from home to bus station (1, 3 and 5 blocks). Respondents were selected in such a way as to form six equal-sized groups, each group consisting of respondents who are located at approximately the same distance from home to an existing bus station. Each group of respondents was administered a different set of nine scenarios which were randomly selected from the 54 scenarios. The reason for selecting only nine out of 54 scenarios is to avoid possible response

---

† Orange County Transit District.

ıtions under
tors postulate
bus operatio
press their pr
plicity evalua
; done in th

the simulatio
de-choices, w
a prelude t
f mode-choic
ould be used t
l in consumer
ived range fo
ondent is rela
eople's choice
attributes ma
d preferences
w system, pas
case, the simu
erhaps the onl
ransit planner.

otained from a
lds through a
ed in Orange
County Transit

licited respon-
under various
ut data source
s adopted the
e survey: 54
ondents, each
ination of bus
ome to nearest
corresponded
0, 75 and 100
.5, 30 and 60
from home to
pondents were
ix equal-sized
pondents who
same distance
n. Each group
different set of
selected from
cting only nine
sible response

---

biases caused by fatigue or boredom from evaluating too many scenarios. A page from the questionnaire eliciting the intended bus usage is reproduced as Figure 1. The nine scenarios presented to each of the six groups are shown in Table I.

METHOD

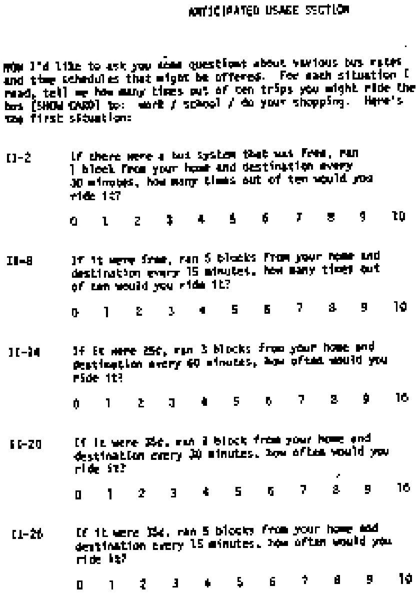
For each combination of bus fare, headway and distance from home to bus station (access distance), the respondents indicated the number of times out of ten that they would choose bus for all travels.

The responses could be interpreted as being either the subjective probability of using a bus (e.g., six out of ten means 60 percent probability of use) or the expected fraction of times a respondent would use a bus (e.g., 60 percent of all future travel will be made by bus).

It is hypothesized that the respondents' stated intentions could be expressed as the average response by all individuals plus main effects due to bus fare (B), headway (H) and access distance (D), and the interaction effects due to interplay of two or more factors. This hypothesis may be expressed as follows:

$$Y_{ijkm} = \mu + H_i + D_j + B_k + (HD)_{ij} + (DB)_{jk} + (HB)_{ik} + (HDB)_{ijk} + e_{ijkm} \qquad (1)$$

where Y_{ijkm} denotes intended bus usage of mth individual when the headway is at the ith level, the access distance at the jth level, and the bus fare at the kth level. The first term μ is the average response of all individuals. The terms H_i, D_j and B_k denote the three main effects when the factors are at the levels of their respective subscripts. The next three terms $(HD)_{ij}$, $(DB)_{jk}$ and $(HB)_{ik}$ denote the two-factor interaction terms while $(HBD)_{ijk}$ denotes the interaction effect of all three factors. The last term e_{ijkm} is a random error term.



ANTICIPATED USAGE SECTION

Now I'd like to ask you some questions about various bus rates and time schedules that might be offered. For each situation I read, tell me how many times out of ten trips you might ride the bus [SHOW CARD] to: work / school / do your shopping. Here's the first situation:

I1-2 If there were a bus system that was free, ran 1 block from your home and destination every 30 minutes, how many times out of ten would you ride it?

