TRANSPORTATION RESEARCH RECORD 637

Forecasting Passenger and Freight Travel

TRANSPORTATION RESEARCH BOARD

COMMISSION ON SOCIOTECHNICAL SYSTEMS
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C. 1977

Transportation Research Record 637
Price $3.80
Edited for TRB by Frances R. Zwanzig

subject areas
03 rail transport
15 transportation economics
55 traffic measurements
84 urban transportation systems

Transportation Research Board publications are available by ordering directly from the board. They may also be obtained on a regular basis through organizational or individual supporting membership in the board; members or library subscribers are eligible for substantial discounts. For further information, write to the Transportation Research Board, National Academy of Sciences, 2101 Constitution Avenue, N.W., Washington, D.C. 20418.

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Forecasting passenger and freight travel.

(Transportation research record; 637)
Includes bibliographical references.
1. Transportation–United States–Passenger traffic–Congresses.
2. Freight and freightage–United States–Congresses. I. Title. II. Series.
TE7.H5 no. 637 [HE332] 380.5'08s
ISBN 0-309-02664-4 [388] 78-6984

Sponsorship of the Papers in This Transportation Research Record

Sponsorship is indicated by a footnote at the end of each report. The organizational units and officers and members are as of December 31, 1976.

Contents

Predicting Car-Pool Demand

Ride Sharing to Work: An Attitudinal Analysis

Abraham D. Horowitz, General Motors Research Laboratories
Jagdish N. Sheth, University of Illinois

A mathematical model of ride sharing was proposed and tested by using data collected in the Chicago area in 1975. The purpose of the model development was to determine how perceived advantages and disadvantages of ride sharing determine behavioral predispositions toward it. The main conclusions are that (a) demographic and travel characteristics are poor indicators and predictors of the choice between driving alone and ride sharing; (b) the study of attitudes toward ride sharing and driving alone provides answers that are relevant to the question of how to develop ride-sharing strategies; (c) with the exception of individuals having a relatively high socioeconomic status, appeals based on public-interest issues of energy, traffic, and air quality have little chance of changing attitudes toward ride sharing; (d) perceptions of drivers toward time loss and the characteristics of convenience and reliability of ride sharing would need to change before their travel behavior would change and perceptions of economic advantages have a minor role in the determination of behavioral predisposition toward ride sharing; and (e) to override negative perceptions about ride sharing, campaigns should address its positive aspects related to the use of travel time and its convenience and reliability.

The literature on ride sharing, which has developed mainly as a consequence of the energy shortage of 1973-1974, is concerned with the travel characteristics of car poolers (10), with ride-sharing matching (3, 6, 15, 20), with the study of incentives for inducing people to share a ride (2, 16, 23), and with clinical-social aspects (1, 4).

Studies on ride-sharing matching and incentives are based on the presumptions that solo drivers can be induced to car pool by offering them direct incentives (for example, parking and traffic priorities) or that driving alone might be discouraged by, for example, increasing the cost of gasoline. Effective promotion of ride sharing requires a direct knowledge of how it is viewed both by commuters who drive alone and by those who share a ride to work.

Attitudes toward ride sharing have been studied by Alan M. Voorhees and Associates (22), by Carnegie-Mellon University (5), and by Dueker and Levin (8). The Voorhees and Associates and Carnegie-Mellon studies showed that there are significant differences in attitudes toward ride sharing between solo drivers and car poolers. However, the structure of attitudes was not studied in depth, nor was there any attempt to identify homogeneous subgroups that might differ in their attitudes. Dueker and Levin examined the way in which the desirability of ride sharing varies as a function of the sex of a rider and whether or not the rider is a prior acquaintance.

Horowitz (11) has developed a theoretical framework for the measurement of attitudes toward ride sharing and driving alone and presented mathematical models relating modal choice to the perceived advantages and disadvantages of ride sharing and to other attitudinal and socioeconomic characteristics. This paper reports the results of testing this framework by a marketing-research survey.

The survey was conducted in 1975 among residents of the Chicago metropolitan area who were contacted through their employers. The main resons for choosing Chicago as the site of the data collection was that it has a wide variety of businesses, both in terms of their type and size and in terms of their locations (city and suburban), and a variety of public transit services.

The personnel departments of 43 firms, chosen randomly from a large list of companies that employ at least 100 people, were contacted. Cooperation was good, and 34 of the 43 (80 percent) agreed to participate in the survey. About 60 percent of these firms are manufacturing companies, and the others are distributors, insurance companies, and other types of organizations. The personnel departments were asked to contact approximately equal numbers of car poolers, solo drivers, and public-transit users, and request them to answer a self-administered mail-back questionnaire that was hand delivered. Two thousand questionnaires were distributed, and 1020 were returned. After eliminating those questionnaires having a large amount of missing data, 822 questionnaires from 323 car poolers, 382 solo drivers, and 117 public transit users remained for analysis.

Because, in this sample, almost all car poolers owned at least one automobile while 75 percent of transit users did not, it was assumed that automobile ownership is a necessary condition for sharing a ride to work. For this reason, only data relating to car poolers and solo drivers were analyzed.

The method of contacting commuters through their employers, a method that is seldom used in transportation research, has certain advantages over traditional methods of data collection. The rate of return is relatively high (about 50 percent) as compared to mail surveys, and the cost for data collection is smaller than that required for home interviews.

Throughout this paper, the two basic modes of travel to work will be called drive alone and ride sharing and the two types of commuters solo drivers and car poolers respectively. The concept of ride sharing is restricted in the present study to the use of privately owned automobiles.

Three types of information were collected through the questionnaires: The first two are socioeconomic and travel characteristics, and the third is attitudinal data with respect to both ride sharing and driving alone.

A few words are desirable to describe the theoretical approach that guided the formulation of the attitudinal questions. There is a consensus among attitude researchers (9, 19, 21) that attitudes consist of one or more of three elements: (a) cognitive evaluations or beliefs, (b) affect (like-dislike emotive tendency), and (c) behavioral intention.

1. Cognitive Evaluations: It is hypothesized that an individual has a set of evaluative beliefs about the ride-sharing and drive-alone modes of travel to work with respect to such factors as cost, time saving, and convenience. Ten such attributes (expensive, comfortable,

pleasant, reliable, saves time, convenient, safe from crime, energy consuming, traffic problems, and pollution) were elicited through informal interviews conducted individually with a few car poolers and solo drivers. Cognitive evaluations of these attributes were measured on a seven-point scale from very low to very high.

2. Affect: This represents the positive or negative emotional predisposition toward an object and is presumed to be unidimensional, although it is possible that there is a complex cognitive structure underlying it. A measure of the affect toward ride sharing was obtained from the replies to the question "All things considered, which statement best describes how you like the idea of you being a member of a car pool?" The possible replies were like (extremely, moderately, or slightly), neither like nor dislike, and dislike (slightly, moderately, or extremely).

3. Intention: Ride-sharing intention refers to the stated plan of an individual to car pool and was measured by the replies to the question "How likely are you to join a car pool within the next two or three months?" The possible replies were definitely will, very likely, somewhat likely, cannot say, somewhat unlikely, very unlikely, and definitely will not. Intention is also hypothesized to be related to the cognitive profile of evaluations. It is a qualified expression of behavior: Given a span of time when behavior is likely to be manifested, the individual estimates at the beginning of the period of time whether or not he or she would behave in a certain manner. Because the shorter the period of time between intention and behavior, the more valid is the intention (21), the time span was limited to the next two or three months. (A theoretical structure of the relation between the cognitive evaluations, affect, and intention will be given later in this paper.)

RESULTS

Demographic and Travel Characteristics

A multivariate analysis of variance (MANOVA) test using Wilks lambda criteria (17) performed on 13 demographic variables showed that solo drivers differ significantly from car poolers [$F = 5.8$, degrees of freedom (df) = 13;691, $p < 0.001$, multivariate-explained variance = 9.8 percent]. [Detailed univariate descriptions have been given by Horowitz and Sheth (12).]

The socioeconomic variable that discriminates most strongly between the two groups is the size of the automobile owned: Car poolers own larger automobiles than do solo drivers. Other discriminant variables, although weaker, indicate that car poolers have worked longer at their present places of employment, are married rather than single, and have lived longer at their present residence. They are somewhat older and have larger families. The following variables do not discriminate between the two groups: (a) number of persons in household with driver's license, (b) number of automobiles owned, (c) age of the automobile that is used for the work trip, (d) sex, (e) income, (f) professional status, and (g) education.

Thus, when compared with those who drive alone in their private automobile, the typical car pooler in the Chicago area has a larger family and a larger car, has lived a longer time at his or her present residence, and has been working longer at the same place of employment. In short, the car pooler may be somewhat later in his or her life cycle than is the solo driver.

A MANOVA test using Wilks lambda criteria performed on seven travel characteristics, as reported by respondents in the survey, also showed that solo drivers differ significantly from car poolers ($F = 22.2$, df = 7;697, $p < 0.001$, multivariate-explained variance = 13.5 percent). The trip-to-work characteristics that discriminate between the two groups are (a) total travel cost for driving alone, (b) gasoline cost for driving alone, (c) travel time, (d) travel time of car poolers if they drove alone, and (e) distance to work. While the car pooler driving alone to work would require an average of 32.3 min, the solo driver needs only 26.5 min. The corresponding average distances from home to work are 26.1 and 17.9 km (16.3 and 11.2 miles). A car pooler spends an average of 34.3 min traveling to work. The characteristics that do not discriminate between the two groups are (a) distance to the nearest public-transportation station [5.9 km (3.7 miles)] and (b) walking time from parking area to work (approximately 3 min).

A few comments are in order: First, a discriminant analysis performed on both the demographic and travel characteristics showed that only 61.7 percent of the 705 commuters were correctly classified by the discriminant function. Since pure chance should give a correctly classified proportion of 50 percent, it follows that the demographic and travel characteristics add in only 11.7 percent of the cases, which is a small and negligible proportion. In summary, the multivariate-explained variance and (not independently) the results of the discriminant analyses indicate that demographic and travel characteristics are poor indicators of whether a commuter to work drives alone or shares a ride.

Second, solo drivers and car-pooling groups are better distinguished from each other by travel characteristics than by socioeconomic characteristics. That the socioeconomic variable that best distinguishes between the two groups is automobile size is consistent with the declining role of socioeconomic variables in the explanation and prediction of consumer choice among the relatively affluent, middle-class population (14, 24).

Finally, the results are partially inconsistent with the Voorhees and Associates study (22) of commuters on the Hollywood Freeway in the Los Angeles area. The only statistically significant discriminant variables the present study and the Voorhees study have in common are distance to work and travel time. The Voorhees study, in contrast to the present one, found that car poolers tend to be somewhat younger than are solo drivers. This discrepancy between the two studies may be attributed to the small number of car poolers (108) in the Voorhees study and to the different locations of the two studies. (It will be shown later, however, that attitudinal differences between car poolers and solo drivers are similar in the two studies and are perhaps more universal than are demographic and travel characteristics.)

Ride-Sharing Cognitive Profile

Of the 10 attributes of the cognitive-evaluation profiles, only the safe-from-crime one was found not to differentiate the two groups or to correlate with any of the other attributes. Figure 1 shows the means of the ride-sharing cognitive profile of the 9 remaining attributes for solo drivers and car poolers. Each attribute was rated on a semantic scale from one to seven where one meant very low, seven very high, and four was the neutral ground. A multivariate test performed on the whole vector of 9 attributes showed that the two groups of respondents differ significantly ($F = 30.6$, df = 9;695, $p \leq 0.001$, multivariate-explained variance = 28.4 percent).

The univariate tests lead to the following observations: First, solo drivers differ significantly from car poolers in their evaluation of ride sharing with respect to convenience, reliability, pleasure, comfort, and time (for each of these attributes, $p < 0.001$), but do not differ in their evaluation of ride sharing with respect to cost, energy, traffic problems, and air pollution.

Regardless of whether the differences between the two ride-sharing attitudinal profiles reflect the cause of commuting behavior or the result of it (dissonance phenomena), the results show the importance of the soft variables, such as convenience and reliability, and the perception of value of time in the perception of driving alone and car pooling.

Second, on the average, solo drivers tend to evaluate car pooling on all nine attributes at or just below the middle ground. This implies that solo drivers have a neutral position of ride sharing and a slight tendency to perceive it as inconvenient or not reliable. If solo drivers had a clearly negative attribute profile toward ride sharing, it might not be easy to change their position but, from a generally neutral position, a change in attitude might be achieved by advertisement and promotional means. [For a discussion of the relation between neutral attitudes and attitudinal change, see Howard and Sheth (13).]

Third, on the average, car poolers evaluate ride sharing as being clearly convenient, reliable, pleasant, comforatble, and economical. To a lesser extent, they perceive ride sharing as time saving and low in creating traffic problems and pollution. In this context, the ride-sharing cognitions of car poolers and solo drivers measured by Voorhees and Associates (22, Figure 12), were compatible with those obtained here despite the differences between the scales used in the two studies. The largest differences between car poolers and solo drivers were found by Voorhees in two semantic scales related to dependence on others.

An additional measure of attitudinal differences between the two groups of respondents based on the car pooling attributes has been obtained through a discriminant analysis. The discriminant function correctly classified 73.6 percent of the respondents; i.e., 23.6 percent in addition to the 50 percent that would be expected to be classified correctly by random assignments to groups, or about twice the discrimination beyond random that was achieved by the socioeconomic and travel characteristics.

Drive-Alone Cognitive Profile

The same nine attributes were also rated in the context of the drive-alone mode. The raw means are shown in Figure 2. A multivariate test performed on the vector of nine attributes showed that the two groups differ significantly, but to a lesser degree than in the case of the ride-sharing evaluation ($F = 10.4$, df = 9;695, $p \leq 0.001$, multivariate-explained variance = 11.8 percent).

An inspection of the individual means and the univariate tests leads to two principal observations. First, both groups of commuters perceive the drive-alone-to-work mode as being high on the qualitative attributes of convenience, reliability, comfort, and time saving. Second, solo drivers are more positive toward their own mode of transportation than car poolers are toward driving alone. This difference is statistically significant for all attributes ($p < 0.001$), with the exception of the public-cost attributes of energy, traffic, and air pollution.

These results suggest that, regardless of whether attitudes determine behavior or vice versa, cost is related to the choice between driving alone and ride sharing, but considerations of energy use, traffic, and pollution are not. (Attention will be given later to the question of how cost considerations differ from considerations of time, convenience, and such, in the determination of the choice between the modes.)

Affect Toward Ride Sharing and Intention to Share a Ride

Figure 3 shows the distribution of affect toward ride sharing for the solo drivers and the car poolers.

There is no need for statistical tests to show that the two groups are significantly differentiated by the affect measure. Solo drivers are divided along the continuum from like extremely to dislike extremely, with about 20 percent being neutral, but almost all car poolers are positive toward ride sharing.

Figure 4 shows the car-pooling-intention distribution for solo drivers. About 9 percent of the 376 solo drivers who answered the question stated a positive intention. Intention is a noncommitment behavior and may be grossly exaggerated, and so it can be assumed that less than 9 percent of the solo drivers surveyed intend to car pool regularly. However, the relation between the affect and intention measures and those of the cognitive perceptions could contribute to understanding the process through which modal choice is determined.

Differences Between Ride-Sharing and Drive-Alone Cognitive Profiles

The attribute evaluations were measured with respect to ride sharing and separately for the drive-alone mode. To obtain a more comprehensive grasp of the ride-sharing cognition and to relate it to both the affective and the intentional components when the drive-alone mode serves as a baseline, the difference between the drive-alone and ride-sharing evaluations was used as a measure of evaluation on each attribute. This difference was computed by subtracting the individual measures shown in Figure 1 from those in Figure 2 and will be denoted by δ_i, $i = 1, \ldots, 9$, where $\delta_i = x_{i,\text{drive-alone}} - x_{i,\text{ridesharing}}$ and $x_{i,\text{mode}}$ is the evaluation of the attribute i on the corresponding mode. The null hypothesis that the δ_i values are not different from zero has been rejected for both solo drivers and car poolers for all attributes ($p < 0.001$) with the exception of pleasant in the car-poolers group.

The δ_i values were, on the average, positive. That is, driving alone is perceived by a commuter, regardless of his or her actual mode of travel, as being more convenient, reliable, . . . , expensive, energy consuming, and such, than is ride sharing. This result warrants classification of the attributes into two groups: negative ride-sharing cognitions (δ_i, $i = 1, \ldots, 5$ for convenient, reliable, pleasant, comfortable and saves time respectively) and positive ride-sharing cognitions (δ_i, $i = 6, \ldots, 9$ for expensive, energy consuming, traffic problems, and pollution respectively). The profiles that determine the δ_i values, which are extracted from Figures 1 and 2, are shown in Figures 5 and 6 for solo drivers and car poolers respectively.

Another reason for the organization of the attributes into two groups arises from the results of factor analyses performed on the δ_i, $i = 1, \ldots, 9$ measures for each group of commuters. These analyses showed that two factors emerged and that they match the negative and positive cognitions. The factor that included the negative cognitions was labeled time-convenience and denoted T, and the factor that included the positive cognitions was labeled private and public cost and denoted C. These factors are summarized below.

Factor	i	Attribute
T (time-convenience)	1	Convenient
	2	Reliable
	3	Pleasant
	4	Comfortable
	5	Saves time

Factor	i	Attribute
C (private and public costs)	6	Expensive
	7	Energy consuming
	8	Traffic problems
	9	Pollution

Figure 1. Cognitive evaluations of ride sharing.

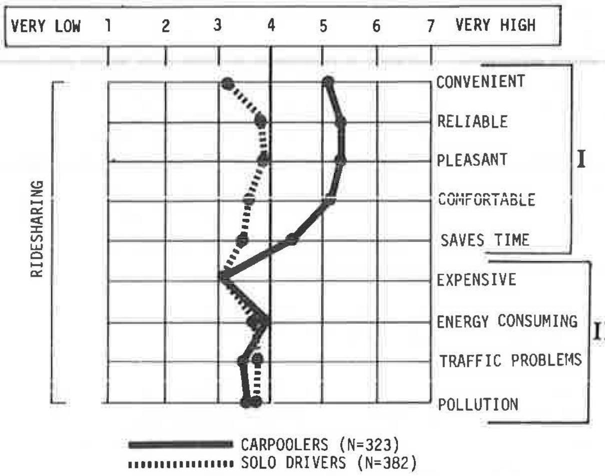


Figure 2. Cognitive evaluations of drive alone.

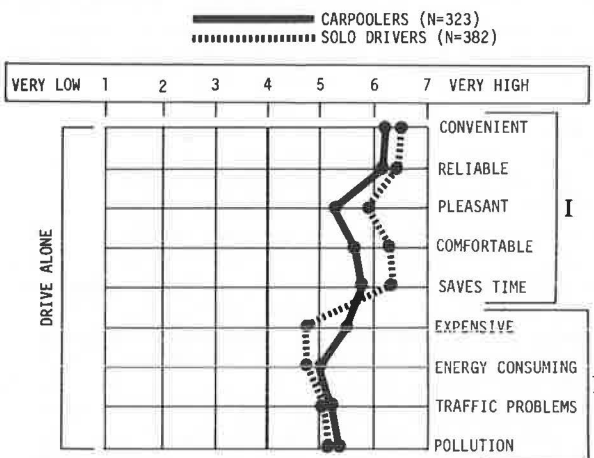
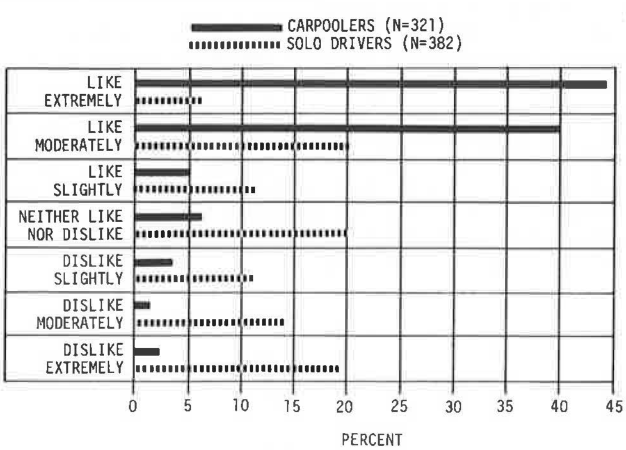


Figure 3. Affect toward ride sharing.



The output from the factor analyses are two factor scores for each individual, one for each factor. A factor score is a weighted average of the δ_i-measures of the corresponding factor.

MODELS RELATING COGNITION FACTORS TO AFFECT AND INTENTION

Research in both social psychology and consumer psychology has indicated that there is a linear additive relation between evaluations (cognition) and affect and intention (9, 21). This implies that positive and negative evaluations compensate for one another. Recently, however, several workers have expressed concern that a linear additive presumption may be a serious limitation to understanding attitudinal structure (7, 18).

Horowitz (11) has suggested an attitudinal ride-sharing model that allows for a noncompensatory relation: "Is it possible that evaluations interact among themselves so that a negative evaluation can reduce the intention to car pool regardless of the magnitude of the positive evaluation?"

To describe the model, assume that each individual is rated as either high or low on each of the two factors, according to whether his or her respective factor scores are higher or lower than the average score. The continuum could be divided into more than two parts, but this is sufficient for model testing. Then, each group (car poolers and solo drivers) will be segmented into four subgroups according to the combination of the two factors, as shown in Figure 7.

Consideration of the meaning of the two factors in relation to ride sharing and solo driving leads to the following interpretation of the cells. Cell [1, 2] includes those individuals who are more positive than the average toward ride sharing along both factors, cell [2, 1] includes those individuals who are negative toward ride sharing on both factors, and the other two cells include the obvious combinations of positive and negative factors scores.

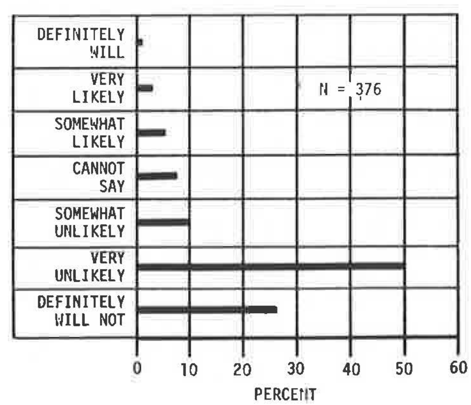
By following the notation introduced above and taking the position that affect is determined by the factors T and C, a linear-interactive model for affect is

$$A_{ijk} = \mu + T_i + C_j + \gamma_{ij} + \epsilon_{ijk} \qquad (1)$$

where

A_{ijk} = individual k's affect toward ridesharing, where his or her T-factor score is i (low, high) and his or her C-factor score is j (low, high),

Figure 4. Intention to car pool for solo drivers.



μ = mean affect over all four cells,
T_i = contribution of factor T to affect at level i,
C_j = contribution of factor C to affect at level j,
γ_{ij} = interaction between T_i and C_j levels, and
ϵ_{ijk} = individual k's error in cell [i, j].

An ordinary 2 × 2 analysis of variance (ANOVA) can be used to test the model. The use of the ANOVA depends on the statistical assumption that the ϵ_{ijk} are independent random variables normally distributed with constant variance. In the present application of ANOVA, these statistical assumptions are not a problem because the number of observations is relatively large. Use of the ANOVA requires independence between observations. Hence, it is necessary that different individuals belong in different cells. This assumption is clearly satisfied in the present design. The ANOVA allows simple, powerful tests for each of the T_i, C_j, and γ_{ij} terms separately.

A similar model can be written for intention, i.e.,

$$I_{ijk} = \mu + T_i + C_j + \gamma_{ij} + \epsilon_{ijk} \quad (2)$$

where I_{ijk} denotes individual k's intention to share a ride and the other terms are analogous to those in the affect model but refer to intention.

Test of Affect Model

Each respondent of the survey was assigned to one of the four cells according to his or her T and C-factor scores.

Figure 8 shows the affect means for each cell for solo drivers and car poolers separately and has two main results. First, the time-convenience factor, i.e, whether a respondent is categorized as low or high on T, is related to his or her affect to a larger extent than is the C-factor. This is seen by a comparison of the slopes of the lines and the distances between the lines for the car poolers and solo drivers. Second, that the lines are non-parallel suggests an interaction between the factors, especially for solo drivers.

Table 1 summarizes the test of the ANOVA model. The contributions of T and C are significant for both groups, but the F-ratios for T are markedly higher than those for C. The interaction term (T × C) is significant

Figure 5. Cognitions: solo drivers.

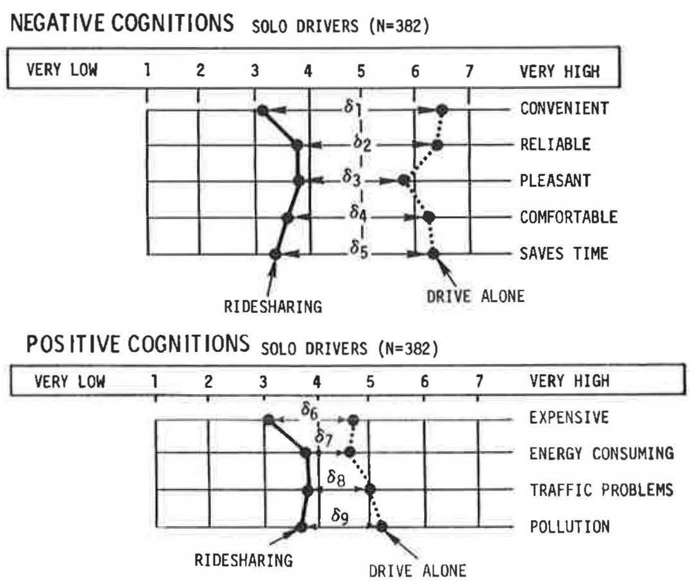
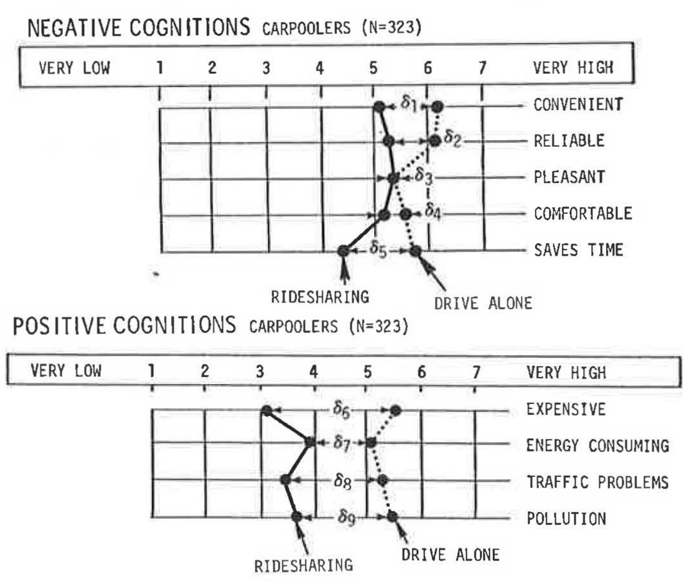


Figure 6. Cognitions: car poolers.



for the solo drivers, but not for the car poolers. The interpretation of the significant interaction is that those solo drivers who are high on T (a relatively large perceived difference in the time-convenience attributes between the two modes) have average affect toward ride sharing (dislike slightly), regardless of their perception of the private and public costs of the two modes. An interaction suggests a noncompensatory model. The parameters of the model for solo drivers are obtained directly from the cell means and are $\mu = 3.8$, $T_1 = 0.8$, $T_2 = -0.8$, $C_1 = -0.2$, $C_2 = 0.2$, $\gamma_{11} = \gamma_{22} = -0.2$, and $\gamma_{12} = \gamma_{21} = 0.2$. For carpoolers, the parameters are $\mu = 6.0$, $T_1 = 0.5$, $T_2 = -0.5$, $C_1 = -0.2$, $C_2 = 0.2$, $\gamma_{11} = \gamma_{22} = 0.1$, and $\gamma_{12} = \gamma_{21} = 0.1$.

The ratios between the absolute values of T and C are 4.0 and 2.5 for solo drivers and car poolers respectively. Since these ratios are larger than 1.0, it follows that the perceptions of ride-sharing disadvantages are more important, especially for solo drivers, than are the advantages in the determination of their attitude (affect) toward ride sharing to work.

Test of Intention Model

Despite the very skewed distribution of the intention-to-car-pool variable toward very unlikely, as shown in Figure 4, an additive compensatory model (Figure 9 and Table 1) was developed. The two lines are parallel, which suggests that there is no interaction. The factors T and C, however, significantly determine the intention, with factor T having a larger influence than C. The values of the parameters of the model are $\mu = 2.3$, $T_1 = 0.3$, $T_2 = -0.3$, $C_1 = -0.2$, $C_2 = 0.2$, and $\gamma_{11} = \gamma_{12} = \gamma_{21} = \gamma_{22} = 0$.

Market-Segmentation Technique

An aspect of enormous interest in the promotion of ride sharing is the identification of homogeneous market segments among solo drivers for whom different promotional methods will be desirable. Specifically, which socioeconomic variables are characteristic of solo drivers whose cognitive perceptions of ride sharing are maximum along its advantages (factor C) and minimum with respect to its disadvantages (factor T), i.e., those who are assigned to cell [1,2] of the cognitive factorial design? Recall that among the four cells of the design, cell [1,2] includes those respondents with the highest positive attitudes toward ride sharing with respect to affect and intention.

To answer this question, the univariate version of the MANOVA program was used with the same 2 × 2 factorial design as above, with the socioeconomic variables (including the distance to work) serving as the dependent variables (one variable was used for each analysis). All socioeconomic variables and the distance variable were tested.

The results showed that there are significant socioeconomic differences among the four cells. First, solo drivers who are more positive toward ride sharing than the average with respect to factor T (cells [1,1] and [1,2]) are from larger households, have worked a shorter time at their last place of employment, and have lived at their present residence a shorter time than

Figure 7. Segmentation based on cognitive differences.

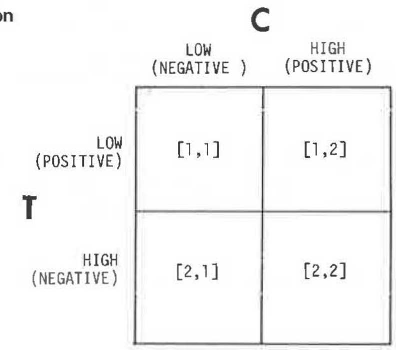


Figure 8. Affect toward ride sharing versus cognitive factors.

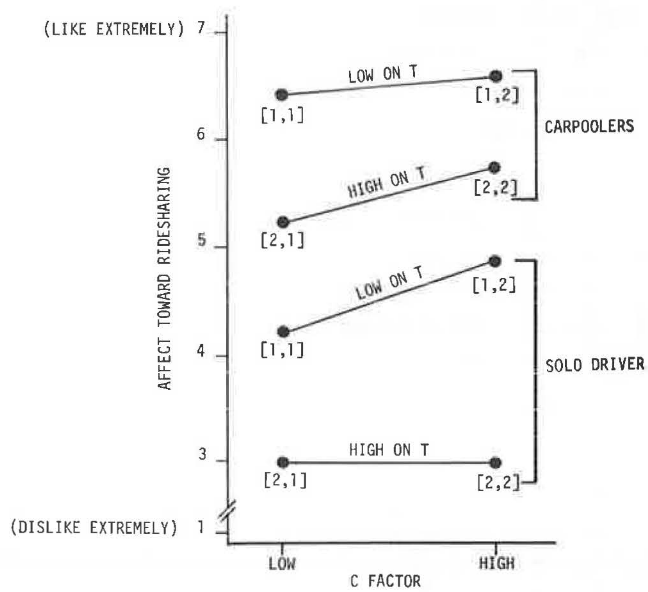
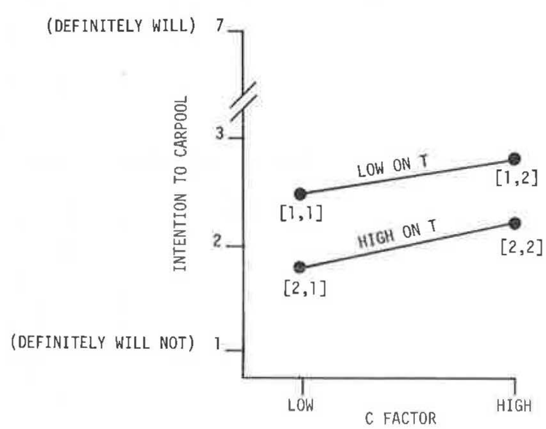


Table 1. ANOVA test results.

Model	Factor	Sum of Squares	df	Mean Square	F-Ratio	Prob ≤
Affect (solo drivers)	T	224.3	1	224.3	72.3	0.001
	C	12.3	1	12.3	3.9	0.05
	T × C	12.1	1	12.1	3.9	0.05
	Error	1171.8	378	3.1	–	–
	Total	1420.5	381	–	–	–
Intention (solo drivers)	T	40.8	1	40.8	26.9	0.001
	C	10.2	1	10.2	6.7	0.01
	T × C	0.1	1	0.1	0.0	N.S.
	Error	574.6	378	1.5	–	–
	Total	625.7	381	–	–	–
Affect (car poolers)	T	78.7	1	78.7	56.8	0.001
	C	8.3	1	8.3	6.0	0.015
	T × C	2.2	1	2.2	1.6	N.S.
	Error	442.0	319	1.4	–	–
	Total	531.2	322	–	–	–

Figure 9. Intention to car pool versus cognitive factors.



the other solo drivers. Second, those solo drivers who are more positive toward ride sharing with respect to factor C (cells [1, 2] and [2, 2]) typically live farther from their work, are males from households with more driver's licenses, and have higher educations, incomes, and occupation levels than the other solo drivers.

The picture of the ride-sharing target market, i.e., cell [1, 2], that emerges is one that includes employed individuals who have high socioeconomic status, as measured by education, income, and occupation; are from relatively large households; and have worked and lived at their last places of employment and residence respectively for a shorter time than the other solo drivers. These types of individuals are sensitive to the private and public costs of solo driving. A ride-sharing promotional campaign could address this segment of the population with issues related to both factors T and C, but the optimal strategy toward other types of commuters, the large majority of solo drivers, should be concentrated on issues related to the time-convenience factor.

CONCLUSIONS

1. For individuals who travel to work by private automobile, demographic and travel characteristics are poor indicators and predictors of the choice between driving alone and ride sharing.
2. The study of attitudes toward ride sharing and driving alone can provide results that relate to the question of how to develop ride-sharing strategies.
3. Solo drivers generally have a neutral attitude toward ride sharing, and a change in attitude might be achieved by proper promotional techniques.
4. With the exception of individuals having a relatively high socioeconomic status, appeals based on public-interest issues of energy, traffic, and air quality have little chance of changing attitudes toward ride sharing.
5. The perception of drivers toward time loss and the characteristics of convenience and reliability about ride sharing would have to change before their travel behavior would change. Perceptions of economic advantages have only minor roles in the determination of behavioral predispositions toward ride sharing.
6. To override negative perceptions toward time loss, convenience, and reliability about ride sharing, campaigns should emphasize positive aspects related to those characteristics that are unknown to the general public: The time spent for travel to work in a car pool as a passenger can be used for reading, sleeping, relaxing, or any other recreational activity that does not require much space and equipment. This type of approach toward the use of travel time has a good chance of success because of the increasing public awareness of the benefits of relaxation. Careful study is required to find ways to promote these ideas among both solo drivers and car poolers. A major component in the perceived inconvenience of car pooling is the difficulty of establishing contact with potential pool mates. These difficulties could be overcome by assistance in taking the initiative to form car pools and organization of car poolers on a face-to-face basis at the place of work. Ride-sharing promotion could use information about the longevity of car pools and the satisfaction of car poolers with the punctuality (reliability) of their pool mates.

REFERENCES

1. B. Barkow. Car Pooling, the Worm's Eye View. Paper presented at the Annual Convention of the American Psychological Association, Washington, D.C., 1976.
2. M. E. Ben-Akiva and T. J. Atherton. Choice-Model Predictions of Car-Pool Demand: Methods and Results. Paper presented at the Annual Convention of the American Psychological Association, Washington, D.C., 1976.
3. W. L. Berry. On the Economic Incentives for Commuter Car Pooling. Graduate School of Business Administration, Harvard Univ., Cambridge, Mass., PhD thesis, 1975.
4. D. P. Blankenship. Utilizing Focus Group Dynamics to Ascertain Rules for Social Interaction for Car Poolers. Orange County Transit District, Santa Ana, Calif., 1975.
5. An Evaluation of the SPRPC Car-Pool Public-Transit Program. School of Urban and Public Affairs, Carnegie-Mellon Univ.; Southwestern Pennsylvania Regional Planning Commission, Pittsburgh, 1975.
6. F. W. Davis and others. Ride Sharing and the Knoxville Commuter. Office of Environmental Affairs, U.S. Department of Transportation, Rept. TCUT-1-75, 1975.
7. G. S. Day. Evaluating Models of Attitude Structures. Journal of Marketing Research, Vol. 9, 1972, pp. 279-286.
8. K. J. Dueker and I. P. Levin. Car Pooling: Attitudes and Participation. Paper presented at the Annual Convention of the American Psychological Association, Washington, D.C., 1976.
9. M. Fishbein. Attitude and the Prediction of Behavior. In Readings in Attitude-Theory Measurement (M. Fishbein, ed.), Wiley, New York, 1967.
10. R. Herman and T. Lam. Car Pools at a Large Suburban Technical Center. Transportation Engineering Journal, Vol. 101, 1975, pp. 311-319.
11. A. D. Horowitz. An Attitudinal Model of Car-Pooling Behavior. General Motors Research Laboratories, Res. Publ. GMR-1969, Aug. 1975.
12. A. D. Horowitz and J. N. Sheth. Ride Sharing to Work: A Psychosocial Analysis, General Motors Research Laboratories, Res. Publ. GMR-2216, Jan. 1977.
13. J. A. Howard and J. N. Sheth. The Theory of Buyer Behavior. Wiley, New York, 1969.
14. G. Katona. Psychological Economics. Elsevier, New York, 1975.
15. D. C. Kendall. Car Pooling: Status and Potential. Transportation Systems Center, U.S. Department of Transportation, Cambridge, Mass., Final Rept. DOT-TSC-OST-75-23, June 1975.
16. J. B. Margolin and M. R. Misch. Incentives and Disincentives to Ride Sharing. Paper presented at the Annual Convention of the American Psychological Association, Washington, D.C., 1976.
17. D. F. Morrison. Multivariate Statistical Analysis. McGraw-Hill, New York, 1967.
18. P. S. Raju and J. N. Sheth. Nonlinear, Noncompensatory Relationships in Attitude Research. College of Commerce and Business Administration, Univ. of Illinois, Working Paper 176, April 1974.
19. M. J. Rosenberg. A Structural Theory of Attitude Dynamics. Public Opinion Quarterly, Vol. 24, 1960, pp. 319-340.
20. S. Rosenbloom and N. J. Shelton. Car-Pool and Bus Matching Program for the University of Texas at Austin. Graduate Program in Community and Regional Planning, Univ. of Texas at Austin, Res. Rept. 11, Sept. 1974.
21. J. N. Sheth. A Field Study of Attitude Structure and the Attitude-Behavior Relationship. In Models of Buyer Behavior (J. N. Sheth, ed.), Harper and Row, New York, 1974.
22. A Study of Techniques to Increase Commuter-

Vehicle Occupancy on the Hollywood Freeway. Alan M. Voorhees and Associates, Nov. 1973.
23. A. M. Zerega and R. B. Ross. Application of Conjoint Measurement Techniques in Evaluating Car-Pooling Policies. Paper presented at the Annual Convention of the American Psychological Association, Washington, D.C., 1976.
24. D. Yankelovich. New Criteria for Market Segmentation. Harvard Business Review, Vol. 42, 1964, pp. 83-90.

