Evaluation of Alternative Market Segmentations for Transportation Planning

Gregory C. Nicolaidis, Martin Wachs,* and Thomas F. Golob, Transportation and Urban Analysis Department, General Motors Research Laboratories, Warren, Michigan

In transportation planning, market segmentation is the division of a total population of travelers into groups (segments) that are relatively homogeneous with respect to certain personal characteristics (the segmentation base). It is desirable that the segments be distinct in terms of travel behavior and their reactions to changes in the travel environment, such as the introduction of new transportation services. This paper describes a comparison of market segmentation using six different bases—two based on demographic variables, two on travel choice constraints, and two on attitudinal variables. The six segmentations were compared with respect to five criteria judged to be important considerations in transportation planning: measurability (data availability), statistical robustness, substantiality (size and importance of the resulting segments), relation to travel behavior, and relation to planning of service options. The comparisons showed that no single segmentation base was superior, according to all criteria, but that the segmentation based on multivariate choice constraints satisfied more of the criteria than did the other segmentations. Segmentations of the traveling population based on attitudes were found to have certain specific uses, but to be inferior to choice-constraints segmentation for most planning purposes.

In an effort to better match transit service to the wide range of needs and expectations that may exist in a given community, planners have recently been giving increased attention to the concept of market segmentation. Market segmentation is a procedure for dividing the (travel) market into homogeneous subsets of customers (segments), where any subset may be selected as a target market for a distinct combination of service characteristics, price levels, or promotional strategies. The primary object of market segmentation is to increase consumer appeal by meeting consumer desires (4). Thus far, it has been applied to only a limited extent in the public sector, where the objectives of such agencies as transit operators are not as clearly defined as are those of private firms (11, 13).

One of the critical aspects of segmentation is the selection of an appropriate base. Although the application of market segmentation of public transit is still a new concept, a number of alternative approaches to segmentation that propose or advocate different segmentation bases have already been suggested. Lovelock (11) discusses the relative merits of a number of such bases.

This study compares the results of several segmentations of the same set of travelers using alternative segmentation bases. The bases are compared with respect to five criteria. The first three criteria have been discussed in the marketing literature (4, 10); the other two are relevant to issues of transportation planning.

1. Measurability: The information gained from a segmentation should be cost-effective in terms of the time and money required to collect and process the necessary data.
2. Statistical robustness: The segments should be significantly different from one another in a statistical sense. The between-segment variations should be relatively larger than the within-segment variations. This helps to ensure that the segments are not the result of random variations in the data and improves the assignment of new travelers to the segments.
3. Substantiality: The segments should be large enough to account for a significant proportion of the population under study or should be sufficiently important with respect to planning policy to merit the time and cost of separate attention.
4. Relation to travel behavior: A segmentation that accounts for a large proportion of the variance in manifest travel patterns would be more useful than one that does not. Segments that differ in terms of modal choice, route choice, or trip timing and frequency are similarly desirable.
5. Relation to planning of service options: If particular transportation service packages serve consumers having very different social or economic characteristics, a segmentation base that defines consumer groups compatible with service options would be more useful than a base that does not. Similarly, if promotional activities are best targeted to consumers having certain preferences, perceptions, and desires, a segmentation procedure that identifies those groups would be more useful than one that does not.

By using data from the Ottawa-Hull metropolitan area in Ontario and Quebec, six alternative bases for market segmentation were chosen. Some of the bases divided the population of work-trip commuters into groups based on demographic or socioeconomic characteristics, some on the basis of such travel constraints as automobile ownership and bus availability, and some on the basis of their stated attitudes. Segmentations using each of these types of variables have been advocated in the transportation planning literature and are related to approaches proposed in the marketing literature (4, 6, 11). For example, Nicolaidis and Dobson (12) segmented travelers on the basis of perceived importance of attributes of transportation modes, and Nicolaidis and Sheth (13) did so on the basis of attitudes toward general environmental conditions. Recker and Golob (15) grouped individuals on the basis of what they termed choice constraints of transportation alternatives, and Costantino and others (2) used demographic and socioeconomic measures.

SEGMENTATION BASES

The six segmentation bases compared in this study were developed from data collected in two home-interview surveys administered in the Ottawa-Hull metropolitan area by the Ottawa-Carleton Regional Municipality. The first survey was a conventional home-interview, household trip-inventory survey. The characteristics of all trips made during a 24-h period and the detailed socioeconomic and demographic characteristics of the respondents were recorded. The second survey covered attitudes toward alternative modes of transportation and toward transportation-related issues in general. This attitudinal survey was administered as a home in-

terview to a subsample of the respondents who had completed the trip-inventory survey. In each of the segmentations, only those respondents who indicated that they made a regularly scheduled work trip were segmented. Thus, the analyses reported in this paper are relevant mainly to travel in the Ottawa metropolitan area for work-trip purposes.