0 1 2 3 4 5 6 7 8 9 10

I1-8 If it were free, ran 5 blocks from your home and destination every 15 minutes, how many times out of ten would you ride it?

0 1 2 3 4 5 6 7 8 9 10

I1-14 If it were 25¢, ran 3 blocks from your home and destination every 60 minutes, how often would you ride it?

0 1 2 3 4 5 6 7 8 9 10

I1-20 If it were 25¢, ran 1 block from your home and destination every 30 minutes, how often would you ride it?

0 1 2 3 4 5 6 7 8 9 10

I1-26 If it were 35¢, ran 5 blocks from your home and destination every 15 minutes, how often would you ride it?

0 1 2 3 4 5 6 7 8 9 10

FIGURE 1 A page from the questionnaire.

TABLE I

Survey design

BLOCKS FROM HOME TO BUS STATION DISTANCE, D	ACTUAL HOME TO BUS STATION DISTANCE IN BLOCKS																	
	1-3			4-7			8-11			12-15			16-20			21 OR MORE		
	HEADWAY TIME IN MINUTES, H																	
	(15)	(30)	(60)	(15)	(30)	(60)	(15)	(30)	(60)	(15)	(30)	(60)	(15)	(30)	(60)	(15)	(30)	(60)
1	B1	--	--	B1	--	--	--	--	B1	B2	--	--	--	B2	--	--	--	B2
3	--	B2	--	--	--	B2	B2	--	--	--	B1	--	--	--	B1	B1	--	--
5	--	--	B1	--	B1	--	--	B1	--	--	--	B2	B2	--	--	--	B2	--

BUS FARES : B1 = 0, 35 AND 75¢
B2 = 25, 50, AND 100¢

From the OCTD survey data, the interaction effects involving bus fare cannot be estimated since for any combination of the three service characteristics only one observation is recorded at each price level, as shown in Table I. Therefore, in order to separate price effects from other effects shown in Eq. (1), both main and interaction effects involving bus fare are represented by a term called covariate, denoted by B which is linearly related to Y_{ijk}. A second covariate, X, is also included in the model so as to capture the influence of current distance between a respondent's home and the nearest bus station. The modified model is then

$$Y_{ijk} = \mu + H_i + D_j + (HD)_{ij} + \gamma B_{ijk} + \delta X_{ijk} + e_{ijk} \qquad (2)$$

where B_{ijk} and X_{ijk} are the two covariates corresponding to the ith level of headway, jth level of access distance for the kth individual in the sample and γ and δ are coefficients of the two covariates.

The main and interaction effects as well as the coefficients γ and δ can be estimated by the least-squares technique which minimizes the sum of squares of the error terms. To test for the significance of the main and interaction effects, the variance of a main or interaction effect is compared with the variance of the error term. If the ratio of the variances, is sufficiently large it is concluded that the factor in question influences intended usage response significantly. Similar tests are performed to test for the significance of a covariate.

The variances are computed from an Analysis of Covariance (ANCOVA) table. ("Covariance" in contrast to "Variance" because of additional information about covariates). A modified multivariate analysis of variance (MANOVA) computer program was used to perform ANCOVA.[5]

The model shown by Eq. (2) was first applied to the total sample in order to compare the results of this study with those obtained by Louviere *et al.* A separate analysis was then performed for each of the five identified homogeneous segments.[6] Figure 2 presents the socioeconomic characteristics of the total sample as well as those of each of the five groups. Each group is significantly different from the total in terms of these socioeconomic characteristics.

RESULTS

The average intended usage responses to the 54 scenarios are shown in Table II. The rows indicate different levels of bus fare and the columns indicate different levels of access distance and headway.