Priority Lanes on Urban Radial Freeways: An Economic-Simulation Model

Kenneth A. Small, Department of Economics and Transportation Program, Princeton University

A simulation of the effects of opening a priority lane on a commuter-oriented freeway is carried out by combining a simple deterministic queuing model of traffic flow with a disaggregate model of modal choice. This permits iterative determination of a supply-demand equilibrium and a precise definition of the resulting benefits within the framework of cost-benefit analysis. By varying the assumptions parametrically, illustrative results for a wide variety of cases are obtained. The benefits are substantial for those cases where initial congestion is heavy. The combination of the rigorously derived objective function and the model of modal choice constitutes a proposed methodology for analyzing highway management policies that could be adapted for use in more detailed engineering studies of particular facilities. The results given here, although derived from a highly simplified model of traffic flow over a peak period, suggest the results that can be expected from such applications.

Our understanding of priority-lane operations and car-pooling behavior has grown rapidly in recent years because of the urgent need for public-policy guidelines. Sophisticated traffic-flow models (10, 13, 14) now permit detailed investigation of patterns of traffic flow under various circumstances. A flurry of activity among demand modelers has produced a number of disaggregate modal-choice models that predict the response of voluntary car pooling to various incentives (Ben-Akiva and Atherton, in a paper in this Record).

Each of these sides of the analysis depends on the other: Traffic-flow models make predictions that are contingent on the volume and mix of traffic, and forecasts from demand models must take as given the costs and levels of service encountered by the users of each mode. The use of either procedure alone may be valid only within an unknown and possibly narrow range of conditions.

Therefore, the need is for an integrated model that determines levels of service and levels of demand simultaneously. Such a model consists conceptually of nothing more than that most basic tool of microeconomic analysis, the supply-demand equilibrium. The demand side is provided by the demand-forecasting model, which predicts the quantities of various types of highway services that individuals will choose, given their prices in terms of monetary cost and level of service. The traffic-flow model and the cost information determine the price that must be paid by users to obtain a certain volume of peak-hour highway services and thus constitute the supply side of the equilibrium.

This paper describes such a model and demonstrates its usefulness by analyzing the impact of a priority lane on an idealized radial freeway subject to peak-period congestion by commuters. The model and results are described more fully by Small (12).

A secondary purpose is to show that the incorporation of a disaggregate demand model facilitates a clear and theoretically rigorous definition of user benefits that is consistent with accepted principles of cost-benefit analysis and to calculate these benefits for the policies considered to form some generalizations about the desirability of priority lanes as public policy.

The model focuses on the supply of and demand for the services of a section of radial freeway during the morning and afternoon peak commuting periods. Congestion is explicitly modeled only on the freeway section itself. All characteristics of the access and distribution networks are assumed to remain constant and enter the model as determinants of demand for the freeway study section.

SUPPLY MODEL

The idealized highway section to be considered is a 10-km (6-mile) length of freeway with no entrance or exit ramps that is used only by commuters. All access to this line-haul section is at one end and all egress is at the other, with the direction reversing from morning to evening. The collection of commuters at the access end and their distribution at the egress end take place on a variety of roads that may include extensions of the study section. [Access and egress are described in a disaggregate manner in the next section of the paper; this section describes traffic flow and cost assumptions for the 10-km (6-mile) section itself.]

Traffic flow on this line-haul section is described by assuming a uniform speed of S_0 km/h ($0.62S_0$ mph), except for the delay caused by deterministic queuing behind a a single bottleneck of capacity C vehicles/h. That is, if t_1 is the time of day at which traffic volume [$D(t)$] entering the freeway first exceeds C, then travel time (T) (in minutes) over the section for a vehicle entering at a later time t, providing the queue does not dissipate prior to t, is

$$T(t) = (600/S_0) + (60/C) \int_{t_1}^{t} [D(t') - C]\, dt \qquad (1)$$

In terms of queuing theory, the integral gives the queue length in vehicles, and (60/C) is the service time in minutes.

This model has been used by May and Keller (9) to analyze the San Francisco-Oakland Bay Bridge. To apply it to the typical radial freeway may seem a bit more tenuous, but it duplicates remarkably well the actual travel times observed during the afternoon rush hour on an 18-km (11-mile) section of I-80 on the eastern side of San Francisco Bay. This study used the results of an origin-destination study to compute net demands at 15-min intervals for a particular three-lane subsection that appears to be the chief bottleneck (1, pp. 8 and B-2).

The parameters S_0 and C were adjusted by trial and error to obtain a satisfactory fit, which gave S_0 = 89.7 km/h (53.8 mph) and C = 1770 automobile-equivalents/h/lane. These values are retained for the present study, with a bus assumed to cause congestion equivalent to 1.6 automobiles (3, p. 257). The institution of a priority lane is assumed to result in two separate traffic streams, each governed by this type of queuing analysis.

The primary endogenous service-level variable is taken to be the average travel time ($\overline{T}$), in minutes, over a peak period of duration (W) with uniform demand volume (D). This may be easily calculated from Equation 1 to be

$$\begin{aligned} \overline{T} &= 6.7 \qquad (D \leqslant C) \\ &= 6.7 + [(D/C) - 1](60W/2) \qquad (D > C) \end{aligned} \tag{2}$$

The assumption that commuters ignore variations in queuing delay over the peak period probably does not affect the results significantly, because in reality commuters tend to adjust their times of travel to minimize that variation. The assumption that the peak-period duration is fixed, however, is potentially important and will be discussed later. Also, the benefits of congestion-reducing policies will be somewhat underestimated because the effect on those individuals who arrive after the peak period but before the queue is dissipated is ignored.

The cost of providing automobile or bus service over the line-haul section are estimated as realistically as possible as a function of $\overline{T}$, by relying on methodology developed by Keeler and others (4) and Small (12). All costs are given in 1972 prices and therefore precede the rapid increases in gasoline and labor costs of the past 4 years.

Costs of automobile travel are assumed to include only the maintenance and operating costs for a compact automobile and to vary with average speed proportionally to fuel consumption. This gives a relation that shows costs per vehicle to be approximately constant at 2.3 cents/km (3.8 cents/mile) for freeway speeds between 67 and 100 km/h (41 and 62 mph), and to rise fairly rapidly outside that range. (This fuel-economy relation was measured under actual freeway conditions and hence reflects the increased congestion that causes lower average speeds.)

To adequately describe the changes in bus service resulting from changes in freeway speeds and aggregate ridership levels would require a model of bus operational policy that predicts route density, headway, and fare. Such models have been developed by assuming either some kind of social optimization (4) or a profit maximization subject to fixed fare (5), but it is not clear that either assumption characterizes actual bus agencies. Instead, it is assumed here that buses are added to existing routes in such a way that occupancy and waiting time remain constant, but that fares are adjusted for any cost changes due to changes in $\overline{T}$.

Studies of bus operations (4, 12) gave assumed agency costs of 11.5 cents/vehicle·km for maintenance and operation, \$15.52/vehicle·h for labor on peak-period runs, and \$5898/vehicle·year for capital cost. Each daily peak-period run was then assumed to require a 20-km (12-mile) round trip with a revenue-haul taking time of $\overline{T}$ and an empty back-haul taking time of 6 min. Adding 10 percent to these figures for miscellaneous extra running time, assuming a bus occupancy of 37, and distributing capital costs over 255 working d/year gives costs of (13.18 + 1.056$\overline{T}$) cents/one-way passenger trip.

In summary, the supply model predicts, for a given total passenger volume and modal split, the line-haul travel times and costs faced by automobile and bus commuters using a given section of freeway. To complete the picture requires a demand model that uses these times and costs to predict the modal volumes.

DEMAND MODEL

The demand for use of a section of radial freeway by commuters is modeled here as a problem of modal choice with a fixed number of total trips. The use of the disaggregate demand approach has two steps. The first is the specification and calibration on some survey sample of a behavioral modal-choice model. This will predict the probability that an individual with given observable socioeconomic characteristics and transportation opportunities will choose one of four modes: automobile noncar pool, car pool, bus with walk access, or bus with automobile access.

The second step is a description of the distribution of these socioeconomic characteristics and transportation opportunities among the population of commuters who use the freeway section in question; this is done by a forecasting sample believed to be representative of such commuters. This step is absolutely essential for unbiased forecasts from the disaggregate modal-split model (8); in the present context, it is here that the characteristics of access to and egress from the freeway study section are accounted for. Also, although in the present paper they are overlapping subsets of a common data base, the forecasting and calibration samples may be entirely distinct; the former must provide a representative selection of the underlying preferences of the population whose behavior is in question, whereas the latter must be representative of their socioeconomic and locational situations.

In this paper, both the calibration and forecasting samples are subsets of a sample of 213 commuters in the San Francisco Bay Area, who were surveyed in 1972 as part of the project reported by McFadden (7, pp. 315-319). The project staff combined the survey information with extensive highway and bus transit data to obtain a complete description of the sample individuals in terms of socioeconomic and transportation variables.

For the calibration sample, the full sample was narrowed to 161 by excluding individuals who walked or bicycled to work or who were captive automobile users because of regular use of the household automobile at work or no feasible bus service available. For the forecasting sample, on the other hand, the desire was to find a sample representative of the users of a typical 10-km (6-mile) length of radial freeway. Since it happened that the original sample of 213 was drawn primarily from potential users of major freeway routes, the present purposes were served simply by narrowing it to those whose trips, if taken by automobile, would involve a substantial length of freeway. There was no requirement that actual express bus service on the freeway be available to the individual; in the absence of such service the forecast uses existing local service as the basis for the assumed travel time and fare at which express service could be instituted. The resulting sample, after eliminating a few for whom no bus service of any kind existed, had 118 individuals.

Calibration of Behavioral Modal-Choice Model

The theory behind the behavioral model of modal choice used here is that the ith individual perceives a utility from a trip on the mth mode of

$$V_m^i = W_m(S^i, x_m^i) + \epsilon_m^i \tag{3}$$

where W_m are universal functions of the socioeconomic characteristics S^i and the transportation variables x^i_m, and where ϵ^i_m express the unobservable idiosyncratic tastes of individual i. McFadden (2, 6) has shown that, if ϵ^i_m is assumed to have a Weibull population distribution independent of m, then the probability that an individual will choose mode m, given his observable characteristics S^i and x^i_m, is given by the logit formula:

$$P^i_m = \exp(W^i_m) \Big/ \sum_n \exp(W^i_n) \tag{4}$$

where $W^i_m = W_m(S^i, x^i_m)$. Given observations on S^i and x^i_m for a sample of individuals and on their actual modal choices, the maximum-likelihood procedure described by McFadden (6) may be used to estimate the parameters of the functions W_m.

The functional form of W_m must be specified in advance, which involves a number of complex issues (12, 15) whose resolution is only summarized here. Socioeconomic variables are included only insofar as there is a priori reason to believe that they serve primarily as indicators of tastes and are exogenous over a time span of several years. Costs are entered as a fraction of the marginal posttax wage rate, given by $w = w_0(1 - \tau)$, where w_0 is the actual wage rate and τ is the tax rate of the income tax bracket determined by the total income of the family. The model therefore implicitly estimates values of time as fractions of this wage. Separate travel-time components are distinguished insofar as the sample size and the quality of the data permit.

The car-pool mode is defined as participation (either as driver or passenger) in a trip by automobile containing three or more people. Rather than adding an arbitrary amount of time to the car-pool trip to account for the extra driving, loading, and unloading, a single car-pool dummy variable permits the calibration procedure itself to estimate an implicit time penalty, which includes both actual extra time and a penalty in time-equivalents for whatever other undesirable features, such as scheduling inconveniences and personal incompatibility, that car pooling may have.

The estimated coefficients are shown in Table 1. Their magnitudes and signs agree with intuition and with other results, except for the first-wait time, which has an excessively high value (469 percent of the marginal posttax wage rate) that is partly responsible for the decision not to incorporate headway changes into the supply model. On-vehicle time is valued at 54 percent of the wage and walk time at 83 percent. A transfer (including the time associated with it) is valued at 13.6 min of on-vehicle time, whereas the inconvenience of car pooling (relative to lower occupancy automobile) is, depending on age and the hours of work, as objectionable as 100 to 160 min of round-trip on-vehicle time!

Such a large natural barrier to car pooling should not be construed as evidence of the hopelessness of incentive policies; indeed, it may suggest the presence of important omitted determinants (e.g., advertising and computer matching services) that are subject to policy manipulation. Other variables tried as explanatory for the car-pool versus non-car-pool automobile choice, but found to give statistically insignificant coefficients, were income, children, length of residence in the neighborhood, and number of workers in the family.

Forecasting Aggregate Modal Split

Turning now to the second step in the construction of demand for the freeway study section, Table 2 shows the main features of the forecasting sample. A typical one-way trip by automobile is 28.4 km (17.5 miles) long and takes 31.5 min; by bus, the same trip takes 49 min of on-vehicle time, one transfer, and (for the morning trip) a 19-min initial headway. As expected from the relatively high incomes represented, this largely suburban sample generates modal-split forecasts that are heavily biased toward the automobile.

To forecast modal split as a function of line-haul times and costs for the various modes, the utilities (W^i_m) are first computed from Equation 5 for each forecasting sample member. The travel times and costs assigned to him or her in the forecasting sample are assumed to include (whatever the mode) the equivalent of a 10-km (6-mile) line-haul trip at 67 km/h (40 mph), the approximate average peak-hour speed observed on major Bay Area radial freeways in 1972. W^i_m is then modified for differences in the line-haul times and costs from this base condition, and the individual modal-choice probabilities are calculated from the logit equation 4. These probabilities are then averaged over the forecasting sample and adjusted to account for the captive automobile commuters, who were excluded from the calibration sample, to obtain aggregate modal-choice probabilities, which are now a function solely of the line-haul times and costs.

The resulting modal-split functions are fairly insensitive to line-haul times and costs. Under 1972 base conditions, 75 percent chose one of the automobile modes. Increasing one-way automobile times by 24 min reduces this to 65 percent; alternatively, the same reduction could be achieved by a one-way toll of $1/automobile.

EQUILIBRIUM RESULTS

The supply and demand models described in the previous sections were computerized, and an iterating algorithm was written to determine the equilibrium values of modal-split and line-haul times and costs for a given total passenger volume. Some results are given in Table 3 for three values of passenger volume, chosen to be representative of conditions that lead under base conditions (no priority lane) to moderate, heavy, and very heavy congestion (defined respectively as one-way delays of 6.8, 15.6, and 24.0 min).

Except at conditions of very heavy initial congestion, a bus-and-car-pool lane increases the queuing in the other lanes, in spite of a substantial decrease in traffic due to induced modal shift. A bus-only lane is, of course, even worse in this respect. To evaluate the effect of divisible lanes, the model was run with the assumption that the total capacity could be divided into any fraction desired, that fraction being set so as to just avoid queuing among priority vehicles. The results indicated that nonpriority queuing is much less severe and that ideally only 8 to 12 percent of total capacity should be allocated to priority vehicles.

BENEFITS

The claim that the incorporation of a behavioral demand model permits the definition and computation of rigorous measure of benefits is based on the theory of cost-benefit analysis, in which benefits are defined as the sum over all individuals of the amounts of money each would be willing to pay for the change plus the identifiable money flows to relevant parties, e.g., government tax collections.

To define the willingness of an individual to pay for a change in the transportation environment facing him or her, let V^{i*} represent the utility actually achieved by choosing the best mode. From Equation 3

$$V^{i*} = \max_m V^i_m = \max_m (W^i_m + \epsilon^i_m) \tag{5}$$

It is assumed that the marginal utility of money (λ^i) is given by the coefficient of travel cost in the behavioral demand model, which is equal to the estimated coefficient divided by the individual's wage (w^i). The incremental willingness to pay for a change that alters the utilities (W^i_m) is then

$$dB^i = (1/\lambda^i)dV^{i*} = (1/\lambda^i)dW^i_k \quad (6)$$

where k is the mode actually chosen.

Everything in Equation 6 except k is observable for an individual in the forecasting sample. The nature of this stochastic utility model is such that the mode an individual will choose cannot be predicted with certainty, and thus his or her benefits from changes that affect modes differentially cannot be predicted. However, for aggregate purposes, it is sufficient to know the expectation of benefits:

$$E[dB^i] = (1/\lambda^i)\sum_m P^i_m dW^i_m \quad (7)$$

where P^i_m is given by the logit formula (Equation 4). Integration of Equation 7 gives the following index of direct benefits:

$$B^i = (1/\lambda^i)\ln\left[\sum_m \exp(W^i_m)\right] \quad (8)$$

In the course of computing the modal-choice probabilities for each individual in the forecasting sample, it is simple to compute this benefit index and aggregate it over the sample. To this value, must be added the benefits accruing to captive automobile users, which are calculated directly from Equation 6 by assuming that the mode chosen is non-car-pool automobile. Finally, the changes in gasoline tax revenues are added to obtain the direct benefits given in Table 3.

This measure of social benefits is termed direct because it excludes a number of potentially quantifiable effects that are external to the present model, but that may be quite important for actual policy purposes. These include parking subsidies, subsidies to bus feeder routes, congestion costs outside the central business district, changes in bus headways or route densities, automobile capital and accident costs, and air pollution. Estimates of these indirect benefits (12, pp. 218-224) indicate that including them would greatly reinforce the case for any policy that reduces automobile traffic.

The results for the bus-and-car-pool lane in Table 3

Table 1. Modal-choice models: estimated coefficients.

Independent Variable (round trip)	Coefficient	Standard Error
Cost[a] ÷ marginal posttax wage, ¢/min	-0.0413	0.0116
On-vehicle time, min	-0.0224	0.0120
Walk time, min	-0.0343	0.0162
First-wait time, min	-0.1938	0.0600
Number of transfers	-0.3043	0.1982
Mode 3 dummy	-1.25	0.48
Automobile dummy[b]	-5.23	1.39
Family income[b], $000s (ceiling of 10)	0.310	0.112
Children under 18 living at home (dummy)[b]	-0.645	0.540
Length of residence in neighborhood[b], years	0.119	0.042
Respondent's age ≥ 45 (dummy)[b]	-0.660	0.574
Car-pool dummy[c]	-2.44	0.54
Respondent's age ≥ 45 (dummy)[c]	-1.138	0.691
Respondent works standard hours (dummy)[c,d]	0.098	0.302
Likelihood ratio index[e]	0.448	—
Percentage correctly predicted[f]	71.4	—

Notes: Sample size = 161.
Mode 1 = automobile with < 3 occupants; mode 2 = bus with walk access; mode 3 = bus with automobile access; mode 4 = car pool with > 3 occupants.

[a] Cost for the automobile modes consists of maintenance and operating costs [at 3.3¢/km (5.3¢/mile)] plus tolls and parking, all multiplied by an expected share of 1/1.11 for mode 1 or 1/3.52 for mode 4.
[b] The variable is as described on modes 1 and 4, zero on other modes.
[c] The variable is as described on mode 4, zero on other modes.
[d] If official work start time is 7:45 to 9:15 a.m. and quit time is 4:15 to 5:45 p.m., this variable is 2. If one of the above holds, it is 1. If neither holds or there are no official times, it is 0.
[e] The likelihood ratio index is the percentage increase in the log likelihood when maximized over its value with all coefficients 0.
[f] A case is correctly predicted if the mode actually chosen is the one with the highest predicted probability.

Table 2. Forecasting-sample summary statistics.

Variable	Mean	Standard Deviation
Automobile round trip		
Distance, km	58	20
On-vehicle time, min	63	26
Parking cost, ¢/vehicle/d		
All trips	41	77
Excluding free parkers	134	84
Bus round trip		
On-vehicle time, min	98	30
Walk time, min	28	19
Number of transfers	2.1	1.4
Fare, ¢	142	64
Other		
First bus headway home to work, min	19	16
Family income, $000s/year	14.8	7.0
Marginal posttax wage, $/h	4.21	1.94

Note: 1 km = 0.62 mile.

Table 3. Equilibrium results.

Policy	Total Passenger Volume per Hour per Lane	Fraction of Total Capacity Used for Priority Lane	One-Way Queuing Delay (min): Automobile	One-Way Queuing Delay (min): Bus	Modal Split (%): Automobile Noncar Pool	Modal Split (%): Automobile Car Pool	Modal Split (%): Bus	Priority Lane Capacity Use (%)	Direct Benefits Relative to Base Case (¢/passenger/d)
Base case: no priority									
Moderate congestion	3160	0.33	6.8	6.8	65	10	25	—	—
Heavy congestion	3580	0.33	15.6	15.6	64	11	25	—	—
Very heavy congestion	4000	0.33	24.0	24.0	64	12	24	—	—
Bus priority									
Moderate congestion	3160	0.33	20.0	0.0	53	8	39	9	-65
Heavy congestion	3580	0.33	25.2	0.0	50	7	43	11	-13
Very heavy congestion	4000	0.33	29.7	0.0	47	7	46	14	43
Bus and car-pool priority									
Moderate congestion	3160	0.33	15.6	0.0	52	14	34	29	-28
Heavy congestion	3580	0.33	19.8	0.0	49	16	36	36	32
Very heavy congestion	4000	0.33	23.7	0.0	46	17	37	44	94
Bus and car-pool priority; divisible lanes									
Moderate congestion	3160	0.08	4.3	0.0	62	11	28	100	37
Heavy congestion	3580	0.10	9.4	0.0	57	12	30	100	89
Very heavy congestion	4000	0.12	14.2	0.0	53	14	33	100	143

range from direct benefits of -28 to +94 cents/commuter/d, depending on the degree of initial congestion. The benefits could be much greater, though the induced modal shift would be smaller, if lanes were perfectly divisible, so as to eliminate the waste of capacity in the underused priority lane. There is a large potential payoff for the development of methods to allow a priority queue to bypass without using an entire lane, and substantially more engineering effort should be devoted to this aspect of the problem. Some tentative suggestions have been made by Small (12, pp. 76-82).

It must not be thought, however, that either full use of the priority lane or an absence of increased queuing is a prerequisite for positive benefits from a priority-lane operation. For example, under initially heavy congestion, a bus-and-car-pool lane carries traffic equal to only 36 percent of its assumed bottleneck capacity and increases the one-way queuing delay for nonpriority vehicles by 4.2 min, yet the direct benefits are positive. This is because the queuing delays are reassigned to different vehicles in an economically efficient manner: Those vehicles whose occupants in aggregate possess a higher value of time (per unit of road capacity used) are permitted to go faster at the expense of others because the benefits to the former outweigh the disbenefits to the latter in the impersonal scales of cost-benefit analysis.

SOME PERSPECTIVES ON RESULTS

Several points may be made in interpreting the usefulness of the model and results presented here.

First, are the potential benefits from priority lanes large? Consider the case of a bus-and-car-pool lane for a six-lane facility initially subjected to heavy congestion. If 255 working d/year are assumed, the estimated direct benefits of 32 cents/passenger/d equal \$1.75 million/year. Compared to the total round-trip costs of commuting, benefits of 32 cents/passenger would appear to be significant, although far from overwhelming. Compared to implementation costs, \$1.75 million/year appears very large. The Voorhees and Associates study (16, p. 25) estimated signing costs for a 19-km (12-mile) priority lane on each side of the I-90 Memorial Shoreway in Cleveland at \$235 000 capital expenses plus annual maintenance and operating costs of \$14 000. Even if special entrance ramps of the type built for a contraflow lane on I-495 in New Jersey (11, p. 23) are added on each side, the capital costs are only about \$500 000. If this is annualized liberally with a capital-recovery factor at a 10 percent interest rate and a 15-year lifetime, the total annual costs are \$80 000, which is an order of magnitude below the potential benefits.

Second, the present model understates the benefits of reducing automobile traffic. Furthermore, the conventional priority lane analyzed here is not necessarily the most favorable configuration for all situations. Other alternatives include contraflow lanes, extra lanes in a median strip, and priority metering at entrance ramps. All of these require greater initial expense, but they may provide considerably greater benefits because they cause less disruption to nonpriority flow. With some modification, the model could be applied to the analysis of such policies.

Another alternative is congestion pricing, in which a peak-period toll, equal to the marginal cost that each vehicle inflicts on all other users through increasing congestion, is charged. In the present model, this marginal cost is the value of the additional travel time and running cost imposed by a user on all those behind him or her in the queue. This alternative was analyzed by using the model and had benefits that exceeded those of a priority lane by about \$1/passenger/d. For the heavy-congestion case, a round-trip toll of \$2.22/automobile eliminated queuing delays entirely and resulted in increases of 13 and 1.5 percent in bus and car-pooling frequencies respectively.

Finally, the effects of some of the simplifying assumptions of the model should be explored. First, the neglect of the speed versus flow relation on those parts of the freeway not affected by queuing appears to affect the results very little, because the overall travel time is much more sensitive to queuing than to nonsaturated speed reductions. Second, the model overstates the changes in congestion levels during the peak period by not allowing for alternative routes and times of day. Third, by excluding nonwork travel, the model probably overstates the modal shifts induced.

CONCLUSIONS

This paper is in large part intended as a contribution to the development of methodologies for evaluating urban-highway operating policies. It appears both desirable and feasible to analyze such policies in an equilibrium context, in which the interaction between traffic-flow relations and demand characteristics is explicitly recognized and in which the benefits to individuals can be defined and evaluated in a rigorous way. The particular model described here is one way to approach this goal, and the results suggest what may be expected from more detailed applications. Either the supply or demand sides of the model can be made more complex and case specific.

Other limitations of the model can be removed only with greater difficulty. To incorporate nonwork trips would require more complex demand modeling. To eliminate the assumption of a fixed-duration peak period would require some behavioral description of individual decisions on the timing of their trips. Finally, the assumption of a fixed number and location of commuters prohibits consideration of the longer range effects of various policies on the shape of an urban area. The further development of the present model on these lines would be both challenging and rewarding.

ACKNOWLEDGMENTS

I am indebted to Theodore Keeler, Daniel McFadden, Adolf May, Moshe Ben-Akiva, and anonymous referees for invaluable advice at various stages of this research. The research was supported in part by the Research Applied to National Needs Program of the National Science Foundation and by the Sloan Foundation through a grant to the Princeton University Department of Economics. The results and views expressed are not necessarily those of the sponsoring agencies.

REFERENCES

1. J. F. Aidoo, R. W. Goedhart, and A. D. May, Jr. Bay Area Freeway-Operations Study: On the Eastshore Freeway (I-80) Northbound. Institute of Transportation and Traffic Engineering, Univ. of California, Berkeley, Final Rept., Pt. 3, 1970.
2. T. Domencich and D. McFadden. Urban Travel Demand: A Behavioral Analysis. North-Holland, Amsterdam, 1975.
3. Highway Capacity Manual. HRB, Special Rept. 87, 1965.
4. T. E. Keeler and others. The Full Costs of Urban Transport, Part III: Automobile Costs and Final Intermodal Cost Comparisons. Institute of Urban and Regional Development, Univ. of California, Berkeley, Monograph 21, 1975.
5. P. M. Lion. Transit-Systems Analysis for Opera-

tional Decisions, Transportation Program, Princeton Univ., Rept. 76-TR-5.
6. D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers in Econometrics (P. Zarembka, ed.), Academic, New York, 1973, pp. 105-142.
7. D. McFadden. The Measurement of Urban Travel Demand. Journal of Public Economics, Vol. 3, 1974, pp. 303-328.
8. D. McFadden and F. Reid. Aggregate Travel-Demand Forecasting From Disaggregated Behavioral Models. TRB, Transportation Research Record 534, 1975, pp. 24-37.
9. A. D. May, Jr., and H. E. M. Keller. A Deterministic Queuing Model. Transportation Research, Vol. 1, 1967, pp. 117-128.
10. R. D. Minister and others. A Computer-Simulation Model for Evaluating Priority Operations on Freeways. HRB, Highway Research Record 461, 1973, pp. 35-44.
11. Results of Case Studies and Analysis of Busway Applications in the United States. R. H. Pratt and Associates; U.S. Government Printing Office, 1973.
12. K. A. Small. Bus Priority, Differential Pricing, and Investment in Urban Highways. Institute of Transportation Studies, Univ. of Calfifornia, Berkeley, Ph.D. thesis and Working Paper 7613, 1976.
13. G. A. Sparks and A. D. May, Jr. A Mathematical Model for Evaluating Priority-Lane Operations on Freeways. Institute of Transportation and Traffic Engineering, Univ. of California, Berkeley, 1970.
14. W. A. Stock and others. Priority-Lane Operations on the San Francisco-Oakland Bay Bridge. Institute of Transportation and Traffic Engineering, Univ. of California, Berkeley, 1971.
15. K. Train and D. McFadden. The Measurement of Urban Travel Demand II. Institute of Transportation Studies, Univ. of California, Berkeley, Working Paper 7517, 1975 (revised).
16. Feasibility and Evaluation Study of Reserved Freeway Lanes for Buses and Car Pools. Alan M. Voorhees and Associates, 1971; NTIS, Springfield, Va.

Choice-Model Predictions of Car-Pool Demand: Methods and Results

Moshe Ben-Akiva, Center for Transportation Studies and Department of Civil Engineering, Massachusetts Institute of Technology
Terry J. Atherton, Cambridge Systematics, Inc., Cambridge, Massachusetts

The results of a number of car-pool strategies were predicted by using disaggregate choice models. Car pooling is explicitly considered as an alternative mode only for work trips. However, the effects of car-pooling incentives on interdependent travel choices and vice versa are also predicted. Forecasts are made by applying the models to each household individually, using revised values of the appropriate independent variables to simulate the particular transportation alternative being analyzed. These household predictions are then summed to represent predicted areawide changes in travel behavior. Before and after data from the implementation of car-pooling incentives and transit-service improvements were used to test the validity of the model's forecasts. Three such tests are reported. The results indicate that the work-trip modal-choice model successfully captures the effects of changes in level of service on modal choice. The predicted effects of several significant car-pooling strategies are presented. In general, traveler response to many car-pooling incentives is small. The most significant changes in travel behavior are predicted for those parking-related policies that combine disincentives for driving alone with incentives for car pooling.

Various strategies designed to increase ride sharing have been proposed and several have already been implemented. For example, strategies such as preferential lanes for high-occupancy vehicles, car-pool-matching and promotion programs (both areawide and employer-based), and preferential parking for car pools have existed for several years. This paper applies a methodology based on disaggregate travel-demand models to predictions of changes in travel patterns that will result from car-pooling incentives and from short-range transportation options in general.

The methodology is described briefly, and a number of validation tests that use before-and-after data are presented. Prediction results from case study applications of the methodology to various car-pooling-related policies are discussed. The paper concludes with a summary of major findings. [Both the methodology used and the analysis of prediction results by market segments are discussed in greater detail by Ben-Akiva and Atherton (1).]

METHODOLOGY FOR SHORT-RANGE TRAVEL-DEMAND PREDICTIONS

The methodology for predicting the changes in travel patterns that will result from short-range transportation options (including car-pooling incentives) is based on the application of disaggregate travel-demand models. These models are based on the multinomial logit, probabilistic choice model, which has been discussed by Domencich and McFadden (5) and Richards and Ben-Akiva (7). The data used to estimate the coefficients of these models are taken from home-interview surveys and represent a cross section of households. The dependent variables of the models are the reported travel choices made; the independent variables are the reported socioeconomic characteristics, engineering measures of travel times and costs, and survey estimates of employment and land-use characteristics in the urban area.

The models consider residential locations and work places as being fixed and predict automobile ownership, choice of mode for the work trip, and frequency, destination, and mode for nonwork travel. Car pooling is explicitly considered as an alternative mode only for work trips. However, the effects of car-pooling incentives on interdependent travel choices and vice versa are also predicted.

To apply these models to the forecasting of changes in travel behavior that will result from alternative car-pooling incentives, a sample enumeration technique is used. In this procedure, a randomly selected sample of households is used to represent the entire population of an urban area. Forecasts are made by applying the models to each household individually, using revised values of the appropriate independent variables to simulate the particular policy being modeled. These household predictions are then summed to represent predicted areawide changes in travel behavior.

VALIDATION WITH ACTUAL CAR-POOLING EXPERIENCE

The most significant test of the validity of a predictive model is a comparison of the changes predicted with the actual observations made before and after a change in transportation service. Several such tests of the work-trip modal-choice model are reported by Cambridge Systematics (3). Three relevant tests having reliable before-and-after data are presented here. These are the Shirley Highway preferential-lane project, the Santa Monica Freeway diamond-lane project, and the Minneapolis bus-on-metered-freeway project. In each of these examples no disaggregate sample was available, and the model was applied by using average values for selected market segments.

The tests were conducted by using the incremental form of the logit model (1). This form predicts revised travel behavior that is based on existing travel behavior and changes in level of service, rather than using the full model to recalculate choice probabilities that are based on the full set of independent variables. Data requirements are greatly reduced by using such an approach; no knowledge of detailed socioeconomic and level-of-service data is required. Only existing shares and proposed changes in level of service are necessary. The incremental form of the logit model for a specific market segment is given by Equation 1.

$$P'(i|A) = [P(i|A)\exp(\Delta V_i)] \Big/ \sum_{j \in A} P(j|A)\exp(\Delta V_j) \qquad (1)$$

where

ΔV_i = change in utility for alternative i = $\Sigma \Theta_k \Delta X_{ik}$,
Θ_k = coefficient of the kth variable,
ΔX_{ik} = change in the kth independent variable for alternative i,
$P(i|A)$ = probability of choosing alternative i before change, and
$P'(i|A)$ = the predicted probability of choosing alternative i after the change.

The first test of the work-trip modal-choice model was conducted for the construction of new preferential lanes on Shirley Highway, which extends to the southwest from the center of Washington, D.C. In 1971, express bus service began on these lanes; bus level of service was further improved by increasing coverage and adding new buses designed for greater passenger comfort. In late 1973, car pools with four or more occupants were also permitted to use the preferential lanes. In addition to the improved bus service and the car-pool preferential lanes, there were other factors that affected the modal choices of commuters between 1970 and 1974 (e.g., increases in the price of gasoline). All of these factors were expressed in terms of changes in the level-of-service attributes and introduced into the work-trip modal-choice model to predict changes in corridor modal shares from 1970 to 1974. The results are given in Table 1. The error in the predicted bus ridership in the corridor is -2.9 percent, and the error in predicting the change in bus ridership is -12.9 percent. These results indicate that the model captured the changes in modal-choice behavior due to changes in the level-of-service attributes.

The second test of the work-trip modal-choice model was undertaken for the Santa Monica Freeway diamond-lane project, which was implemented on March 15, 1976, and discontinued on August 13, 1976. When this project began, the use of the median lane was reserved for buses and car pools with three or more persons. Initially, this resulted in severe congestion in the nonpreferential lanes and on alternate parallel arterials in the corridor and long delays at entrance ramps, while the preferential lane was sparsely used. However, by the tenth week of the project, which is the point in time for which the model was applied, conditions had changed considerably. Travel times in the nonpreferential lanes had decreased to their preproject level and the ramp delays had decreased somewhat, but preferential vehicles still enjoyed a 24 to 32-km/h (15 to 20-mph) speed differential.

At the same time that the preferential lane was implemented, significant improvements were made to transit service in the corridor. The improvements included express bus service to the Los Angeles central business district (CBD), improved distribution and feeder service, park-and-ride lots, and extended route coverage. Approximately 50 new buses were put into service to effect these improvements. Each of these factors was expressed in terms of changes in the level-of-service attributes and introduced into the model. In addition, a dummy variable that captures the promotional and awareness aspects of car-pooling programs was set equal to one. The predicted and observed corridor modal shares for the tenth week of the project are given in Table 1. Although the predictions for the transit modal share are accurate, the predicted changes in the two-person and three-or-more-person car-pool shares overestimate those observed. One probable reason for this is because only 10 weeks had elapsed at the time the after data were recorded, travel in the corridor had not yet passed the transition period. It is also possible that seasonal effects, which are not accounted for in the analysis, may have had a significant impact on the observed changes in travel behavior.

The final test presented here was conducted on before-and-after data from the Minneapolis express-bus-on-metered-freeway (I-35) project that was implemented in 1972. In this project, buses are given priority access to the freeway by express bus ramps. At the same time, automobiles are metered onto the freeway so that traffic volumes allow a desired level of service. In addition, express bus service on I-35 to the CBD was gradually improved from 36 trips to 118 trips during the morning peak period over a 2-year period.

Each of these factors was expressed in terms of changes in level-of-service attributes and introduced into the model to predict changes in modal shares from 1972 to 1974. The results of this model application are also given in Table 1. Here again, the relatively small prediction errors suggest that the model has successfully captured the effects of changes in level of service on modal choice.

RESULTS

The base values for Washington, D.C.; Birmingham, Alabama; and Denver used in the model are given in Table 2.

The model was used to analyze the effectiveness of a number of car-pooling policies. These include

1. Employer-based actions (e.g., car-pool matching and promotion),
2. Parking availability and cost (e.g., preferential parking measures),
3. Traffic regulation and control (e.g., preferential traffic control), and

4. Travel cost (e.g., fuel-price increases).

The predicted impacts of these policies for the Washington, D.C., metropolitan area are summarized in Table 3. For each policy, the predicted percentage changes from base values are given for work-trip modal shares (drive alone, shared ride, and transit), work-trip automobile occupancy, vehicle kilometers traveled (VKT) (for work, nonwork, and total), and fuel consumption. (These percentage changes represent the areawide percentage changes that will result from a policy regardless of the actual proportion of the areawide population affected by that policy.) The predicted impacts for a limited number of these policies for Birmingham and

Table 1. Before-and-after validation tests for corridor modal shares.

Example	Mode	Base Modal Share (before)	Actual Modal Share (after)	Predicted Modal Share (after)	Actual Change	Predicted Change	Error in Prediction[a] (%)	Error in Predicted Change[b] (%)
Shirley Highway	Automobile	67.8	57.3	58.5	-10.5	-9.3	2.0	-12.9
	Bus	32.2	42.7	41.5	10.5	9.3	-2.9	-12.9
Santa Monica Freeway	Drive alone	70.1	66.6	66.7	-3.5	-3.4	0.1	-19.9
	2-person car pool	22.3	22.2	21.3	-0.1	-1.0	-4.2	90.0
	≥3-person car pool	6.0	7.6	8.4	1.6	2.4	9.5	33.3
	Bus	1.6	3.6	3.5	2.0	1.9	-2.9	-5.3
Minneapolis bus project	Drive alone	47.2	43.0	42.0	-4.2	-5.2	-2.4	19.2
	Car pool	19.8	18.0	18.0	-1.8	-1.8	0.0	0.0
	Transit	33.0	39.0	40.0	6.0	7.0	2.5	14.3

[a] Given by (predicted - actual)/predicted. [b] Given by [| predicted - base | - | actual - base |] /| predicted - base |.