The first two segmentations divided the population into groups that are homogeneous with respect to demographic characteristics. The underlying assumption is that travel demands and consumer responses to marketing are strongly associated with socioeconomic status or cultural variables. The first demographics segmentation was based on a multidimensional matrix of the following descriptors: language (English versus French), number of vehicles in the household, length of time lived at the current residence, number of residents in the household, type of residence (single family versus attached versus apartment), sex, age, education, possession of driver's license, employment (homemaker and student versus working), occupation (managerial and professional versus clerical and sales versus service and craftsperson), and income. The second demographics segmentation used only a single variable: Language was chosen because the residents of the study area were clearly distinguishable into those of English-speaking and those of French-speaking cultures.

This comparison of single versus multidimensional demographics segmentations was intended to determine whether the increased complexity of the multidimensional approach adds any significant value above that of a more simplistic measure. Also, because many cities include unique cultural or ethnic groups, the inclusion of a single cultural variable was intended to test the effectiveness of such a simple categorization as a market-segmentation base.

The second pair of segmentations dealt with travel-choice constraints. Many authors (8, 15) believe that such variables as automobile ownership and the availability of bus service influence travel behavior and responses to marketing more significantly than do demographics differences. The choice-constraints segmentations were intended to test such hypotheses. The multidimensional choice-constraints segmentation includes data—bus access time (waiting plus walking times), bus transfers needed, and automobile availability—related to automobile availability, accessibility to bus service, and the appropriateness of the routing of buses that were available to the respondents. Because these data are sometimes unavailable and because of the general goal of determining whether simple measures are as effective as complex ones, a segmentation based solely on automobile ownership was also included.

The last two segmentations were based on the respondents' answers to the attitudinal questions included in the Ottawa survey. Attitudes have been proposed as useful bases for market segmentation by a number of authors (12, 13). Two types of attitudinal data were examined. For the first, the segmentations were performed on the basis of general attitudes toward transportation. Six-point Likert scales were used to measure the degree to which the survey respondents agreed with each of the following 16 statements concerning general transportation-related conditions:

1. Traffic congestion in this city is a major problem that must be solved.
2. It is necessary to reduce the use of automobiles in the city by supplying an effective network of rapid public transit.
3. By and large, automobiles have outlived their usefulness except for trips between cities or into the country.
4. Driving in the city is frustrating and can cause anxiety and tension.
5. Riding in public transportation makes people feel awkward or lonely or just part of a crowd.
6. I could feel embarassed taking someone to a social function by public transit.
7. Drastic action must be taken to improve the public transit service in this city.
8. An automobile is more than just transportation; having a nice automobile to drive is appealing in itself.
9. Although automobiles are sometimes necessary, they are also a nuisance. I would just as soon do without one if other transportation met my needs.
10. I enjoy (would enjoy) driving an automobile.
11. Not having an automobile available is like being trapped.
12. The lack of adequate transportation facilities for all leads to family squabbles.
13. The government should actively discourage people from using automobiles in busy sections of the city by making it more difficult to drive and park there.
14. Children need good public transportation or they make too many demands on their parents to drive them around.
15. Your social life definitely suffers if there is no automobile available.
16. To be honest, there is no public transportation system I can picture that would make me give up using my automobile in the city.

Specific attitudes toward transportation alternatives were the final segmentation base. Six-point semantic differential scales (very important to not important) were used to measure the importances travelers placed on the following 25 attributes characterizing bus and automobile modes when making modal choices:

1. Comfortable seating,
2. Dependability of on-time arrival,
3. Availability more or less when you want it,
4. Attractiveness of vehicle,
5. Low noise level in vehicle,
6. Vehicle safety,
7. Smoothness of ride,
8. Privacy from other people,
9. Avoiding exposure to traffic congestion,
10. Minimum exposure to bodily crowding,
11. Low out-of-pocket cost,
12. Low riding time,
13. Low walking time,
14. Low waiting time,
15. Opportunity to meet and talk with other people,
16. Opportunity to relax,
17. Opportunity to read,
18. Continuous ride with few stops,
19. Protection from weather on entire trip,
20. Flexible destination, can go anywhere,
21. Not having to change vehicles,
22. Year-round temperature comfort in vehicle,
23. Assurance of having a seat,
24. Security from undesirable acts of others, and
25. Low level of pollution.

SEGMENTATION PROCEDURE

The six segmentations were performed by using similar procedures with slight variations among the situations where the segmentation base had only one variable, a few variables, or a large number of variables containing possible multicolinearities. For the two unidimen-

sional segmentation bases, language and automobile ownership, the segments were formed by the natural categorization of respondents according to their values of the variables. The version of the procedure applied to the only base having a few variables—multidimensional choice constraints—involved standardizing the base variables to zero mean and unit standard deviation to eliminate scale biases and then clustering the individuals in the space of the standardized variables. The last version of the procedure was applied to the multidimensional demographics, general attitudes, and attribute-importance bases, which all had a larger number of variables containing possible intercorrelations. This procedure involved factor analyzing the variables and then clustering the individuals in the space of the resulting latent factors.

The specific factor-analysis technique used was principal-components analysis applied to the variable correlation matrix, followed by varimax rotation to facilitate factor interpretation (7). An iterative technique described by Recker and Golob (16) was used to determine, and consequently eliminate from factoring, those few variables that might have contained primarily noise and would not add any information to the factor results. The selection of the latent factors that best expressed the variable interrelations was made on the bases of criteria also described by Recker and Golob.