TABLE II

Average intended usage responses

BUS FARE in ¢ B	DISTANCE FROM HOME TO BUS STATION								
	1 BLOCK			3 BLOCKS			5 BL		
	HEADWAY TIME IN MINUTES H								
	(15)	(30)	(60)	(15)	(30)	(60)	(15)	(30)	[illegible]
0	6.11	6.21	5.27	5.76	5.27	4.86	4.49	4.24	2.92
25	5.50	5.43	4.22	4.82	3.83	3.63	3.63	3.36	2.44
35	4.93	4.64	3.56	4.34	3.52	3.27	2.54	2.74	2.16
50	3.54	3.30	2.90	2.45	2.16	2.37	2.38	2.24	1.66
75	2.32	2.27	1.72	2.23	1.80	1.76	1.94	1.49	1.34
100	1.65	1.73	1.40	1.47	1.40	1.42	1.52	1.28	1.24

It is clear from the rows that the average intended usage response decreases as access distance and headway increase. Similarly, the columns indicate that the average intended usage response also decreases as bus fare increases.

Figure 3 shows three plots of the average intended usage response as a function of bus fare when the access distance is one block corresponding to three levels of headway (viz. 15, 30 and 60 minutes) and Figures 4 and 5 show similar plots when the access distance is three and five blocks respectively. All three figures clearly show that the average intended usage response is a decreasing function of bus fare, access distance and headway. However, in Figure 3 the average intended usage response is approximately the same when the headway is 15 or 30 minutes, but is lower when it is 60 minutes. Thus, it appears that when the access distance is one block, people are indifferent to headways of 15 and 30 minutes. Similar examinations of Figures 4 and 5 suggest that when the access distance is three blocks, people would be indifferent to headways of 30 and 60 minutes and that when the access distance is five blocks, they would respond differently to different headways. These apparently dissimilar response patterns lead to the hypothesis that the main as well as the interaction effects of the three service characteristics are significant. The hypothesis is tested using ANCOVA.

The ANCOVA was first employed on the total sample taking bus fare and current distance between home and the nearest bus stop as two covariates. The ANCOVA results in Table III show that for the entire sample varying levels of access distance and headway significantly influenced the intended usage responses. In particular, both main and interaction effects of the two factors were

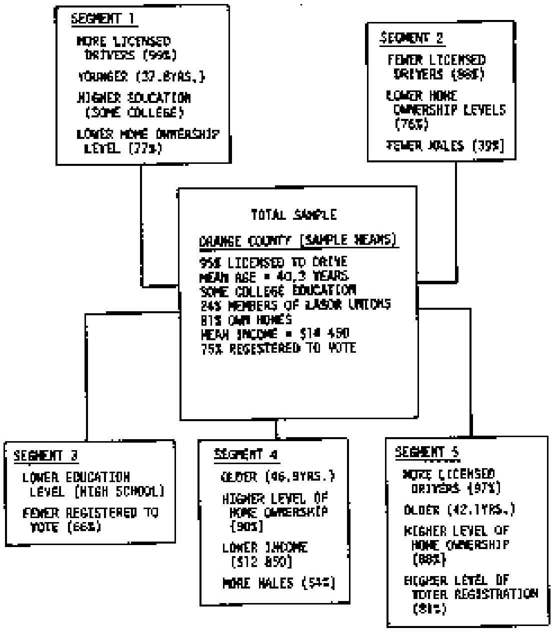


FIGURE 2 Difference among segments—socio economic characteristics.

tatistically significant at 0.05 level, confirming ypotheses generated from the inspection of igures 3, 4 and 5. Since the coefficients of two ovariates were also significant, it was concluded hat bus fare and current distance significantly nfluenced the intended usage response. These esults coincide with those obtained by Louviere *et l.*

ANCOVA results for each of the five homogeneous segments are shown in Tables IV through VIII and summarised in Table IX. The significance of the main and interaction effects is tested by F-values shown in these tables. The last column indicates the level at which the F-value is significant; "NOT SIGNIFICANT" indicates that the F-value is insignificant at five percent level.