Table 2. Base values.

Location	Work-Trip Modal Shares: Drive Alone	Work-Trip Modal Shares: Shared Ride	Work-Trip Modal Shares: Transit	Work-Trip Automobile Occupancy	VKT (km/d): Work (per worker)	VKT (km/d): Nonwork (per household)	VKT (km/d): Total (per household)	Fuel Consumption (L/d/household)
Washington, D.C. (excluding weekend travel)	0.53	0.25	0.16	1.24	17.7	24.1	41.8	10.6
Washington, D.C. (including weekend travel)	0.53	0.25	0.16	1.24	17.7	40.2	57.9	14.8
Birmingham, Ala. (excluding weekend travel)	0.68	0.24	0.08	—	12.9	22.5	35.4	8.9
Birmingham, Ala. (including weekend travel)	0.68	0.24	0.08	—	12.9	37.0	49.9	12.7
Denver (including weekend travel)	0.78	0.19	0.03	—	20.9	30.6	49.9	10.0

Note: 1 km = 0.62 mile; 1 L = 0.26 gal.

Table 3. Predicted areawide impacts.

Policy	Percent Change From Base Value — Work-Trip Modal Shares: Drive Alone	Shared Ride	Transit	Work-Trip Automobile Occupancy	VKT (km/d): Work (per worker)	Nonwork (per household)	Total (per household)	Fuel Consumption (L/d/household)
Employee incentives								
Washington[a]	-1.4	4.4	-2.1	1.1	-0.62	0.17	-0.15	-0.11
Birmingham (all workers regardless of employer size)[a]	-4.4	16.1	-9.9	—	-2.1	0.08	-0.3	-0.2
Denver (employer size ≥50)[a]	-2.9	13.0	-10.0	—	-1.4	—	-1.4[b]	-1.2[b]
Preferential parking								
Washington[a]	-2.7	8.2	-3.6	1.9	-0.74	-0.24	-0.16	-0.13
Denver (including subsidized car pools)	-3.44	15.4	-10.8	—	-1.7	—	-1.7[b]	-1.5[b]
Preferential parking with pricing disincentives								
Washington[a]	-8.6	17.7	1.4	4.8	-4.1	0.88	-1.1	-0.97
Preferential lanes								
Washington[a]	-1.7	2.9	1.1	0.9	-1.3	0.22	-0.39	-0.32
Fuel price doubled								
Washington[c]	-1.8	1.9	2.9	0.8	-1.6	-8.0	-6.1	-5.7
Birmingham[c]	-1.0	1.6	4.1	—	-1.2	-6.7	-5.6	-4.8
Denver[a]	-0.92	2.6	3.6	—	-0.77	-16.0	-10.0	-10.0
Fuel price tripled								
Washington[c]	-3.6	4.0	6.0	1.6	-3.3	-14.9	-11.5	-10.7

[a] Using base value excluding weekend travel. [b] Values for work trips only. [c] Using base value including weekend travel.

Denver are also given in Table 3.

Employer-Based Matching and Promotion

This policy represents the implementation of employer-related car-pool incentive programs, such as intra-company advertising, car-pool-matching assistance, and promotion. Incentives such as these cannot be readily quantified in terms of travel time and travel cost. This makes their representation somewhat of a problem, which is solved here by the use of a dummy variable.

Since car-pooling incentives such as these are feasible only for organizations with a relatively large number of employees, the most logical criterion for determining the availability of such incentives to an individual worker is employer size. For this particular policy, a lower bound of 100 employees was used to differentiate between large and small employers in Washington. This dummy variable was set equal to one for those workers employed by organizations with at least 100 employees and zero otherwise. For Birmingham, however, employer-size data were not available, and this dummy variable was set equal to one for all workers. In Denver, a lower bound of 50 employees was used. As shown in Table 3, the implementation of car-pool-matching assistance and promotion programs by large employers in Washington resulted in a 4.4 percent areawide increase in the number of workers in car pools. However, while the work-trip VKT were reduced by 0.62 percent, the increased number of automobiles remaining at home (and therefore available for use by other family members for other trip purposes) resulted in a 0.17 percent increase in nonwork-trip VKTs, which offset 36 percent of the work-trip VKTs savings.

The potential areawide effectiveness of this policy is muted somewhat by two conditions existing in the Washington area. These are that

1. By restricting the availability of car-pool incentives to those workers employed by large employers, only 68 percent of the work force is affected and
2. Forty-four percent of the work force already has these incentives available.

The result is that only 24 percent of the working population is affected by this policy. In other urban areas, where the initial level of employer-based car-pool incentives is lower or a greater proportion of the work force is employed by large employers, this policy would be more effective in terms of increased car pooling and VKT reduction. In Birmingham, for example, the initial level of employer-based car-pool incentives was assumed to be zero and thus, the entire work force was affected by the policy. As shown in Table 3, the percentage increase in shared ride was almost four times that in Washington. The initial level of employer-based incentives was assumed to be zero in Denver also. In this case, however, the lower bound was 50 employees, which resulted in a percentage increase in shared ride approximately three times that in Washington (Table 3).

Preferential Parking Measures

Two sets of preferential parking measures were analyzed for Washington. These were

1. A program implemented by large employers (i.e., those with more than 100 employees), who gave subsidized, preferential parking locations to car-pool vehicles and
2. Car-pool incentives coupled with areawide parking-price disincentives aimed at single-occupant vehicles.

The first set of measures is represented in the model by setting parking cost equal to zero and decreasing walking time from parking location to final destination for the shared-ride alternative and by increasing walking time for the drive-alone alternative for those workers employed by large employers. The magnitude of these changes in round-trip walking times are -4.27 min for shared ride and +1.64 min for drive alone. These values were calculated on a basis of the cumulative walking-time distribution for parked vehicles during the peak period in the Washington area and the percentage of all automobiles used for car pools. In addition to these employer-based, car-pool parking incentives, the second set of measures included minimum parking charges for the drive-alone alternative of $2.00 in the CBD and $1.00 elsewhere in the metropolitan area (for work trips only).

The predicted results of these two parking policies are given in Table 3. For the first policy, the shared-ride share increases significantly, and the shares of both drive alone and transit decrease. For the second policy, an even greater increase in shared ride is predicted. In this case, however, while the drive-alone share drops markedly, the transit share shows a slight increase. This occurs because the first policy consists primarily of car-pool incentives, and therefore shared ride will draw from both drive alone and transit. In the second policy, however, the drive-alone disincentive dominates, and more commuters will shift from drive alone to transit than from transit to shared ride.

A preferential parking measure similar to the first one discussed for Washington was analyzed in Denver. Here, however, parking costs were assumed to be subsidized by the employer (essentially, this affects only CBD workers) and the lower bound for employer size was set at 50 employees rather than 100. The results (based on the analysis of work trips only) given in Table 3 show a percentage increase in ride sharing that is almost double that in Washington.

Preferential Traffic Control

This policy was analyzed only in Washington and consisted of preferential lanes for multiple-occupancy vehicles. It was analyzed on an areawide basis by identifying those trips that would use facilities for which a preferential-lane policy would be feasible. This approach, rather than one of analyzing one specific facility, was taken because of the relatively small sample used in forecasting (i.e., 800 households for the entire Washington metropolitan area), which results in an extremely small and statistically unreliable subsample of observed work trips on any given facility.

After potential locations of preferential lane and ramp treatments were identified, differential time savings were estimated as follows for three broad categories of work trips. For trips from outside the beltway to the inner core, a differential of 16 min was used (8 percent of the sample). For trips (a) from outside to inside the beltway, (b) along the beltway, or (c) from inside the beltway to the inner core, a differential of 8 min was used (31 percent of the sample). For all other trips, e.g., outbound commute or circumferential within the beltway, no time savings were assumed (61 percent of the sample). No time savings were assumed for nonwork trips. These travel time differentials were based on the following assumptions: (a) an average base speed of 56 km/h (35 mph) on all facilities, (b) an average preferential lane speed of 81 km/h (50 mph), (c) an average nonpreferential lane speed of 48 km/h (30 mph), and (d) average preferential-lane lengths of 16 and 8 km (10 and 5 miles) respectively for the first two categories of work trips described above.

The predicted impacts of this preferential lane policy are given in Table 3. The shared-ride modal share increased 2.9 percent, while that of transit increased by only 1.1 percent. The reason for this difference becomes evident if one looks at the results for workers residing in suburban areas. Here, the difference is even greater, which suggests that for those workers for whom preferential lanes are most attractive, transit availability is somewhat limited. Here again, the decrease in work-trip VKT is partially offset by increased nonwork travel.

Fuel Price Increases

Two policies in which the price of gasoline was increased were analyzed for Washington. In one case, the base price was doubled, and in the other, the base price was tripled. Unlike the preceeding policies, these policies are directed not only towards work trips, but also directly affect weekday and weekend nonwork travel.

These policies are represented in the model by increasing the portion of automobile-travel cost that is represented by fuel costs. In this case, fuel costs represent 70 percent of automobile-operating costs: On the average, automobile-operating costs represent 50 percent of automobile-travel costs, the remainder being parking costs. Therefore, doubling the price of gasoline, for example, would result in an approximate 35 percent increase in out-of-pocket travel costs.

The impacts predicted for these policies are given in Table 3. In both cases, the reduction in nonwork travel is greater than that predicted for work trips despite the increased automobile availability for nonwork trips. This agrees with the idea that nonwork trips, being more discretionary in nature, are more sensitive (more elastic) to changes in travel costs. While travel costs will increase by the same amount for both drive-alone and shared-ride vehicles, the assumption that costs are borne equally among car-pool members results in a significantly smaller cost increase per person for shared ride. Because of this, the shared-ride modal share increases, although transit shows the larger gain in modal share.

Similar impacts are predicted for Birmingham and Denver for the policy of doubling the price of gasoline, although in Denver the much higher sensitivity of nonwork travel to fuel-price increases relative to the other two cities results in a significantly larger decrease in total fuel consumption.

Summary

In addition to the results tabulated here, strategies have been analyzed for both Washington and Birmingham by using an earlier version of the model system (2). Similar studies evaluating the effectiveness of alternative car-pooling policies have also been conducted by others (4, 6). While a direct comparison of the results from these studies is difficult because of the differences in data used and policy definitions, it appears that the results reported here are more conservative than others.

CONCLUSION

The use of the methodology developed in this study was demonstrated by an analysis of car-pooling strategies, but the methodology can also be used to analyze a wide variety of short-range transportation alternatives. Before-and-after tests have shown the validity of applying the work-trip modal-choice model to the prediction of impacts of short-range transportation policies on travel behavior in several urban areas.

Several car-pooling strategies were analyzed. A specific car-pool strategy would be either an incentive, i.e., a direct inducement to workers to share rides by improving the level-of-service attributes available to car pools, or a disincentive, i.e., an indirect inducement for car pooling by worsening the level-of-service that attributes for solo drivers. The major conclusions are

1. Car-pooling incentives will attract transit as well as drive-alone commuters and, because the potential areawide increase in ride sharing is small, the decrease in VKTs will be small;
2. Automobile disincentives are much more effective than car-pooling incentives in increasing ride sharing and transit use, but these policies are less acceptable to the public and therefore less likely to be implemented;
3. A coordinated program of both incentives and disincentives could effectively increase car pooling and reduce congestion, VKT, and fuel consumption, and significant parking incentives and disincentives appear to be the most effective ways to increase car pooling;
4. Car-pooling strategies directed at work trips result in increased automobile travel for nonwork purposes because of increased automobile availability during work hours for nonworking members of a household, and the increased nonwork VKTs offset by approximately one-third the reduction in work VKTs; and
5. The effectiveness of a particular car-pooling strategy will vary significantly among urban areas because of differing base conditions.

REFERENCES

1. M. Ben-Akiva and T. Atherton. Methodology for Short-Range Travel-Demand Predictions: Analysis of Car-Pooling Incentives. Journal of Transport Economics and Policy, in preparation, 1977.
2. Car-Pool Incentives: Analysis of Transportation and Energy Impacts. Cambridge Systematics; Federal Energy Administration, 1976.
3. Tests of Transferability and Validation of Disaggregate Behavioral Demand Models for Evaluating the Energy-Conservation Potential of Alternative Transportation Policies in Nine U.S. Cities. Cambridge Systematics; Federal Energy Administration, 1977.
4. Estimating the Effects of Urban Travel Policies. Charles River Associates; Office of the Secretary, U.S. Department of Transportation, 1976.
5. T. Domencich and D. McFadden. Urban Travel Demand: A Behavioral Analysis. North-Holland, Amsterdam, 1975.
6. A Marketing Approach to Car-Pool-Demand Analysis. Peat, Marwick, Mitchell, and Co. and Market Facts; Federal Energy Administration, 1976.
7. G. Richards and M. Ben-Akiva. A Disaggregate Travel-Demand Model. Saxon House, D. C. Heath, England, 1975.

Discussions

David S. Gendell, Office of Highway Planning, Federal Highway Administration

In the past several years, we have seen a change in emphasis in transportation planning from long-range to a shorter time frame and a corresponding change in the alternatives being analyzed from capital-intensive fixed-guideway systems to those having a more operational character. An implication of this is that transportation modelers may be no longer needed because these newer alternatives might best be evaluated through intuitive procedures. I think the papers presented here give us a clear indication that technical analysis or modeling is very much a part of planning in this era of transportation-system management.

Horowitz and Sheth have demonstrated that psychological scaling techniques can be meaningfully applied to the determination of the attributes of transportation supply that have an impact on travel behavior. They successfully use these techniques to develop initiatives to encourage ride sharing and show why several previous promotional techniques have had minimal success.

Small presents the results of a pioneering effort to bring together modeling techniques in the areas of traffic-flow and travel-demand analysis. Previous works dealing with the effects of priority-lane allocation on traffic flow have suffered from the lack of acknowledgment of the effects that changes in performance level have on travel demand. At the same time, previous travel-demand models have typically begun with some assumed change in system performance in determining the effect of priority lanes on modal use. Small has demonstrated that these disciplines can be brought together to analyze the desirability of priority-lane allocation. This is clearly a worthwhile activity in planning for such facilities. Small also develops the concept of benefit versus cost, which is a commendable addition to transportation-system-management planning procedures.

Ben-Akiva and Atherton apply an advanced model system to the evaluation of a wide range of strategies aimed at encouraging increased ride sharing. They skillfully model a number of behavioral effects that are often ignored in work of this nature. Their paper will be useful to policy officials because it illustrates the comparative effectiveness of various strategies in reducing travel and fuel consumption. They have also contributed to transportation-system-management planning methodology by demonstrating the use of sample-by-sample enumeration techniques in quantifying the impacts of various strategies on different geographic areas and income groups. Also, the planning practitioner will benefit from learning about the incremental form of the logit model, which should facilitate policy analysis in areas where the enumeration approach is not practical.

While I am pleased by the research represented here, I do have some concerns. The data used by Horowitz and Sheth were not based on a random sample, but on an uncorrected choice-based sampling procedure in which a roughly equal number of car-pool and single-occupant automobile commuters were surveyed at their places of employment. This resulted in a biased sample that may have influenced their results. For example, the average trip distance for car-pool commuters was 26.2 km (16.3 miles), while that for single-occupant automobile commuters was 17.8 km (11.2 miles). Thus, the respondents are not from the same sampling frame and are not thinking of the same trip when responding to the questionnaire. For example, when a single-occupant automobile respondent views a car pool as inconvenient, he or she is doing so for an 17.8-km (11.2-mile) trip. He or she would probably view it as less inconvenient for the 26.2-km (16.3-mile) trip evaluated by the average car pooler. Thus the differences between the means of the various variables thought to explain behavior in part reflect the difference in sampling frames.

The papers by Small and by Ben-Akiva and Atherton use logit-model formulations that contain demographic, travel-time, and travel-cost measures. Both resulted in relatively large negative car-pool bias coefficients, indicating that there are probably important behavioral variables that are not included in the model specifications. Perhaps incorporation of some of the softer attributes studied by Horowitz would reduce the size of these coefficients and lead to additional policy sensitivity.

All of the papers interpreted the policy implications of their findings well, although I believe that Small might have gone further in testing a range of occupancies qualifying for use of the priority lanes. The paper would also have been improved by a case-study application of the techniques to an operational priority lane to demonstrate their ability to reproduce an observed response.

Of the strategies tested by Ben-Akiva and Atherton, the doubling and tripling of the price of fuel had the greatest effect on travel and fuel consumption. The results of these particular tests should probably be interpreted with caution because, as the authors point out, the models do not predict the probable shift in automobile-size distribution that would result. Perhaps more moderate tests would have been advisable with the current formulation.

While I have discussed several areas in which the research reported in these papers might have been improved, I would like to conclude by noting that these papers represent a significant contribution to the knowledge of the effects of a wide range of strategies on car-pool demand. In addition, they demonstrate the effective application of analytical planning procedures to evaluating transportation-system-management strategies. As such, they represent an area of research that should assist the practicing planner.

Daniel Brand, Massachusetts Executive Office of Transportation and Construction

These papers by Horowitz and Sheth, Small, and Ben-Akiva and Atherton improve our ability to predict both car-pool demand and demand for travel by all other modes. Forecasts of impacts of policies to promote car pooling are of little use if they do not include impacts on other travel modes and their associated effects on such factors as energy consumption, traffic congestion, and air quality.

I would like to congratulate Horowitz and Sheth for a well written and neatly packaged piece of work. Their paper introduces some complicated concepts in simple and easily understood fashion. However, they have used data on stated general intentions to change from solo driving to car pooling rather than behavioral data on persons who actually changed modes in reaction to a change in a choice situation that they confront; for example, a new preferential lane for car pools or an employer-based car-pool-matching program. In addition, the attitudinal data on intentions simply asked how likely it was that the respondent would join a car pool within the next two or three months. I would much rather see them gather policy-related attitudinal responses to such questions as "How likely are you to join a car pool if you were offered a list of fellow employees who live near you to car pool with?" Data on such attitudes could help isolate promising markets for car-pool-matching programs and other car-pool-promoting policies. A useful extension of this paper would include behavioral data, not only for what it would tell us about car-pooling behavior, but also for investigating the relations between stated intentions or attitudes and actual behavior. Attitudinal data are, of course, usually much easier to collect than behavioral data.

The first conclusion of the paper, namely that "demographic and travel characteristics are poor indicators and predictors of the choice between driving alone and ride sharing" is subject to misinterpretation and needs restatement. This conclusion can be taken to mean that similar travel patterns (same origins and destinations as produced in a car-pool-matching program) do not lead to car-pool formation. That is not what is meant at all, and the authors might be more specific.

Examination of the cognitive profiles presented in the paper for actual car poolers and solo drivers in terms of attribute-by-attribute evaluations of ride sharing and solo driving is quite fascinating. Our objective in promoting car pooling is to change people's evaluations of car pooling by promotional campaigns that have the effect of bringing these attribute-evaluation profiles closer together in the belief that this will bring about actual car-pooling behavior. Generally, car pooling is regarded by solo drivers as less convenient than solo driving, but not substantially cheaper or less energy consuming or air polluting. The conclusion one draws from these data is that we ought to be promoting car pooling as a convenience, rather than as being cost saving. This is, I think, a very significant finding. It is a sobering conclusion, but consistent with the results of transit travel forecasting that transit use is less likely to grow from fare decreases than from making use of the private automobile less convenient.

Finally, Horowitz and Sheth reach some interesting conclusions about the identity of groups most likely to car pool. They identify employees of high socioeconomic status and those from large households who have relatively recently moved their places of residence or employment as being the most likely persons to change to car pooling on the basis of their attribute-evaluation profiles. This contrasts with their findings that actual car poolers have been at their present residences and employment locations longer than solo drivers. This disparity between attitude and behavior is again an indication of the usefulness of before-and-after behavioral data as opposed to data about general attitudes and intentions. The persons who finally begin to car pool in response to a multitude of factors may not be the persons who state that they are likely to car pool.

Small's paper is a neatly done equilibrium analysis of the consequences for bus, private automobile, and car-pool use and users of reserving a freeway lane for buses or for buses and car pools. Varying the number of freeway lanes and the before conditions on degree of congestion with no reserved lanes are added features of this paper.

In general, this paper shows that congestion relief on the general-purpose (unreserved) lanes actually occurs when a lane is reserved for buses and three-or-more-occupant car pools when the before conditions are characterized by heavy and very heavy congestion. This is truer when one lane is taken from a five or four-lane facility rather than from a three-lane facility. That is, in the case of a three-lane facility, (slight) congestion relief occurs on the two unreserved lanes only when the before conditions are characterized by very heavy congestion. These conclusions are consistent with calculations done in Massachusetts in preparation for the spring 1977 opening of a reserved lane for buses and three-or-more-occupant car pools on the heavily congested four-lane (each way) Southeast Expressway in Boston. In the Boston calculations, it was predicted that congestion on the unreserved three lanes would not significantly increase with the reservation of a bus and car-pool lane and, of course, delays to buses and car pools would be eliminated.

The third and final paper discussed here is Ben-Akiva and Atherton's excellent and comprehensive set of predictions of the use of several modes, including car pooling, that will result from many proposed alternatives for increasing car pooling. The authors use their multinomial logit, disaggregate, demand forecasting model, which was estimated and tested by using data from several urban areas in different parts of the country. The model provides quite reasonable forecasts of the consequences of the specific policies tested and documented in the paper. The conclusions of the paper on the impacts of various car-pooling strategies appear consistent with the results of previous studies. However, the use of a demand-model system that includes work and non-work-trip models enables the authors to go beyond previous findings in certain ways; e.g., leaving a car home and car pooling to work results in increased nonwork vehicle kilometers of travel (VKT) by nonworking members of the household that offsets by approximately one-third the reduction in work-trip VKT from car pooling.

All three specific comments on the paper relate to the details of the way in which the authors have modeled car-pooling behavior. The first comment relates to a possible difficulty in forecasting car-pool use resulting from preferential lane strategies by the use of a model estimated using cross-sectional data. That is, longer trips tend to exhibit a higher car-pooling modal split. If this bivariate cross-sectional relation is reflected in the multinomial-logit model, car-pooling use in response to preferential lanes that reduce car-pool travel times may tend to be underestimated. The model used in the paper did underestimate the car-pool formation that resulted from the Santa Monica Freeway reserved (diamond) lane. Can this effect be the explanation for this underestimation?

My second comment on the model relates to the way in which employer-based car-pool-matching and promotion programs are represented. Such employer-based car-pool programs are certainly the most widely applied car-pool-promotion strategies at the present time. The authors indicate that car-pool-promotion programs are represented in the work-trip modal-choice model by a dummy variable that picks up the effects of employer-based car-pool programs. It is not stated whether the data used to estimate the model included an indication of the presence or absence of an employer car-pool program at the place of work for each work-trip observation. If these data are available, they represent a substantially improved data base over that normally available.

My third and final comment relates to the property of the logit model by which users of a new mode (e.g., car pooling) are drawn from the existing modes in proportion to their existing modal shares. Is this property operative in the forecasts being made in this paper, and if so, do the authors think that this (separability) property is reasonable? The issue is very important to transit officials who are concerned that car-pool-promotion programs may divert transit users into car pools, particularly in areas that have very high transit use. Putting their minds at ease would help to generate support for car-pool programs.

Authors' Closures

Abraham D. Horowitz and Jagdish N. Sheth

We thank Gendell and Brand for their excellent analyses and insightful comments and appreciate the many positive comments they have made.

We shall concentrate here on their criticisms and methodological questions. At issue, it seems, there are the following three main points.

1. The nonrandom sampling procedure, especially the oversampling of car poolers, has resulted in biased data, such as significant differences in the average trip distance between solo drivers [17.8 km (11.2 miles)] and car poolers [26.2 km (16.3 miles)]. This may explain differences in car-pooling attitudes between solo drivers and car poolers rather than any fundamental psychological differences.

The answer to this criticism is relatively simple and straightforward. First, there are two different and independent populations, those who car pool and those who drive to work. Over or undersampling a particular population cannot affect the parameter (average trip distance) of another population. Second, given the budget constraints, a simple random-probability sample would have produced very few car poolers. A smaller sample size based on random probability would have resulted in an estimator with a higher variance. It was necessary to oversample car poolers precisely to reduce the variance of the estimator. Finally, the shorter trip distance among solo drivers, coupled with their attitude that car pooling is more inconvenient, indicates that it is not the trip distance, but some other psychological factor that produces a greater negative attitude toward car pooling. Otherwise, we should expect just the opposite.

2. It is much better to conduct an experiment and measure its impact by before-and-after changes in the actual behavior of solo drivers rather than collect survey data and stated general intentions and preferences.

We agree with this comment in principle, but not in practice. First, experiments are extremely difficult to carry out in the area of public services because of the large numbers of legal and political constraints. Second and more importantly, to design experiments one needs very good hypotheses and theories, which we simply do not as yet have in the area of urban transportation. It is still a matter of learning more about the realities of urban transportation. Otherwise, the experiments will be carried out on wrong policy variables and result in no main effects and numerous side effects.

We think that it is a better strategy to first conduct a survey and identify potential areas of policy variables and then design proper experiments to measure their behavioral impacts. For example, our study clearly indicates that experiments based on economic incentives would fail to motivate solo drivers to switch to car pooling.

3. It is much better to ask the respondent whether he or she will join a car pool if a certain situation arises than to ask about general intentions to join a car pool within the next two or three months.

This is a very good criticism and we fully agree with it. In fact, it is a major weakness of all attitudinal research that specific situational aspects are not taken into consideration (21). We have collected data on several if-and-what scenarios, such as increased gasoline prices, parking restrictions, and environmental pollution. Unfortunately, we did not have the space to analyze and report the situation-bound intentions in this report.

Kenneth A. Small

The thorough and perceptive comments by the discussants of these papers leave little to explain or refute. At the same time, they raise some important points that are shared by my own paper and that of Ben-Akiva and Atherton.

The two papers contain some basic similarities in both methodology and results. Both use disaggregate models and a sample-enumeration (forecasting-sample) procedure. Both find that traveler response to typical policies is rather moderate, that disincentives to low-occupancy automobiles are more powerful than incentives to car pools or transit, and that the overall effect of car-pooling incentives is muted by their tendency to divert transit riders to car pools in significant numbers.

Furthermore, as Gendell notes, the calibration of our modal-split models results in a large pure-mode bias against car pooling. Why? I believe the results are correct; i.e., given equal values of those time, cost, and socioeconomic variables that we have identified as influencing travel behavior, car pooling is perceived as much less desirable than lower-occupancy automobile. But this opens to consideration two possibilities: Perhaps this bias can be altered by such policies as matching services and promotional campaigns, and perhaps it can be explained by including such additional variables as attitudinal measures. The latter would be most helpful in directing the policies undertaken with regard to the former, and it is here that a synthesis with the attitudinal approach exemplified by Horowitz and Sheth would appear especially productive.

Moshe Ben-Akiva and Terry J. Atherton

Gendell and Brand have provided very useful discussions that raise several important concerns.

1. The magnitude of alternative specific constants: Gendell is concerned with the large negative constant in the car-pool utility function of the modal-choice model used in our study. He states that the existence of a large negative car-pool constant indicates that important behavioral variables were omitted from the model. Therefore, he suggests that some of the softer attributes studied by Horowitz and Sheth be included in the model.

The existence of an alternative specific constant is due to omitted variables, but is also influenced by the definition of alternative choices (e.g., modes) that is used in the estimation and application of a choice model. This is demonstrated below for the logit model.

Consider a logit model with no alternative specific constants as follows:

$$P(i|C) = \exp V_i \Big/ \sum_{j\epsilon C} \exp V_j \tag{2}$$

where $P(i|C)$ is the choice probability of alternative i, given choice set (C), and V_i is the systematic utility of alternative i. Suppose that some of the alternatives in C have identical systematic utilities and consequently equal choice probabilities. Partition C into nonoverlapping subsets of alternatives with equal systematic utilities as follows:

$$C = \{1, 2, \ldots, i, \ldots, J\} = \{A_1, A_2, \ldots, A_k, \ldots, A_L\} \tag{3a}$$

and

$$V_i = V_k \quad (i = 1, \ldots, n_k) \tag{3b}$$

where

J = number of elemental alternatives in the choice set,
L = number of subsets of identical elemental alternatives ($L \leq J$),
A_k = subset of identical alternatives,
V_k = systematic utility of alternatives in A_k, and
n_k = number of elemental alternatives in A_k.

We can now rewrite the model as shown by Lerman (13):

$$P(A_k|C) = n_k \exp V_k \Big/ \sum_{l=1}^{L} n_e \exp V_e$$

$$= \exp(V_k + \ln n_k) \Big/ \sum_{l=1}^{L} \exp(V_l + \ln n_l) \tag{4}$$

If n_k is not known, its natural logarithm serves as an alternative specific constant. In this case, the constants result not from omission of attributes from the utilities but from lack of knowledge about the true elemental alternatives in choice sets of individuals and the need to define aggregate or combined alternatives for practical purposes.

There are a variety of car-pooling arrangements, such as with or without cost sharing and door-to-door collection and distribution versus common meeting point. All possible elemental car-pooling alternatives are represented in the modal-choice model by a single car-pool mode. Because it is impossible to collect sufficient information and model explicitly all possible car-pooling arrangements, the model must include a car-pool specific constant even if all of the relevant attributes are included. There is no unique natural count of the number of elemental car-pooling alternatives.

The often used example of the red and blue bus alternatives is derived from the same problem of an arbitrary definition of alternatives. In this case, an individual can be expected to perceive one alternative bus mode and the analyst is assumed to specify two alternative bus modes. These are identical modes except for color, which is assumed to have no effect on the choice of mode. Thus, if the true model without mode-specific constants is expressed as

$$P(\text{bus}) = \exp V_{\text{bus}}/(\exp V_{\text{bus}} + \exp V_{\text{automobile}}) \tag{5}$$

then the definition of two identical bus modes implies a model with constants as follows:

$$P(\text{red bus}) = \exp(V_{\text{bus}} + \ln 0.5)/[\exp(V_{\text{bus}} + \ln 0.5) + \exp(V_{\text{bus}} + \ln 0.5) + \exp V_{\text{automobile}}] \tag{6}$$

These constants are set such that the sum of the red and blue bus probabilities is the same as the bus probability in the true model.

However, evidence in the form of large constants is not required to observe that important attributes that influence modal choice are missing from the existing models. Attributes, such as comfort, reliability, and safety, cannot easily be measured and because of the lack of data are omitted from the existing models. This problem, however, cannot be solved by a direct substitution of reported attitudinal data into the existing choice models.

Choice behavior is determined by perceptions and attitudes. These, in turn, are determined by physical characteristics and past and present behavior. The existence of this reverse causal link from behavior to attitudes and beliefs (the phenomenon of cognitive dissonance) was noted in the context of travel decisions recently by Golob and others (11) and by Brown (9). For policy analysis, in which the consequences of changes in policy variables and physical attributes need to be predicted, this second relation must also be modeled. Thus, the incorporation of softer attributes requires more complex models that have not yet been studied. Models explaining behavior in terms of attitudes and beliefs have been estimated, but have not been used effectively in transportation analyses because models predicting perceptions and attitudes in terms of physical characteristics and previous behavior have not been developed. Such demand models involving simultaneous equations or a dynamic structure with an explicit history of behavior should be the subject of future research to find ways in which softer variables could be incorporated.

2. Automobile-type choices: Gendell points out that models predicting probable shifts in automobile-size distribution were not included in our analysis, and therefore our tests of the effects of large fuel-price increases should be interpreted with caution. The reported results should be interpreted only as a short-run trend under the present average fuel economy.

An automobile-type choice model was not available when the study was conducted. Recent work by Lave and Train (12) and an ongoing project at Cambridge Systematics are directed toward the development of disaggregate automobile-type choice models that could be added to the model in future applications.

In the course of the study, we did use an aggregate automobile-type model developed by Chamberlain (10) in some of the tests. This model is sensitive to aggregate economic variables and fuel price and was used to predict the higher average fuel economy that would be caused by a fuel-price increase. When this predicted change in average fuel economy is used, the predicted travel-demand changes are markedly different from those reported in the paper. There was less mode changing and, as a result, a significantly smaller decrease in work-trip and nonwork vehicle kilometers of travel was predicted to result from a fuel-price increase. However, a significantly larger decrease in fuel consumption was also predicted. For the policy of doubling the fuel price, a reduction in the range of three to four times greater than that reported in the paper was predicted. Thus, it is obvious that automobile-type choice is highly sensitive to fuel price and should be explicitly included in such an analysis.

3. Downward-biased travel-time coefficient: Brand comments that the predicted increase in car pooling due to reduced car-pool travel times may be underestimated. It is always possible that the model will be misspecified such that the coefficient of travel time will be biased. If the coefficient is biased toward zero, a change in travel time will be predicted to have a smaller effect than would actually occur. Brand supports this possibility by the existence of a positive correlation in cross-sectional data between trip length and car pooling and by the underestimation of car pooling in the Santa Monica Freeway before-and-after test.

However, there could be other reasons why the car-pooling predictions are downwardly biased. There could be simultaneous changes in factors other than travel time that affected modal choice, but they were not considered in predictions. One such factor in the case of the Santa Monica Freeway preferential lanes was the added car-pooling publicity and promotion during the first weeks of operation. Repeating the Santa Monica predictions, but including this effect by the use of the car-pool-promotion-and-awareness dummy variable, we obtained an almost perfect prediction of the increased car pooling. This, of course, could be a coincidence. Nevertheless, it provides a plausible explanation to the initial car-pool underprediction that is not based on a downwardly biased travel-time coefficient, as suggested by Brand.

The coefficient of the car-pool-promotion-and-awareness dummy variable in the modal-choice model was estimated by using data taken from a home-interview survey of the Metropolitan Washington Council of Governments in 1968. These data are less than ideal, but they are among the best available. In particular, this dummy variable was defined on a basis of the limited car-pooling promotion and matching available at that time to employees in the large federal office buildings.

4. Uniform cross elasticities: Brand's final comment is directed to the logit model's property of uniform cross elasticities. Car-pooling incentives will cause the choice probabilities of all other modes to decrease by the same proportion. However, this property is valid only for disaggregate predictions. It is not valid for aggregate predictions, as the results reported in the paper demonstrate. [The difference between disaggregate and aggregate elasticities has been shown by Ben-Akiva (8).] It is an unreasonable property for aggregate predictions, but there is no empirical evidence to reject it, if the model is otherwise well specified, for the disaggregate predictions.

Thus, given the successful before-and-after tests of the modal-choice model, there is no apparent reason to suspect the validity of the predicted diversions from transit to car pools. The only way to avoid a shift from transit to car pools is to accompany car-pooling incentives with transit-service improvements in areas having heavily used transit services.

REFERENCES

8. M. Ben-Akiva. Structure of Passenger Travel Demand Models. Department of Civil Engineering, MIT, Cambridge, PhD dissertation, 1973.
9. H. P. Brown. Attitudinal Measures in Models of Mode Choice. Paper presented at the 3rd Australian Transport Research Forum, Melbourne, 1977.
10. C. Chamberlain. A Preliminary Model of Automobile Choice by Class of Automobile: Aggregate State Data. Transportation Systems Center, U.S. Department of Transportation, Cambridge, Mass., Discussion Paper, 1974.
11. T. F. Golob, A. D. Horowitz, and M. Wachs. Attitudinal-Behavior Relationships in Travel Demand Modeling. Resource paper prepared for the 3rd International Conference on Behavioral Travel Modeling, Tanunda, South Australia, 1977.
12. C. Lave and K. Train. A Behavioral Disaggregate Model of Automobile-Type Choice. Univ. of California, Irvine; Energy Research and Development Agency, 1977.
13. S. R. Lerman. A Disaggregate Behavioral Model of Urban Mobility Decisions. Department of Civil Engineering, MIT, Cambridge, PhD dissertation, 1975.

Publication of these papers sponsored by Committee on Passenger Travel Demand Forecasting.

Analysis and Prediction of Nonwork Travel Patterns of the Elderly and Handicapped

David T. Hartgen, Planning Research Unit, New York State Department of Transportation
Stephen M. Howe,* North-Central Texas Council of Governments
Mark Pasko,* Lancaster County Planning Commission

This paper summarizes a recent survey of 165 randomly selected elderly and handicapped persons in the Albany, New York, standard metropolitan statistical area. The respondents were administered a 6-min questionnaire on nonwork travel habits, perceived barriers to travel, and intended travel if barriers were removed. Four disaggregate models were constructed relating total travel and modal choice to system, demographic, mode availability, and physical handicap factors. The results show that, contrary to present thinking, the elderly and handicapped vary widely in mobility problems and travel patterns and there is no homogeneity within each group; travel mobility is primarily a function of physical disability, availability of an automobile, and the individual's ability to use it; specific bus-service improvements will not materially affect transit demand, but will ease the travel burden; and improvements concentrating on service availability and direct pickup appear to be the most promising.

In recent years, public transportation systems have been encouraged (and mandated) to give special attention to the services provided to the elderly and handicapped. Off-peak transit fares for these persons are now required as a condition for federal transit-operating assistance under section 5 of the Urban Mass Transportation Act of 1974; federal regulations also require full consideration of these persons in transit system design and operation. The unified work programs prepared annually by metropolitan planning organizations also include similar requirements. These activities are generally consistent with the attitudes of the citizens of New York, 85 percent of whom support reduced fares and special services for the elderly and handicapped (1).

The study discussed here was undertaken in the Albany, New York, standard metropolitan statistical area (SMSA) to determine the factors influencing nonwork travel demand by the elderly and handicapped and to develop a method of estimating their nonwork travel demand. Further results are given by Hartgen and others (2).

DATA

Numerous studies, as well as common sense, suggest that nonwork travel by the elderly and handicapped will probably depend on a number of basic factors. It is hypothesized that, in addition to the traditional socioeconomic factors, such as income and automobile ownership, travel by these special groups (both the number of trips and the mode chosen), is affected by the characteristics of the individual's disability, the characteristics of buildings and other potential destinations, and features of the various means of transportation (e.g., barriers).

In this study, cost and time considerations dictated a small-sample daytime telephone survey. Because some handicapped and elderly persons work, such a sample will not include many of these. Hence the study is essentially limited to the nonwork travel patterns of those elderly and handicapped who do not work. Two groups were studied:

1. The transportation handicapped (self definition), which includes some elderly persons, and
2. The elderly (age ≥ 65) who are not handicapped.

To calibrate disaggregate models for these groups, sample sizes of at least 30 to 50 individuals are required. The handicapped are the rarer group: about 4.7 percent of all persons in the large metropolitan areas of New York State have a physical handicap that inhibits travel (1). Therefore, this group controlled the sample design: About 150 to 200 households would be needed to yield 30 to 50 households with a handicapped person and 120 to 150 households with one or more elderly persons.