The specific cluster-analysis technique used in both of the latter two versions of the procedure involved a customized algorithm closely related to the ISODATA algorithm and to the class of cluster-analysis techniques referred to as K-means clustering (1). For a given number of segments, the algorithm assigned each respondent to the segment with the property that the distance between that respondent and the centroid of the segment was smaller than the distance between the respondent and the centroid of any other segment; new centroids were then computed and the process repeated. A procedure was used for determining the final number of segments from successive analyses with different numbers of segments based on matrices of generalized distances between segments in the factor spaces (5) and on the summary compactness indexes for each clustering [Λ-statistic due to Wilks (18)].

SEGMENTATION RESULTS

Factor analysis of the 12 demographic and socioeconomic variables in this segmentation base gave four latent factors. These factors accounted for approximately 62 percent of the variance in the original variables and were labeled as social rank, life cycle, occupation, and household size through interpretation of the correlations among the factors and the original variables. Cluster analysis of the 324 survey respondents for which full data were available gave three segments. These three segments were given subjective labels based on interpretation of the positions of their centroids in the space of the four factors. This information and the proportions of the samples that were assigned to each segment are given in Table 1, which also includes the results of the other five segmentations. The sample sizes (Ns) of the segments are given below.

Base	Segment	N
Multidimensional demographics	French speaking	107
	Younger and more renters	96
	Older and more males	121
Language	English speaking	180
	French speaking	98
Multidimensional choice constraints	Mobile	211
	Inappropriate bus routing	99
	Poor bus accessibility	94
	Automobileless	91
	Busless	48
Automobile ownership	None	17
	One	191
	Two or more	86
General attitudes	Not automobile dependent	124
	Driving conditions acceptable	143
	Public transit acceptable	151
	Transportation improvements needed	121
Attribute importances	Ambivalent	82
	Service versus personal environment	186
	Total environment versus travel convenience	122
	Travel convenience versus service	114

The work-trip travelers surveyed in the Ottawa metropolitan area were divided into English-speaking and French-speaking segments. There was complete data on these variables for 278 survey respondents.

Cluster analyses of the 543 respondents for whom perceived constraints on choice of mode were available gave five segments. Interpretations of the positions of the segments in the three-dimensional choice-constraint space gave the segment labels shown in Table 1. Segments four and five were labeled automobileless and busless respectively to reflect that the majority of respondents did not own an automobile or did not have bus service available to them.

The 294 survey respondents for whom complete data on automobile ownership were available were divided into no-automobile, one-automobile, and two-or-more-automobiles segments.

Factor analysis of the levels of agreement with the 16 statements measuring general attitudes gave four factors that accounted for 63 percent of the variance in the factored variables. These factors were subjectively labeled as anticongestion, automobile dependence, new intraurban transport is needed, and public transit is depersonalizing.

Cluster analysis of the 539 respondents for whom full data on the base variables were available gave four segments.

Factor analysis of the importance ratings of the 25 modal attributes gave five factors that accounted for 56 percent of the variance in the factored variables. These factors were labeled as service, vehicle comfort, system environment, travel convenience, and personal environment. Cluster analysis of the 505 respondents for whom full information was available gave four segments.

TESTS OF SEGMENTATION REDUNDANCIES

The question arises as to whether or not the six segmentations merely represent six ways of dividing the total population into the same basic groups. This question can be rephrased in a statistical sense for each pairwise comparison of segmentations: If it is known into which segment a particular respondent is classified in one segmentation, can it be predicted with significantly better than random probability into which segment this same respondent will be classified in another segmentation? It can be expected that the two segmentations using demographic bases (the multidimensional demographics and language segmentations) will be highly related, that the two choice-constraint segmentations (the multidimensional choice-constraints and automobile-ownership segmentations) will be related, and that possibly the two attitudinal segmentations will be re-

Table 1. Segmentation results.

Type of Variables	Segmentation Base	Segment	Percentage of Segmentation Total
Demographic	Multidimensional	French speaking	33
		Younger and more renters	30
		Older and more males	37
		Total	100
	Language	English speaking	65
		French speaking	35
		Total	100
Choice constraint	Multidimensional	Mobile	39
		Inappropriate bus routing	18
		Poor bus accessibility	17
		Automobileless	17
		Busless	9
		Total	100
	Automobile ownership	None	6
		One	65
		Two or more	29
		Total	100
Attitudinal	General attitudes	Not automobile dependent	23
		Driving conditions acceptable	27
		Public transit acceptable	28
		Transportation improvements needed	22
		Total	100
	Attribute importances	Ambivalent	16
		Service versus personal environment	37
		Total environment versus travel convenience	24
		Travel convenience versus service	23
		Total	100

lated. However, if there are strong relations among segmentations using different types of bases, these relations must be considered when interpreting the results of the comparative analyses given in this paper.

The Pearson χ^2 test of association through contingency tables of segment membership (9) was used to identify possible redundancy in the segmentations. As expected, the two demographics and the two-choice constraints segmentations are associated at the 95 percent confidence level, but the two attitudinal segmentations are not significantly associated (even allowing much lower confidence bounds on acceptance of random occurrences). Of the 12 comparisons of pairs of segmentations from different types of segmentation bases, only the multidimensional demographics versus automobile-ownership pair was found to be significantly associated. This result is consistent with the models calibrated in many transportation planning studies to distribute and forecast automobile ownership (3).