TABLE III

Analysis of covariance for total sample

SOURCE OF VARIATION	DEGREES OF FREEDOM	SUM OF SQUARES	MEAN SQUARES	F	SIGNIFICANCE
BUS FARE	1	24369.26	24369.26	2846.66	.001
ACTUAL DISTANCE	1	35.91	35.91	4.20	.001
HEADWAY	2	1822.06	911.03	106.42	.001
ACCESS DISTANCE	2	3810.75	1905.38	222.56	.001
HEADWAY x ACCESS DISTANCE	4	593.45	148.62	18.38	.001
ERROR	16225	138906.38	8.56	-------	----
TOTAL	16235	596017.60	-------	-------	....

TABLE IV

Analysis of covariance, segment 1

SOURCE OF VARIATION	DEGREES OF FREEDOM	SUM OF SQUARES	MEAN SQUARES	F	SIGNIFICANCE
BUS FARE	1	7672.59	7672.59	882.97	.001
ACTUAL DISTANCE	1	1.53	1.53	0.18	NOT SIGNIFICANT
HEADWAY	2	438.75	219.38	25.52	.001
ACCESS DISTANCE	2	908.37	454.18	56.73	.001
HEADWAY x ACCESS DISTANCE	4	258.31	64.60	7.43	.001
ERROR	4534	39310.57	8.31	-----	----
TOTAL	4544	48457.12	-------	------	....

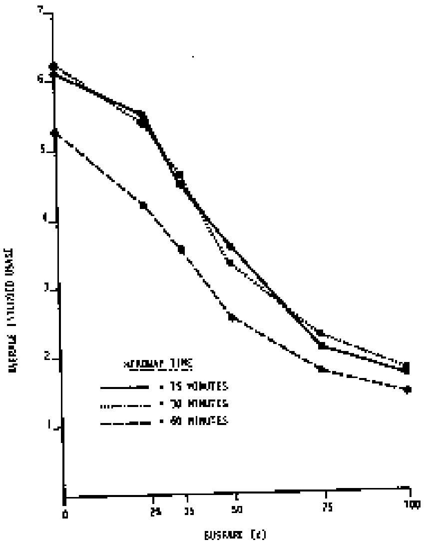


FIGURE 3 Average intended usage—access distance of 1 block.

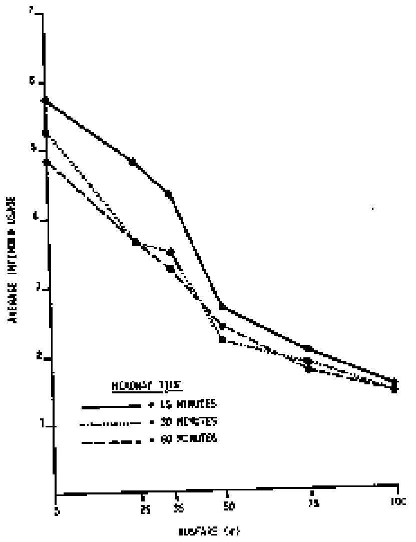


FIGURE 4 Average intended usage—access distance of 3 blocks.

In segments 1, 2 and 5, the main and interaction effects of headway and access distance are significant. However, in segments 3 and 4, only the main effects are significant; the interaction between headway and access distance is insignificant.

The coefficients of bus fare and current distance are significant in all but segment 4. Both covariates are insignificant in segment 4, which consists of older, predominantly male population with higher home ownership and lower income.

TABLE V

Analysis of covariance, segment 2

SOURCE OF VARIATION	DEGREES OF FREEDOM	SUM OF SQUARES	MEAN SQUARES	F	SIGNIFICANCE
BUS FARE	1	8424.12	8424.12	850.08	.001
ACTUAL DISTANCE	1	24.64	24.64	2.51	NOT SIGNIFICANT
HEADWAY	2	586.78	293.39	29.61	.001
ACCESS DISTANCE	2	1332.23	666.12	67.22	.001
HEADWAY × ACCESS DISTANCE	4	61.46	15.37	1.55	NOT SIGNIFICANT
ERROR	3490	34587.46	9.91	------	----
TOTAL	3500	45023.01	-------	------	----