The sample was drawn in March and April 1976 from Albany, Schenectady, Rensselaer, and a part of Saratoga counties by using the Albany and Rensselaer Metropolitan Telephone Book (1976). A systematic sampling strategy with a random starting point was used. Of 743 residences contacted, 578 had neither an elderly nor a handicapped person. The remaining 165 residences constitute the sample and are distributed as shown below.

Category	Number of Respondents			Percentage of Contacts
	Control Group	Modeling Group	Total	
Handicapped (includes some elderly)	6	29	35	4.7
Elderly (not handicapped)	15	115	130	17.5
Total	21	144	165	–

A 6-min questionnaire (2) was administered to these respondents about their travel patterns, perceptions of barriers, and travel characteristics.

TRAVEL CHARACTERISTICS

This section describes the travel habits and patterns of the elderly and handicapped respondents of the survey. The analysis also shows comparison statistics from other major studies, where possible. The major source of comparison statistics is the National Health Survey (NHS) (3). However, such comparisons can only be rough, because the NHS report summarizes only chronically disabled persons and uses slightly different question formats.

Table 1 shows the demographic data of the handicapped sample. The results show that the handicapped population is about equally divided into elderly and nonelderly groups and between those who require aid and those who do not. The sample also agrees well with the NHS data. Their physical problems are reflected directly in the mobility levels of the handicapped: Seventy-seven percent of the sample has at least some difficulty in getting about outside the home. Physical handicaps clearly imply special problems in transportation mobility as well. But at the same time, the handicapped as a group are not homogeneous; there exist wide differences in disability and extent of mobility within the group, which leads to quite different transportation problems and hence (probably) different solutions.

The sample screening procedure is such that persons interviewed as elderly are not also handicapped. Table 2 summarizes the demographic data of the elderly nonhandicapped and handicapped persons. The sample overestimates the younger elderly (65 to 70 years) and women. These discrepancies are probably due to the daytime telephone-interviewing procedure. However, the sample is generally consistent with the conventional wisdom in that women make up a higher proportion of the elderly than do men and that the incidence of physical handicaps generally increases with age. Table 3 summarizes the family sizes of the elderly and handicapped and the automobile-ownership characteristics of these families. The sample clearly demonstrates a smaller than average family size for the elderly and handicapped than for the general population. But, while the elderly and handicapped have mobility problems related to their physical situations, apparently they nevertheless have automobiles available through other members of their families.

As expected, travel by the elderly and handicapped is primarily nonwork oriented. Table 4 shows the frequency of travel for nonwork purposes. The elderly and handicapped make about 7.0 and 5.3 nonwork trips/week respectively. Work trips account for an additional 17 percent of their trips (Table 5). Only 11 percent of the respondents use transit for nonwork travel; however, transit-use rates range from 1 to 11 trips/week. Reliance on the automobile is heavy. However, as shown below, the handicapped are far more reliant on automobiles driven by others than are the elderly: Two-thirds of all nonwork trips by the handicapped are as automobile passengers, but only one-third of such trips by the elderly are as automobile passengers.

Mode Used	Percentage of Trips	
	Elderly	Handicapped
Automobile driver	60	30
Automobile passenger	26	52
Bus	12	15
Taxi, walk, or other	2	3

Further, as Table 6 shows, the private automobile is generally available to the elderly and handicapped, and most individuals can either drive or be driven in automobiles. For transit services, however, the picture is uneven. Regular bus service is generally perceived to be available, but special (e.g., client-agency) bus service is not. The limited awareness of special bus service also reduces its effectiveness.

The survey also asked respondents to identify any problems or barriers encountered in using bus service. Table 7 summarizes the responses and shows that many respondents were unable to identify any particular problem of a typical bus trip. These low numbers reflect the voluntary nature of the response: In most barrier studies, respondents are given a list of barriers and are asked to respond, so that their responses include yea-and-nay effects. While the handicapped generally perceive more barriers and view them as more severe, both groups generally stress the same items: climbing bus steps, lack of handrails, crossing streets and curbs, and seat comfort. The picture, then, is one of a wide range of perceptions of transportation barriers, with only a few such barriers being perceived as important by the group as a whole. Thus, specific improvements to the transportation system will probably not significantly reduce barriers for most travelers, and their effect on travel demand will probably be small. Further analysis shows that, if all transportation barriers were removed, the percentage of the elderly and handicapped who use the bus for at least 1 trip/week would increase from 11 to 21 if regular bus service were improved and to 27 if special bus service were improved. However, this is a noncommitment response; experience shows that actual increases would be only one-half to one-third as much.

MODELS OF TRAVEL

Four linear disaggregate models to estimate total nonwork trips and transit use were constructed from the data base. These were

1. Handicapped: total nonwork trips per week,
2. Handicapped: percentage of nonwork trips via transit,
3. Elderly: total nonwork trips per week, and
4. Elderly: percentage of nonwork trips via transit.

Table 1. Comparative data for the handicapped.

Descriptor	Percentage of Sample	Percentage of NHS Survey	Descriptor	Percentage of Sample	Percentage of NHS Survey
Sex			Age, years		
Female	59	56	≥65 (elderly)	49	54
Male	41	44	45 to 64	33	32
Special aid			17 to 44	18	11
Needed	49	51	≤17	0	3
Not needed[a]	51	49			

[a] Includes those confined to house.

Table 2. Comparative data for the elderly.

Descriptor	Sample: Nonhandicapped	Sample: Handicapped No.	Sample: Handicapped %	Sample: Total	Sample: % of Total Sample	Albany SMSA (1970) No.	Albany SMSA (1970) %	% Handicapped in NHS Survey (1972)
Sex								
Female	89	11	11	100	68	48 497	60	18.5
Male	39	8	17	47	32	32 389	40	16.2
Age, years								
65 to 70	59	5	8	64	44	26 502	33	–
71 to 75	35	5	13	40	27	22 058	27	–
76 to 80	13	5	28	18	12	16 119	20	–
>80	21	4	16	25	17	16 207	20	–
Total	128	19	147	13	100	80 886	100	17.6

Table 3. Size and automobile-ownership characteristics of families having elderly and handicapped members.

Descriptor	Sample No.	Sample %	Albany SMSA (1970) No.	Albany SMSA (1970) %	Descriptor	Sample No.	Sample %	Albany SMSA (1970) No.	Albany SMSA (1970) %
Family size					Automobile ownership				
1	43	27	45 120	20					
2	59	37	67 504	29	0	38	24	41 392	18
3	25	16	39 113	17	1	80	50	123 389	54
4	17	11	33 852	15	2	33	20	57 336	25
≥5	15	9	44 895	19	≥3	9		8 377	4
Total	159	100	230 484	100	Total	160	100	230 484	100

Note: The following mean values were obtained: for the sample population—family size of elderly = 2.49, family size of handicapped = 2.75, and automobile ownership = 1.08/family; for the Albany SMSA population—family size = 3.22 and automobile ownership = 1.14/family.

Table 4. Nonwork trip frequency.

Category	Nonwork Trips: n	Nonwork Trips: Average Trip Rate	Nonwork Trips: S.D.	Transit Use: n	Transit Use: Average (%)	Transit Use: S.D.
Handicapped	29	5.34	4.21	27	0.145	0.34
Elderly (nonhandicapped)						
Urban	60	6.73	4.12	–	0.21	–
Suburban	41	7.29	5.60	–	0.03	–
Rural	14	7.00	5.07	–	0	–
Total	115	6.96	4.79	110	0.13	0.32
Total complete samples	144	6.63	–	137	0.13	–
Total sample	165	5.80	–	–	–	–

The data for these models were analyzed separately by using stepwise linear regression methods and assuming the following general structures:

$$T_i = b_0 + b_1x_1 + b_2x_2 + \ldots \quad (1)$$

$$\%\,Tr_i = c_0 + c_1x_1 + c_2x_2 + \ldots \quad (2)$$

where

T_i = trips per week (nonwork for person i),
$\%Tr_i$ = percentage of trips via transit for person i,
$x_1, x_2, \ldots, x_n$ = independent variables,
$b_1, b_2, \ldots, b_n$ = coefficients, and
$c_1, c_2, \ldots, c_n$ = coefficients.

Nonlinear (e.g., logit) models were not attempted, but are a possibility for later analysis. Table 8 shows the variables available for input to each model.

The results of this analysis are shown in Table 9 and the statistical indexes of the values developed are shown in Table 10. Although numerous variables were available

Table 5. Detailed travel data.

Item	For Work (%)	For Nonwork (%)
No. of one-way trips per week		
0	88	9
1 to 3	1	18
4 to 7	2	43
8 to 11	8	17
12 to 15	1	12
≥16	0	1
Mean	1.06	5.80
Mode used		
Automobile driver	79	53
Automobile passenger	11	32
Bus	5	11
Taxi, walk, or other	5	4
Frequency of bus use, trips per week		
0	90	83
1 to 3	5	6
4 to 7	5	9
8 to 11	0	2

Table 6. Availability to elderly and handicapped of modes.

Item	No.	%
Availability of automobile		
Always	112	71
Most of the time	15	10
Occasionally	11	7
Never	19	12
Total	157	100
Ability to use automobile		
Drive with no difficulty	85	55
Drive with some difficulty	6	4
Ride only, but no difficulty	46	30
Ride only, with some difficulty	11	7
Ride only and need help	5	3
Total	153	100
Availability of regular bus service		
Nearby and frequent	77	48
Nearby but infrequent	25	16
None nearby	50	31
Don't know	7	4
Total	159	100
Availability of special bus service		
Yes	16	10
No	62	39
Don't know	81	51
Total	159	100

Table 7. Barriers to using bus service.

	Handicapped			Nonhandicapped Elderly		
Barrier	Percentage Per Perceiving[a]	Avg Barrier Level[b]	Barrier Index[c]	Percentage Per Perceiving[a]	Avg Barrier Level[b]	Barrier Index[c]
Transportation						
Reading schedules	18.5	1.4	25.9	1.9	1.0	1.9
Reading maps	22.2	1.5	33.3	2.8	1.0	2.8
Getting information over telephone	18.5	1.4	25.9	4.7	1.0	4.7
Uneven ground and slopes	40.7	1.27	51.7	8.5	1.11	9.4
Street crossings and curbs	40.7	1.18	48.0	8.5	1.0	8.5
Bad weather	33.3	1.44	48.0	14.2	1.07	15.2
Fear of crime	14.8	1.25	18.5	3.8	1.0	3.8
Distance to vehicle	29.6	1.25	33.0	6.6	1.43	9.4
No shelter	7.4	1.50	11.1	3.8	1.25	47
Wait too long	25.9	1.29	33.4	4.7	1.2	56
Climbing steps	59.3	1.31	77.7	17.0	1.11	18.9
No handrails	40.7	1.27	51.7	9.4	1.10	10.3
Crowding or rushing	29.4	1.25	36.8	4.7	1.0	4.7
Handling change or tokens	14.8	1.25	18.5	0.9	1.0	0.9
Cost	7.4	1.00	7.4	0.9	1.0	0.9
Not enough time to sit down	37.0	1.20	44.4	6.6	1.14	0.75
Getting to seat near back	14.8	1.25	18.5	1.9	1.0	0.19
No space for wheelchair, crutches, or such	20.2	1.40	28.3	0	0	–
Seats not right	18.5	1.0	18.5	7.3	1.38	1.01
Lack of comfort	18.5	1.0	18.5	1.9	1.00	0.19
Swaying and lurching	7.4	1.0	7.4	4.5	1.67	0.75
Travel time too long	11.1	1.0	11.1	6.4	1.14	0.73
Pull cord	19.2	1.0	19.2	1.9	1.0	0.19
Pushing door open	23.1	1.0	23.1	4.7	1.0	0.47
Place						
Uneven ground and slopes	26.9	1.29	34.7	6.6	1.0	0.66
Street crossings and curbs	30.8	1.13	34.8	6.6	1.0	0.66
Climbing steps	50.0	1.38	69.0	14.8	1.13	1.67
Opening doors	30.8	1.13	34.8	3.6	1.0	0.36
Unfamiliar areas	11.5	1.00	11.5	1.8	1.0	0.18
Cannot go very far or fast	38.5	1.20	46.2	9.4	1.1	1.03

[a] Percentage of respondents mentioning a given barrier.
[b] Average severity level (1 = some problem and 2 = severe problem) for those perceiving this barrier.
[c] Percentage perceiving times average level.

Table 8. Variables used in model building.

Variable	Form	Variable	Form
Personal		Trip	
Age	In years	Mode	Transit or other
Sex	M or F	Travel time	Minutes perceived
Disability	Type of ailment (1 to 9 scale)[a]	Trip length	Miles perceived
		Automobile availability	1 to 4 scale[a]
Aid	Type of aid used (1 to 6 scale)[a]	Ability to use automobile	1 to 6 scale[a]
Extent of disability	Degree of disabledness (1 to 5 scale)[a]	Regular bus availability	1 to 3 scale[a]
Family size	1, 2, ...	Special bus availability	1 to 3 scale[a]
Automobiles owned by family	0, 1, 2, ...	Barriers listed in Table 7	0 to 2 scale[a]

[a] Increasing severity.

Table 9. Summary of travel-demand models.

	Handicapped				Nonhandicapped Elderly			
	Trip Generation		Modal Split		Trip Generation		Modal Split	
Variable	Value	t-Statistic	Value	t-Statistic	Value	t-Statistic	Value	t-Statistic
Constant	5.07	—	0.15	—	10.11	—	-0.06	—
Family size	1.42	14.54	—	—	+0.56	4.83	—	—
Physical-aid index	—	—	-0.11	5.84	—	—	—	—
Ability to use automobile	-1.45	7.86	—	—	—	—	—	—
Automobile unavailability	—	—	+0.15	6.82	-0.97	5.53	+0.21	106.17
Bus unavailability	—	—	—	—	-1.24	8.36	-0.079	12.83
Bus-steps barrier	—	—	—	—	-3.41	13.6	—	—

Table 10. Statistical indexes of models.

	Handicapped				Nonhandicapped Elderly			
	Trip Generation		Modal Split		Trip Generation		Modal Split	
Variable	Value	t-Statistic	Value	t-Statistic	Value	t-Statistic	Value	t-Statistic
Statistical index								
R^2	0.45	—	0.39	—	0.21	—	0.57	—
Standard error of estimate	3.32	—	0.28	—	4.33	—	0.21	—
F-ratio	10.71	—	7.61	—	7.44	—	70.26	—
n	29	—	27	—	115	—	110	—
Means								
Constant	5.34	4.21	0.145	0.34	6.96	4.79	0.13	0.32
Family size	2.75	1.64	—	—	2.49	1.62	—	—
Physical-aid index	—	—	2.07	1.20	—	—	—	—
Ability to use automobile	2.51	1.18	—	—	—	—	—	—
Automobile unavailability	—	—	1.46	0.94	1.60	1.02	1.59	1.01
Bus unavailability	—	—	—	—	1.87	0.97	1.84	0.95
Bus-step barrier	—	—	—	—	0.20	0.44	—	—

Table 11. Elasticities.

	Handicapped		Nonhandicapped Elderly	
Variable	Trip Generation	Modal Split	Trip Generation	Modal Split
Family size	0.72	—	0.56	—
Physical-aid index	—	-1.56	—	—
Ability to use automobile	-0.68	—	—	—
Automobile unavailability	—	1.53	-0.22	2.56
Bus unavailability	—	—	-0.33	-1.12
Bus-steps barrier	—	—	-0.43	—

Note: Calculated at the mean.

for entry into the models, only a few did so. These were primarily automobile and bus unavailability, ability to use an automobile, physical disability (reflected in the aid index), and family size. With one exception (elderly modal choice), barriers did not enter the models. The influence of automobile availability is clear: increasing levels of automobile availability increase total nonwork travel, but decrease the propensity to use transit. For the elderly, increasing levels of transit availability influence both total travel and modal choice. The effect of family size is to increase total travel by the elderly and handicapped by providing other household members as chauffeurs and increasing family ties for the elderly or handicapped individual. The absence of barriers in these models suggests that, generally, travel patterns of the elderly and handicapped depend primarily on the availability of transportation service and not on the degree to which such service, when available, is barrier free. These findings confirm the conclusions of other studies [e.g., the Institute for Public Administration Planning Handbook (4) and Knighton and Hartgen (5)], which emphasize service rather than vehicle in transit system design. The high modal-split elasticities for service availability (Table 11) underscore these results.

DISCUSSION AND SUMMARY

This report describes a recent study in the Albany-Schenectady-Troy SMSA of nonwork travel habits of the elderly and handicapped. The study was based on a sample of 165 elderly and handicapped persons, who were telephoned at random. The significant results of the study are that

1. The elderly and handicapped are not a homogeneous group, either separately or together: There are wide variations in travel behavior and mobility problems within each group;
2. The elderly and handicapped average about 7.0 and 5.3 one-way nonwork trips/week respectively;
3. Automobile availability to the elderly and handicapped is not signficantly less than that to the general population;
4. Travel of these groups is primarily by automobile, either as a passenger or a driver with bus travel constituting only about 13 percent of their nonwork trips;
5. For the handicapped, travel mobility is primarily a function of personal disability and the ability of the individual to use an automobile: Bus service improvements would appear to change this picture only slightly;
6. Specific barriers on the public bus system do not materially affect either total nonwork travel or modal split, but the availability of bus transportation affects both;
7. Bus systems that emphasize availability (coverage and frequency) as well as direct pickup appear to be the most promising for increasing the mobility of the elderly and handicapped; and
8. The widely divergent needs of these individuals imply that very specialized solutions will probably be required to solve their transportation problems.

A set of small-sample disaggregate models was developed to enable prediction of elderly and handicapped nonwork travel and modal choice. The models are generally sensitive to automobile and bus availability, family size, and the level of disability of the individual.

REFERENCES

1. E. P. Donnelly and others. Statewide Public-Opinion Survey on Transit Operating Assistance: Technical Report. Planning Research Unit, New York State Department of Transportation, Preliminary Res. Rept. No. 80, 1975.
2. D. T. Hartgen, M. Pasko, and S. M. Howe. Forecasting Nonwork Travel Demand for the Elderly and Handicapped. Planning Research Unit, New York State Department of Transportation, Preliminary Res. Rept. No. 108, 1976.
3. C. S. Wilder. Limitation of Activity and Mobility Due to Chronic Conditions, U.S. 1972. National Center for Health Statistics, Rockville, Rept. 10:96, Nov. 1974.
4. Planning Handbook: Transportation Services for the Elderly. Institute for Public Administration, Administration on Aging, Nov. 1975.
5. R. G. Knighton and D. T. Hartgen. Incorporating Barrier Effects in Elderly and Handicapped Demand Forecasts. New York State Department of Transportation, Draft Rept., Oct. 26, 1976.

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

**Dr. Howe and Mr. Pasko were with the New York State Department of Transportation when this research was performed.*

Policy-Contingent Travel Forecasting With Market Segmentation

Frederick C. Dunbar, Charles River Associates, Cambridge, Massachusetts

Market segmentation of travel data gives a data base that is easy to use and interpret. This paper develops methods for tabulating travel data so that disaggregate travel-demand models can be applied to market segments. These methods result in improved accuracy of travel forecasts because aggregation bias is reduced. The approach also allows nearly immediate computation of demand elasticities. These procedures can be applied to most urban travel-data files by using cross-tabulation software. To demonstrate the methods and their accuracy, the work-trip modal split is simulated on Nationwide Personal Transportation Survey data by using a disaggregate logit model. Travel demand is forecast under a variety of transportation policies that involve automobile controls and transit level-of-service improvements.

An approach to the use of market segments with existing disaggregate demand models has been developed. The advantages of such an approach include accurate travel-demand forecasts with minimal data and computational resources. In the present case, the effects of a policy scenario can be calculated by most programmable calculators or within a few hours by hand.

The use of market segments is not a new technique. Usually, market segments are defined by the characteristics of the trip maker rather than by those of the trip. However, travel data are sometimes cross tabulated by distance and time as well as the socioeconomic characteristics of the trip makers. This format has been useful in segmenting the travel market so that the impact of policies on particular socioeconomic groups can be emphasized (1). Market segmentation has the additional advantage of reducing aggregation error when such data are analyzed with disaggregate logit models.

The application of multinomial logit models to market segments is actually an extension of the early development of logit analysis. Models of binary choice were originally developed from the application of statistical tools to contingency tables (2). These models gave the probability that a response to a stimulus would occur within a specified range. For a simple univariate model, a table giving the proportions of the sample that will respond at each level of stimulus will have sufficient information for the estimation of the model. Similarly, given a model such as an estimated logit equation, the proportion of a sample that will respond to stimuli within given ranges can be predicted.

This approach can be generalized to the common specification of disaggregate modal-split models. If only two modes are considered, then the response will be the proportion of trips by a given mode, for example, automobile. The approach becomes computationally more complex as the number of different types of stimuli (independent variables such as modal attributes) increases. Rather than a column of numbers representing the sample at each level of stimulus, a multidimensional array representing the number of travelers who have alternative levels of service among modes becomes necessary.

DATA AND MODEL PREPARATION

The data base used to test this approach is the journey-to-work trip record from urban households in the Nationwide Personal Transportation Survey (NPTS) (3). Thus, the results of simulations with these data can be viewed as representing the effects of national policies. Alternatively, the data can be used to reflect the effects of ubiquitous transportation level-of-service changes in an average urban area. Independent of the interpretation of the results of policy scenarios, the use of market segments with NPTS data can be replicated with data available from the transportation planning activities of most urban areas.

The original home-interview tape from the survey was cross tabulated into market segments suitable for application of the original Charles River Associates work-trip modal-split model (4). This was a three-stage process: (a) the relevant variables are identified from the demand model, (b) the market segments are formed from the home-interview tape, and (c) the variables representing market segments are constructed for application of the demand model.

MODAL-SPLIT MODEL AND VARIABLES

The general form of the logit modal-split model is as follows:

$$P(a) = 1 \bigg/ \left\{ 1 + \sum_{\substack{i=1 \\ i \neq a}}^{n} \exp[-\alpha(x_a - x_b) - \beta_y] \right\} \tag{1}$$

$$P(i) = \exp[-\alpha(x_a - x_i) - \beta_y] \bigg/ \left\{ 1 + \sum_{\substack{j=1 \\ j \neq a}}^{n} \exp[-\alpha(x_a - x_j) - \beta_y] \right\} \tag{2}$$

$$\ln[P(a)/P(i)] = \alpha(x_a - x_i) + \beta_y \qquad (i \neq a) \tag{3}$$

where

$P(a)$ = probability of automobile drive alone being the chosen mode;
$P(i)$ = probability of alternative i being the chosen mode;
x_a = vector of costs and times for making the trip by the automobile-drive-alone mode;
x_i = vector of costs and times for making the trip by mode i;
y = vector of socioeconomic variables and mode-specific constants; and
α and β = estimated vectors of coefficients for the time, cost, and socioeconomic variables and for the mode-specific constants.

For the purposes of exposition, Equation 3 will be used. The estimated model is given by Equation 4.

$$\ln[P(a)/P(b)] = -4.77 - 2.24(C_a - C_b) - 0.411(T_a - T_b) - 0.114(S_a - S_b) + 3.79Y \tag{4}$$

where

$P(b)$ = probability of transit being the chosen mode;
C = costs of making the round trip by automobile (a) or transit (b) ($);
T = in-vehicle and wait times for the round trip by automobile (a) or transit (b) (min);
S = access walking time for the round trip by automobile (a) or transit (b) (usually assumed to be zero for automobile trips) (min); and
Y = automobiles per worker in the household.

Because the model and its estimation are described in detail in other places, it will not be evaluated here except to note some of its t-statistics (4). The t-statistics of the coefficients in Equation 4 are given below.

Value	t-Statistic	Value	t-Statistic
4.77	3.88	0.114	2.69
2.24	4.53	3.79	4.06
0.0411	1.96		

For the sample size used to estimate Equation 4, which was 115 observations, t-statistics of 1.96 and 2.33 indicate that a parameter is significantly different from zero at the 2.5 and 1 percent levels of significance respectively for a one-tailed test, which means that all of the estimated parameters are highly significant. Another test of the model is whether the predicted probability of the selected mode for individuals is greater than 0.5. Equation 4 performed well in this respect also. The model predicted the correct choice of mode for 107 of the 115 observations that were used in its estimation, which is an accuracy level of 93 percent.

Construction of NPTS Market Segments

To construct the NPTS market segments, the work-trip records from the home-interview survey of urban areas were cross tabulated across three variables: trip distance, access distance to transit, and automobile availability. In the data base, there were 1774 such trips recorded. Of these, 221 were eliminated on error checks, usually because there were insufficient data on the record. Another 101 trips were eliminated because they involved more than one mode of travel. The remaining 1452 trip records form the basis of the market segments used for the analysis.

The market-segment categories are described below.

1. Distance—trip distance was divided into two categories with the following ranges: (a) short trips are less than 14.6 km (9.1 mile) and (b) long trips are greater than or equal to 14.6 km (9.1 mile). Several different methods could have been used for the determination of the ranges for the short and long-trip categories. For example, the dividing line could have been the median trip distance or that distance for which the total vehicle kilometers of travel (VKT) in each category are equal. In the ranges actually used, the mean trip distance [14.6 km (9.1 mile) on a round trip basis] was used as the dividing line; this number is between those that result from using the other two rules. The average distance characteristics for each mode category—where the mode categories are defined as (a) drive alone is automobile, truck, or motor-cycle; (b) transit is bus, streetcar, commuter train, subway, or elevated; and (c) car pool is automobile with other persons—are given below (1 km = 0.62 mile).

Mode	Distance (km) Short Trips	Distance (km) Long Trips	Time (min) Short Trips	Time (min) Long Trips
Automobile drive-alone	13.89	53.52	29.04	64.53
Transit	14.05	59.10	50.95	109.83
Car pool	12.81	58.73	31.33	73.85

2. Transit accessibility—the transit accessibility categories were determined by the distance from home to the nearest public transportation that could be used for the journey to work. The data were originally coded in blocks and were later transormed to kilometers [one block equals approximately 0.13 km (0.083 mile)]. These categories and their ranges in distance are as follows: (a) high transit accessibility is zero to two blocks, (b) middle transit accessibility is three to six blocks, and (c) low transit accessibility is more than six blocks. These ranges, which were selected after an examination of more refined breakdowns, showed the groupings that would tend most to equalize the number of trips among categories. The trip characteristics for these categories are summarized below (1 km = 0.62 mile).

Transit Access	Distance (km)	Trips (%)
High	0.109	37.7
Middle	0.574	16.3
Low	2.132	46.1

3. Automobile availability—household automobile availability was divided into the following two categories: (a) automobiles per worker are less than or equal to 0.5 and (b) automobiles per worker are greater than 0.5.

These categories arise naturally from the bimodal distribution of the data; most work trip makers have either 0 or 1 automobile/worker in the household. These characteristics are summarized below.

Automobile Availability	Automobiles per Worker	Trips (%)
<0.5	0.020	89.7
>0.5	1.753	10.3

Table 1 gives the modal splits, number of trips, and VKT for each of the twelve market segments. The modal splits and total trips were computed directly from the data, but some assumptions were necessary to compute the VKT. That part of the VKT that can be attributed to the automobile-drive-alone mode is the sum of the round-trip distances for each trip made by this mode. However, the information in the data base does not allow a direct compuation of the VKT incurred by car pools because the distribution of car-pool sizes, i.e., the number of passengers per vehicle, is not known and therefore the number of vehicles used for this mode is not known. To derive an estimate of the VKT that can be attributed to car pools, a distribution of one, two, and three-passenger car pools was created, and each person in the car pool was credited with an equal share of the car-pool VKT. The distribution of car-pool sizes is derived from the predictions of the modal-split model. This distribution varies from cell to cell, but its aggregate ratio is 0.78:0.17: 0.04 for one-passenger: two-passenger: three-passenger car pools respectively. The NPTS distribution, tabulated from a different part of the survey, is that, for all travel, the ratio of car-pool sizes is 0.72:0.17:0.11 for one-passenger: two-passenger: three-passenger carpools respectively. Thus, the two independent estimates of passengers per automobile are in reasonably close agreement.

There is a large amount of flexibility in deciding the number of variables to be cross tabulated, the number of categories to be used, and the ranges to be applied. The decisions made about each of these issues reflected a desire to minimize the number of market-segment cells and, at the same time, capture the essential information of the modal-split model in data points having small associated variances. New variables and more refined breakdowns of the variables already chosen increase the number of cells multiplicatively rather than additively; for example, if in addition to the variables already chosen, a cross tabulation that used two categories of trip time was performed, the number of market-segment cells would increase from 12 to 24. Unless broad ranges of categories are created and relatively few variables are selected, the data base can easily become overly cumbersome, which loses the advantage of using market segments.

Although the choices of ranges and variables are basically rather arbitrary, there were some rules and reasons behind the decisions actually made. Some of the more important of these (in addition to those already presented) are listed below.

1. The variables were selected to conform to the independent variables in the logit model. Both access time to transit and automobiles per worker are direct inputs to the model, and the model treats line-haul costs and times as functions of trip distance, which makes this variable an obvious choice on which to make a cross tabulation.

2. Although trip-time data are available and are an input of the model, trip time is so closely proportional to trip distance that it was deemed unnecessary to create an extra variable for cross tabulating by time or trip.

3. Those variables that contribute most to the aggregation problem require more refined categories. Earlier research has indicated that automobiles per worker and access to transit cause more variation in logit-model log-odds functions than do other variables (5). Subdivision into eighteen market segments did not substantially increase the accuracy of model predictions.

Thus, the market segments created are dictated by the requirements of the model and the empirical testing of its performance. In this sense, the market segments presented here are intended to suggest things that can be done for the application of nonlinear disaggregate models. Because models and data bases vary, the cross tabulations performed by others for policy-evaluation purposes will also vary. In particular, the classic purpose of market segmentation is to emphasize socioeconomic groupings rather than trip characteristics. The methodology for market segmentation in such a case would be quite different.

Construction of Mode-Specific Variables

The independent variables required for application of the modal-split model must be constructed from the variables used for creating the market-segments data. The variables of the model, in the two-mode case of automobile drive alone and transit, are given in Equation 4. The variables available from the data have been given above. In addition to the two modes represented in Equation 4, it is also useful to construct data that represent automobile with passenger modes.

The formulas for constructing the mode-specific variables are given below:

1. Automobile drive alone: C_a = 0.035 × automobile-drive-alone distance, T_a = automobile-drive-alone

time, and $S_a = 0$;

2. Transit: $C_b = 0.4928$, T_b = transit time, and $S_b = 2 \times 19 \times$ distance to transit.

3. Car pool with k passengers: $C_{ck} = C_a[1 + (k/3)]/(k + 1)$, T_{ck} = [(k/3) × automobile-drive-alone distance/(car-pool distance for short trips ÷ car-pool time for short distance)] + automobile-drive-alone distance/(car-pool distance ÷ car-pool time) + (20 × k), and $S_{ck} = 0$;

4. Driver serve passenger: $C_d = 2 \times C_a$, $T_d = 3 \times T_a$, and $S_d = 0$; and

5. Automobiles per worker: y = 0 if <0.5 automobiles/worker or 1 if >0.5 automobiles/worker.

Most of these equations are self-explanatory, but the following assumptions should be noted:

1. The cost for an automobile trip is \$0.022/km (\$0.035/mile) [1969 data (6)].
2. Transit fare is the 1969 national average of \$0.4928 for a round trip (7).
3. Walking speed to transit is 12 min/km (19 min/mile).
4. For each distance category, the average automobile-drive-alone trip distance increases by one-third for picking up and dropping off each potential car-pool passenger (8).
5. Car-pool passengers make arrangements to share costs equally.
6. Car-pool line-haul speeds for picking up and dropping off passengers are equal to the speed for car-pool trips in the short-distance category.
7. The schedule delay associated with each potential car-pool passenger is 20 min.
8. The driver-serve-passenger mode involves a household member who drives the trip maker to work and returns home for the first leg of the round trip and then drives from home to the workplace and returns with the passenger for the second leg.
9. The automobile-per-worker variable was set to zero or one, and the model was calibrated on the automobile-per-worker coefficient to obtain a value of 4.60.

Although most of the above assumptions represent straightforward interpretations of the data, the heuristic nature of the construction of the automobile-with-passenger variables deserves further comment. In the absence of adequate level-of-service data on the availability of car-pool alternatives to non-car-pool, work trip makers, some judgments about this mode are necessary. When the problem of designing an optimal household survey to collect car-pool data is considered, it is easy to see why such data do not exist. Meanwhile, the modeling of shared rides will continue to be one of the weakest parts of the total travel-demand system. The major justifications for using the approach described here are that the assumptions are consistent with intuition about car pools and that the model predicts car-pool modal split reasonably well.

MODEL APPLICATION AND PERFORMANCE

The use of the model to predict modal splits and VKT has the following steps:

1. Each of the mode-specific variables for each of the 12 market segments is constructed by using the formulas and data given above.
2. For each market segement, a log-odds function for the automobile-drive-alone mode versus each of the other modes is calculated by using Equation 4, the variables constructed in the previous step, and 4.60 substituted for the coefficient on y.
3. For each market segment, the probability of an individual choosing a given mode, other than automobile drive alone, is computed by using Equation 2. The automobile-drive-alone probability is computed from Equation 1.
4. The modal splits for each market segment are computed as follows: (a) automobile-drive-alone modal split = automobile-drive-alone modal-choice probability; (b) transit modal split = transit modal-choice probability; and (c) car-pool modal split = sum of one-passenger, two-passenger, and three-passenger car-pool and driver-serve-passenger modal-choice probabilities.
5. The VKT for each market segment is the sum of the following VKT calculations for each mode: (a) automobile-drive-alone VKT = automobile-drive-alone modal-choice probability × automobile-drive-alone distance × total trips, (b) k-passenger car-pool VKT = k-passenger car-pool modal-choice probability × [1 + (1/3)] × automobile-drive-alone distance × total trips, and (c) driver-serve-passenger VKT = driver-serve-passenger modal-choice probability × 2 × automobile-drive-alone distance × total trips.
6. The aggregate modal split is computed as the weighted average of the predicted modal splits for each market segment.
7. The aggregate VKT is computed as the sum of the VKT across the market segments. With these procedures, the model was used to predict the modal splits and VKT for each of the cells in the NPTS market-segment data base. The actual values and the aggregate predictions are given below (1 km = 0.62 mile).

Value	Modal Split: Automobile Drive Alone	Transit	Car Pool	VKT (without driver serve passenger)	VKT
Actual	0.637	0.160	0.202	31 849	–
Predicted	0.635	0.159	0.206	31 335	32 040

The predicted modal splits conform closely to the actual modal splits. The first VKT (that without driver-serve passenger values) corresponds to the VKT that can be calculated from the data and does not include the VKT that are attributable to one-half of the driver-serve-passenger trips (that half which is traveled by the driver without a passenger is not captured by the NPTS data). The second VKT includes all of the VKT associated with driver-serve-passenger trips as well as with other automobile-oriented trips. When the first VKT is used as a basis for comparing the predicted to the actual, the model predicts the VKT within 1.6 percent. For most applications of the model, this error is well within the range of predicted effects and within the errors that might have other causes, such as data errors or parameter estimation errors. In general, the model performs well in replicating the aggregate figures from the data.

The modal split and VKT estimate for each market segment are given in Table 1. A comparison of the actual and predicted values indicates potential biases and the areas of greatest error. As expected, the error associated with any given market segment is greater than the aggregate error. The highest errors are those associated with the market segments that have the fewest total trips (basically, the six market segments in which automobiles/worker <0.5). There appears to be some tendency for the model to overpredict driver-serve-passenger trips for short-distance trips and underpredict car-pool trips for long-distance trips. In general, the errors associated with individual market segments tend to cancel when aggregated.

FORECASTING EFFECTS OF TRANSPORTATION ALTERNATIVES

The procedures developed above were applied to a variety of transportation policy scenarios to forecast the effects of these policies on trip-making behavior. The approach to investigating a particular policy is relatively straightforward: The policy is examined from the question of how it would affect the independent variables in the logit model. This effect is quantified by changing the values of the independent variables from those that they were in the base case. With the new values of the variables, the logit model is applied to the NPTS market-segments data and modal splits, and the VKT are forecast. The predicted modal splits and the VKT with the policy effects are then compared to the base-case predictions to forecast the impact of the policy.

Gasoline Tax

The model was used to predict the effects of a 100 percent gasoline tax [in addition to the existing gasoline taxes (which are assumed to be 7 percent to the state and 4 percent to the federal government)]. One of the purposes of this exercise is to compute the implied price elasticity of gasoline. This provides a test of the approach because the result can be compared to other, independent gasoline-price-elasticity estimates.

The effect of a 100 percent gasoline tax will be to increase automobile operating costs per kilometer by 50 percent. The pump price of gasoline is increased by 69 percent when a 100 percent tax rate is applied to the pretax cost of gasoline.

The forecasts of aggregate modal split and VKT under the assumption of a 100 percent gasoline tax are given in Table 2. The elasticities of the VKT are -0.256, -0.184, and -0.128 with respect to automobile operating costs, the pump price of gasoline [which is within the range of short-run elasticities estimated by econometric studies of gasoline demand (9)], and the pretax cost of gasoline respectively. The predicted 12.8 percent decrease in VKT is predicted to occur as a one-third increase in transit trips, an 11.2 percent decrease in automobile-drive-alone trips, and an 8.7 percent increase in car-pool trips.

Because of space limitations, the effects by market segment, which have been discussed by Charles River Associates (8), are not given here. The gasoline tax has its greatest impact on long trips with good to fair transit access. This is to be expected because, on a per-trip basis, the tax will have its highest dollar impact on long trips. The result is that the model predicts a higher incentive for mode switching on long trips for this scenario.

Transit Speed

In this scenario, it is assumed that a combination of shorter headways and faster transit will cause a uniform 10 percent decrease in transit line-haul-plus-wait time per trip. Access time to transit is assumed to be unchanged. This scenario was modeled by multiplying the transit line-haul-plus-wait time by and then applying the logit model to the NPTS market segments.

The predicted aggregate effects of this policy are given in Table 2. The predicted decrease in VKT was 3.22 percent, and the predicted increase in transit trips was 12.6 percent. One of the interesting results of this exercise is the relatively high elasticity of transit modal split with respect to transit speed (1.26). The biggest impacts occur on relatively long trips with good to medium transit access. As with the case of a gasoline tax, the effect of a uniform percentage decrease in transit time will have its largest absolute impact on long trips. Consequently, those trip makers who have longer trips have the most incentive to switch modes. The 10 percent decrease in transit time implies a saving of about 10 min for long trips, but only about 5 min for short trips. Also, as would be expected, the transit-speed policy has little predicted effect on trip makers who have poor access to public transit.

Transit Access:

Uniform Improvement

Because the weights that trip makers place on access time to transit are higher than the weights that they place on line-haul time, it can be assumed that the effect of decreasing access time would be greater than would be the effect of decreasing line-haul-plus-wait time. The results of various transit access scenarios indicate that this hypothesis deserves more consideration.