Thus, the conclusion of the redundancy test is that comparisons among segmentations using different types of bases need be qualified only when the demographics versus automobile-ownership pair of segmentation is involved. Furthermore, comparisons between the two attitudinal segmentations are valid also without qualification.

EVALUATION OF SEGMENTATIONS

The results of the evaluations of the six segmentation bases on the five criteria are summarized in Table 2.

Measurability

The three types of segmentation bases are clearly distinguishable with respect to measurability. The demographic data are the most readily available; these data are collected in almost every origin-destination home-interview survey, on-board transit-user survey, or other traveler survey designed to gather information about individual respondents and their households. Consequently, if an acceptable sample of such survey responses is available for the population to be segmented, demographics segmentation bases are cheapest in terms of the time and cost of data collection and processing.

Choice-constraints data are of the type needed for estimating disaggregate travel-demand models. It can thus be expected that these data will be collected in future origin-destination surveys. However, at present, data on travelers' perceptions of constraints on choices of modes, routes, and trip times are limited. On the other hand, collection of choice-constraint data requires that only a few questions be asked of respondents. [For the specific data used in the analyses reported here, these questions have been given by Recker and Golob (16).] This relative simplicity makes possible the use of cost-effective data-collection techniques such as telephone surveys.

Attitudinal data usually require a separate survey. Moreover, because of the complexity of the explanations of the questions and the monitoring of responses, these surveys usually must be administered as home interviews. Their higher data-collection costs can be offset by the use of the data obtained in providing non-segmentation planning information, but discussion of such uses is beyond the scope of this paper. Thus, in the present evaluation, the two attitudinal segmentations share a common burden of costs that must be offset by gains in planning information over and above the level of information provided by the competing demographics and choice-constraints segmentations.

Statistical Robustness

Two tests of the degree to which the different segmentations succeeded in identifying distinct structure in the segmentation-base data were conducted for those four segmentations that involved multiple base variables (i.e., the multivariate segmentations). The first test focused on an overall statistic measuring the effectiveness of a cluster analysis in determining segments that

Table 2. Evaluation results.

Type of Variables	Segmentation	Evaluation Criterion: Measurability	Statistical Robustness	Substantiality	Relation to Travel Behavior	Relation to Planning of Service Options
Demographic	Multidimensional	+	+	0	0	0
	Language	+	–	0	0	-
Choice constraint	Multidimensional	0	+	0	+	0
	Automobile ownership	+	–	0	+	0
Attitudinal	General attitudes	-	-	0	-	+
	Attribute importances	-	-	0	-	+

Note: + indicates that a segmentation was judged to be significantly superior on a criterion; 0 indicates that a segmentation was judged to be neither superior nor inferior on a criterion; - indicates that a segmentation was judged to be significantly inferior on a criterion.

are both compact and significantly different from one another. For reasons discussed by Friedman and Rubin (6), the overall test statistic chosen was the Wilks Λ-criterion. This statistic, which is invariant under linear changes of scales on which variables are measured, is defined as the ratio of the determinant of the pooled within-segment variance (a measure of the compactness of the clusters) to the determinant of the between-segment variance (a measure of the dispersion of cluster centroids in the variable space). By using a variance-ratio transformation proposed by Rao (13), the possibility that the clusters could occur in randomly structured data was evaluated for each of the segmentations.

For each of the four multivariate segmentations, the hypothesis that a segmentation structure was due to random variation in the data was rejected at a very high confidence level. Thus, no one segmentation performed better than the others, and the segmentations were judged to be indistinguishable in terms of the Wilks Λ-criterion.

The second test of statistical robustness involved how successfully the observations could be assigned to their correct segments. Such assignments are commonly done by using linear functions of the base variables determined through multigroup discriminant analysis (17). These functions were calculated for each of the multivariate segmentations, and the discriminant classifications for each of the travelers in the original sample were compared with their cluster-analysis segment assignments. The percentage of correct classifications for each group in each segmentation and for each segmentation are shown in Table 3.

The four multivariate segmentations are distinguishable in terms of their percentages of correct discriminant classifications. The multidimensional choice-constraints and multidimensional demographics segmentations showed the most successful classifications. Moreover, with the sole exception of the relatively small busless segment in the choice-constraints segmentation, the individual segments in these two segmentations were uniformly high in correct classifications. Such balanced classification success is deemed to be desirable in the absence of independent information about the differential planning importances of various segments. The general-attitudes segmentation showed a modest classification performance that was balanced among its four segments. Finally, the attribute-importances segmentation showed the poorest classification performance, both in terms of overall success and of balance among the four segments.

Substantiality

Market segmentation can contribute to the efficiency of planning and marketing when the segments are substantial in size and when the distribution of segment sizes contains few extremes. For example, a segmentation that included more than 90 percent of the travelers in one cluster and only 1 or 2 percent of the travelers in another would be difficult to use in the planning or marketing of services. The distribution of cluster sizes cannot be considered in the abstract, however, but must also be analyzed in terms of the significance of particular segments to transportation programs. Thus, if the current policy emphasizes transit planning for the elderly, it might be useful to isolate the elderly in a segmentation procedure even though they might constitute a very small proportion of the total pool of travelers.