TABLE VI

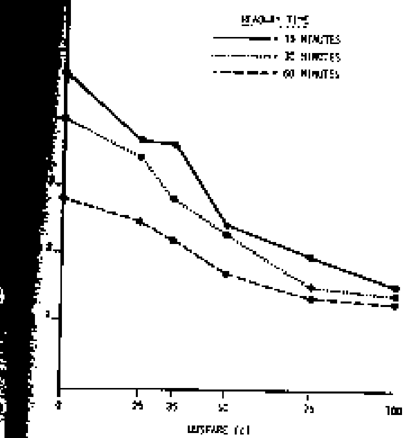
Analysis of covariance, segment 3

SOURCE OF VARIATION	DEGREES OF FREEDOM	SUM OF SQUARES	MEAN SQUARES	F	SIGNIFICANCE
BUS FARE	1	3966.62	3966.62	440.05	.001
ACTUAL DISTANCE	1	35.54	35.54	3.95	.05
HEADWAY	2	227.24	113.61	16.98	.001
ACCESS DISTANCE	2	515.75	257.86	28.43	.001
HEADWAY × ACCESS DISTANCE	4	31.61	7.90	0.72	NOT SIGNIFICANT
ERROR	2557	23000.22	9.01	-------	----
TOTAL	2561	24475.15	-------	-------	----

TABLE VII

Analysis of covariance, segment 4

SOURCE OF VARIATION	DEGREES OF FREEDOM	SUM OF SQUARES	MEAN SQUARES	F	SIGNIFICANCE
BUS FARE	1	2275.34	2275.34	436.09	.001
ACTUAL DISTANCE	1	105.66	105.66	18.20	.001
HEADWAY	2	186.48	93.42	11.32	.001
ACCESS DISTANCE	2	[illegible]	140.75	22.81	.001
HEADWAY × ACCESS DISTANCE	4	26.87	6.87	0.79	NOT SIGNIFICANT
ERROR	2052	16907.56	8.24	------	----
TOTAL	2060	19841.33	-------	------	----



GURE 5 Average intended usage—access distance of 5 cks.

TABLE VIII
Analysis of covariance, segment 5

SOURCE OF VARIATION	DEGREES OF FREEDOM	SUM OF SQUARES	MEAN SQUARES	F	SIGNIFICANCE
BUS FARE	1	3662.17	3662.17	428.80	.001
ACTUAL DISTANCE	1	16.18	16.18	1.89	NOT SIGNIFICANT
HEADWAY	2	475.88	237.94	27.81	.005
ACCESS DISTANCE	2	773.12	386.56	45.18	.001
HEADWAY × ACCESS DISTANCE	4	72.08	18.02	2.11	NOT SIGNIFICANT
ERROR	3236	27654.62	8.54	------	....
TOTAL	3244	32868.97	-------	-----	....

DISCUSSION

The Analysis of Covariance tests show that changes in bus fare, headway time and access distance significantly affect individuals' intended usage of bus. However, intended usage is relatively insensitive to the three service characteristics when they are already at extreme high or extreme low levels. For example, a change in bus fare from 0 to 25 cents did not produce as much decline in the usage response as an increase from 25 cents to 35 cents. The latter increase in bus fare, however, produced a larger decline in the usage response than an increase from 75 cents to one dollar. These results lend evidence to the existence of lower and