The first of a series of scenarios for the improvement of transit access involved decreasing transit-access time by a uniform 10 percent for all market segments. In the base-case projections, the access times to transit for high, middle, and low-access categories were 2.58, 14.25, and 50.35 min respectively. Thus, only short transit trips with poor access would have time savings for equal percentage declines in access time that were equivalent to those for equal percentage declines in line-haul-plus-wait time. In all

Table 1. Actual and predicted modal splits and VKT for NPTS market segments.

Market Segment			Observed Values					Predicted Values				
			Modal Split					Modal Split			VKT	
Automobiles per Worker	Trip Length	Transit Access	Automobile Drive Alone	Transit	Car Pool	Total Trips	VKT	Automobile Drive Alone	Transit	Car Pool	Without Driver Serve Passenger	With Driver Serve Passenger
>0.5	Short	High	0.603	0.197	0.200	315	3 002	0.598	0.140	0.261	3 041	3 286
>0.5	Short	Middle	0.589	0.113	0.298	124	1 228	0.667	0.041	0.292	1 335	1 442
>0.5	Short	Low	0.780	0.006	0.214	355	4 285	0.695	0.001	0.303	3 984	4 304
>0.5	Long	High	0.644	0.178	0.178	135	5 337	0.559	0.345	0.096	4 404	4 407
>0.5	Long	Middle	0.711	0.132	0.158	76	3 231	0.749	0.122	0.129	3 323	3 324
>0.5	Long	Low	0.798	0.027	0.175	297	14 158	0.851	0.002	0.147	14 758	14 767
<0.5	Short	High	0.014	0.836	0.151	73	77	0.040	0.942	0.018	48	51
<0.5	Short	Middle	0.083	0.917	0.000	24	27	0.132	0.811	0.058	51	55
<0.5	Short	Low	0.000	0.083	0.917	12	63	0.650	0.056	0.284	126	137
<0.5	Long	High	0.042	0.792	0.167	24	167	0.016	0.981	0.003	23	23
<0.5	Long	Middle	0.083	0.833	0.083	12	82	0.057	0.933	0.010	40	40
<0.5	Long	Low	0.400	0.000	0.600	5	191	0.696	0.184	0.120	203	203

Note: 1 km = 0.62 mile.

Table 2. Aggregate effects of travel-forecasting scenarios.

Scenario	Modal Split: Automobile Drive Alone	Transit	Car Pool	VKT
Base case	0.635	0.159	0.206	32 040
100% gasoline tax	0.564	0.212	0.224	27 940
10% transit-speed increase	0.619	0.179	0.209	31 010
10% transit-access-time decrease	0.631	0.163	0.205	31 780
Low transit-access improvement	0.600	0.202	0.198	29 795
Low and middle transit-access improvement	0.583	0.225	0.193	28 735

Note: 1 km = 0.62 mile.

other cases, the time savings from a 10 percent reduction in line-haul-plus-wait time would be much greater than the time savings from a 10 percent reduction in access time. This helps to explain some of the results given below.

The aggregate effects of this policy are given in Table 2. The decrease in the VKT caused by this policy is predicted to be 0.7 percent, and the predicted increase in transit trips is 2.5 percent. Both the VKT and the transit-ridership elasticities are much lower for access times than for line-haul-plus-wait times.

The market segments having the greatest impact are those where access to transit is in the middle category; those with good transit access are relatively insensitive to further improvements, and those with poor access would not find a 10 percent improvement sufficient inducement to switch modes.

Low-Transit-Access Improvement

The results of the previous section indicate that making transit available to everyone would induce significant increases in transit ridership. Therefore, this scenario assigned to the low-transit-access market segment the same access time that the middle-access group currently has. All other variables were unchanged although it is unlikely that any real transit service design that provided such a large change would not also affect accessibility in other market segments and line-haul and wait times in all market segments.

The aggregate results are given in Table 2. The change in average access for the whole population is 62.9 percent, and the decrease access time for the market segment that previously had low transit availability was 71.7 percent. This change caused a decrease in VKT of only 7 percent for an elasticity of 0.111 and a transit-patronage increase of 27 percent for an elasticity of -0.429. These elasticities are higher than those in the previous access-time scenario. For households in which the number of automobiles per worker is greater than 0.5, the predicted change in VKT is 10.2 percent. The effect of the policy on households with low automobile-ownership rates is quite dramatic, but because these contribute relatively little to the VKT, they have a small impact on the aggregate effects.

Low and Middle-Transit-Access Improvements

To evaluate the effect of improving transit access for those in the middle-transit-access market segment the low access category is again assigned the same access that the middle-access group currently has and the middle-access group is assigned the access time that the high-access group currently has.

The aggregate results of this policy are given in Table 2. The percentage change in VKT is 10.3, and the implied elasticities are somewhat higher for both VKT and transit ridership. The conclusion that may be drawn from this series of scenarios is that improvements in transit access are more effective when moderate service is made better than when poor service is made only adequate.

CONCLUSION

The preceding results show that the use of market segments with behavioral demand models is promising for quick policy contingent forecasting. The examples presented are somewhat simplistic and indicate that a module that translates complex policy issues and planning alternatives into quantifiable demand-model inputs is needed. This module could be a manual activity that uses existing planning resources to determine the effects of a policy or system on the level of service for the relevant market segments. Moreover, this approach would allow quick parametric representations of level-of-service changes that are consistent with Urban Mass Transportation Administration guidelines for alternatives analysis. Other areas of future research include applying the approach to nonwork trips and linking the demand effects with cost models to determine the cost-effectiveness of and trade-offs among policies.

ACKNOWLEDGMENTS

The research for this paper was performed under a contract from the Transportation Systems Center, U.S. Department of Transportation. Donald Ward of the Transportation Systems Center, the technical monitor for the project, gave appreciated critical guidance to the project. Geoffrey Lowney performed the invaluable and underrated tasks of decoding the NPTS data files, cross tabulating the data, and simulating the demand model.

REFERENCES

1. C. H. Lovelock. A Market-Segmentation Approach to Transit Planning, Modeling, and Management. Proc., Transportation Research Forum, Vol. 16, No. 1.
2. D. R. Cox. The Analysis of Binary Data. Methuen, London, 1970.
3. Home-to-Work Trips and Travel. Federal Highway Administration, Nationwide Personal Transportation Survey Rept. No. 8, Aug. 1973.
4. T. A. Domencich and D. McFadden. Urban Travel Demand: A Behavioral Approach. North-Holland, Amsterdam, 1975.
5. F. C. Dunbar. Quick Policy Evaluation With Behavioral Demand Models. TRB, Transportation Research Record 610, 1976, pp. 47-49.
6. Statistical Abstract of the United States. U.S. Bureau of the Census, 1971, Table 854.
7. Transit Fact Book. APTA, March 1975, p. 20.
8. Estimating the Effects of Urban Travel Policies. Charles River Associates, Cambridge, Mass., April 1976.
9. Policies for Conserving Fuel. Charles River Associates, Cambridge, Mass., in preparation.

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

Effects of Small Sample Origin-Destination Data on Transportation Study Results

Ralph E. Foster, Northeastern Oklahoma State University, Tahlequah
Vergil G. Stover and Jimmie D. Benson, Texas Transportation Institute, Texas A&M University

This paper discusses the effects of using small-sample origin-destination survey data as the basis for urban transportation planning models. The study was based primarily on home-interview survey data collected in the 1969 San Antonio-Bexar County Urban Transportation Study. The analysis demonstrates the ability of small sample origin-destination data to produce travel estimates that are in close agreement with model results obtained by using traditional large-sample survey data. The survey data for 12 477 dwelling units (i.e., a 5 percent sample) were used as a data base from which repeated geographically stratified random samples of 6400, 3200, 1600, 800, 400, and 200 observations were drawn. Two samples for each sample size, representing the 10th and 90th percentiles of distributions of sample estimates of total automobile-driver travel, were selected for evaluation. This procedure provided a 0.8 probability that samples of similar size will produce travel estimates as good as or better than those obtained here. The selected samples were used to develop inputs to trip-generation, trip-distribution (gravity), and traffic-assignment models. Samples of 400 or more dwelling units were found to produce acceptable results.

Traditional urban transportation studies invest a significant amount of time and effort in data collection. The home-interview survey generally accounts for a major portion of these costs, and reductions in the sampling rates used in these surveys will obviously result in significant cost savings. This paper reports an investigation to determine the reduction that could be made in the size of dwelling-unit samples without producing significantly different travel predictions after applying forecasting techniques. Data from the San Antonio-Bexar County, Texas, urban area transportation study were used as the data base.

PREVIOUS RESEARCH

There has been a limited amount of research directed toward the determination of minimum sample-size requirements, and much of this has been directed toward sample-size requirements for the calibration of trip-generation and distribution models. Little research has been devoted to determining the smallest acceptable sample for the entire transportation process; i.e., from trip generation through traffic assignment.

Sosslau and Brokke (1) used one-half, one-third, and one-tenth subsamples from the 1-in-15 dwelling-unit sample collected in the 1957 Phoenix origin-destination (O-D) survey to estimate the root mean square error (RMSE) as a function of sampling rate. The results were extrapolated for a range of sampling rates.

Heanue, Hamner, and Hall (2) tested cluster sampling as a way to reduce sample size in a 1960 study of the Pittsburgh urban area. They concluded that cluster sampling was unsatisfactory because of the bias introduced by the location of the clusters.

Parsonson and Cribbins (3) used systematic subsamples from the Raleigh, North Carolina, Urban Transportation Study to compare several approaches to trip generation. They observed differences in the responses of different trip-generation models to decreasing sample size and concluded that models based on small samples estimate the full origin-destination (O-D) data better than do models based on O-D data expanded from the small samples.

By using 100 percent survey data from three zones in San Antonio, Stover, Benson, and Ringer (4) found that large variances of estimates can be expected when traditional sampling rates are used to estimate both trips per dwelling unit and total trip ends for a zone. This research is significant because the sample estimates could be compared to known population values. They concluded that regression models or cross-classification rates provide better estimates of the total number of trip ends by zone than do expanded O-D survey data.

Further analysis of the San Antonio data by Benson, Pearson, and Stover (5) showed that, with the traditional sampling rates, a large majority of the interchange volumes of 1 to 10 trips were undetected while those detected were substantially overestimated. Sampling rates of more than 25 percent would generally be required to estimate nonzero interchange volumes of fewer than 50 trips within ±100 percent at 95 percent confidence.

In an investigation of the sensitivity of traffic assignment, Stover, Benson, and Buechler (6) showed the power of the assignment process to mask major inaccuracies in estimates of zonal trip ends and zonal interchange volumes. They concluded that reasonably accurate assignment results may be anticipated if the preceding modeling phases produce reliable estimates of the total trips and trip-length frequencies. Only reasonable (i.e., relatively coarse) estimates of the geographic distribution of trip ends are needed. They also found (7) that the trip-length frequency distribution is functionally related to the mean trip length and the maximum interzonal separation and developed a procedure for estimating the trip-length frequency distribution.

METHOD OF STUDY

This investigation used data for automobile-driver trips from the San Antonio-Bexar County Urban Transportation Study (SABCUTS). The study covered an area of 3230 km^2 (1247 $miles^2$), which was divided into 778 internal zones with 24 external stations. The population of the area at the time (1969) was 825 800 and comprised 256 640 households. Samples of 200, 401, 803, 1606, 3212, and 6425 observations (for convenience, these sample sizes are rounded to the nearest hundred when referred to in the text) were drawn from the 12 477 home interviews completed and collected in the nominal 5 percent home-interview survey. These samples represent sampling rates ranging from 0.08 to 2.56 percent.

The complete processing and evaluation of even a single sample is a costly and time-consuming process;

the evaluation of a large number of samples is prohibitively expensive. A study design was required that would make possible valid conclusions from the results of evaluating a limited number of samples for a given sample size. Previous research (6) has shown that the total number of trips and the trip-length frequency distribution are the dominant variables in terms of the results obtained from the traffic-assignment procedure. The product of the two variables—total number of trips times mean trip length—is an estimate of total travel. This estimate of total travel (in vehicle-minutes) was selected as the indicator of the traffic-assignment results that a given sample should produce when used as an input to urban transportation planning models.

The procedures for sampling from the full data set were designed so that the various samples selected would reflect, to the extent possible, the observations that would have resulted if the various small samples had been selected by using traditional sampling procedures. Therefore, the number of observations to be made at each sampling rate was calculated by multiplying the desired nominal sampling rate (e.g., $\frac{1}{8} \times 5$ percent = 0.625 percent) times the number of households (256 640) in the study area. The specific households in each sample were then selected in a manner similar to the traditional sampling procedures used in home-interview surveys. This ensured that the geographical distribution of the observations selected would be proportional to the distribution of households in the study area.

One thousand samples were drawn at each sample size to assess the sampling distributions of mean trip length, total trips, and total travel for automobile-driver trips. A sample selected near the mean of the distribution of the total-travel estimates would be expected to produce good assignment results because it would provide a good estimate of total travel. The selection of a sample at random, however, would not allow analysis of the results that might be expected from some other sample of the same size. Therefore, at each sample size, the samples representing the 10th and 90th percentiles of the expanded estimates of total automobile-driver travel were selected for use in the modeling procedures. These two samples can provide a basis for estimating the probability that other samples of the same size will perform as well as or better than the samples evaluated. If both samples produce acceptable travel estimates, other randomly selected samples of the same size will have an approximately 0.8 probability of producing travel estimates as good as or better than the samples evaluated. If either sample for a given sample size fails to produce acceptable travel estimates, the sampling level should be considered unsatisfactory.

The full set of survey data and the sample sets of data were processed independently to develop inputs for trip-generation, trip-distribution, and traffic-assignment models. The trip-generation analysis used disaggregated trip-generation rates. Rates were developed for three trip purposes (home-based work, home-based nonwork, and non-home-based) by using income and automobile ownership as independent variables. A minimum of 25 observations in any cell were used as the criterion for combining cells before the calculation of trip-generation rates.

Because the estimates of truck and taxi travel, external-local travel, and external-through travel are based on surveys other than the home-interview survey, the trip generation and trip distribution for these trip purposes were performed once, and the resulting trip tables were merged with the gravity-model trip tables developed for each set of home-interview data. Gravity-model trip distributions and all-or-nothing assignments were performed by using the Federal Highway Administration battery of programs.

The relative zonal attractions used were those developed in the urban transportation study. The implicit assumption is that small-sample home-interview data are not used for estimating zonal attractions. Instead, attraction rates are developed by using special surveys or secondary data sources.

Data from the SABCUTS study provided a dwelling-unit count for each analysis zone. However, because very small samples are not an adequate basis for estimating zonal distributions of dwelling units by income or by automobile-ownership level, tract data from the 1970 census of population and housing were used to estimate these. These population estimates were used when applying the trip-generation rates to the calculation of the total zonal trips by purpose for the full data set and all sample sets.

When samples of a given size give satisfactory results, it is reasonable to assume that larger samples will also give satisfactory results, and unsatisfactory results indicate that smaller samples will also be unsatisfactory. Samples of 1600 observations (approximately one-eighth of the full data set) were evaluated first to determine whether further processing should use larger or smaller samples. Because the results using samples consisting of 1600 observations were satisfactory, further processing of larger samples was discontinued.

ANALYSIS

Analyses using 100 percent survey data (4) have shown that much of the difference between observed and estimated trip ends is due to sampling error in the number of observed trips. This indicates that expanded O-D data will be of limited value in evaluating results obtained from small samples. Therefore, the results obtained from models generated or calibrated by using the full set of survey data for trip generation, distribution, and assignment were used as the standard of comparison for evaluating sample results. Each sample was evaluated as to its representation of the full data set and its performance in trip-generation, trip-distribution, and traffic-assignment models. Traffic-assignment results were also compared with actual traffic counts.

Samples

As expected, the variation in the distribution of estimates of total travel, total trips, and mean trip length increased as the sample size decreased. The 80 percent probability limits for sample estimates of total automobile-driver travel, total trips, and mean trip length are given below.

	Percentage of Full-Data Values		
Number of Observations	Total Travel	Total Trips	Mean Trip Length
6400	±1	±1	±1
3200	±2	±2	±1
1600	±3	±3	±2
800	±4	±4	±2
400	±6	±6	±5
200	±8	±9	±8

Thus, for each sample size, there is a 0.8 probability that a randomly selected sample will estimate the parameter within the indicated range of the population value as estimated from a sample of 12 477 observations. (ranges were rounded to the nearest integer value).

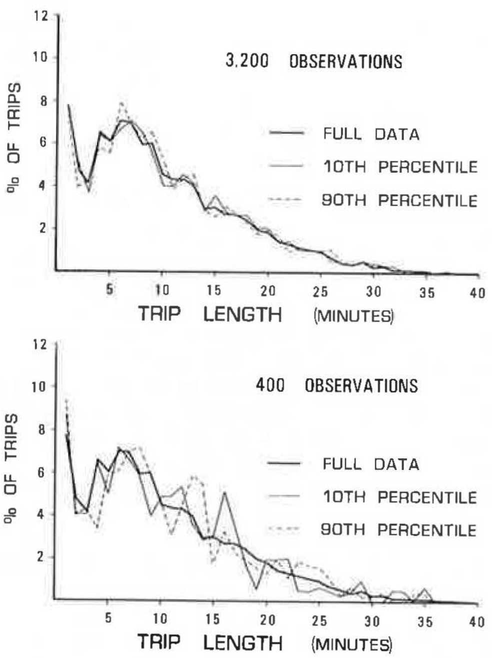
For example, a sample consisting of 800 dwelling-unit observations will estimate the total automobile-

Table 1. Sampling means and standard deviations by sample size.

Sample Size	Total Travel (vehicle·min) Amount	Total Travel (vehicle·min) Percent[a]	Standard Deviation (% of mean)	Total Trips Amount	Total Trips Percent[a]	Standard Deviation (% of mean)	Mean Trip Length (vehicle·min) Amount	Mean Trip Length (vehicle·min) Percent[a]	Standard Deviation (% of mean)
12 477	13 922 025		—	1 265 923		—	11.00		—
6 400	13 775 250	98.95	1.21	1 251 225	98.84	1.20	11.01	100.09	1.10
3 200	13 780 320	98.98	1.69	1 250 041	98.75	1.65	11.02	100.18	1.28
1 600	13 759 140	98.83	2.31	1 243 970	98.27	2.29	11.06	100.55	1.53
800	13 822 820	98.29	3.31	1 240 522	97.99	3.36	11.14	101.27	2.01
400	13 813 390	99.22	4.65	1 239 805	97.94	4.67	11.14	101.27	2.57

[a]Based on amount for full-data set.

Figure 1. Comparison of origin-destination trip-length frequency distributions.



driver trips within 4 percent of the population value at the 80 percent probability level. These data indicate that the dispersion of extreme values does not become significant until the sample consists of 400 or fewer total dwelling-unit observations. As shown by the data in Table 1, the means of the sampling distributions for each of the three estimates (total travel, total trips, and mean trip length) for all of the sample sizes investigated were within 2 percent of the full data values. For samples consisting of 6400 observations, the standard deviations for the estimates of total travel, total trips, and mean trip length were about 1.2 percent of the mean values. Beyond 800 observations, further decreases in sample size resulted in significantly greater increases in the standard deviation expressed as a percentage of the mean. The results indicate that a relatively small sample has a high probability of yielding good estimates of study-area travel parameters.

The trip-length frequency distributions for non-home-based trips, shown as examples in Figure 1, compare two samples of 3200 observations and two samples of 400 observations with the distribution from the 12 477 observations in the full data set. As expected, the distributions become less smooth as the number of observations is reduced; however, they are considered satisfactory approximations. Fitting a smoothed curve to the data points for the full data set or to data points for the smaller samples produces essentially identical curves. Similar results are observed for the other trip purposes.

The percentage distributions of dwelling units by income and by automobile-ownership levels for each sample were compared with those for the full data set. None exactly matched the full data-set distributions, but all of the samples of 400 or more observations produced acceptable comparisons. The samples with only 200 observations exhibited distributions that were judged to be significantly different from the full data-set distributions. This agrees with the previous research (5) that found that at least 400 observations were necessary to estimate mean trip length for automobile-driver trips with acceptable accuracy.

The income-level and automobile-ownership distributions estimated by the samples of 400 observations or more agreed closely with the distributions from the full data set. Consequently, in view of the previous research (4), it was concluded that small samples selected in the traditional manner can provide reliable estimates of the socioeconomic variables used in transportation studies.

Trip Generation

The dwelling-unit records from the full and sample data sets were cross-classified by using four automobile-ownership and five income levels. After all of the dwelling-unit records were assigned to the appropriate cells, the cells were grouped as necessary so that no cell would contain fewer than 25 observations by combining cells that, a priori, might be expected to exhibit similar trip-generator characteristics. The basic criteria for cell combination were to

1. Make the fewest possible combinations;
2. Consider that, at zero automobile ownership, income is the less important variable and make combinations across income levels; and
3. Create combined cells that were rectangular rather than L-shaped.

Each of the three cross-classification matrices of trip productions (one for each trip purpose) for each sample was compared to the four by five matrices generated from the full data set by using the calculated chi square (χ^2) and RMSE values summarized in Table 2. These calculated values tended to increase with decreasing sample size. However, the pattern has several exceptions. For example, the RMSE values of the 90th percentile sample for 1600 observations are less than those for the 90th percentile sample for 3200 observations.

Because a very small difference may make a large contribution to the calculated χ^2 value, this statistic is of questionable practical significance. Moreover, the total trip matrices differ to a much lesser degree because errors by individual trip purpose tend to cancel out. As shown in Table 3, the range in total number of trips between the 10th and 90th percentiles increases significantly with fewer than 400 observations. With two exceptions (the samples consisting of 200 and 400

Table 2. Chi square and root-mean-square error comparison of trip generations for three trip purposes.

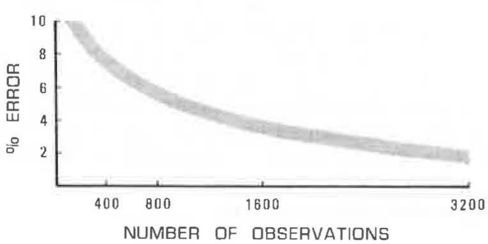
Sample		χ^2			RMSE[a]		
Size	Percentile	Home-Based Work	Home-Based Nonwork	Non-Home-Based	Home-Based Work	Home-Based Nonwork	Non-Home-Based
3200	10th	83.6	155.3	81.3	347	549	233
	90th	97.7	70.1	234.7	277	2333	477
1600	10th	102.9	353.3	447.7	334	990	670
	90th	36.1	215.6	274.4	241	792	481
800	10th	251.4	581.7	768.1	602	1181	1019
	90th	352.0	649.8	1604.9	637	1462	1758
400	10th	1466.7	513.1	315.4	1961	1268	371
	90th	351.2	644.5	1592.7	860	1365	2348

[a] Trips in each cell for full data were used as expected; trip generation rates estimated from each sample data set times total dwelling units in corresponding cells for full data were used to calculate observed trips.

Table 3. Trip productions generated for study area by different size samples.

Sample		Productions			
Size	Percentile	Home-Based Work	Home-Based Nonwork	Non-Home Based	Total
12 477	—	413 705	664 473	343 165	1 421 343
3 200	10th	421 030	658 222	359 674	1 439 926
	90th	408 557	682 849	363 468	1 454 874
1 600	10th	423 212	686 262	318 740	1 428 214
	90th	409 722	679 531	365 792	1 455 045
800	10th	413 490	712 069	356 537	1 482 096
	90th	418 386	682 837	395 211	1 496 434
400	10th	382 460	673 825	325 383	1 381 668
	90th	395 493	692 481	399 240	1 487 214
200	10th	461 431	605 431	324 277	1 391 178
	90th	441 511	757 019	362 445	1 560 975

Figure 2. Expected error in number of total trips (0.8 probability level).



observations at the 10th percentile), the total trips generated by the samples slightly overestimated total driver trips as compared to total trips generated by the full data.

The frequency distribution of trip productions by zone for each sample was compared to those for the full data set. Samples of 400, 800, 1600, and 3200 observations had distributions essentially identical to that for the full data set. Distributions produced by samples of 200 observations were significantly different from that for the full data set.

Trip productions by trip purpose and by income levels were compared using 14 geographic sectors. The comparisons by geographic sectors indicated no significant differences between the full data set and the samples of 400 or more observations in either patterns of trip making or geographic patterns of travel.

Based on the combined analysis of the several sample data sets and previous research (4), the maximum expected error for total trips at the 0.8 probability level was determined as shown in Figure 2. The average error is estimated to be slightly less than 7 percent for a sample of 400 dwelling units and decreases to about 3 percent for a sample of 1600. Increasing the sample size from 1600 to 3200 observations decreases the maximum error to approximately 1.0 percent, which indicates that larger sample sizes contribute only marginally to the accuracy of travel estimates. Therefore, samples of 400 or more observations are adequate to produce acceptable trip-generation results.

The presample determination of cross-classification cells was compared to the use of a more detailed matrix and postsample combining of cells to provide a minimum of 25 observations per cell. The two samples of 400 observations, the smallest sample size that produced acceptable trip-generation results, were used in the analysis. The results achieved by the procedure of reducing the number of cells after the sample had been selected and the observations assigned to the appropriate cross-classification cell (four automobile-ownership and five income level) were compared to the following cross-classifications, which were established prior to the sample selection: (a) two automobile-ownership levels (zero and one plus) and no breakdown by income level, (b) three automobile-ownership levels (zero, one, and two plus) and no breakdown by income level, and (c) two automobile-ownership levels (zero and one plus) and two income levels (low and medium to high). To achieve at least 25 observations per cell, the four by five matrix was reduced to the following five cells: (a) zero automobiles by all income levels, (b) one plus automobile by low income, (c) one plus automobile by medium to low income, (d) one plus automobile by medium income, and (e) one plus automobile by medium to high and high-income levels. All of the presampling classification schemes produced results that were inferior to those of the cross-classification involving postsampling reduction in the number of cells. Furthermore, the analysis indicated that better results were produced as the number of cross-classification cells increased. This suggests that combining cells to achieve a minimum number of observations per cell after the sample has been selected is the better procedure to follow when a small sample is selected by using traditional sampling procedures.

Trip Distribution

The trip-length frequency distributions by trip purpose obtained from the survey data were used as gravity-model calibration criteria. The calibrated mean trip lengths were all within 3 percent of the target O-D data values; most differed from the O-D mean trip length by less than 1.0 percent. The calibrated trip-length frequency distributions for the samples were in close agreement with the comparable distributions for the full set of survey data.

The trip table is a more deterministic measure of the adequacy of models calibrated from small samples. Analysis of the zonal-interchange-volume distributions indicated no significant differences between the results

Figure 3. Limits of expected error for traffic-assignment parameters (0.8) probability level) (a) total VKT, (b) screen line volumes, (c) cutline volumes, and (d) link volumes.

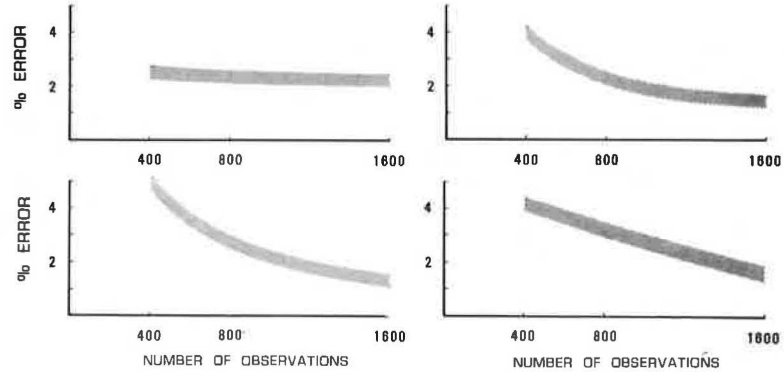


Table 4. Cutline volumes.

Sample		RMSE (comparison to counted volumes)	Comparison to Full-Data Set	
Size	Percentile		RMSE	Percent RMSE
12 477	–	11 084	–	–
1 600	10th	11 489	1653	3.6
	90th	10 806	734	1.6
800	10th	10 668	1456	3.1
	90th	10 643	1638	3.5
400	10th	11 138	3999	8.6
	90th	11 945	2314	5.0

Table 5. Link-volume differences.

Sample		Mean Volume Difference	Standard Deviation	RMSE
Size	Percentile			
1600	10th	-28	185	187
	90th	39	88	97
800	10th	68	116	135
	90th	85	131	156
400	10th	-78	214	407
	90th	102	407	420

using the small samples and the full data. In each comparison, more than 50 percent of the cells in the total trip table were estimated as having exactly the same interchange volume, and approximately 95 percent of the cells were within ±2 trips of the full data results. Although there was some tendency for the variations in the trip tables to increase as the sample sizes decreased, the differences in the trip tables were not significantly affected by the sample sizes used in the calibration. (The trip table that most closely resembled that using the full data set was produced by the 10th percentile sample of 400 observations, while the 90th percentile sample of 400 observations provided the poorest comparison.)

Traffic Assignment

Traffic assignments that used the modeled trip tables based on samples of 400, 800, and 1600 observations were compared to the traffic assignments based on the full set of survey data and to the counted traffic volumes. The expected errors (at the 0.8 probability level) for the total vehicle kilometers of travel (VKT), screen-line volumes, and cutline volumes are shown in Figure 3.

Estimates of the total VKT obtained with the models calibrated by using the small samples agreed with that using the full data set within 2.7 percent. When tabulated by the 14 geographic sectors and compared to the full data-set value, the RMSEs for the samples were less than 1.0 percent. Thus, a sample consisting of 400 observations is adequate to produce acceptably accurate estimates of the total VKT and the geographical distribution of the VKT.

A rail right-of-way that essentially bisects the study area was used as a major screen line. The screen-line volumes given in the table below are in close agreement with those for the full data.

Sample		Volume	
Size	Percentile	Value	Percent of Full Data
12 477	–	444 339	100.0
1 600	10th	436 368	98.2
	90th	448 684	101.2
800	10th	452 566	101.9
	90th	454 549	102.3
400	10th	428 647	96.5
	90th	462 154	104.3

Although the expected error begins to increase for samples of less than 800 observations, the samples of 400 observations produced estimates of screen-line volumes that are within 4.5 percent, or less, of the estimated volume based on the full data set. Therefore, sample sizes of 400 observations or more produce acceptably accurate estimates of screen-line volumes.

Twenty-eight cutlines were used to compare the assigned volumes in various travel corridors. The mean differences from the full data-set assignment were 1.2, 2.6, and 4.1 percent for 1600, 800, and 400 observations respectively. Although the maximum expected error tends to increase with decreasing sample size (Figure 3), the magnitude of the error in cutline volumes with samples of 400 observations is considered to be within acceptable limits. As indicated in Table 4, when the assigned volumes are compared to the counted volumes, the samples of 400 or more observations produced RMSEs that are not appreciably larger than that resulting with the full data set. All of these assigned cutline volumes were within 10 percent of the counted volume, and three-quarters were within 5 percent.

Comparisons of the individual cutline volumes indicate that the small samples produce results that are frequently as good as or better than that produced by the full data set. For example, when the 90th percentile sample of 800 observations is used, 17 of the 28 cutlines have assigned volumes that are closer to the counted volume than that produced by full data set. These analyses suggest that factors other than the number of dwelling-unit observations have equal, or greater, impacts on the cutline results.

The network links were classified for comparison into 15 volume groups on the basis of the full-data-set traffic assignment. Over 95 percent of the link volumes for the sample assignments were within the predetermined acceptable error ranges of the volume for the full-data-set assignment. The remainder were only slightly outside the acceptable range of error. The mean percent volume differences (based on the mean volume difference as a percentage of the midpoint of each volume range) were 1.0, 2.2, and 2.9 percent with samples of 1600, 800, and 400 observations respectively.

As indicated by Table 5, the calibration of models with vary small sample sizes does not contribute to a serious deterioration in overall results, although the mean differences and variances in assigned link volumes increase with decreasing sample size. A comparison of these parameters by volume group found, for example, that the 10th percentile samples of the 1600 and 400 observations produced nearly identical results. The best overall results by volume group were those of the sample of 800 observations. This, together with the observed pattern of trip generations by income and automobile-ownership levels, suggests that random variations are more significant in affecting the assignment results than is sample size. Therefore, although the maximum expected error (Figure 3) for assigned link volumes tends to increase with smaller sample sizes, assignments developed from samples of 400 observations are within the acceptable limits of error.

Examination and comparison of the posted traffic assignments did not identify any significant differences between the assignment based on the full data set and the assignments based on the samples. Similarly, comparisons based on 12 selected major routes did not identify significant differences in assigned volumes. The full data set resulted in overassignment on 7 routes and underassignment on 5 routes when compared to the counted volumes. For samples of 800 and 1600 observations, the assignments on each of the routes were within 2.5 percent of those produced by the full data set. The samples of 400 observations produced assignments that were within 6.0 percent of those from the full data set.

EVALUATION AND IMPLICATIONS

The results of this study indicate that urban transportation models calibrated for three trip purposes (home-based work, home-based nonwork, and non-home-based) will produce acceptably accurate travel estimates from data based on as few as 400 dwelling-unit interviews. The analyses indicate that there is only a modest decrease in the precision of the estimates as the sample size is reduced to 400 observations, but that thereafter the reliability of the estimates deteriorates rapidly. This indicates that the collection of the larger samples that results from traditional sampling rates is not cost-effective in the traffic-assignment results.

This and other research have established that it is the number of observations in the sample, rather than the percentage of dwelling units surveyed, that determines whether urban planning models produce acceptably accurate traffic assignments. Thus, the results of this study can be applied to study areas having both smaller and larger populations. The minimum sampling rate will vary inversely with the study-area population, but the minimum number of observations required is a constant.

Previous research has shown that the development of acceptably accurate traffic assignments requires relatively precise estimates only for total trips for the study area and for the mean trip length. Estimates of trip-length frequencies and the geographic (zonal) distribution of trip ends need only be reasonable approximations (and errors in the geographic distribution of trip ends will be offset).

Although samples consisting of 400 dwelling-unit observations are an adequate basis for the calibration of transportation models to produce acceptably accurate traffic assignments, this does not imply that 400 observations is an adequate sample size for defining other, more detailed travel characteristics. A larger number of observations or specially designed studies are necessary to define travel characteristics such as the number of trip attractions or the relative attractions for individual zones; travel patterns for trip interchanges between zones; the temporal stability of trip-generation rates; trip priorities, for example, which trips would not be made under adverse conditions such as fuel shortages; modal choice among alternative means of travel; measures of trip-generation characteristics of specific population segments, such as by dwelling-unit type or persons per dwelling unit; and the number of trip ends in a specific geographic area on the basis of O-D data expansion.

The modification of urban transportation study procedures to efficiently use small samples will permit significant cost savings in data collection and reduction without causing measurable effects on the assignment results. The effective use of small-sample survey data suggests the following modifications in transportation study procedures:

1. Disaggregate trip-generation techniques must be used—cross-classification with the number of cells to be used in the trip-generation analysis should be determined after the collection of the O-D data and
2. Stratified cluster sampling might be used to further simplify and reduce the cost of the dwelling-unit inventory and sample selection tasks.

Traditional record-keeping systems and procedures for developing the dwelling-unit inventory might be used for the geographic subdivisions not selected for data collection, for which only the number of dwelling units in each area would be required.

The urban transportation study has been a valuable tool for the transportation analyst in the development of transportation demand forecasts, the evaluation of land-use and transportation system alternatives, and the identification of major inconsistencies between proposed activity patterns and the transportation network. The use of a much smaller sample of home-interview data will permit the transportation study to continue to be a valuable, but more cost-effective tool.

ACKNOWLEDGMENT

The research leading to the development of this paper was sponsored by the Federal Highway Administration, U.S. Department of Transportation. The paper is based on a research report entitled Consequences of Small Sample O-D Data Collection in the Transportation Planning Process.

The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

REFERENCES

1. A. B. Sosslau and G. E. Brokke. Appraisal of Sample Size Base on Phoenix O-D Survey Data. HRB, Bulletin 253, 1960, p. 114-127.
2. K. E. Heanue, L. B. Hamner, and R. M. Hall. Adequacy of Clustered Home-Interview Sampling for Calibrating a Gravity-Model Trip-Distribution Formula. HRB, Highway Research Record 88, 1965, pp. 116-136.
3. P. S. Parsonson and P. D. Cribbins. Estimating Urban Trip Production and Attraction. Journal of the Highway Division, Proc., ASCE, Nov. 1967, pp. 115-134.
4. V. G. Stover, J. D. Benson, and L. J. Ringer. Accuracy of Travel-Pattern Estimated From the Home-Interview Survey. Texas Transportation Institute, Texas A&M Univ., College Station, Res. Rept. 167-7, Aug. 1973.
5. J. D. Benson, D. F. Pearson, and V. G. Stover. Accuracy of Travel-Pattern Estimates From the Home-Interview Survey. Texas Transportation Institute, Texas A&M Univ., College Station, Res. Rept. 167-8, March 1974.
6. V. G. Stover, J. D. Benson, and J. Buechler. Sensitivity Evaluation of the Traffic Assignment Process. TRB, Transportation Research Record 610, 1976, pp. 50-52.
7. D. F. Pearson, V. G. Stover, and J. D. Benson. A Procedure for the Estimation of Trip-Length Frequency Distributions. Texas Transportation Institute, Texas A&M Univ., College Station, Res. Rept. 17-1, May 1974.

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

An Application of Diagnostic Tests for the Independence From Irrelevant Alternatives Property of the Multinomial Logit Model

Daniel McFadden and Kenneth Train, University of California, Berkeley
William B. Tye, Charles River Associates, Inc., Cambridge, Massachusetts

Statistical tests are proposed to diagnose the validity of the independence from (of) irrelevant alternatives property of the multinomial logit model. Application of the tests is illustrated by the use of actual travel data representing urban modal choice in the San Francisco area. The property as it applies to travel demand forecasting is discussed, and the common misconception that the property holds for market shares in heterogeneous populations is shown by examples to be incorrect. The relation of the property to the basic assumptions of the model is described, and it is shown that the validity of the property in disaggregate modeling is an empirical issue that depends on the model specification and data in a particular application. A series of diagnostic tests for the property are developed and applied to actual travel data.

The most widely used functional form for choice probabilities in disaggregated transportation-demand analysis is the multinomial logit (MNL) model,

$$P(i|C) = \exp V(x^i, s) \Big/ \sum_{j \in C} \exp V(x^j, s) \tag{1}$$

where

C = finite choice set,
P(i|C) = choice probability for alternative $i \in C$,
x^i = vector of the observed characteristics of alternative i, and
s = vector of the observed characteristics of the decision maker and the choice environment.

The scale function $V(x^i, s)$ may be interpreted as the representative utility of alternative i and is normally assumed to be linear in the parameters. The MNL model has significant advantages over the available alternatives in terms of flexibility and computational efficiency and permits a simple behavioral interpretation of the parameters of the scale function.

The MNL model also has the property that the ratio of the probabilities of choosing any two alternatives

$$P(i|C)/P(k|C) = \exp V(x^i, s)/\exp V(x^k, s) \tag{2}$$

is independent of the attributes or the availability of a third alternative (j), which is termed the independence from (of) irrelevant alternatives (IIA) property. This property greatly reduces the complexity of estimation and forecasting and in this respect is quite useful. However, it imposes restrictions on the structure of choice probabilities and cross elasticities; these restrictions may be invalid in some applications. Hence, tests of the validity of the IIA property should be made whenever a violation of the assumption is suspected.