Table 1 showed the proportions of the samples that were assigned to each cluster in each of the six segmentations. The multidimensional demographics and general-attitudes segmentations divided the sample into segments of approximately equal size and containing few extremes. However, both of these segmentations gave rise to segments that could not be identified with transportation policy questions. All of the other segmentations gave rise to cluster-size distributions that were quite acceptable. The smallest proportion assigned to any segment was the 9 percent of the travelers who were assigned to the busless segment in the multidimensional choice-constraints segmentation. The busless, however, constitute a group of high salience with respect to current transportation planning policy and, in a segmentation that gave rise to five clusters, this proportion seems quite acceptable. On balance, then, the segmentations are relatively indistinguishable with respect to the substantiality criterion.

Relation to Travel Behavior

One of the important criteria by which the usefulness of a market-segmentation base can be judged is the extent to which the resultant segments are distinguishable in terms of travel behavior. If the clusters of travelers resulting from a segmentation have significantly different travel demands and trip patterns, this segmentation is more useful for planning and marketing than one that gives clusters that are undifferentiated in terms of travel. Three steps were involved in comparing the six segmentation bases with respect to this criterion. First, the segmentation bases were compared to determine which of them resulted in segments having different frequencies of modal choice for the journey to work. Second, the groups in each segmentation were compared to determine whether they differed in terms of such trip characteristics as trip length, access and egress times, number of transfers and other reported work-trip characteristics. Finally, a multidimensional logit model of modal choice was applied to each segment in each segmentation to determine whether certain ones resulted in better goodness of fit of the demand model. The independent variables in these choice models consisted of the satisfaction ratings of the survey respondents on attributes of automobile and bus.

The data used to determine differences in terms of the frequency of modal choice were the reported frequencies with which respondents traveled to work as automobile drivers, automobile passengers, or bus passengers during the 4-week period preceding the survey day. Multiple-group discriminant analysis was used, and discriminant functions were determined for each of the six segmentation bases by using the three modal-choice frequencies as independent variables. If no statistically significant discriminant function could be found for a particular segmentation, it was concluded that the segmentation did not distinguish on the basis of modal-choice frequencies. If a satisfactory discriminant function could be computed, it was concluded that those independent variables that had significant coefficients (using an F-test and a 0.99 significance level) varied significantly among the segments produced by that base. Table 4 shows the results of these discriminant analyses. For each segmentation, mean modal-frequency values are shown only for those cases in which the means are significantly different from one another.

No significant discriminant function could be found for either the general-attitudes or the attribute-importances segmentations. Thus, it was concluded that these attitudinal segmentations have little statistical association with modal-choice frequencies. On the other hand, the segmentation based on multidimensional choice constraints resulted in segments having significantly different frequencies with which the work trip was made as an automobile driver and as a bus passenger. Segmentation on the basis of automobile ownership, rather than the more complex choice-constraints base, also produced segments that differed significantly with respect to two of the three modal-choice frequencies. Segmentations based on demographics and on the single dimension of language produced clusters that differed from each other in terms of only one modal-choice variable—the frequency of bus use. In summary, then, the segmentations based on choice constraints best described travelers' modal-choice frequencies, while the segmentations based on attitudes were the poorest discriminators of modal choice.

The analysis to determine which of the segmentations produced groups that differed in terms of trip characteristics was based on eight reported, or perceived, trip characteristics. By following a procedure similar to that used in the analysis of modal-choice frequencies, discriminant functions were estimated by using as independent variables any of the eight trip characteristics that could explain the various segmentations in a statistically significant manner.

Table 5 shows the results of these discriminant analyses. Once again, it was not possible to construct a statistically significant discriminant function for segmentation based on general attitudes or attribute importances. Similarly, the multidimensional choice-constraints segmentation gave a discriminant function that included three of the eight variables. The choice-constraints segmen-

Table 3. Results of segmentation discriminant analyses.

Type of Variables	Segmentation	Percentage of Overall Correct Classification	Segment	Percentage of Correct Classification
Demographic	Multidimensional	95	French speaking	95
			Younger and more renters	95
			Older and more males	96
	Language	—	—	—
Choice constraint	Multidimensional	98	Mobile	99
			Inappropriate bus routing	98
			Poor bus accessibility	100
			Automobileless	100
			Busless	85
	Automobile ownership	—	—	—
Attitudinal	General attitudes	87	Not automobile dependent	87
			Driving conditions acceptable	84
			Public transit acceptable	87
			Transportation improvements needed	89
	Attribute importances	77	Ambivalent	65
			Service versus personal environment	81
			Total environment versus travel convenience	72
			Travel convenience versus service	85

Table 4. Results of discriminant analyses based on modal-choice frequencies.