TABLE IX
Summary of ANCOVA results

	SOCIOECONOMIC CHARACTERISTICS	HEADWAY	ACCESS DISTANCE	INTERACTION	COVARIATES	
					BUS FARE	CURRENT ACCESS DISTANCE
TOTAL SAMPLE	95% LICENSED DRIVERS AVERAGE AGE : 40.1 YEARS SOME COLLEGE EDUCATION 24% MEMBERS OF LABOR UNIONS 83% HOME OWNERS MEDIAN INCOME, $14 450 75% REGISTERED VOTERS	*	*	*	*	*
SEGMENT 1	MORE LICENSED DRIVERS YOUNGER HIGHER EDUCATION LOWER HOME OWNERSHIP	*	*	*	*	*
SEGMENT 2	FEWER LICENSED DRIVERS LOWER HOME OWNERSHIP FEWER MALES	*	*	*	*	*
SEGMENT 3	LOWER EDUCATION FEWER REGISTERED TO VOTE	*	*		*	*
SEGMENT 4	OLDER HIGHER HOME OWNERSHIP LOWER INCOME MORE MALES	*	*			*
SEGMENT 5	MORE LICENSED DRIVERS OLDER HIGHER HOME OWNERSHIP MORE REGISTERED TO VOTE	*	*	*	*	*

*: INDICATES SIGNIFICANCE AT THE 5% LEVEL

TABLE X

Percent of variation

SOURCE OF VARIATION	SEGMENT					TOTAL SAMPLE
	1	2	3	4	5	
BUS FARE	15.7	18.7	17.2	13.5	11.1	14.4
ACTUAL DISTANCE	0.0	0.1	0.2	0.6	0.0	0.0
HEADWAY	0.9	1.3	1.1	1.1	1.4	1.1
ACCESS DISTANCE	2.0	3.0	2.6	2.2	2.4	2.3
HEADWAY X ACCESS DISTANCE	0.6	0.1	0.2	0.2	0.2	0.2

upper threshold points corresponding to each service characteristic and to the hypothesis that individuals' responses are sensitive to the levels of the service characteristics only when they vary between these upper and lower threshold points.

Two of the five segments yielded insignificant interaction between headway time and access distance. The absence of the interaction can be interpreted to mean that headway time and access distance influence intended usage response in a linear compensatory manner. In other words, disutilities resulting from increases in the level of one service characteristic can be offset by utilities gained from decreases in the level of the other service characteristic. One of the two segments which yielded no interaction effect is the older, predominantly male segment, which is also characterized by higher home ownership and lower income. The other segment consists of fewer registered voters with lower education.

Although ANCOVA yields significant F-ratios for main effects as well as the interaction effects, the percents of variations explained by these sources are small. From Table IX the percent of all variations is practically accounted for by the variations in bus fare alone.

The approach used in this research is not to be viewed as a forecasting tool for planning purposes. Rather, it is an explanatory tool for identifying the variables which influence transit usage, and the ranges of the variables at which the influence is maximum. The approach allows a transit planner not only to assess people's sensitivities to three service characteristics but also to identify segments of the sample which exhibit different sensitivity structures. This knowledge can lead to the formulation and structuring of travel demand models which would include appropriate causal variables and account for interactions among these variables. This can be potentially useful for designing transportation systems which will provide service to different segments of population.

REFERENCES

1. J. L. Louviere, L. L. Beavers, K. L. Norman, and F. C. Stetzer, "Theory, methodology and findings in mode choice behaviour," Working Paper Series 11, The Institute of Urban Regional Research, The University of Iowa, Iowa City, Iowa, July, 1973.
2. N. H. Anderson, "Integration theory and attitude change," *Psychological Review* 78, 1971, pp. 171–206.
3. G. C. Nicolaidis, and J. N. Sheth, "An application of market segmentation in urban transportation planning," GMR-2149, General Motors Research Laboratories, Warren, Michigan, 1975.
4. G. C. Nicolaidis, and K. S. Krishnan, "A methodology for planning bus service," *Transportation*, 1977, pp. 249–263.
5. E. Cramer, "Revised MANOVA program," University of North Carolina, Chapel Hill, North Carolina, 1970.

OTHER REFERENCE

D. Brand, "Travel demand forecasting, some foundation and a review," Highway Research Board, Special Report 143, pp. 239–282, 1973.