This paper analyzes the IIA property and discusses several diagnostic tests. Complete descriptions of the tests and thorough instructions for construction of the test statistics can be found in a National Cooperative Highway Research Program (NCHRP) report (1) and McFadden, Tye, and Train (2).

INDEPENDENCE FROM IRRELEVANT ALTERNATIVES PROPERTY OF THE MULTINOMIAL LOGIT MODEL

In applications of the MNL model to individual modal choice, the IIA property (Equation 2) requires that if two modes are available and a new mode is introduced, the ratio of the probabilities of the two preexisting modes will be unchanged regardless of the probability of choice for the new mode. For example, if the new mode will be chosen with a probability of 0.10 and each preexisting mode had a 0.50 probability before the introduction of the new mode, the probability of each of the preexisting modes will be 0.45 after the new mode is introduced, thus preserving the one-to-one ratio of probabilities of the preexisting modes.

The IIA property also greatly facilitates the forecasting problems associated with new modal-choice predictions. If 100 persons have the same observed characteristics of alternatives, the same observed characteristics of the decision maker, and the same choice set, i.e., they have the same $V(x^i, s)$'s, the demand for a new mode can be calculated by adding another term to the denominator of Equation 1 and recomputing all choice probabilities. The new probabilities can then be multiplied by 100 to estimate the demand for each mode. If the old modes formerly shared the market equally and the probability of the new mode is 0.10 for each individual, the predicted modal demands will be 45, 45, and 10.

An example of a choice setting in which the IIA property is inappropriate is the classic blue automobile versus red automobile case. Assume that the bus mode and the blue automobile each capture 50 percent of a given travel market as shown in the first column of the table below.

	Modal Choice (%)		
Mode	True and MNL (binary choice)	Predicted MNL (3 modes)	True (3 modes)
Bus	50	33.3	50
Blue automobile	50	33.3	25
Red automobile	0	33.3	25
Total	100	100	100

Assume then that a new automobile mode is introduced with exactly the same service attributes as the blue automobile mode except that the automobile is painted a different color, e.g., red (patrons are assumed to be indifferent to color). Assume also that the red automobile is leased for this trip only, to remove questions of automobile ownership and competing demands for the automobile. The true modal shares will now be 50, 25, and 25 percent, for bus, blue automobile, and red automobile respectively; i.e., no bus users will switch to the new mode and automobile users will split evenly between the two automobile modes. However, the MNL model will forecast that each of the three modes captures one-third of the market, as shown by the second column of the table above, because the IIA property requires that the ratio of the bus share to the blue automobile share be unaffected by the introduction of the red automobile. In this example, the ratio is 1.0: When the red automobile is introduced, the ratio of the blue automobile share to the red automobile share is 1.0 (because patrons are assumed indifferent to color). The only shares that allow both the ratio of bus share to blue automobile share and the ratio of blue automobile share to red automobile share to equal one are one-third shares for each mode. Thus, the MNL model predicts shares of 33, 33, and 33 percent when the actual shares are 50, 25, and 25 percent for bus, blue automobile, and red automobile, respectively.

If the problem were confined to this simple example, it would be trivial. The new automobile mode is clearly irrelevant and should not be introduced as a mode. However, this extreme case points to a gray area, where the demand forecast for a new mode could be seriously compromised by incorrectly applying the IIA property.

In the MNL model, the IIA is a property of individual probabilities and market shares in homogeneous populations, but not a property of market shares in heterogeneous populations. Much unwarranted criticism of the MNL model has been based on the erroneous application of the IIA property to market shares in heterogeneous populations. It should be emphasized that the MNL model does not predict that the ratio of market shares in a heterogeneous population will be invariant with the introduction of a new alternative.

To take a specific example, the MNL model does not in general predict that, if a new mode is introduced to a population composed of different market segments that have different observed socioeconomic characteristics and level-of-service attributes [different $V(x^i, s)$'s for the individuals], the percentage of automobile drivers who will use the new mode is equal to the percentage of transit users who will shift.

This principle may be illustrated by an example. Table 1 presents the case of a population composed of two market segments of 100 persons each. Each segment is composed of homogeneous individuals; i.e., each person in the segment assigns the same representative utility to each alternative. Assume that the choice environment of observed attributes is identical for all persons within each segment and that it differs significantly between the two market segments. Segment 1 is automobile oriented, splitting 90 to 10 in favor of the automobile, and segment 2 is transit oriented, splitting 90 to 10 in favor of transit.

A new mode—dial-a-bus—is introduced. The MNL model predicts that it will capture 5 percent of segment 1 and 15 percent of segment 2. As indicated in Table 1, the ratio of the automobile market share to the bus market share is preserved within each homogeneous market segment. However, the ratio of the bus modal share to the automobile modal share is not constant after the new bus mode is introduced, but decreases from 1.0 to 0.91 for the entire population (86 ÷ 94 = 0.91).

Although the percentage diversions from the bus and the automobile to the dial-a-bus are the same within each homogeneous market segment (e.g., in segment 1,

Table 1. Effect of IIA property on a forecast of behavior in a population of heterogeneous market segments.

	Modal Share					
	MNL (binary choice)			Predicted MNL (3 modes)		
Mode	Market Segment 1	Market Segment 2	Total Market	Market Segment 1	Market Segment 2	Total Market
Bus	10	90	100	9.5	76.5	86.0
Automobile driver	90	10	100	85.5	8.5	94.0
Dial-a-bus	0	0	0	5.0	15.0	20.0

5 percent of both bus and automobile patrons switch), the predicted diversions from the automobile and the bus are not the same for the population as a whole. Of the 100 total bus patrons in the binary-choice situation, 14 percent (100 - 86) were predicted to switch to the dial-a-bus, but only 6 percent of the total automobile users were predicted to switch to the dial-a-bus.

The IIA property is obviously a key assumption of the MNL model. Previous studies of it have tended to discuss its reasonableness or unreasonableness on logical grounds. This paper argues that the issues raised by the property are essentially empirical. The convenience of the IIA property in estimating and forecasting makes it extremely attractive to use when it is valid. But the undesirable consequences of assuming the IIA property when it is invalid are reason for caution in applying the MNL model without assurances of the reasonableness of the IIA property.

This dilemma is addressed here by the development of statistical tests that can identify whether or not the IIA property is reasonable in the particular circumstances. These tests are comparable to the standard statistics that are routinely calculated as part of regression programs to identify whether or not the assumptions of the least-squares model are reasonable.

SOURCES OF VIOLATION OF THE INDEPENDENCE OF IRRELEVANT ALTERNATIVES PROPERTY

In developing statistical tests to determine the validity of the IIA property it is useful to consider the basic assumptions of the MNL model and the ways in which they might be violated. The utility for the individual of the ith alternative is assumed to be a function of the observed characteristics of that alternative, the observed characteristics of the decision maker and choice environment, and an unobserved component that represents the effects of omitted random taste variations, choice attributes, and socioeconomic variables.

$$U_i = U_i(x^i, s, \mu_i) \tag{3}$$

where

- U_i = utility of the ith alternative,
- x^i = vector of observed characteristics of the ith alternative, $(x_{i_1}, \ldots, x_{i_N})$,
- s = vector of observed characteristics of the decision maker, and
- μ_i = vector of unobserved characteristics of the decision maker and the ith alternative.

Without loss of generality, U_i can be separated into two parts: $V(x^i, s)$, a function of the observed data, and ϵ_i, a random component that is not observed; i.e.,

$$U_i = V(x^i, s) + \epsilon_i \tag{4}$$

The nonstochastic term is called the representative utility and is specified to be linear in the parameters:

$$V(x^i, s) = \beta Z(x^i, s) \tag{5}$$

where Z = vector-valued function of x^i and s and β = vector of the parameters.

Assume that alternative i is chosen if, and only if, it has greater utility than any other alternative; i.e., if $U_i > U_k$ for all $k \neq i$. Because the ϵ_i are random variables, the event $U_i > U_k$ for all $k \neq i$ is also random. The probability that the ith alternative is chosen is given by

$$P(i|C) = P(U_i > U_k) \quad (k \neq i) \tag{6}$$

and, from Equation 4,

$$P(i|C) = P[\epsilon_k - \epsilon_i < V(x^i, s) - V(x^k, s)] \quad (k \neq i) \tag{7}$$

To determine the probability that U_i satisfies Equation 7, we must know the probability distribution of ϵ_i. Assume that ϵ_i has a reciprocal exponential (Weibull) distribution, distributed identically and independently across all alternatives; i.e.,

$$P(\epsilon_i \leq t) = \exp(-e^{-t}) \tag{8}$$

Given this assumption, it is possible to derive the MNL model (Equation 1) and the IIA property (Equation 2) (3, 4, 5).

Any significant violation of the assumptions of the MNL model will usually cause the IIA property to fail to be valid. Generally, the violations may be traced to the MNL assumption that the unobserved-utility component is independent across alternatives and independent of the observed attributes (1, 2).

Because the unobserved terms are defined simply as the difference between the true utility and the representative utility, the independence or nonindependence of the ϵ_i's depends on the specifications of the representative utility. In a given choice situation, two different specifications of representative utility will result in two different sets of ϵ_i's. One set of ϵ_i's might be independent, while the other might not. Thus, the IIA property might be valid for one specification of representative utility and not for another, even though both specifications relate to the same choice situation. This means that the IIA property is or is not valid for a particular specification of representative utility in a logit model of a particular choice situation, not for the choice situation itself. Consequently, it is meaningless to say, for example, that the IIA property is or is not valid for a traveler's choice of mode. It is only possible to state that the IIA property is or is not valid for a particular specification of the representative utility of the various modes.

Intuitively, the IIA property plays a role in the MNL model that is analogous to the assumption of independent-error terms in least-squares regression. The IIA property implies that the factors omitted from the analysis (the ϵ_i's) are independent random variables.

APPLICATION OF DIAGNOSTIC TESTS FOR THE INDEPENDENCE OF IRRELEVANT ALTERNATIVES PROPERTY

Suppose a set of qualitative choice data is hypothesized to satisfy a particular specification of an MNL form. If the hypothesis is valid, the data and fitted models should have internal consistency properties; these can form the basis for diagnostic tests of the IIA property.

Model Specification

Table 2 presents an MNL model of the choice of mode for the work trip. The estimation was performed by the maximum likelihood method described by McFadden (5) on a sample of 641 workers in the San Francisco-Oakland Bay Area.

The model considers seven alternative modes: two automobile modes (automobile alone and car pool), two bus modes (one with walk access to bus and the other with automobile access), and three Bay Area Rapid Transit (BART) modes (with walk, bus, and automobile access). Most of the independent variables are self-explanatory, and their coefficients are readily interpreted. For example, the negative coefficient of on-vehicle time indicates that, when time spent in the vehicle for a particular mode increases, the probability of that mode being chosen decreases, all else being constant. Since the ratio of the on-vehicle time coefficient to the cost divided by wage coefficient is 0.43, the estimated value of on-vehicle time is 43 percent of the wage. Three income variables are included to allow for a nonlinear relation between income and the representative utility of the automobile-alone alternative. [These variables can be understood most readily by reference to Train (8).]

The model given in Table 2 seems particularly appropriate for testing violations of the IIA property. Because some of the alternative modes are similar, unobserved attributes of each mode may be correlated across modes. For example, the comfort on on-vehicle travel is similar for bus with walk access and bus with automobile access, and yet no comfort variable is included in the model. Failure of the IIA property could also result from the attributes of the alternatives not being exogenous. If the choice of how many automobiles to own is related to the work-trip modal choice, then the automobiles per driver variables are endogenous, and if the choice of where to live is related to the work-trip modal choice, then the cost and time variables are also endogenous.

Table 2. Work-trip modal-choice model.

Independent Variable	Estimated Coefficient	t-Statistic
Cost divided by posttax wage, ¢ ÷ ¢/min	-0.0380	6.83
On-vehicle time, min	-0.0162	1.91
Walk time,[a] min	-0.1006	4.25
Transfer-wait time,[a] min	-0.0122	0.923
Headway of first bus,[a] min	-0.0341	3.51
Automobiles per driver (ceiling of one)[b]	2.38	6.16
Automobiles per driver (ceiling of one)[c]	1.48	1.92
Dummy if person is head of household[b]	0.494	2.62
Number of persons in household who can drive[b]	0.5242	4.18
Number of persons in household who can drive[c]	0.7567	3.82
Family income (ceiling of $7500),[b] $/year	-0.000 308	2.18
Family income minus $7500 (floor of $0 and ceiling of $3000),[b] $/year	0.000 139	1.05
Family income minus $10 500 (floor of $0 and ceiling of $5000),[b] $/year	-0.000 096 6	1.78
Automobile-alone dummy[d]	-1.84	1.74
Bus-with-automobile-access dummy[e]	-5.38	5.69
BART-with-walk-access dummy[f]	1.94	3.18
BART-with-bus-access dummy[g]	-0.159	0.285
BART-with-automobile-access dummy[h]	-4.06	4.38
Car-pool dummy[i]	-2.39	5.28

Notes: Likelihood ratio index = 0.4119; log likelihood at zero = -982.6; log likelihood at convergence = -577.9; degrees of freedom = 2460; percent correctly predicted = 64.27; values of time saved as a percentage of wage: on-vehicle time = 43, walk time = 265, and transfer wait time = 32 respectively.
All cost and time variables calculated for round trip. Dependent variable is alternative choice (1 for chosen alternative, 0 otherwise).
Sample size = 641.

[a] Variable is 0 for automobile-alone and car-pool alternatives and takes dummy value for other alternatives.
[b] Variable takes dummy value for automobile-alone alternative and is 0 otherwise.
[c] Variable takes dummy value for bus-with-automobile-access and BART-with-automobile-access alternatives and is 0 otherwise.
[d] Variable is 1 for automobile-alone alternative and 0 otherwise.
[e] Variable is 1 for bus-with-automobile-access alternative and 0 otherwise.
[f] Variable is 1 for BART-with-walk-access alternative and 0 otherwise.
[g] Variable is 1 for BART-with-bus-access alternative and 0 otherwise.
[h] Variable is 1 for BART-with-automobile-access alternative and 0 otherwise.
[i] Variable is 1 for car-pool alternative and 0 otherwise.

Universal Logit Method

The universal logit (UL) model is a more general model than is the MNL model; it takes advantage of the fact that every choice model with positive probabilities can be written in apparent MNL form, except that the scale function of alternative i will depend on attributes of other alternatives.

To use the universal logit method, a model is specified that includes all the variables in Table 2 plus some variables that are defined so that the attributes of one alternative are allowed to enter the representative utility of another. The hypothesis that the coefficients of all of the extra variables are zero is tested. If the hypothesis of zero coefficients is rejected, then the joint hypothesis of the MNL form and the specification in Table 2 is rejected.

The more general model includes the variables given in Table 2 and the following others:

1. Cost divided by posttax wage of automobile alone, with bus-with-walk-access and BART-with-walk-access alternatives having the values given in Table 2 and other alternatives having the value 0;
2. Cost divided by posttax wage of bus with walk access with automobile-alone and BART-with-walk-access alternatives having the values given in Table 2 and other alternatives having the value 0;
3. Cost divided by posttax wage of BART with walk access wage, with automobile-alone and bus-with-walk-access alternatives having the values given in Table 2 and other alternatives having the value 0;
4. Total weighted time (sum of on-vehicle time, 2.5 times walk time, 1.25 times transfer-wait time, and 1.25 times first headway) of automobile alone, with bus-with-walk-access and BART-with-walk-access alternatives having the values given in Table 2 and other alternatives having the value 0;
5. Total weighted time of bus with walk access, with automobile-alone and BART-with-walk-access alternatives having the values given in Table 2 and other alternatives having the value 0; and
6. Total weighted time of BART-with-walk-access, with automobile-alone and bus-with-walk-access alternatives having the values given in Table 2 and other alternatives having the value 0.

The null hypothesis that the interaction effects are zero (i.e., the MNL model and the IIA are true) can be tested by using the likelihood ratio; when the MNL model and the UL model are fitted by maximum likelihood estimation, the likelihood-ratio statistic

$$\chi^2 = 2(\text{log likelihood of UL model} - \text{log likelihood of MNL model}) \quad (9)$$

is asymptotically distributed chi square (χ^2) with degrees of freedom (df) equal to the number of parameter restrictions imposed by the null hypothesis. [This has been discussed in the NCHRP report (1) (Vol. 2, pp. C-172-175)].

The log likelihood at convergence for the more general model is -567.6. The log likelihood at convergence for the model in Table 2 is -577.9. Therefore, the test statistic (from Equation 9) is 20.6. The critical (0.05-level) value of χ^2 with six df is 12.6. The joint hypothesis that the MNL form and the specification of Table 2 are correct is rejected.

The signs of the coefficients of the extra variables are consistent with the hypothesis that the value of automobile on-vehicle time is higher than that of transit on-vehicle time. Variables 5 and 6 entered with negative signs (the latter with a t-statistic of 3.0), and the coefficient of variable 4 was estimated to be positive. Train (6) found the value of automobile on-vehicle time to

Table 3. Tests based on conditional choice.

Statistic	Alternatives Included in Subset: All Except BART Modes	All Except Bus Modes	All Except Bus and BART-With-Automobile-Access Modes	All Except BART-With-Walk-Access Mode	All Except Car-Pool Mode
Log likelihood at convergence for subsample choosing an alternative within subset of alternatives	-452.6	-400.1	-452.3	-557.8	-230.7
Log likelihood with coefficients restricted to values given in Table 2	-454.6	-403.3	-455.0	-557.9	-247.3
test	4.0	6.4	5.4	0.2	33.2
df	16	17	17	18	18
Critical (0.05 level) value of χ^2 with appropriate df	26.3	27.6	27.6	28.9	28.9

Table 4. Tests of association.

	Alternative											
	Automobile Alone			Car Pool			Bus With Walk Access			Bus With Automobile Access		
	No. of Residuals		Avg Probability	No. of Residuals		Avg Probability	No. of Residuals		Avg Probability	No. of Residuals		Avg Probability
Cell	Positive	Negative	for Cell	Positive	Negative	for Cell	Positive	Negative	for Cell	Positive	Negative	for Cell
1	17	5	0.93	6	16	0.060	14	3	0.66	1	16	0.090
2	21	1	0.89	9	13	0.055	10	7	0.52	0	17	0.053
3	17	5	0.87	11	11	0.050	7	10	0.43	2	15	0.045
4	20	2	0.85	8	14	0.046	7	10	0.37	2	15	0.038
5	16	6	0.82	6	16	0.042	4	13	0.31	2	15	0.034
6	17	5	0.81	6	16	0.038	7	10	0.26	0	17	0.030
7	20	2	0.80	4	18	0.034	1	16	0.22	0	17	0.028
8	19	3	0.78	4	18	0.030	3	14	0.19	1	16	0.025
9	17	5	0.77	5	17	0.028	2	15	0.17	0	17	0.022
10	18	4	0.75	6	16	0.025	2	15	0.15	0	17	0.020
11	14	8	0.74	4	18	0.021	3	14	0.13	1	16	0.018
12	19	2	0.72	2	19	0.018	1	16	0.11	0	17	0.016
13	14	7	0.69	5	16	0.015	3	14	0.10	0	17	0.014
14	18	3	0.67	6	15	0.010	2	15	0.085	0	17	0.013
15	13	8	0.65	2	19	0.004	0	17	0.072	0	17	0.012
16	14	7	0.63	5	16	0.17	2.	15	0.059	0	17	0.011
17	12	9	0.62	6	15	0.164	0	17	0.051	0	17	0.009
18	12	9	0.60	7	14	0.158	0	17	0.044	0	17	0.008
19	14	7	0.57	1	20	0.154	1	16	0.038	0	17	0.008
20	11	10	0.53	1	20	0.147	0	17	0.034	0	17	0.007
21	11	10	0.49	4	17	0.141	1	16	0.028	0	17	0.006
22	7	14	0.47	4	17	0.135	0	17	0.022	0	17	0.005
23	10	11	0.42	6	15	0.129	0	17	0.019	0	17	0.005
24	11	10	0.38	3	18	0.122	0	17	0.015	0	17	0.004
25	3	18	0.35	1	20	0.116	0	17	0.013	0	17	0.003
26	5	16	0.32	5	16	0.110	0	17	0.010	0	17	0.003
27	5	19	0.27	4	17	0.104	0	17	0.008	0	17	0.002
28	5	19	0.20	2	19	0.085	0	16	0.006	0	16	0.001
29	4	17	0.13	0	21	0.080	0	16	0.004	0	16	0.001
30	3	18	0.05	4	17	0.054	0	16	0.001	0	16	0.000

Table 5. Tests of association (continued).

	Alternative								
	BART With Walk Access			BART With Automobile Access			BART With Bus Access		
	No. of Residuals		Avg Probability	No. of Residuals		Avg Probability	No. of Residuals		Avg Probability
Cell	Positive	Negative	for Cell	Positive	Negative	for Cell	Positive	Negative	for Cell
1	0	12	0.057	6	6	0.390	1	4	0.266
	0	12	0.037	4	8	0.289	1	4	0.148
2	1	11	0.029	2	10	0.222	1	4	0.119
3	0	12	0.025	5	7	0.195	1	4	0.103
4	1	11	0.021	3	9	0.177	1	4	0.088
5	0	12	0.019	1	11	0.160	0	5	0.080
6	0	12	0.017	2	10	0.138	0	5	0.067
7	1	11	0.016	2	10	0.124	1	4	0.060
8	0	12	0.014	1	11	0.113	0	4	0.052
9	0	12	0.012	0	12	0.103	0	4	0.044
10	1	11	0.011	2	10	0.093	0	4	0.034
11	0	12	0.010	1	11	0.086	0	4	0.031
12	0	12	0.008	1	11	0.077	0	4	0.027
13	0	12	0.007	2	10	0.069	0	4	0.023
14	0	12	0.007	1	11	0.065	0	4	0.018
15	0	12	0.006	0	12	0.060	0	4	0.016
16	0	12	0.005	0	12	0.055	0	4	0.014
17	0	12	0.005	0	12	0.050	0	4	0.014
18	0	11	0.005	0	11	0.046	0	4	0.012
19	0	11	0.004	0	11	0.042	0	4	0.010
20	0	11	0.004	0	11	0.038	0	4	0.008
21	0	11	0.003	0	11	0.034	0	4	0.007
22	0	11	0.003	0	11	0.030	0	4	0.006
23	0	11	0.002	0	11	0.028	0	4	0.005
24	0	11	0.002	0	11	0.025	0	4	0.005
25	0	11	0.002	0	11	0.021	0	4	0.004
26	0	11	0.001	0	11	0.018	0	4	0.003
27	0	11	0.001	0	11	0.015	0	4	0.002
28	0	11	0.001	0	11	0.010	0	4	0.001
29	0	11	0.000	0	11	0.004	0	4	0.000
30				0	11				

be higher than that of bus on-vehicle time and gave the explanation that, while automobiles are more comfortable than transit, the difficulty of driving an automobile during rush-hour congestion makes automobile on-vehicle time more onerous than transit on-vehicle time. The model of Table 2 requires that automobile and bus times be valued equally; this constraint may contribute to the failure of the model of Table 2 in the test against the more general model.

Tests Based on Conditional Choice

If two dependent modes are included in the calibration sample, a different set of model coefficients will be generated than those generated from a model in which one of the the dependent modes is eliminated; i.e., violation of the IIA property will cause the maximum-likelihood parameter estimates to be biased. If the IIA property is valid, however, the coefficients estimated from the full choice set will coincide with the coefficients for a smaller choice set. An obvious test of the validity of the IIA property is whether or not the coefficients estimated from a reduced choice set are statistically different from those estimated from the full choice set.

In applying this test, the estimation is performed on the subsample of individuals who chose an alternative in the subset of alternatives to be tested for dependence. The coefficients of representative utility are estimated on the subsample and the log likelihood at convergence is calculated; the log likelihood is also calculated on the subsample with the coefficients restricted to the values given in Table 2. By using the likelihood-ratio test statistic analogous to that applied to the UL model test (Equation 9), the hypothesis that the coefficients estimated on the subsample are the same as those given in Table 2 is tested. The results of the tests for various subsets of alternatives are given in Table 3. The subsets chosen for testing were those that seemed most probable to cause rejection of the hypothesis of equal coefficients. For example, models similar to that of Table 2 estimated on a sample taken before BART was providing service greatly overpredict the use of BART with walk access; hence, the subset consisting of all alternatives except BART with walk access seemed particularly relevant for testing models based on conditional choice.

The hypothesis that the coefficients estimated on the subsample are the same as those of Table 2 (the hypothesis that the property IIA is valid) is accepted for each subset of alternatives except the subset that includes all alternatives except car pool. The failure of this test for this subset is probably the result of measurement errors in the observed attributes of the car-pool alternative. The exact attributes of the car-pool mode depend on such factors as the number of persons in the car pool, each person's home and work locations, and the allocations of costs among car-pool members. Because these variables cannot be determined for persons who do not choose car pool, crude estimates were used in calculating car-pool attributes.

Residuals Tests

Violations of the IIA property will cause systematic errors in the predicted choice probabilities. The difference between the observed choices and the predicted choice frequencies (the residuals) will therefore depend on whether the IIA property is valid or not (1).

To illustrate the way in which the residuals may be used to test the validity of the IIA property, suppose that an MNL model is estimated. Then the residuals

$$D_{jn} = (S_{jn} - R_n P_{jn})/(R_n P_{jn})^{1/2} \tag{10}$$

can be defined, where

n = 1, . . . , N is a sample,
P_{jn} = $P(j \mid C, X_n, s_n)$ for $j \epsilon C$ is the estimated choice probability,
R_n = number of repetitions (possibly one) of sample point n, and
S_{jn} = number of choices j.

To avoid statistical dependence in the above residuals, it is sometimes more convenient to work with the transformed residuals,

$$Y_{jn} = D_{jn} - D_{ln}(P_{jn})^{1/2}\,[1 - (P_{ln})^{1/2}]/(1 - P_{ln}) \tag{11}$$

where $l \epsilon C$ is a fixed alternative and $j \neq 1$. Under the hypothesis that the estimated model is correct, the residuals D_{jn} have, asymptotically, zero mean, unit variance, and covariance, $ED_{ln}D_{jn} = -(P_{ln}P_{jn})^{1/2}$. The residuals Y_{jn} are asymptotically independent and have zero mean and unit variance. Further discussion of these residuals and their properties has been given by McFadden (5).

Tables 4 and 5 present tests of association of the residuals and estimated probabilities of the model in Table 2. For each alternative, a contingency table is constructed as described by McFadden, Tye, and Train (2): The estimated probabilities for the alternative are ranked and classified into 30 cells, with each cell containing approximately the same number of cases, and the numbers of positive and negative residuals associated with the probabilities in a cell are counted. (The number of positive and negative residuals summed over all cells for a particular alternative is different for different alternatives because the number of persons in the sample who have a given alternative available varies among alternatives.)

If the MNL form and the specification of Table 2 are accurate, then the number of positive residuals is expected to be higher for low-numbered cells than for high-numbered cells (a positive residual is generated if the alternative was actually chosen). This pattern emerges for each alternative.

The goodness-of-fit test

$$\chi^2 = \sum_{m=1}^{M} (S_m - N_m \overline{P}_{jm})^2 / N_m \overline{P}_{jn} \tag{12}$$

where

m = index of cell,
M = total number of cells,
S_m = number of positive residuals in cell,
N_m = total number of observations in cell,
$\overline{P}_{jm}$ = average probability for alternative j in cell m, and
$\overline{P}_{jn}$ = average probability of alternative j for total sample,

has an asymptotic distribution bounded by χ^2 distributions with M - 1 and M - K - 1 df, where K is the number of estimated parameters. These test statistics are not independent across alternatives.

The test statistic for each alternative is given below.

Alternative	Test Statistic
Automobile alone	17.51
Bus with walk access	14.14
Bus with automobile access	15.75
Car pool	38.63
BART with automobile access	13.81
BART with walk access	15.83
BART with bus access	5.32

Since there are 30 cells and 19 parameters, the test statistic has an asymptotic distribution, under the hypothesis that the MNL form and the specification of Table 2 are correct, bounded by χ^2 distributions with 29 and 10 df. The critical (0.05-level) value of χ^2 with 29 df is 42.56; that with 10 df is 18.31. The values of the test statistic for all alternatives except car pool are below the lower of the two bounding critical values, and therefore the hypothesis is accepted for those alternatives. For the car-pool alternative, the test statistic falls between the two bounding critical values: The test is therefore inconclusive. As in the failure of the test based on conditional choice, measurement errors in the car-pool attributes are probably the reason that the car-pool alternative cannot pass the test of association unambiguously.

Other Tests

Other tests using the properties of residuals are the means test and the variance test. Other tests that may be used are the saturated model test, which was found not to be powerful, and tests using two data sets. Tests using two data sets were found to be particularly powerful in identifying violations of the IIA property (2). For example, a before-and-after data set involving the introduction of a new mode offers a particularly powerful test of the independence of the mode. Both likelihood-ratio and residuals tests can be used. Another alternative that deserves consideration is to test the MNL model against the multinomial probit model with an explicit structure of dependence of unobserved attributes, which is practical if the number of choice alternatives is four or less (7).

Modifications of the Modal Choice Model to Correct for Violations of the Irrelevance of Independent Alternatives Property

The model given in Table 2 failed two of the tests of the IIA property. First, it failed the universal-logit test against a more general model with six extra cross-alternative variables. The probable reason for this failure is that the model constrains the value of automobile and transit time to be equal. Second, it failed the test of equality of coefficients across choice sets when the car-pool alternative was eliminated. The probable reason for this failure is that the car-pool data were poor.

A new model of work-trip modal choice has been given by Train (8). This model is more general than the model given in Table 2 in that, among some other generalizations, automobile and transit on-vehicle times are allowed to have different coefficients and socioeconomic variables are allowed to enter the car-pool alternative.

This more general model passed both of the diagnostic tests that the MNL model (Table 2) failed:

1. The universal logit test—the log likelihood of Train's model is -519.9. The log likelihood of the more general model (which includes the six cross-alternative variables) is -515.5. Therefore, the test statistic is 8.8. The critical (0.05-level) value of χ^2 with six df is 12.6. The model passes the test.

2. The test of equality of coefficients across choice sets—the log likelihood of Train's model with the car-pool alternative removed is -191.0. The log likelihood of the model with the car-pool alternative removed and the parameters restricted to those obtained with all alternatives included is -199.4. The test statistic, therefore, is 16.8. The critical (0.05-level) value of χ^2 with 23 degrees of freedom is about 35. The hypothesis of equal parameters is accepted, and the model passes the test.

These results illustrate that the passing or failing of the diagnostic tests depends on the specification of the model for a particular choice situation, not on the choice situation itself; i.e., the IIA property is or is not valid for a particular model, not for a particular choice situation. These results indicate the way in which the diagnostic tests can be used to find problems in the specification of the model. The diagnostic tests applied to the model in Table 2 indicated that there were problems in the on-vehicle-time variable and the car-pool alternative; these problems were corrected in Train's model.

ACKNOWLEDGMENT

The development of the concepts and statistical tests of this research was supported by the National Cooperative Highway Research Program. The application of the model calibration and of the tests was supported by the National Science Foundation Research Applied to National Needs Program. The opinions and conclusions expressed or implied are those of the authors. They are not necessarily those of the Transportation Research Board, the National Academy of Sciences, the Federal Highway Administration, the American Association of State Highway and Transportation Officials, the individual states participating in the National Cooperative Highway Research Program, or of the National Science Foundation, the University of California, or Charles River Associates, Inc.

REFERENCES

1. The Independence of Irrelevant Alternatives Property of the Multinomial Logit Model. In Disaggregate Travel Demand Models, Charles River Associates, Cambridge, Mass.; NCHRP, Phase 1 Rept., Vols. 1 and 2, 1976.
2. D. McFadden, W. Tye, and K. Train. Diagnostic Tests for the Independence From Irrelevant Alternatives Property of the Multinomial Logit Model. Institute of Transportation Studies, Univ. of California, Berkeley, Working Paper No. 7616, Aug. 1976.
3. A Disaggregated Behavioral Model of Urban Travel Demand. Charles River Associates, Cambridge, Mass., 1972.
4. T. Domencich and D. McFadden. Urban Travel Demand: A Behavioral Analysis. North-Holland, Amsterdam, 1975.
5. D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers in Econometrics (P. Zarembka, Ed.), Academic, New York. 1973.
6. K. Train. Work-Trip Mode Split Models: An Empirical Exploration of Estimate Sensitivity to Model and Data Specification. Institute of Transportation

Studies, Univ. of California, Berkeley, Working Paper No. 7602, 1976.
7. J. Hausman and D. Wise. A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences. Department of Economics, MIT, Cambridge, Working Paper No. 173, April 1976.
8. K. Train. A Post-BART Model of Mode Choice: Some Specification Tests. Institute of Transportation Studies, Univ. of California, Berkeley, Working Paper No. 7620, 1976.

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

Effects of Parking Costs on Urban Transport Modal Choice

David W. Gillen, Department of Economics, University of Alberta, Edmonton

The effects of parking costs on urban modal choice are investigated by using a standard binary-choice model and estimated by using the logit technique. Previous studies have misspecified the form of the parking-cost variable and the model normally estimated. After estimating the traditional and correctly specified models, the claim that parking taxes are an effective substitute for roadway pricing in influencing congestion is only partially supported. Aggregate elasticities for four policy-oriented variables are calculated. The elasticities provide a measure of the bias from misspecification and indicate the most effective policy variable for the reduction of automobile use.

This paper is concerned with estimating the effects of parking costs on individual choice of transportation mode for trips within urban areas. It has three basic objectives:

1. The determination of how to characterize the parking variable and incorporate it into a model of modal choice,
2. The calculation of the elasticity of modal choice with respect to parking costs, and
3. The determination of the way in which changes in one of the characteristics that determine the modal choices of individuals will affect the expected proportion of individuals taking the choice being considered.

{The third objective is implemented by examining the way in which changes in individual characteristics affect the mean of the distribution of population probabilities [cf. Westin (11)]}.

The first section introduces a model of individual choice of transportation alternatives that treats parking as a commodity, the demand for which is derived from the choice of the automobile as the transit mode. The second section describes the data and the implications of this model for the structural forms of the estimating equations. The third section presents the empirical results for an application of this model to data for Toronto. The fourth section presents the derivation of the elasticity of modal choice with respect to instrumental variables and empirical results for aggregate and individual elasticities.

BASIC MODEL

The variable to be explained is the individual's choice of transportation mode (automobile versus public transit). The econometric model used in this paper to represent this binary-choice problem is derived from a choice-theoretic framework, based on a microeconomic behavioral model developed by DeSerpa (2), in which individuals maximize utility in choosing among alternative goods and the times allocated to them, subject to income and time-resource constraints. In this model, the choice of any amount (X_i) of commodity i places only a lower bound on the amount of time (T_i) the individual must use in consuming X_i; a change in relative prices of either goods or times causes the individuals to substitute among goods of various time intersities and, therefore, to implicitly substitute among alternative uses of time. Others (5, 6, 7, 9, 10) have used similar theoretical approaches to demonstrate the relation between the microeconomics of choice behavior and binary-choice economitric models. These models suggest, that modal choice is a function of two categories of variables, transportation-system characteristics that affect the money and time costs of travel and user characteristics that serve as proxies for objective comfort characteristics.

Traditionally, modal-choice studies have simply added the costs of parking to the automobile running costs (7, 10). This procedure implicitly assumes that parking services and automobile use enter into the individual's production function in fixed proportions. It also implies that the decision about where to park is independent of modal choice, so that parking-location decisions are unaffected by variations in time costs.

In this paper, parking is defined as a commodity that is complementary to automobile trips. The individual is assumed to maximize a utility unction [$U(C_i)$], where $C_i = F(X_i, T_i)$, subject to income and time-resource constraints. The explicit specification of the production functions that determine C_i is important for understanding the role of parking use. For transit, the service consumed is generated by the production function

$$C_T = F_T(X_T, T_T) \quad (1)$$

where X_T = transit service purchased and T_T = time spent in using X_T; for automobile use, the service consumed is generated by the function

$$C_A = F_A(X_A, T_A, X_P, T_P) \quad (2)$$

where

X_A = automobile service consumed,
T_A = time spent in using the automobile,
X_P = parking services used, and
T_P = time spent in moving from parking facility to final destination.

The consumption of parking is location specific because each location is characterized by a unique money and time cost and alternative parking locations are substitutes. The individual, in choosing his or her mode, is aware of some average parking cost about the destination point, but has some discretion with respect to the final location chosen. Although the individual faces a binary choice with respect to mode, he or she does not with respect to parking-location decisions.

Changes in parking costs cause both parking relocation and modal switching. As parking prices change, individuals tend to reallocate their money and time resources to less expensive substitute commodities in two ways: (a) by reallocation of time and money within the automobile mode by changing parking locations and (b) by reallocation of time and money by substituting other modes. The former effect is not captured in traditional modal-choice models.

The suppression of the parking-relocation decision in the traditional modal-choice models implies that such models will have a biased prediction of the responsiveness of changes in mode to changes in relative parking or modal running costs. The bias may enter in a number of ways. First, if parking costs are added to automobile running costs, one will find that, as the distance traveled increases, the automobile running costs increase and parking costs become a lower proportion of total automobile costs and could then conclude that an increase in parking costs will have little effect on the modal choices of more distant individuals. However, the lack of effect could also be attributed to the increasing differential between other service attributes of the two modes as distance increases. Second, if higher income individuals do consume more housing and thus tend to live at more distant locations, the lack of a parking-price effect may be attributed to an income effect rather than to a price effect. If one does not treat parking costs separately and stratify the data by income level, one cannot separate these effects.

ESTIMATING EQUATIONS

The general form of the equation used for estimating modal-choice behavioral models is

$$P_i = f(TSC_j, UC_k) \tag{3}$$

where

P_i = probability of selecting mode i,
TSC_j = jth transportation-system characteristic, and
UC_k = kth user characteristic.

Since the P_i are limited to the internal (0, 1) choice, the function f is nonlinear.

There are alternative methods for estimating probabilistic modal-choice models. The method used here is logit analysis (6, 8, 9). This technique estimates the probability of a particular choice of mode as a nonlinear function of the explanatory variables. The relationship is assumed to be S-shaped, which is logistic sigmoid (i.e., the cumulative distribution function for the logistic). The form of the logit model is

$$P = \left\{1 + \exp\left[-G(x)\right]\right\}^{-1} \tag{4}$$

where G(x) = linear function of explanatory variables and P = probability of choosing to use the automobile mode conditional on x.

Logit models will not be discussed in detail here because they are well documented in the literature (1, 3, 5, 6, 9). A computer program described by Nerlove and Press (8) was used in all of the estimations.

The data were taken from the 1964 Metropolitan Toronto and Regional Transportation Study (MTARTS) home-interview summary. The estimates used in this research are based on a subsample of 515 work trips. The dependent variable is the probability of choosing the automobile for all of the equations estimated.