	Segmentation		Mean No. of Work Trips Made During Last 4 Weeks		
Type of Variables	Base	Segment	Automobile Driver	Automobile Passenger	Bus User
Demographic	Multidimensional	French speaking	NS	NS	1.9
		Younger and more renters	NS	NS	6.9
		Older and more males	NS	NS	5.7
	Language	English speaking	NS	NS	6.0
		French speaking	NS	NS	2.3
Choice constraint	Multidimensional	Mobile	13.3	NS	4.8
		Inappropriate bus routing	16.4	NS	3.3
		Poor bus accessibility	15.0	NS	2.1
		Automobileless	0.5	NS	12.9
		Busless	17.5	NS	0.0
	Automobile ownership	None	NS	5.1	13.1
		One	NS	3.3	4.7
		Two or more	NS	2.6	3.1
Attitudinal	General attitudes	All	NS	NS	NS
	Attribute importances	All	NS	NS	NS

Table 5. Results of discriminant analyses based on work-trip characteristics.

Type of Variables	Segmentation		Trip Characteristics							
	Base	Segment	Trip Time by Mode Actually Used (min)	Trip Time by Bus (min)	Access Time to Bus, Walking (min)	Egress Time for Mode Actually Used (min)	Egress Time for Bus (min)	Number of Transfers, Bus	Type of Bus Service Available[a]	Avg Wait for Bus (min)
Demographic	Multidimensional	French speaking	31.3	NS	2.1	NS	NS	NS	NS	NS
		Younger and more renters	24.8	NS	3.0	NS	NS	NS	NS	NS
		Older and more males	24.4	NS	2.7	NS	NS	NS	NS	NS
	Language	English speaking	24.6	40.6	NS	NS	NS	NS	2.0	NS
		French speaking	31.0	49.0	NS	NS	NS	NS	1.8	NS
Choice constraint	Multidimensional	Mobile	NS	22.0	1.7	NS	NS	NS	NS	2.4
		Inappropriate bus routing	NS	64.3	4.5	NS	NS	NS	NS	7.8
		Poor bus accessibility	NS	47.8	6.4	NS	NS	NS	NS	10.9
		Automobileless	NS	44.5	4.3	NS	NS	NS	NS	5.9
		Busless	NS	90.4	6.2	NS	NS	NS	NS	12.3
	Automobile ownership	None	NS	NS	NS	NS	NS	0.4	NS	NS
		One	NS	NS	NS	NS	NS	0.3	NS	NS
		Two or more	NS	NS	NS	NS	NS	0.7	NS	NS
Attitudinal	General attitudes	All	NS	NS	NS	NS	NS	NS	NS	NS
	Attribute importances	All	NS	NS	NS	NS	NS	NS	NS	NS

[a]Types of bus service were coded as follows: 1 = no bus available, 2 = local bus service, 3 = express bus service only, 4 = local and express bus service available. The scale was assumed to be interval in this analysis.

tation gave clusters that had significantly different values of bus trip time, varying from a mean of 22 min for the mobile cluster to 90 min for the busless cluster. Access time varied significantly among the groups produced by the choice-constraints segmentation, from 1.7 min among the mobile segment to 6.4 min among the segment having poor bus access. Average waiting times for buses also varied significantly among the choice-constraints clusters, from only 2.4 min for the mobile cluster to 12.3 min for the busless. Also, the simplified choice-constraints segmentation, based solely on automobile ownership, distinguished travelers more poorly than did the multidimensional choice-constraints segmentation, since only the number of transfers entered the discriminant function for the automobile-ownership classification. The segmentation based on language distinguished among clusters on the basis of three of the eight trip characteristics, while that based on multidimensional choice-constraints discriminated on the basis of only two of the eight variables. Overall, the segmentations based on attitudinal variables distinguished as poorly among trip characteristics as they did among modal-choice frequencies. The segmentations based on multidimensional choice constraints and on language were the most effective at discriminating on the basis of trip characteristics.

The final test of association between the various segmentation bases and travel behavior involved the fitting of a separate travel-demand model for each segment produced in each segmentation. If models applied to segments drawn from one base yield goodness-of-fit measures that are consistently superior to those produced by a different segmentation, the first segmentation base is deemed superior to the second for purposes of demand modeling.

The methodology underlying the estimation of modal-choice models for each of the segments in the four segmentations has been described in detail by Recker and Golob (16). Briefly, it involved factor analyzing the attribute-satisfaction ratings for the two alternative modes to remove multicolinearity and calibrating a probabilistic choice model using attributes representing the latent perception factors as explanatory variables. The choice model used was the logit model with maximum likelihood estimation of parameters, and the dependent variable was the respondent's modal choice on the survey day. Only those explanatory variables were used in the final estimation that had coefficients significantly different from zero at the 95 percent confidence level.

Table 6 lists two goodness-of-fit indexes for the choice models. The first—pseudo R^2—is analogous to the coefficient of determination in linear models and is expressed as the ratio of explained log likelihood to total log likelihood. Unfortunately, this measure has no known distributional properties and can be shown to have a maximum value significantly less than 1.0. The second is calculated from posterior probability estimates. The primary disadvantage of this statistic is that cases in which only slight errors in probability are made (e.g., 51 percent posterior probability in favor of choosing the mode not actually chosen) are treated the same as cases in which gross errors are made.

There is little differentiation among the results for the six segmentations. The overall weighted averages of the indexes for the multidimensional choice-constraint segmentation were slightly higher than the others, but the difference was not sufficient to permit conclusions to be drawn. Moreover, the results for the choice-constraints segmentation were mediated by the fact that the sample size for one segment was insufficient to permit estimation of the choice model, and another segment exhibited very poor results.

Fourteen of the 19 segments for which choice models were estimated exhibited goodness-of-fit indexes that were substantially better than the indexes of the aggregate model. Thus, choice-model descriptive power can be increased through the use of segmentation, although there is little evidence to favor the use of one segmentation basis over the others for modeling travel choices.

Considering the three measures of the association between segmentation bases and travel behavior that were used here, it must be concluded that the choice-constraints segmentation is superior to the others in distinguishing travel behavior. However, the differences among the segmentations are not always large, and the objectives of a particular segmentation study might justify the use of other segmentation bases. The attitudinal bases are clearly the weakest for distinguishing travelers on the basis of travel behavior.

Relation to Planning of Service Options

Demand models were estimated for each segment produced by all six segmentation bases. These models were evaluated to assess their potential contribution to

Table 6. Results of goodness of fit of choice model.

Type of Variables	Segmentation		Goodness of Fit of Choice Model	
	Base	Segment	Pseudo R^2	Percentage of Correct Classification
Demographic	Multidimensional	French speaking	0.61	92
		Younger and more renters	0.30	79
		Older and more males	0.34	77
	Language	English speaking	0.35	75
		French speaking	0.55	91
Choice constraint	Multidimensional	Mobile	0.47	83
		Inappropriate bus routing	0.66	91
		Poor bus accessibility	0.74	94
		Automobileless	0.29	80
		Busless	—	—
	Automobile ownership	None	—	—
		One	0.39	81
		Two or more	0.60	89
Attitudinal	General attitudes	Not automobile dependent	0.19	77
		Driving conditions acceptable	0.48	83
		Public transit acceptable	0.52	85
		Transportation improvements needed	0.50	86
	Attribute importances	Ambivalent	0.35	81
		Service versus personal environment	0.37	78
		Total environment versus travel convenience	0.49	84
		Travel convenience versus service	0.44	83

the design and evaluation of transportation plans. On the basis of arguments advanced by Recker and Golob (16), it was assumed that modal-choice models using satisfaction ratings on automobile and bus attributes were appropriate demand models to use. Because a system attribute, such as bus waiting time, would be affected by the implementation of plans that change headways on bus routes, it is desirable that satisfactions with bus waiting time be a significant explanator of modal choice for some of the segments used in the study. It is, in other words, desirable that the segmentations used in a particular planning study divide the total population such that some of the groups are sensitive to changes in key attributes. If market segmentation is to be effective, it is desirable that models estimated on segments of the population be more sensitive to service variables than are models estimated on the total population.

This comparison is dependent on the particular planning study under consideration. Three types of planning studies were chosen: those involving primarily changes in bus service characteristics, those involving changes in automobile costs and traffic congestion, and those involving primarily changes in bus comfort and amenity factors.

An attribute describing the bus or automobile mode was judged to be a significant explanator of modal choices for a particular segment if statistical tests showed a significant coefficient for that attribute in a modal-choice model. These tests were performed by using the attitudinal choice models described above; the dependent variable in the model estimated for each segment was the binary modal-choice variable, and the potential independent variables were the satisfaction ratings of the attributes of the bus and the automobile. The attributes are the same 25 for which the importance ratings were obtained. Care was taken not to allow independent-variable multicolinearity to affect the tests of significance.

For planning studies involving bus service characteristics, the segmentations were indistinguishable. An example is the planning of new bus routes for areas not being served. Four attributes of travel included in the attitudinal survey of workers in Ottawa were proposed as attributes that could conceivably be affected directly by such plans: bus availability, bus walking time, bus waiting time, and number of bus transfers. The segmentations were indistinguishable with respect to differential sensitivities to these four attributes.

For planning studies involving automobile costs and traffic congestion, the two demographics segmentations were deficient. The multidimensional choice-constraints segmentation and the two attitudinal segmentations produced groups that were homogeneous with respect to their differential sensitivities to the three proposed direct-effect attributes.

Finally, for planning studies involving bus comfort and amenity factors, the multidimensional demographics and general attitudes segmentations were the most appropriate. The evaluation of the possible introduction of a new transit vehicle, such as one based on new standards for bus design issued by the U. S. Department of Transportation, is an example of such a study. The results indicate that the multidimensional demographics and general attitudes segmentations are the most appropriate for use in such a study.

CONCLUSIONS

The evaluation results are summarized in Table 2. Different segmentation bases will be useful for different purposes; the findings of this study do not recommend a single segmentation base as superior. Most of the criteria that were used in this study failed to differentiate among the four nonattitudinal segmentations (i.e., multidimensional demographics, language, multidimensional choice constraints, and automobile ownership).

The choice-constraints segmentations performed as well as or better than all others on most criteria. Having clear advantages in measurability and statistical robustness over the attitudinal segmentations, the choice-constraints segmentations displayed the strongest and most easily interpreted associations with travel behavior. These results argue for the use of these segmentation bases in many planning programs.