The following criteria were used to select the subsample:

1. The trips originated within the boundaries of metropolitan Toronto;
2. The zone of origin was the same as the home zone because we are interested in home-based trips;
3. The trip destination was within the central business district;
4. The mode used for the trip was either private automobile or public transit; and
5. Each family had the use of an automobile, and at least one member of the family possessed a driver's license.

From the total sample of 84 064 trips, a final subsample of 3012 trips was selected. This number was significantly reduced to 515 trips because of missing or miscoded information.

All the data used in the estimation were taken from the computer tape provided by MTARTS. Some variables used in the estimation were calculated from information provided by MTARTS.

The real-route distance was not coded. Distance was calculated as the straight-line distance between centroids of the home and destination zones. The relation between the real or route distance (D_R) and the straight-line distance (D_{SL}) is $D_R = 1.4D_{SL}$ (12).

The travel-time variable used in the model is T_c/T_T. The data provided information on the departure and arrival times of the mode used. It also provided information on the excess vehicle time at the beginning and end of the trip.

The following procedure was used to estimate the time of travel for the mode not used for the given trip: A regression of the travel time on the straight-line distance for each mode was computed, and the regression estimate was then used as the proxy for the time required for the journey to work if that mode is used. The regressions were computed over the whole sample; the results for the automobile mode are:

$$T_{C_1} = 0.30 + 0.037D_{SL}, \quad T_{C_2} = 0.23 + 0.038D_{SL} \tag{5}$$

(R = 0.383 and N = 341.)

where

T_{C_1} = total travel time by automobile (h),
T_{C_2} = in-vehicle travel time by automobile (h),
D_{SL} = straight-line distance between origin and destination zone,
R = correlation coefficient, and
N = number of observations.

The regression results for public transit are

$$T_{P_1} = 0.31 + 0.072D_{SL}, \quad T_{P_2} = 0.16 + 0.072D_{SL} \tag{6}$$

($R = 0.59$ and $N = 430$.)

where

T_{P_1} = total travel time by transit (h) and
T_{P_2} = in-vehicle travel time by transit (h).

The t-statisties for the coefficients in Equations 5 and 6 are shown below.

Value	t-Statistic	Value	t-Statistic
0.30	9.54	0.31	12.67
0.037	7.62	0.072	14.9
0.23	7.27	0.16	6.06
0.38	6.38	0.72	14.3

The differences between T_{c_1} and T_{c_2} and between T_{P_1} and T_{P_2} are walking times at the beginning and ending of the trips. The constant term decreases from T_{c_1} to T_{c_2} and from T_{P_1} to T_{P_2}. The change is greater for transit than for automobile because the removed components are a larger proportion of transit.

The small decrease from T_{c_1} to T_{c_2} is due to the discretion that the individual has with respect to scheduling and location. The rather high constant term in the equation for T_{c_2} may be interpreted as the average effect of congestion on in-vehicle time because a large percentage of the sample trips were made in the peak periods and may also include the average time used looking for a parking location.

Explanatory Variables

The following transportation-system and user characteristics were used as explanatory variables:

T_c = overall travel time by automobile for a given trip;
T_T = overall travel time by transit for a given trip;
C_c = money cost of emphasizing the automobile for a given trip, which are here set equal to the marginal (equal to the average) running costs of the automobile (fuel, oil, tires) per kilometers traveled plus the parking fee;
C_T = money cost of using transit for a given trip;
PC = parking fee paid for a given trip;
F_c = money cost of using the automobile minus the parking fee ($F_c = C_c$ - PC);
T_T = C_T;
EPC = inclusive price of parking (parking fee plus time costs), which is equal to PC plus the marginal value of time times walking time from parking location to destination (the time component is deducted from T_c when EPC is entered into the estimating equation);
Y = personal income of the trip maker;
A = dummy variable for age of the trip maker (1 if age is between 20 and 55 years, 0 if otherwise);
SX = sex (1 if female, 0 if male); and
SS = occupation-status dummy (1 if the individual is a white collar worker, or a middle or professional manager, 0 if otherwise).

The dependent variable is coded 1 if automobile is used, 0 if transit is used.

Travel Costs

The questions in the MTARTS survey relating to direct costs or money saved or lost by alternative modes were unusable or qualitative, which is disappointing because perceived cost is the relevant cost in this behavioral model.

For transit riders, the perceived cost is the fare, but the answer is less straightforward for automobile users. Individuals normally purchase an automobile for many uses, of which work is one; hence intuitively it seems reasonable to conclude that individuals will consider only the marginal running costsof a trip. This includes fuel, oil, tires, and maintenance at most and fuel at least. (The inclusion of parking costs should be entered as a separate cost variable.)

The running costs are calculated on an average automobile being driven at the average speed:

$$C_i = \bar{C} \times 1.4D_{SL} \tag{7}$$

where

C_i = cost of ith trip,
$\bar{C}$ = average running cost of automobile [approximately 2.5 cents/km (4 cents/mile)], and
1.4 = proportionality factor relating straight line to real distance.

The parking-cost variable was coded for those individuals who use the automobile, but not for transit riders. A parking-cost variable for that group was constructed from knowledge of their walking time at the end of their trip to their destination point, their zone of destination and its associated parking-rent gradient, and the trip purpose of the individual.

There are a number of methods that can be used to construct the EPC for transit users. The method here was to use the longer of the average distance walked by automobile users in that census tract or the distance walked from the transit terminus to the destination point. This establishes an outer bound for the transit user.

User Characteristics

The MTARTS home-interview survey also provides the following socioeconomic information: sex, age, occupation (social status), number of wage earners, household income, trip purpose, travel mode, driver's license, number of vehicles owned or leased, and number of persons in the household and their age distribution.

Parking Data

The data about parking fees and the fee structure across the urban area were not available from the MTARTS tape. This information was provided by the Toronto Parking Authority, City Parking Limited, and the Metropolitan Toronto Department of Public Works. These agencies provided information about hourly and daily rates of each parking facility and the location of each facility within the central business district. All parking data were for 1964 to be compatible with the MTARTS data.

Full Price of Parking

The value of time for use in the full parking-price variable was calculated from the data set by using the technique developed by McFadden (6) and was 58 percent of the wage rate.

Model Specifications

Three models were specified and estimated to test the effects of different specifications of the effect of parking costs on modal choice. Model 1

$$G(x) = b_0 + b_1 T_c/T_T + b_2 C_c/C_T + b_3 A + b_4 SX + b_5 SS + b_6 Y \tag{8}$$

where $b_0, \ldots, b_n$ are estimated parameters, is a variant of the modal-choice models used in other studies. It will serve as a benchmark against which the estimates of this data can be compared with those of other studies. It incorporates parking costs with the modal running-cost variable: Because $C_c = F_c + PC$, the relative cost coefficients can be compared with those in models 2 and 3 in which modal money and parking costs are separated. Model 2

$$G(x) = b_0 + b_1 T_c/T_T + b_7 F_c/F_T + b_8 PC/F_T + b_3 A + b_4 SX + b_5 SS + b_6 Y \tag{9}$$

separates parking costs and modal running costs. This permits examination of the separate influences of changes in modal money costs and changes in parking fees on modal choice and the determination of whether or not individuals capitalize money costs at the same rate despite differences in the service purchases; i.e., whether $b_7 \leq b_8$. It also allows comparison of the relative magnitudes of the relative cost and relative fare variables. If $b_2 > b_7$ and b_8 is significant, the effect of a change in parking costs measured in model 1 will be upwardly biased. Model 3

$$G(x) = b_0 + b_1 T_c/T_T + b_7 F_c/F_T + b_9 EPC/F_T + b_3 A + b_4 SX + b_5 SS + b_6 Y \tag{10}$$

is a variation of model 2 that incorporates an inclusive parking price, a formulation suggested by McFadden (6). EPC combines the money and time costs associated with parking. The time element associated with the terminal end of the trip, which was previously included in the overall trip time, is now included in the full price rather than in T_c. This permits an estimation of the effects of changes in parking fees on the full price of parking and a comparison of the coefficient estimates of models 2 and 3. A priori, one would expect that $b_9 < b_8$ and that the elasticities calculated from these variables would have the same relation because parking costs as a proportion of full costs decrease with distance.

EMPIRICAL RESULTS

The parameter estimates of models 1, 2, and 3 are given in Table 1. The estimates of all three models are of the expected signs and the magnitudes agree with those of previous modal-choice studies. The likelihood ratio, $-2 \log \lambda$, tests the hypothesis of dependence between the dependent and explanatory variables. The values of the likelihood ratio for all three models indicate that the null hypothesis of independence between dependent and explanatory variables can be rejected. Because the focus of this paper is on the relative cost and parking variables, only these will be discussed.

The values of b_2 and b_7 differ significantly when parking costs are and are not included in modal running costs. The high estimate of b_2, the coefficient of the relative-cost variable (in which parking fees are included), would seem to indicate that, for a given change in relative modal costs, the effect on modal switching will be perceived to be the same regardless of the source of the change in relative costs. The low estimate of b_7, the coefficient of the relative fare variable, and the relatively higher values of b_8 and b_9, the coefficients of the parking costs variables, indicate that changes in relative modal fares have a negligible effect on modal switching, but changes in parking costs have a higher effect on inducing individuals to switch to transit. If one attempts to measure the effects of changes in parking costs on modal choice through the relative-cost variable, the results will be upwardly biased.

The sign of the coefficient for the parking variable indicates that parking services and automobile use are complements in the case of the work trip. These two services, parking and automobile use, may well be substitutes in the non-work-trip case in which variable parking duration is allowed. The magnitudes of b_8 and b_9 indicate that individuals are relatively more responsive to changes in parking costs than to changes in relative modal fares in their modal-choice decision. However, the effect of changes in parking costs on automobile use is not as great as some have previously believed. This is because changes in parking costs allow parking relocation as well as modal switching effects.

The value of b_9 is expected to be less than that of b_8 because changes in either parking fee or time costs will cause changes in full price according to its' proportion of full price.

Table 2 gives the results of estimates of models 1 and 2 when the data are stratified by income class. The relative magnitudes of the relative cost, relative-fare, and parking-costs variables are similar to those shown in Table 1.

ELASTICITY OF MODAL CHOICE

A measure of the sensitivity of the change in the probability of choosing the automobile with a change in one of the explanatory variables is provided by the point elasticity. The elasticity may be defined as

$$\xi = (\partial E[P]/\partial X_i)(X_i/E[P]) \tag{11}$$

Table 1. Parameter estimates using binary logit models.

Parameter	Explanatory Variable	Model 1 Value	Model 1 t-Statistic	Model 2 Value	Model 2 t-Statistic	Model 3 Value	Model 3 t-Statistic
b_0	Constant (automobile preference effect)	1.84	5.87	1.44	6.51	1.07	6.14
b_1	T_c/T_T	-2.00	8.30	-1.73	7.05	-1.83	7.46
b_2	C_c/C_T	-0.159	4.09	—	—	—	—
b_7	F_c/F_T	—	—	-0.013	1.46	-0.021	1.87
b_8	PC/F_T	—	—	-0.244	6.81	—	—
b_9	EPC/F_T	—	—	—	—	-0.205	6.12
b_3	A	0.47	3.93	0.396	3.15	0.419	3.38
b_4	SX	-0.85	6.00	-0.84	6.16	-0.85	6.43
b_5	SS	-0.02	0.19	0.049	0.38	0.008	0.068
b_6	Y	0.096	3.07	0.103	3.13	0.167	4.82

Notes: Sample size = 515; dependent variable is automobile transit choice (0 if transit is used, 1 if automobile is used) for the work trip; $-2 \log \lambda$ = 246.9, 284.7, and 275.1 for models 1, 2, and 3 respectively.

Table 2. Parameter estimates by income group.

Income Class ($ 000)	-2 log λ	Constant (Automobile Preference Effect) Value	Constant t-Statistic	Age Value	Age t-Statistic	Sex Value	Sex t-Statistic	T_c/T_T Value	T_c/T_T t-Statistic	F_c/F_T Value	F_c/F_T t-Statistic	C_c/C_T Value	C_c/C_T t-Statistic	PC/F_T Value	PC/F_T t-Statistic
4 to 6	42.01	3.24	4.52	0.238	0.818	-1.12	3.14	-3.91	4.63	-0.041	1.22	—	—	-0.652	1.49
6 to 8	39.65	4.03	4.61	0.027	0.091	-1.01	3.11	-3.66	3.99	-0.037	1.12	—	—	-0.259	3.59
8 to 10	42.23	2.65	4.25	0.438	1.53	-0.61	1.99	-2.24	3.48	-0.069	1.17	—	—	-0.237	3.32
>10	27.29	6.06	4.08	1.06	2.75	-1.49	3.37	-4.41	4.47	-0.025	1.32	—	—	-0.291	3.18
4 to 6	43.25	3.38	4.54	0.295	1.04	-1.09	3.22	-3.88	4.01	—	—	-0.091	1.48	—	—
6 to 8	46.57	3.66	4.09	0.156	0.575	-0.97	3.34	-3.58	4.29	—	—	-0.178	1.88	—	—
8 to 10	46.56	2.65	3.91	0.443	1.65	-0.78	2.34	-2.46	3.90	—	—	-0.17	2.21	—	—
>10	35.23	4.39	4.15	1.03	3.12	-0.93	2.85	-4.09	5.09	—	—	-0.283	2.02	—	—

where ξ = elasticity of choice of mode with respect to variables X_i and $E[P]$ = expected probability of choosing the automobile.

One can simplify this expression by substituting $E[P] = \exp G(X)/[1 + \exp G(X)]$ and differentialing. Then

$$\begin{aligned}\xi &= \left(\partial\left\{\exp G(x)/[1+\exp G(x)]\right\}/\partial G(x)\right) \times \partial G(x)/\partial X_i \times \\ &\quad [X_i/\exp G(x)]/[1+\exp G(x)] \\ &= \left\{\exp G(x)/[1+\exp G(x)]^2\right\} \times b_i \times [X_i/\exp G(x)]/[1+\exp G(x)] \\ &= (1 - E[P])b_iX_i \end{aligned} \tag{12}$$

ξ becomes intractable if the transportation-system characteristics are expressed in terms of differences because at some point X_i may be zero. The elasticity is a linear function of the X_i's and is proportional to the probability of choosing the other mode.

The elasticities may be calculated in three different ways. The probability may be estimated for a specific individual by using individual values of the explanatory variables. The elasticity would then be estimated by Equation 12. The second method, which is referred to below as the individual elasticity (ξ), is to estimate the logit function across individuals for the set of explanatory variables. The probability is then estimated from the sum of the parameter estimates times the mean value of each explanatory variable, given the distribution of characteristics over the population. The expected probability (E[P]) is then used in Equation 12 to determine ξ. The third method, developed by Westin (11), is to determine the probability for each category or class within the distribution for each explanatory variable. These probabilities are then aggregated, and E[P] is determined from these weighted probabilities and used in Equation 9 to estimate ξ', which is referred to as the aggregate elasticity. Westin has shown that the difference between (ξ) the individual elasticity and ξ' is determined by the ratio $E[P](1 - E[P])/E[P]\ (1 - P)$. The difference between ξ and ξ' occurs because ξ is determined from the mean value of the explanatory variables, which are aggregated, and the probability P is then determined, but ξ' is the mean of the aggregate of probabilities, which are estimated for each class in the distribution of characteristics. One would expect ξ and ξ' to be very close.

The table below gives both the individual and the aggregate elasticities for C_c/C_T, F_c/F_T, PC, and EPC.

Elasticity	C_c/C_T	F_c/F_T	PC	EPC
ξ_i	-0.34	-0.068	-0.38	-0.25
ξ_i'	-0.29	-0.059	-0.31	-0.19

The decrease in the relative-cost elasticity when parking costs are and are not included, i.e., in C_c/C_T and F_c/F_T respectively, demonstrates the nonresponsiveness of modal switching to changes in transit fares. It is also evident that the relative-cost elasticity upwardly biases the expected change in mode when relative running costs change but downwardly biases the expected change in modes when parking fees change. The parking-fee elasticity is larger than both the relative-cost and relative-fare elasticities, but is somewhat smaller than previously suggested (4).

CONCLUSIONS

The models developed and the empirical results obtained in this research point to the following conclusions:

1. The form of the parking-cost variable is crucial if one hopes to obtain unbiased estimates of the effects of charges in parking costs on modal choice. The response to modal switching will be overestimated if one simply adds the parking costs to the modal running costs. The parking variable was entered in the estimations in two forms. The parking-cost variable (PC) allows one to estimate the effects on modal choice of any changes in the many costs of parking. The full-price variable for parking (EPC) is somewhat more flexible than PC because it allows one to estimate the effects of a number of parking-policy changes, whether they relate to changes in parking fees (such as a parking tax) or changes in time costs (such as parking restrictions).
2. The sign of the coefficient of the parking variable indicates that parking services and automobile use are complements for the work trip. The relatively low elasticities, especially in those of the EPC, can be attributed to the fact that changes in parking costs result in parking relocation as well as modal switching. It is only those individuals who are on the margin of relocating or switching modes who tend to switch. These individuals are generally those who have located some distance from their destination point and who represent a small proportion of the total parking. This point is more precisely captured in the EPC variable.

The elasticity of the probability of automobile use with respect to parking costs indicates that a 10 percent increase in parking fees will result in a 3.1 percent decrease in automobile users as a percentage of current automobile users. If automobile and transit uses are split 50:50 for the work trip, 3.1 percent of the automobile users will switch modes. If this split is 75 percent automobile and 25 percent transit, a change in parking costs may cause a 10 or 15 percent increase in transit use as a percentage of current transit use. Such large changes would require a coordinated transit investment; i.e., a parking pricing policy.

If a parking authority attempts to introduce parking fee changes of a local nature, i.e., at one or two points, the effect on automobile use will be negligible because the affected individuals will simply relocate.

3. The use of the technique suggested by Westin for making aggregate predictions from individual data leads to calculated aggregate elasticities that are slightly smaller than the individual elasticities. These elas-

ticities are valid only for small changes in the explanatory variable.

REFERENCES

1. F. X. deDonnea. Consumer Behavior, Transport Mode Choice, and the Value of Time. Regional and Urban Economics, Vol. 1, No. 4, 1972.
2. A. C. DeSerpa. The Theory of the Economics of Time. Economic Journal, Dec. 1971.
3. D. W. Gillen. An Economic Analysis of Parking and the Effects of Alternative Parking Policies on Modal Choice and Congestion. Department of Political Economy, Univ. of Toronto, unpublished PhD dissertation, 1975.
4. I. Heggie. Urban Transport Congestion: Solution by Restraint. Proc., Transportation Research Forum, 1973, pp. 27-44.
5. C. A. Lave. A Behavioral Approach to Modal-Split Forecasting. Transportation Research, No. 3, Dec. 1965.
6. D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers of Econometrics (P. Zarembka, ed.) Academic, New York, 1974, Chap. 4.
7. R. G. McGillivray. Demand and Choice Models of Modal Split. Journal of Transport Economics and Policy, Vol. 4, No. 2, May 1970.
8. M. Nerlove and S. J. Press. Univariate and Multivariate Log-Linear and Logistic Models. Rand Corporation, Santa Monica, Calif., R-1306-EDA/NIH, 1972.
9. S. Reichman and P. Stopher. Towards Disaggregate Stochastic Models of Travel Mode Choice. HRB, Highway Research Record 369, 1972, pp. 91-104.
10. A. Talvitie. Comparison of Probabilistic Modal-Choice Models: Estimation Methods and Systems Inputs. HRB, Highway Research Record 392, 1972, pp. 111-121.
11. R. B. Westin. Predictions From Binary-Choice Models. Journal of Econometrics, Vol. 2, 1974, pp. 1-16.

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

Analysis of Predictive Qualities of Disaggregate Modal-Choice Models

Thomas E. Parody,* Charles River Associates, Cambridge, Massachusetts

The results of a study that examined the predictive accuracy and ability of a set of disaggregate, behavioral demand models of modal choice are presented. Although other issues such as sample size, value of time, demand elasticity, and policy predictions are discussed, the primary objective was to test the validity of disaggregate logit models in forecasting. The analysis is structured around a carefully designed before-and-after study of individual travel behavior as affected by significant, short-term changes in the transportation system. Various specifications of disaggregate modal-choice models are calibrated by using as input data the actual responses of individuals from the before phase of the travel-behavior surveys. This was followed by a series of prediction and validation phases by using the after data that was generated by changes in the transportation system. Because the actual modal shares are known from the longitudinal data, it is possible to assess accurately the predictive qualities of the calibrated logit models. The results of the empirical analysis indicate that disaggregate models, especially those that include a full range of transportation level-of-service and socioeconomic variables, can be used to predict future behavior with acceptable levels of performance.

Traditionally, disaggregate demand models have been evaluated on the basis of how well they calibrate (or of how well they replicate existing behavior) rather than on their ability to forecast adequately changes in travel demand. Such analyses are severely limited when the sets of data that were used in the model calibration are also used for its validation. As expressed by Pratt (14), "there have been all too few rigorous comparisons of modeled travel demand with actual before-and-after data." Yet, if the primary function of a modal-choice model is to predict the impact of changes in the transportation system on travel behavior, then an essential characteristic of such a model is its ability to predict accurately. However, because of data restrictions, it is generally not possible to evaluate the accuracy of forecasts obtained from disaggregate demand models.

An opportunity to evaluate and document the capabilities of disaggregate models arose when the University of Massachusetts at Amherst was selected to participate in a demonstration of a free bus service accompanied by increases in parking fees and associated parking restrictions. The demonstration, which was funded by a grant from the Urban Mass Transportation Administration, involved expanding a limited, 3-bus, campus shuttle system by the addition of 10 new buses so that commuter trips could also be served. Later, the fleet was expanded to 16 vehicles by the purchase of 3 more buses.

The purpose of the demonstration was to determine the extent to which the availability of a free, fairly convenient bus service, coupled with the introduction of restrictive parking policies and increased parking prices, might cause a shift away from commuting by automobile to commuting by bus. Later, during the oil embargo and the accompanying energy crisis, it was apparent that further changes in travel patterns and behavior were occurring. Thus, although not part of the original scope of the work, the effect of rapidly increasing gasoline costs was introduced into the study framework. The main focus of the demonstration, however, was on the institution of an attractive and convenient free bus service; extensive data collection and a carefully delineated experimental design centered around changes in the parking policy.

To determine the impact of each change in the transportation system, the demonstration and data-collection phases were divided into four separate events. A before survey of travel patterns was conducted during the semester preceding the demonstration (i.e., survey period 1, the fall 1972 semester). Thereafter, observations of behavior and user characteristics were made: (a) after the introduction of the expanded free bus service (i.e., survey period 2, the spring 1973 semester), (b) after the introduction of changes in parking prices and regulations (i.e., survey period 3, the fall 1973 semester), and (c) during the approximate peak of the energy crisis (i.e., survey period 4, the spring 1974 semester).

To monitor travelers' responses to changes in the transportation system, telephone travel-behavior surveys were conducted that involved reinterviewing of the same sample of individuals over time and collecting information on the characteristics of alternative modes. Only individuals associated with the university (i.e., students, staff, and faculty) were surveyed. The data obtained from these longitudinal or time-series surveys were left in their disaggregate form.

To evaluate the accuracy of model forecasts at least two sets of data are required: one to calibrate the model and the other(s), preferably made after some quantifiable change to the transportation system, to test the prediction. This is seldom possible and the model and the forecasts usually must be tested on the same data set (15). In this study, disaggregate modal-split models were calibrated with data from one time period and validated with data from subsequent time periods. By using a longitudinal data base that was generated by planned transportation changes over time, the calibrated models could be used to forecast changes in modal choice that should result from distinct short-range changes in the transportation system. Because the actual amount of modal shifting was known, the forecasts obtained from the model could be precisely evaluated. A favorable evaluation of the model in prediction implies a properly specified and calibrated model and increases confidence in it.

MODEL

The logit form was chosen as the statistical technique to be used in calibrating the models. Detailed descriptions of this model form are given elsewhere (1, 2, 5). The following equations express the form of the multinomial logit model:

$$P(m{:}M_t) = \exp G(m, t) \Big/ \sum_{m'=1}^{M} \exp G(m', t) \tag{1}$$

where

$P(m{:}M_t)$ = probability that traveler t, out of the total sample of T travelers, will select mode m from the M set of available modes and
$G(m, t)$ = utility of mode m to traveler t.

G(m, t) can be expressed more specifically in the following form:

$$G(m, t) = G_m(Z_{m,t}, S_t) = \theta_k X_{mtk} \tag{2}$$

where

$Z_{m,t}$ = vector of variables describing the level-of-service characteristics of mode m for traveler t (e.g., time and cost),
S_t = vector of socioeconomic characteristics for traveler t (e.g., income, sex, and automobile ownership),
θ_k = K×1 vector of coefficients $(\theta_1, \theta_2, \ldots, \theta_k)$, and
X_{mtk} = K×1 vector of variables $(x_{mt1}, x_{mt2}, \ldots, x_{mtk})$.

Because the primary purpose of the free bus demonstration project was to assess the diversion of travelers from automobile to bus that would result from the implementation of free bus service and increased parking fees (and later increased gasoline costs), only the automobile and bus modes are considered here. For this two-mode situation, the model can be rewritten as

$$\begin{aligned} P(a) &= \exp \theta_k X_{atk} / (\exp \theta_k X_{atk} + \exp \theta_k X_{btk}) \\ &= 1/[1 + \exp - \theta_k (X_{atk} - X_{btk})] \end{aligned} \tag{3}$$

or

$$\ln[P(a)/P(b)] = \theta_k (X_{atk} - X_{btk}) \tag{4}$$

Variables Considered in the Model

The selection of variables was based on a combined analysis of those typically considered in previous modal-choice models [Parody (12)] and those that were included in the telephone-survey questionnaires conducted during the free bus demonstration. They are described in Table 1.

Data Selection

The first task in selecting the sample population for use in the analysis was the identification and eventual discarding of improper data. For example, the main purpose of this study was to evaluate the choices travelers will make when there are changes in the transportation system. Therefore, all captive mode users were eliminated from the sample. The individuals eliminated included those for whom bus service was not available (as determined by their proximity to the nearest bus stop) and those who responded negatively to the automobile-availability question. By using these criteria, about 70 percent of the total sample population were considered to be choice automobile or bus users.

Telephone survey 2, taken toward the end of the first semester of the free bus operation (i.e., spring 1973), was chosen as the initial data base for calibration, from which all forecasts and predictions were to be made. Thus, for calibration purposes, survey 2 was the before time period, and survey 3 (after the introduction of higher parking fees) was the forecasted time period. Consequently, in selecting the individuals to be included in the sample, it was important to ensure that the respondents answered at least surveys 2 and 3 and that they could be considered choice travelers in terms of the automobile or bus modes.

Initially, it was intended that only the reported responses of the individuals would be used as inputs to the model. However, it soon became apparent that some individuals did not record some data (usually travel times) for the alternative mode. This was almost exclusively limited to those for whom the bus was the alternative mode: Very few bus users did not estimate automobile-drive time. Of the approximately 400 individuals surveyed in each time period, 91 responded completely to all of the appropriate questions. These individuals comprise data set 1-1.

To enlarge the sample size, estimates were made for those automobile users who did not report bus travel-time information during the survey. Because a large fraction of the sample population lived in approximately 15 apartment complexes surrounding the university, the automobile and bus travel times reported by those individuals who resided in the largest of the apartment complexes were analyzed, and the results of this analysis were used as a guide in estimating omitted bus travel-time data. Then, the data on the individuals residing in apartment complexes, for whom it had been necessary to estimate missing responses to the bus travel-time question, were combined with the previous data set to form data set 1-2. The use of this procedure increased the sample size to 128 individuals.

For individuals not residing in apartment complexes, estimates of bus travel times were based on residential location and, to the extent possible, on data generated from nearby apartment complexes. The sample size of this third data set (i.e., data set 1-3) was 164.

These three data sets were used to evaluate the stability and statistical reliability of the model variables as a function of increasing sample size. The three data sets also were used to test the predicting capability of models, each drawn from the same population, but based on various levels of formulated data because the first data set was based entirely on the responses of the surveyed individuals, the second data set contained the best manual estimation of some sample points, and the third data set contained additional estimates of travel time values.

Model Calibration

The first specification of model variables was limited to the time and cost differences between automobile and bus. By using this model specification as a starting point, other combinations of variables were then used to test a number of models. The coefficients (estimated on the basis of maximum likelihood techniques), t-statistics, and levels of significance (LOS) for each variable are given in Table 2 for the three-variable model specification and in Table 3 for the seven-variable model specification. Other summary statistics are given in Table 4.

From a priori knowledge of travel behavior, it appears at first that all of the variables except sex have the correct sign. Past studies have usually indicated either that stratifications by sex cannot be used to differentiate modal preference (10) or that females, on a relative basis, select the bus mode more often than do males (5, 16). However, the sex-variable sign given in Table 3 indicates that males select the bus more frequently than do females. A closer examination of data set 1-3, for example, shows that males are split 58 percent automobile users and 42 percent bus users, and females are split 79 percent automobile users and 21 percent bus users. Thus, in the Amherst-University of Massachusetts area, males, on a relative basis, are about twice as likely to choose the bus mode than are females. (All females included in the sample must have responded affirmatively to the question on automobile availability.)

Statistical Reliability

An initial comparison of the magnitudes of the estimated coefficients for the three models in Table 3 shows a high degree of stability. A more exact way to examine both the stability and statistical reliability is to compare the relative magnitudes of the standard error terms for the various coefficients. As expected, this comparison (12) shows that the standard errors decrease and that all variables, except possibly the sex variable, approach a fairly stable condition as the sample size increases. As in other studies (3), travel time is the most stable coefficient.

From this analysis, it can be inferred that there will be relatively high standard errors if the sample size is too small, even if it is based entirely on the responses of individuals. But what can be said about the size of the sample necessary for calibration? Watson (16) has reported, for example, that 100 data records should be the absolute minimum required for calibration. Ben-Akiva and Richards (3) concluded that a desirable sample size is between 300 and 400 observations and that at least 600 observations are desirable if comparisons are to be made between two independent random samples. From the changes in standard error as a function of increasing sample size given here, it appears that most coefficients (except sex) appear to stabilize at sample sizes of 175 to 200 observations. [Additional details have been given by Parody (12).]

MODEL PREDICTION AND VALIDATION

Test 1

Possibly the best way to examine the efficacy of a model is to evaluate its ability to predict changes in travel behavior that occur as a result of actual changes in the transportation system. To make such an evaluation, all three data sets and the three- and seven-variable specifications of the models were used for the first before-and-after prediction analysis. The before case was represented by the period of the expanded free bus service and regular parking charges (i.e., survey period 2). The after case was represented by the time period following the introduction of significantly higher parking fees (i.e., survey period 3).

In the before period, a flat \$5/year parking fee was charged for all lots. In the next time period, a convenience fee based on the approximate location of the desired lot with respect to the center of the campus was charged in addition to the \$5 base fee. Three categories of convenience fees were established: (a) core lots at \$36.00/year (except \$50.00 for one lot adjacent to the administration building), (b) edge lots at \$12.00/year and (c) peripheral lots at \$0.00/year. (Shuttle bus service was provided to these outlying parking lots.) It was proposed originally that parking fees be increased up to \$75.00, with reserved parking spaces prices at \$125.00 but, because of strong, adverse, and vocal reactions by students and union employees, the rates of increase were scaled down.

From the revenues and capacity figures for each lot, it was determined that parking fees increased an average of \$21.00. If it is assumed that most spaces are purchased on a 9-month basis, this represents an average daily increase of 11 cents. Thus, for the after analysis, data were generated by adding 11 cents to each individual's automobile cost. The calibrated model coefficients, which were assumed to remain constant, were applied to the vector of forecasted variables for each traveler to generate new probabilities of modal selection. Finally, modal shares were developed by summing choice probabilities across all individuals in the data set. The use of this forecasting procedure makes the analysis free from aggregation bias.

Table 5 gives the actual and predicted modal shares and the percentages by which the predicted values differed from the actual ones for both the three- and seven-variable models for the three data sets. For example, the percentage error associated with data set 1-1 with the three-variable specification is (34.7 - 28.6)/28.6, or an overestimation in the change of automobile use of 21.5 percent.

Table 5 shows that the seven-variable model specification performs better in prediction than does the three-variable model. Similarly, models calibrated with data set 1-3 perform better than models calibrated with data sets 1-1 and 1-2 in terms of their accuracy in forecasting modal switching. This conclusion is consistent with the statistical stability results given above.

Test 2

The data sets for the first prediction test were compiled on the basis of surveys of travel behavior before and after the parking-fee increase. The second prediction test extended the time-series analysis by one more period to analyze, by using the disaggregate logit model, those changes in modal split that resulted from the in-

Table 1. Description of variables.

Variable	Description
Const	Modal constant (automobile = 1, bus = 0)
BusW	Walk time at origin to nearest bus stop (min); automobile walk time at origin assumed to be 0
FOS	Level-of-service variable representing frequency of bus service [low service frequency (headways of about 20 min) = 1, high service frequency (headways of about 10 min) = 0]
TimeDif	Total travel time by automobile minus total travel time by bus (min)
CostDif	Total travel cost by automobile minus total travel cost by bus (¢); in general terms can be expressed as A_1 × automobile fixed cost + A_2 × automobile operating cost + A_3 × (parking + toll) - transit fare[a]
Sex	Binary variable (male = -1, female = 0)
SEcon	Occupation status of traveler (graduate student = 1, undergraduate student = 2, nonprofessional staff = 3, professional staff = 4, faculty = 5)[b]

[a] In this application, A_1, is assumed to be 0, which is common to most analyses [except Wigner (16)]; operating cost is based on link trip distance × 3 ¢/km (5¢/mile) and A_2 is assumed to be 1; A_3 is also assumed to be 1 (5), but drops out in calibration because parking cost was 0 [Wigner (16) assigns A_3 a value of 0.5; limited empirical evidence (6) suggests that travelers are more responsive to a unit change in parking cost than to a unit change in operating cost; equivalent observation for in-vehicle and out-of-vehicle times is given by Kraft and Domencich (8)].
[b] Variable may also be regarded as a surrogate for income.

Table 2. Coefficients and levels of significance for three-variable model specification.

	Data Set 1-1			Data Set 1-2			Data Set 1-3		
Variable	Coefficient	t-Statistic	LOS	Coefficient	t-Statistic	LOS	Coefficient	t-Statistic	LOS
Const	-1.215 7	-2.246	0.03	-0.187 19	-0.469 67	NS	-0.007 252	-0.018 052	NS
TimeDif	-0.478 65	-4.875	0.01	-0.463 71	-5.365 3	0.01	-0.505 10	-6.019 7	0.01
CostDif	-0.060 25	-1.89	0.06	-0.063 08	-2.433 9	0.02	-0.070 68	-2.797 20	0.01

Note: N = 91, 128, and 164 for data sets 1-1, 1-2, and 1-3 respectively.

Table 3. Coefficients and levels of significance for seven-variable model specification.

	Data Set 1-1			Data Set 1-2			Data Set 1-3		
Variable	Coefficient	t-Statistic	LOS	Coefficient	t-Statistic	LOS	Coefficient	t-Statistic	LOS
Const	-4.169 7	-2.6475	0.01	-2.970 7	-2.6309	0.01	-2.606 5	-2.4905	0.02
TimeDif	-0.544 32	-3.9813	0.01	-0.521 06	-4.5335	0.01	-0.556 43	-4.9055	0.01
CostDif	-0.121 87	-2.1518	0.04	-0.126 93	-2.9958	0.01	-0.141 73	-3.4131	0.01
BusW	0.473 57	1.5034	0.14	0.550 86	2.1263	0.03	0.462 71	2.0422	0.05
Sex	0.921 48	1.1987	NS	0.788 41	1.3813	0.16	0.942 25	1.6875	0.10
FOS	2.096 3	1.932	0.06	1.820 2	2.337	0.02	1.958 7	2.586	0.01
SEcon	0.640 19	1.9725	0.06	0.571 03	2.5648	0.01	0.579 17	2.6290	0.01

Note: N = 91, 128, and 164 for data sets 1-1, 1-2, and 1-3 respectively.

Table 4. Summary statistics for three- and seven-variable model specifications.

	Three-Variable Model			Seven-Variable Model		
Variable	Data Set 1-1	Data Set 1-2	Data Set 1-3	Data Set 1-1	Data Set 1-2	Data Set 1-3
$-2\log\lambda$	53.1637	59.4995	97.3101	72.4601	87.7606	131.6842
LOS	0.01	0.01	0.01	0.01	0.01	0.01
$L^*(0)$	-63.076	-88.723	-113.676	-63.076	-88.723	-113.676
$L^*(\hat{\theta})$	-36.495	-58.973	-65.021	-26.841	-44.842	-47.834
ρ^2	0.42	0.34	0.43	0.57	0.49	0.58
$\bar{\rho}^2$	0.40	0.32	0.42	0.53	0.47	0.56

Note: N = 91, 128, and 164 for data sets 1-1, 1-2, and 1-3 respectively.

Table 5. Disaggregate predictions for the three and seven-variable model specifications.

				Three-Variable Model				Seven-Variable Model			
		Actual		Automobile		Bus		Automobile		Bus	
Data Set	Survey Period	Automobile (%)	Bus (%)	%	Difference[a] (%)	%	Difference[a] (%)	%	Difference[a] (%)	%	Difference[a] (%)
1-1	2	42.9	57.1	42.9	—	57.1	—	42.9	—	57.1	—
	3	28.6	71.4	34.7	21.5	65.3	8.6	31.5	10.40	68.5	4.2
1-2	2	57.8	42.2	57.8	—	42.2	—	57.8	—	42.2	—
	3	37.5	62.5	47.2	25.8	52.8	15.5	42.0	12.1	58.0	7.2
1-3	2	65.9	34.1	65.9	—	34.1	—	65.9	—	34.1	—
	3	48.8	51.2	55.6	13.9	44.4	13.2	50.9	4.2	49.1	4.0

[a] Based on actual modal shares.

Table 6. Coefficients and levels of significance for data set 2 and six-variable model specification.

Variable	Coefficient	t-Statistic	LOS
Const	-3.845 5	-3.0566	0.01
TimeDif	-0.578 48	-3.9406	0.01
CostDif	-0.150 97	-2.5985	0.01
BusW	0.676 72	2.7409	0.01
FOS	2.163 3	2.0849	0.05
SEcon	0.632 39	2.3275	0.03

Table 7. Disaggregate predictions for the six-variable model specification and data set 2.

	Actual		Predicted			
			Automobile		Bus	
Survey Period	Automobile (%)	Bus (%)	%	Difference[a] (%)	%	Difference[a] (%)
2	68.3	31.7	68.3	–	31.7	–
3	52.9	47.1	53.1	0.4	46.9	-0.4
4	48.3	51.7	46.0	-4.8	54.0	+4.5

[a] Based on actual modal shares.

crease in gasoline prices after the embargo of oil by the Organization of Petroleum Exporting Countries.