Segmentations based on attitudinal variables performed poorly when compared with segmentations based on choice-constraints variables on several criteria. The availability of attitudinal data is usually lower because cost of its collection is higher. However, attitudinal segmentations did perform more satisfactorily than the other bases in assessing demand sensitivity to potential comfort or amenities improvements in transit service (e.g., bus shelters or bus interiors). It may well be that attitudinal segmentations are useful primarily for

certain sophisticated marketing purposes, including product design refinements and the development of promotional strategies.

The demographics segmentations gave results that were somewhere between the choice-constraints and the attitudinal segmentations with respect to most criteria. The demographics segmentations gave rise to satisfactory, although not outstanding, associations with potential service improvements. The major advantage of the demographics segmentations is, of course, a high level of measurability. One surprising result was that demand models calibrated for the demographics segmentations were not found to be sensitive to changes in automobile-travel costs or traffic congestion; demand models estimated for the demographics segments could not be used to estimate responses to possible changes in such automobile-travel attributes.

DIRECTIONS FOR FUTURE RESEARCH

First, the data set used in this study was limited to work trips, and a broadening of the analysis to trips for other purposes would be extremely useful. Second, this was a case study in one metropolitan area, and the extent to which the unique characteristics of the Ottawa area may have influenced the findings is unknown. Further study using data from other cities would be desirable. Finally, the criteria that were used in the evaluation of the segmentation bases are still considered preliminary. It was difficult to formulate operational measures to match each criterion, and the list of criteria may be too short. The criteria generally emphasized relations between segmentations and travel-behavior variables or demand modeling. Modal choice, however, was the primary measure of traveler behavior used, and other measures including automobile ownership and household-location decisions might also be included. The list of criteria used could also be strengthened by including additional measures tailored more precisely to the evaluation of segmentations in terms of their contributions to promotional efforts in transit marketing.

REFERENCES

1. G. H. Ball and D. J. Hall. A Clustering Technique for Summarizing Multivariate Data. Behavioral Science, Vol. 12, 1967, pp. 153-155.
2. D. P. Costantino, R. Dobson, and E. T. Canty. Investigation of Modal Choice for Dual-Mode Transit. TRB, Special Rept. 170, 1976, p. 67, and General Motors Research Laboratories, Warren, Mich., Publ. 1587, 1974.
3. H. D. Deutschman. Automobile Ownership Revisited: A Review of Methods Used in Estimating and Distributing Automobile Ownership. HRB, Highway Research Record 205, 1967, pp. 31-49.
4. R. E. Frank, W. F. Massy, and Y. Wind. Market Segmentation. Prentice-Hall, Englewood Cliffs, N.J., 1972.
5. H. P. Friedman and J. Rubin. On Some Invariant Criteria for Grouping Data. Journal of the American Statistical Association, Vol. 62, 1967, pp. 1159-1178.
6. R. I. Haley. Benefit Segmentation: A Decision-Oriented Research Tool. Journal of Marketing, Vol. 32, 1968, pp. 30-35.
7. H. H. Harman. Modern Factor Analysis. Univ. of Chicago Press, 1967.
8. D. T. Hartgen. Attitudinal and Situational Variables Influencing Urban Mode Choice: Some Empirical Findings. Transportation, Vol. 4, 1974, pp. 377-392.
9. W. L. Hays. Statistics. Holt, Rinehart, and Winston, New York, 1963.
10. P. Kotler. Marketing Management: Analysis, Planning, and Control. Prentice-Hall, Englewood Cliffs, N.J., 1972.
11. H. Lovelock. A Market Segmentation Approach to Transit Planning. Proc., Transportation Research Forum, Vol. 16, 1975, pp. 247-258.
12. G. C. Nicolaidis and R. Dobson. Disaggregated Perceptions and Preferences in Transportation Planning. Transportation Research, Vol. 9, 1975, pp. 279-295.
13. G. C. Nicolaidis and J. N. Sheth. An Application of Market Segmentation in Urban Transportation Planning. General Motors Research Laboratories, Warren, Mich., Publ. 2149, 1976.
14. C. R. Rao. Linear Statistical Inference and Its Application. Wiley, New York, 1973.
15. W. W. Recker and T. F. Golob. A Behavioral Travel-Demand Model Incorporating Choice Constraints. Advances in Consumer Research, Proc., Annual Conference of the Association for Consumer Research, Vol. 3, 1976, pp. 416-424.
16. W. W. Recker and T. F. Golob. An Attitudinal Modal-Choice Model. Transportation Research, Vol. 10, 1976, pp. 299-310.
17. M. M. Tatsuoka. Discriminant Analysis: The Study of Group Differences. In Selected Topics in Advanced Statistics, Institute for Personality and Ability Testing, Champaign, Ill., No. 6, 1970.
18. S. S. Wilks. Certain Generalizations in the Analysis of Variance. Biometrika, 1932, pp. 471-494.

Publication of this paper sponsored by Committee on Traveler Behavior and Values.

Mr. Wachs is now at the School of Architecture and Urban Planning, University of California, Los Angeles.