The before time period was identical to that used in the first prediction test. The two after cases were represented by the periods following increases in parking fees and in gasoline costs respectively. In selecting the sample for data set 2, the procedures that were used to establish the data sets in the first prediction tests were followed with the exception that individuals had to respond to at least surveys 2, 3, and 4 to be included in this new data set. With this procedure, a sample (containing both actual and estimated responses) of 104 individuals was selected for the second prediction analysis. About one-third of the individuals who had responded to surveys 2 and 3 could not be reached at the time of survey 4 or had made modal choices other than automobile or bus.

To keep the evaluation as independent as possible from the first prediction analysis, logit models were again calibrated for data set 2 [a cross-prediction test in which models were not recalibrated can be found in Parody (12)] using the same seven-variable model specification. For this model, all variables (except sex) were significant at appropriate confidence levels. Therefore, a model was calibrated without the sex variable; its coefficients are shown in Table 6. All six variables are significantly different from zero at the 0.05 level or better. Summary statistics for the entire model are given below.

Term	Value
$-2 \log \lambda$	86.972
LOS	0.01
$L^*(0)$	-72.087
$L^*(\hat{\theta})$	-28.602
ρ^2	0.60
$\bar{\rho}^2$	0.58

Forecasts were then made with this calibrated model for two different time periods. The first forecast, to the period represented by increased parking fees, followed the procedures outlined for the first prediction test. For forecasts into the second period, it was first necessary to compute the amount that gasoline prices (and thus automobile costs) had increased over the base time period. In the Amherst area, the pre-embargo price of gasoline was about 10.6 cents/L (40 cents/gal). During the embargo, it increased to approximately 14.5 cents/L (55 cents/gal), an increase of 37.5 percent. [A study of the embargo period in the Chicago area calculated that gasoline prices there increased more than 40 percent (13) and, in the New York-New Jersey region, gasoline prices increased 35 to 40 percent (11).] On the assumption that most drivers consider only the operating cost of traveling, the 3.1 cents/km (5 cents/mile) driving cost used in calibration was increased by 37.5 percent to 4.3 cents/km (6.9 cents/mile). This new operating cost, in addition to the increased parking cost, was used to determine the automobile traveling cost for the second prediction period. This new vector of data for each individual was used with the calibration coefficients to give a second set of new probabilities and their resulting modal splits. Table 7 gives the actual and forecasted modal shares and the percentage by which they differed from each other for the before and the two after time periods.

For the period representing increased parking fees, the predictions obtained from the logit model were exceptionally accurate. For the period representing increased gasoline prices, the model overpredicted the amount of switching away from the automobile mode by about 5 percent. Given the large changes that occurred in modal splits from the base time period, the errors associated with these predictions were considered to be very reasonable.

In terms of the way in which modal use changed over time, two points deserve note. First, as was also shown by the first prediction test, a fairly significant number of automobile users switched to the bus mode after the introduction of higher parking fees. Second, the higher gasoline and driving costs caused by the oil embargo caused only a minimal amount of modal switching. For example, for data set 2, automobile users were 68.3 percent of the sample population during the base period; this decreased to 52.9 percent after the parking fee was increased, and to 48.3 percent during the embargo.

Other studies have arrived at comparable results with regard to changes in travel behavior during the embargo. For example, a Northwestern University study showed that the increased price of gasoline had little effect on the work-trip traveling habits of individuals (13). A more important factor was the availability of gasoline.

ADDITIONAL ANALYSIS

Value of Time

For the model calibrated with data set 1-3, the implied price of time was $2.36/hour. The use of the identity given by Kendall and Stuart (8) for the ratio of two random variables gives a standard error of estimate of $1.05/hour. Consequently, the value reported is significantly different from zero at the 95 percent confidence level.

Elasticity

To make comparisons with other studies or to examine the sensitivity of model variables, it is useful to consider the elasticity of demand with respect to price and time.

Direct and cross elasticities of demand were computed based on the aggregate elasticity identities derived for logit models by using data set 1-3 [see Parody (12)]. The automobile time and cost elasticities given below compare favorably with elasticities reported elsewhere.

Aggregate Elasticity	Value
Direct	
Automobile cost	-0.2772
Automobile time	-0.8622
Bus time	-2.0679
Bus walk	-0.322
Cross	
Bus with respect to automobile time	1.6993
Bus with respect to automobile cost	0.5346
Automobile with respect to bus time	1.0722
Automobile with respect to bus walk	0.1670

For example, in a study by Charles River Associates (CRA) on free transit (9) the automobile-cost (line-haul) direct elasticity for work trips was found to be -0.49, and the automobile-cost elasticity determined here is -0.28; CRA found an automobile-in-vehicle-time elasticity of -0.82, which is very close to the -0.86-value given here. CRA also computed elasticities for in-vehicle and out-of-vehicle transit time. However, because the bus time used in this study is actually a door-to-door travel time, the components of time elasticities must be summed to make an equivalent comparison. Thus, their total transit-time elasticity is -1.1 and their transit access time is -0.71. In the present study, the equivalent values are -2.1 and -0.32. The larger difference in the transit elasticities may be due to the fact that, in the CRA study, all transit modes—commuter rail, subway, bus, and streetcar—were combined into a single modal classification, which reduced the accuracy of their estimated transit-demand relationships.

Several observations can be made from the elasticities given above. First, automobile users appear to be about three times more responsive to changes in automobile time than to changes in automobile cost (CRA cost-to-time ratio for automobile travel is about 4.5). With regard to the inelastic nature of the automobile-cost variable, it is evident that a significant increase in automobile cost had to occur (which was the case with the five-fold increase in parking cost) to observe such a large shift away from the automobile mode. Second, because travel appears to be very sensitive to changes in automobile time, it could be expected that there will be a large amount of mode shifting if the university phases out the center parking lots in an attempt to have a pedestrian campus by the process of introducing automobile-free zones.

Last, the demand for bus travel is quite elastic with respect to bus time, in marked contrast to bus walk time, which is highly inelastic. Initially one might expect the elasticity of bus walk time to be greater than that of total transit time, which includes in-vehicle time. Although the respective coefficients given in Tables 3 and 6 indicate that travelers weight a minute of bus walk time about the same as a minute of total transit time, the elasticity for bus walk time is less because the values of bus walk time are generally much smaller than the values of total transit time. Consequently, a 1 percent change in bus walk time is much less in absolute terms than a 1 percent change in total transit time. This observation has also been noted in a recent study by CRA (4). It appears, therefore, that a very productive way to attract passengers would be to improve service by increasing bus frequencies and better scheduling. Conversely, there will be substantial reductions in bus ridership if headways are increased or if schedules become unreliable.

CONCLUSIONS

By using a longitudinal data base of individual travel behavior that was a product of planned, phased changes in the transportation system, the accuracy of forecasts from disaggregate modal-choice models was evaluated. In terms of the actual setting from which the data were drawn, the disaggregate, behavioral models of modal choice were able to forecast future modal shares with reasonable and acceptable levels of accuracy. Only a relatively small sample of specially collected data was required to estimate the models. Because these results are only a single test in one setting of the forecasting capability of disaggregate models, additional research efforts are desirable and may be particularly appropriate in a more complex, urban environment.

ACKNOWLEDGMENTS

I would like to acknowledge the advice and guidance of Paul W. Shuldiner of the University of Massachusetts. Review and comments were also provided by Lee A. Webster of Curran Associates, William W. Boyer and Frank C. Kaminsky of the University of Massachusetts, and William B. Tye of Charles River Associates. I also thank Moshe Ben-Akiva and Thomas Adler of the Massachusetts Institute of Technology Department of Civil Engineering for furnishing a copy of the computer-estimation routines.

REFERENCES

1. M. E. Ben-Akiva. Alternative Travel-Behavior Structures. TRB, Transportation Research Record 526, 1974, pp. 26-41.
2. P. R. Stopher. Discussion of paper, Alternative Travel-Behavior Structures, by M. E. Ben-Akiva. TRB, Transportation Research Record 526, 1974, pp. 56-57.
3. M. E. Ben-Akiva and G. Richards. Disaggregate Multimodal Model for Work Trips in the Netherlands. TRB, Transportation Research Record 569, 1976, pp. 107-123.
4. Estimating the Effects of Urban Travel Policies. Charles River Associates, Cambridge, Mass.; Transportation Systems Center, April 1976.
5. Disaggregate Travel Demand Models. Charles River Associates, Cambridge, Mass.; NCHRP, Phase I Rept., Vol. 2, Feb. 1976.
6. C. Difiglio and M. F. Reed, Jr. Transit Sketch-Planning Procedures. TRB, Transportation Research Record 569, 1976, pp. 1-11.
7. D. W. Gillen. Parking Prices and Parking Restrictions: The Effects for Urban Mode Choice and Congestion. Proc., Transportation Research Forum, Vol. 16, No. 1, 1975.
8. M. G. Kendall and A. Stuart. The Advanced Theory of Statistics. C. Griffin and Co., London, 3rd Ed., Vol. 1, 1969, p. 232.
9. G. Kraft and T. Domencich. Free Transit. Charles River Associates, Cambridge, Mass., 1968.
10. C. A. Lave. Modal Choice in Urban Transportation: A Behavioral Approach. Stanford Univ., PhD Thesis, 1968.
11. E. J. Lessieu and A. Karvasardky. Recent Observations on the Effect of Gasoline Price and Supply. TRB, Special Rept., 153, 1975.
12. T. E. Parody. A Disaggregate Prediction Analysis of Mode Choice. Department of Civil Engineering, Univ. of Massachusetts, PhD dissertation, 1976.
13. R. L. Peskin, J. L. Schofer, and P. R. Stopher. The Immediate Impact of Gasoline Shortages on Urban Travel Behavior. U.S. Department of Transportation, April 1975.
14. R. H. Pratt. Demand Forecasting for Short-Range

and Low-Capital Options. HRB, Special Rept., 143, 1973.
15. P. L. Watson. Predictions of Intercity Modal Choice From Disaggregate, Behavioral, Stochastic Models. HRB, Highway Research Record 446, 1973, pp. 28-35.
16. P. L. Watson. The Value of Time. Behavioral Models of Modal Choice. Lexington Books, 1974.
17. M. F. Wigner. Disaggregated Modal-Choice Models of Downtown Trips in the Chicago Region. HRB, Highway Research Record 446, 1973, pp. 49-65.

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

**Dr. Parody was a student at the University of Massachusetts, Amherst, when this research was performed.*

Adaptable-Zone Transportation-Assignment Package

Carter Brown, Tri-State Regional Planning Commission
Yehuda Gur, John Hamburg and Associates, Inc., Chicago

The adaptable-zone transportation-assignment package is a computer software system of traffic assignment for subregional analysis. The development of the package was completed in early 1976; it is currently being used by the Tri-State Regional Planning Commission. It combines a windowing operation that creates a zonal structure and network focused on the area of interest with relatively conventional traffic-distribution techniques in an integrated software package. These features make it possible to analyze subregional plans and projects faster and at a lower cost than other operational planning packages.

An increasingly large portion of the highway planning activities of the Tri-State Regional Planning Commission (TSRPC) deals with the design and evaluation of subregional plans or individual projects within the New York metropolitan area. To perform subregional planning, it is necessary to have tools that permit detailed presentation and analysis of an area of interest. At the same time, it is necessary to consider the regional effects of the area of interest and the relation of the particular project to the regional plan. Other important requirements of subregional planning tools are understandable techniques that allow cooperative planning with subregional agencies, reasonable costs, and timely responses.

The need for efficient subregional planning tools is common to most public planning agencies, but this requirement is more pronounced in the case of TSRPC, because of the size of the metropolitan New York region and the many subregional agencies that require technical support. The TSRPC region has an area of about 23 000 km (9000 miles^2) and a population of more than 18 million. The standard data files include about 7000 zones, which range in size from 0.65 to 41 km^2 (0.25 to 16 miles^2) and about 16 000 links (excluding centroid connectors). Most of the available highway-assignment programs cannot handle a problem of this size or can do it only at a prohibitively high cost. Thus, subregional analysis tools are not only desirable, but are absolutely necessary for TSRPC.

Because of the lack of tools having thre required capabilities, TSRPC began in mid-1974 to develop them. The project, which was completed in early 1976, resulted in the adaptable-zone transportation-assignment package (AZTAP). This paper gives an overall description of the package. A more detailed description of AZTAP has been given by John Hamburg and Associates (1).

MAJOR FEATURES OF AZTAP

AZTAP is structured along the lines of the conventional urban transportation planning process for highway planning process for highway planning. It includes the stages of trip distribution, assignment, and assignment summaries. Within this framework, it also includes special features that answer the needs of subregional analysis.

Windowing

AZTAP provides for windowing of subareas. There are two aspects of windowing: areal aggregation and network culling. Within a user-specified area of interest, the analysis uses all of the available data, including zones that might be as small as 0.65 km^2 (0.25 mile^2) and all the coded network. As the distance from the area of interest increases, the level of detail given for other regions diminishes gradually, reaching zones as large as 662 km^2 (256 miles^2) and networks that are only a skeleton of major regional arteries. The construction of the windowed data files is done by computer, based on structured data files that include detailed information about the entire region and a user-specified window configuration.

Structured Input Data

Both the network and the zonal data have built-in hierarchical structures, with all links assigned to one of three network levels and all zones classified by one of four sizes. Both sets of data also share a common geographical reference using a rectilinear coordinate system. These two features greatly simplify the processing required to relate the networks and zones to the specified window configuration.

Streamlined Application

The package was designed to permit streamlined, fast-response applications to specific study areas with a minimal need for manual intervention in the process. This objective has been achieved through a number of built-in features, including the selection of models and software capabilities. The models used in the package are vigorous in their performance and relatively easy to calibrate.

Figure 1. Adaptable-zone transportation-assignment package program sequence.

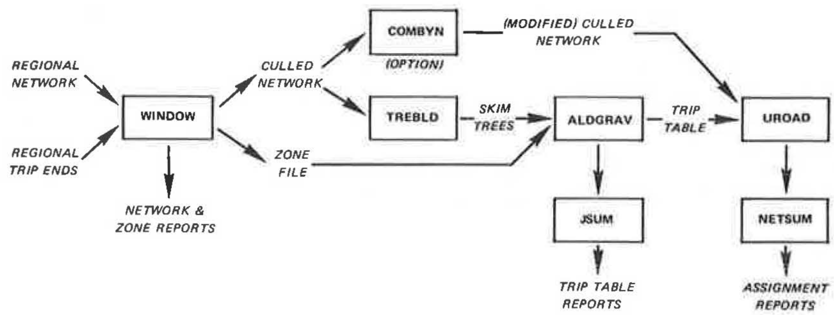
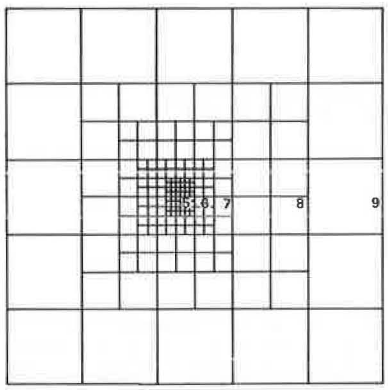


Figure 2. Typical hierarchical zonal structure.



Streamlined applications are also facilitated by special software, including a package driving program (DRIVER), job-control procedures, and prespecified job programs. Many tasks normally done manually are carried out during processing. The most important is the creation of approach links that is performed in the WINDOW program. This capability is the key element in the fast creation of zones and networks tailored for specific projects. Intermediate evaluation reporting is also performed within the package.

Compatability With Other Systems

Built-in procedures permit the loading of relevant data on the Interactive Transportation Analysis System (INTRANS) (2) data files as an integral part of the analysis process. Most of the data files in the package are also compatible with the U.S. Department of Transportation Urban Transportation Planning System (UTPS) (3), which permits the use of various programs in UTPS for a multitude of auxiliary operations.

AZTAP DATA INPUTS

The input data used by AZTAP consists of the TSRPC highway network and regional, vehicle trip ends. Both the trip-ends data file and the network file conform to the rectilinear grid system used by TSRPC. Other than that, the two files can be manipulated independently; the common coordinate system provides a geographical interface during operation of the model.

Trip ends are input to AZTAP as uniform grid cells of 0.65, 2.5, 10, and 41 km^2 (0.25, 1, 4, and 16 miles^2) as delineated by the coordinate system. Current-year trips are based on 1963 home-interview surveys, 1970 census data, and annual state vehicle-registration data. Forecast trips are created by applying trip-generation rates developed from the above data to land-use demographic forecasts.

The network file contains 15 771 links representing 27 056 route km (16 285 route miles). The network includes all freeways, principal and minor arterials within the region, freeways and major highways outside the region, and collectors or selected local streets in higher density areas within the region. The future highway network also includes all proposed roadways in the regional transportation plan (4). The network uses a link-name identification system with each node identified by the rectangular grid coordinate.

The highway network is coded in a three-level hierarchy. The highest level (level 1) includes all freeways and major highways and any additional links required to ensure network continuity. Level 2 includes links that generally bisect the areas bounded by level 1 links; levels 1 and 2 links form a fully connected network. All remaining links are classified as level 3.

AZTAP OPERATIONS

AZTAP consists of 20 programs of varying importance and complexity. The basic interrelationships among the most important of these programs WINDOW, TREBLD, ALDGRAV, COMBYN, and UROAD) is shown in Figure 1: The first main program is WINDOW, which creates a zonal file and a modified network file with approach links. TREBLD generates a zone-to-zone friction table. ALDGRAV creates a trip table that uses the friction table and the zonal file. COMBYN is used to modify the network files so that more than one network can be used to reflect various improvement alternatives. UROAD then carries out a conventional traffic assignment by using networks from either WINDOW or COMBYN and the ALDGRAV trip table.

WINDOW: ZONAL CONFIGURATION

Because of the size of the TSRPC area [the base zonal trip file covers 40 122 km^2 (15 491 miles^2)], there is always a trade-off in assignment work between the need for detailed networks and small zones and computer costs and core-storage requirements. WINDOW has been created to deal with this inherent conflict.

WINDOW performs three individual tasks: areal aggregation, network culling, and construction of approach links that connect the zones to the network.

Areal Aggregation

The areal aggregation is done by the program and based on the user-specified window configuration. Through a set of control cards, the desired minimum zone size is specified in each part of the region. Those sizes range from 0.65 to 2.5, 10, 41, 166 and 662 km^2 [0.25 to 1, 4, 16, 64 and 256 (16 by 16) miles^2]. The program structures the analysis zones according to the user's requests. In particular, it aggregates the zones from the input trip file into larger analysis zones. A typical zonal configuration is shown in Figure 2, where the zones have the areas given below (1 km^2 = 0.39 $mile^2$).

Type	Size (km^2)	No. of Zones	Type	Size (km^2)	No. of Zones
5	2.5	48	8	166	24
6	10	52	9	662	16
7	41	32			

The aggregation process has a number of limitations. First, the program never increases the level of detail in the input trip file; thus, if an area is described in the input file on a level of 10 km^2 (4 $miles^2$), it remains this way even if the request in WINDOW is for smaller subdivisions. The second limitation is that the zones (both those in the input file and those created by WINDOW) must always be perfect squares that conform to the grid of the coordinate. To do this, the zones must always be nested within the larger size zone, with smaller zones occupying at least the area of the next larger size. Consequently, the analysis zones do not necessarily coordinate with political or geographical boundaries. (These limitations are implemented automatically by the program without undue complications in the method by which the window is specified by the user.)

The flexibility and convenience in specifying the zonal structure gives the package tremendous power. The structure is uniquely specified for each problem and tailored for the analysis needs. The zones are typically structured in the shape of concentric rings, with small zones [0.65 and 2.5 km^2 (0.25 and 1 $mile^2$)] in the immediate study area and progressively larger zones as the distance from the area of interest increases; wherever appropriate, the ring structure is adapted to unique conditions.

Network Culling

The culling of the network is done by WINDOW and based on user-specified rules. Through a set of control cards, the user specifies the lowest level network to be included in each part of the region. The program reads the entire network file and builds a new file that includes only those links that pass the selection test. During this process, the link records are restructured to fit the needs of the network-analysis programs. In particular, node numbers are assigned to the ends of the links to augment the link-based identification scheme used in the TSRPC network.

Typically, the network culling program follows the zonal aggregation; e.g., wherever large zones are requested, only level 1 links are included, and vice versa.

Building Approach Links

Approach links are built by WINDOW after the stages of zonal aggregation and network culling to ensure that connections are made to nodes that exist in the network and that the connections are compatible with the zone size [e.g., 0.65 km^2 (0.25 $mile^2$) zones are not connected to expressways].

A search routine checks each quadrant of each zone for network nodes. The best nodes are selected to serve as end points for the approach links. At least one approach link is generated for each zone, but as many as four links may be generated. Approach links are added to the network file to create a culled network that is structured as a conventional file for analysis. These generated approach links become the interface between the zonal information and the network, which are independent before the generation of the approach links. This network and the zonal file comprise the major output of WINDOW and are the major inputs to the subsequent steps in the process.

ALDGRAV: TRIP DISTRIBUTION

The ALDGRAV program generates trip-interchange tables for the zones created in the WINDOW program. This trip-distribution model was developed by Schneider (5) and was refined in a combined project by the Chicago Area Transportation Study (CATS) and Creighton, Hamburg and Associates (6). The ALDGRAV function is similar to the gravity model in that the number of trips between two zones (V_{ij}) is a function of the number of trips in the origin zone (V_i), the attraction of the destination zone ($A_j \times B_j$) relative to all other zones, and a travel function between the two zones (G_{ij}):

$$V_{ij} = V_i \times A_j \times B_j \times G_{ij} \Big/ \sum_k A_k \times B_k \times G_{ik} \qquad (1)$$

ALDGRAV differs from the conventional gravity model in that the travel function (G_{ij}) is a Bessel function. The travel function is calculated as $G_{ij} = GB(BTR \times F_{ij})$, where F_{ij} is the travel friction between zones i and j as measured in the TREBLD program, GB is a Bessel function, and BTR is a parameter of the travel function. An attractive feature of the Bessel function is that the shape of the trip-length distribution that it produces fits the observed data well.

The ALDGRAV program permits convenient and effective control of the key measures of the trip-distribution table. By the use of the BTR parameter of the Bessel function and the parameters that control the calculations of the intrazonal travel function, it is possible to easily adjust the average trip length and the total vehicle kilometers of travel (VKT) assigned. Correct values for these key parameters and a correct trip-length distribution ensure satisfactory assignment results. The close control on these factors obviates the need for cumbersome calibration procedures that try to duplicate observed interchange volumes.

This observation of the relative insensitivity of the assignment to certain inaccuracies in the trip table agrees with the results of Miller and Nihan (7) and Stover and others (8). Of particular relevance here is the work by Stover, which shows that good assignment results may be obtained as long as the trip table produces the correct VKT and the correct numbers of trip productions and attractions are preserved.

UROAD: TRAFFIC ASSIGNMENT

The UTPS program UROAD (6) is used for traffic assignment. To use the program as part of AZTAP required additional programming through user-coded subroutines. A new network-input procedure was implemented to allow direct use of the WINDOW network file. With this procedure, it is possible for the network to contain as many legs from each node as is desired. This freedom is particularly important in conjunction with the automatic-approach link assignment. In addition, externally calculated capacity and zero-volume speeds for each link are read from the network file, rather than being estimated by the program, which permits the use of off-line procedures for the calculation of these variables. The network-output procedure has also been modified to permit the use of the assignment results with a number of evaluation and summary programs, which are produced in addition to the standard UROAD reports.

The major technique used for assignments within AZTAP is the CATS option of UROAD. In this form of restrained assignment, the process begins with zero-volume speeds assigned to all links. Zones are loaded in batches, and link impedances are updated after the loading of each batch. During the updates, delay penalties are added to the input impedance of each link. These penalties are computed based on the volume-to-capacity ratios that result from the sequential batch loading. Extensive modifications have been made to the default volume-delay curves of UROAD. The modified curves are different for arterials and for expressways and are based on the results of the West Side Highway study (9).

Figure 3. Approach nodes, contrasting two types of zone.

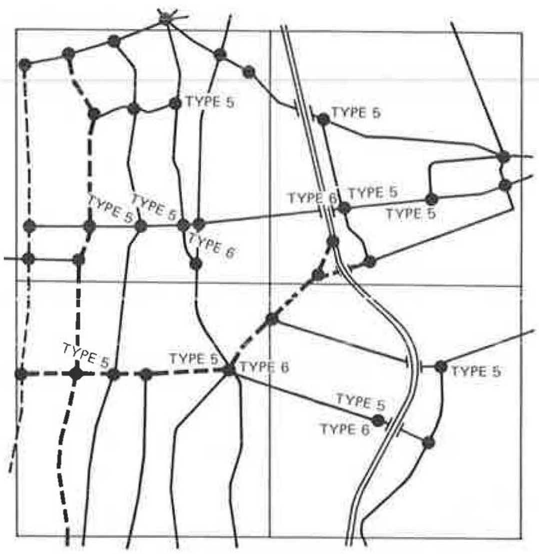
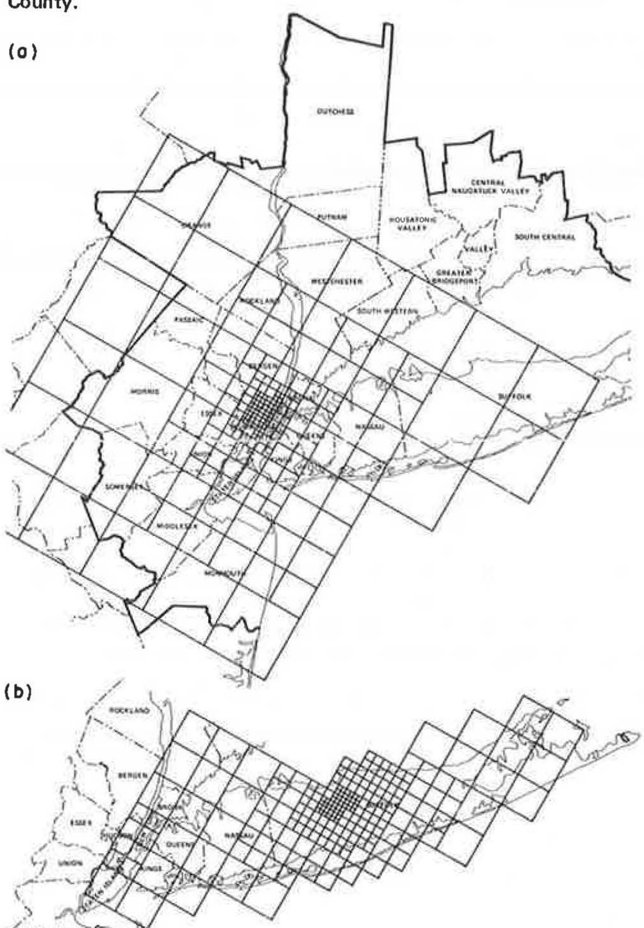


Figure 4. Zonal configurations: (a) Hudson County and (b) Suffolk County.



TESTING, APPLICATION, AND OPERATION

The primary impetus for the development of AZTAP was the need to carry out traffic assignments quickly and easily in many locations throughout the New York metropolitan area. The package was tested by applications in the planning environment by applying it to ongoing traffic studies as early in the development process as was possible. In parallel, testing was carried out to prove the mechanical performance of the software and to establish the proper parameters for ALDGRAV and the impact of the variable zonal structure.

The ALDGRAV program was originally developed as a multimodal trip-distribution model. However, in AZTAP it is used for vehicular trips only. Consequently, it was necessary to establish the proper range for the travel-function parameters by using cost values appropriate for current conditions. The initial testing established that BTR values of between 0.014 and 0.020 $cents^{-1}$ with a cost of time of 6.4 cents/min gave acceptable trip distributions. Variations in zonal structure showed that the program is relatively insensitive to zone size. For example, one zone of 10 km^2 (4 $miles^2$) generated the same VKT as four zones of 2.5 km^2 (1 $mile^2$) in the same location (discounting trips among the four zones). Variations in zone sizes did cause aberrations in the trip-length distribution, but the variations coincided with the network and approach links and were dissipated in the assignment stage.

The WINDOW structure is crucial for the assignment. As in conventional assignments, the zonal structure affects loading. Unlike conventional assignments, however, the structure can be varied to match study needs. Figure 3 shows the approach nodes for one location when two different zonal structures are used.

In the first case, the area was subdivided into four 2.5-km^2 (1-$mile^2$) zones (type 5) with 10 approach links; in the second case, it was described as one 10-km^2 (4-$mile^2$) zone (type 6) with four approach links. To obtain good assignment results, it was necessary to match the zonal structure in and around the area of interest (and the resulting number of approach links) to the density of the coded network.

Symmetry in the zonal configuration around the area of interest was more important than coverage for the sake of consistency with political boundaries. When zones were added in an unbalanced manner to include all of the region, there was very little change in the assigned volumes, but there was a substantial increase in the size of the problem. On the other hand, lack of symmetry in the zonal configuration caused reduced travel toward the thin side. However, reliable assignments could be made with 200 or fewer zones.

AZTAP has been used in a wide range of study conditions including high-density areas in the center of the region (Hudson County, New Jersey) and low-density areas in the outer suburbs (Suffolk County, Long Island). In these varying conditions, AZTAP has proved to be very satisfactory. Zonal configurations for the two studies are shown in Figure 4. There are certain differences between the configurations: The Suffolk County configuration does not include the entire region because it is far from the core, but the Hudson County configuration covers a larger area because of its central location. The hierarchial pattern is broken for Manhattan in the Suffolk zonal structure, to account for effects of the unique network structure and the high densities.

Comparisons to detailed ground-count data were used as the major means to assess assignment quality. For most studies (with Manhattan the exception), the assigned VKT for the area of interest was within 2 percent of ground-count VKT. It was necessary to adjust the trip-distribution parameters for different sectors of the region. Studies in suburban New York have required longer trip lengths (BTR = 0.014 cents^{-1}), and studies in northern New Jersey have required shorter trip lengths (BTR = 0.020 cents^{-1}). It has not yet been determined whether the need for adjustment is due to real differences in travel behavior or whether it is due to model-specification errors. Such errors might make the model unduly sensitive to variations in trip-end densities and measures of separation.

Although the ALDGRAV and WINDOW programs have proven to be easily adaptable to various locations and conditions, the results of the capacity-restrained UROAD assignments have not been uniformly succesful. Because aggregate measures of assignment accuracy, such as screen lines, cordons, and cutlines, have been quite good (many within 3 percent of the ground counts), it appears that trip tables that are generated by the process are acceptable. However, in urban areas with close roadway spacing, the assignment accuracy suffers (root-mean-square = 45.5 percent, for example). Apparently, there is no easy substitute for meticulous network coding in urban areas where variables, such as signal timing and roadside characteristics, affect the impedance more than do physical roadway characteristics. Further improvements in volume-delay curves are also needed.

Operationally, AZTAP has proven to be quite efficient. A full package run for a new study area can be programmed in 1 working d. More frequently, jobs are run in a series of steps that takes advantage of the modular nature of package plans. A typical run sequence for a study of future-year alternatives and representative computer run times are shown below (central processing unit time on IBM 158; problem size of 200 zones and 12 000 links).

Operation	Programs Used	Time (min)
Current year initial run	WINDOW, TREBLD, ALDGRAV, COMBYN, UROAD	20
Trip-distribution adjustment	ALDGRAV, UROAD	11
Network adjustment	COMBYN, UROAD	9
Future no build	WINDOW, TREBLD, ALDGRAV, COMBYN, UROAD	20
Future alternative	COMBYN, UROAD	9

DIRECTIONS FOR FURTHER WORK

A comprehensive system for subregional analysis is clearly needed. The adaptable zone package is only one element in the set of tools: The most obvious others are those for summaries and evaluation, for multimodal planning, and for analysis of low-capital, small to medium-scale projects.

The need for fast summary and evaluation of assignment results is most pressing. The use of interactive graphics (INTRANS) for assignment evaluation is promising, but its main role to date has been to supplement standard (mostly manual) techniques. There is a need for a significant research and development effort to learn how to fully exploit the power of interactive graphics in assignment evaluation. The effective use of this powerful tool could reduce the time required for analysis and evaluation by orders of magnitude.

Traffic assignments are the basic input data for plan-evaluation work and environmental studies. Efficient development of these postassignment tasks requires additional software to convert link volumes and speeds to useful evaluation measures. The broader evaluation issues, such as estimates of social and environmental impacts and economic analysis, must be brought into the systems-evaluation procedure.

At present, AZTAP serves only highway planning. There is no comparable tool that performs windowing on transit networks. Analyses that are based on the flexible zonal structure of AZTAP are incompatible with transit analyses that use fixed zonal structures and network descriptions. At present, this precludes multimodal, subregional analysis. Thus, a logical immediate step is to expand AZTAP to include the processing of transit networks (10) or to implement a similar tool for transit analysis.

The increased importance of small and medium-scale projects in transportation system development makes increased detail and accuracy of assignments necessary. AZTAP is the first step in this direction, in that it permits fast, inexpensive, and detailed analysis of subareas. A natural extension of AZTAP are facilities that will permit an increase in the level of detail within the area of interest, along the lines of micro assignment. There is a high probability that such a tool could support comprehensive and meaningful evaluation of such projects as improvements in traffic-control strategies and widening of major streets.

ACKNOWLEDGMENTS

This report describes a project that was performed under the sponsorship of and in cooperation with Tri-State Regional Planning Commission. It was financed in part by the Federal Highway Administration and the Urban Mass Transportation Administration, by a comprehensive planning grant from the U.S. Department of Housing and Urban Development, and in cooperation with the states of Connecticut, New Jersey, and New York. The findings and opinions expressed herein are not necessarily those of the sponsoring agencies.

REFERENCES

1. Tri-State Highway Planning Package: User Manual. John Hamburg and Associates, Bethesda, Md.; Tri-State Regional Planning Commission, New York, March 1976.
2. Y. Gur. Intractive Graphics as a Tool in Plan Evaluation. TRB, Transportation Research Record 491, 1974, pp. 60-68.
3. Urban Transportation Planning System Reference Manual. U.S. Department of Transportation, June 1976.
4. Maintaining Mobility. Tri-State Regional Planning Commission, New York, Nov. 1976.
5. M. Schneider. Access and Land Development. HRB, Special Rept. 97, 1968, pp. 164-177.
6. The Access and Land Development Model: The Calibration of the ALD Travel Function. Chicago Area Transportation Study, Technical Rept., Vol. 2, 1975.
7. D. G. Miller and N. Nihan. The Subarea Focusing Concept for Trip Distribution in the Puget Sound Region, TRB, Transportation Research Record 610, 1976, pp. 37-43.
8. V. G. Stover, J. Benson, and J. Buechler. Sen-

sitivity Evaluation of the Traffic-Assignment Process. TRB, Transportation Research Record 610, 1976, pp. 50-52.
9. A Description of the West Side Highway Project Traffic-Assignment Process. Creighton, Hamburg, Delmar, N. Y.; July 1974.
10. Y. Gur. Directions for the Development of Transit Analysis Capabilities for TSRPC. Creighton, Hamburg, Delmar, N. Y.; Tri-State Regional Planning Commission, July 1975.

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

Abridgment

Propensities to Ship Manufactures by Rail Within Four U.S. Traffic Flows

Thomas A. Stone, Northampton, Massachusetts

The use of a freight transportation service at two levels of geographic aggregation, represented by nonoverlapping component traffic flows and the aggregate that combines them, is modeled. The basic aim is estimators of demand parameters. Propensity is used to denote the nonrandom part of the actual percentage of goods shipped via the specified service. Under certain conditions, propensity statistics that remove impacts of shipment distance and weight are a means of estimating impacts of demand determinants that are assumed to vary systematically among component flows (service quality might be one such determinant).

The general approach is by three hypotheses: The first states that the expected value of use in the aggregate is a weighted average of expected use in the components. The second states that expected use in each component equals expected aggregate use plus a constant. The third states that all of the random variables describing service use have the same form, whether aggregate or component. This form varies from one version of the model to another. The simplest model is extended by explicit inclusion of variables for relative price and other determinants that are assumed to be aggregate in the sense that they influence demand in the same way everywhere. The slopes of the aggregate determinants and the geographic propensity statistics can then be estimated simultaneously.

An illustrative study of rail shipments by manufacturers uses two-digit data from the Public Use Computer Tapes, 1967 Census of Transportation, in four component traffic flows—east to east, east to south and west, south and west to east, and south and west to south and west. No aggregate determinants are included. Differences among estimated propensity statistics are often small and of mixed sign. However, greater commodity detail and the inclusion of at least a relative price are required for a conclusive empirical study.

Publication of this paper sponsored by Committee on Passenger and Freight Transportation Characteristics.

Approach to Measurement of Modal Advantage

Ronald D. Roth, Department of Economics, American Trucking Associations, Inc.

Shipper preferences for truck or rail were determined by using shipment mass and distance as criteria. Data on individual manufactured commodities from the 1972 Census of Transportation Commodity Transportation Survey were arranged in a matrix-table format and analyzed for the extent of involvement by these primary modes in carrying cargo of various weight brackets moving over a series of distance blocks. As expected, trucks dominated the movement of lighter weight shipments and rails dominated the movement of heavy shipments. Competition, or involvement by both modes, was limited to cargo in the medium-weight range, which was not an extensive amount. Amounts of cargo either modally dominated or competitive were determined and correlated to actual overall percentages of cargo by mode. Changes in overall modal percentages of cargo over time have a strong relationship to changes in the size of shipments over time, although not as strong and direct a relationship for the rail mode as for trucking. Major shifts of cargo between modes will not occur in the absence of artificial or arbitrary obstacles to market forces.

The national transportation policy of the United States has long recognized the importance of a balanced transportation system to satisy the distribution demands of manufacturers and the consumption demands of the public. The system that the policy promotes and maintains generally provides alternative means of transport, available on demand, to all shippers. The rationale of this system is that each mode possesses inherent or relative advantages that allow it to handle some portion of the total freight traffic at the least possible cost to the economy. In basic terms, it recognizes that trucks can handle short hauls of limited quantities at a cost that could not be equaled by any other mode, but that for long hauls of large quantities, rails are generally the lower cost mode, although water carriers also provide low-cost bulk movements for shippers located along waterways. Air carriers provide transport services for products that are extremely time-elastic.

While few would disagree with the basic premise that there are relative advantages in transportation, an effective and responsive policy must recognize where the advantage of one mode ends and that of the other begins. That is, how is the freight traffic that is represented by the gray area between short hauls of limited quantities and long hauls of large quantities allocated between the modes? This analysis of this question is based on actual shipper preferences, i.e., shippers' perceptions of relative modal advantages, as they existed in 1972.

PURPOSE

The 1972 Census of Transportation provides data on the movement of manufactured commodities by mode of transportation within specific weight and distance brackets. These source data, when arranged in a matrix format, show discernible and consistent trends in modal use that allow one to address a number of pertinent questions concerning transportation:

1. What are the limits of the relative advantages of truck and rail most suited to satisfying the demands of shippers, based on distances traveled and weight of shipments carried? For what shipments do truck and rail compete?
2. What is the extent of competition between truck and rail for manufactured freight: (a) relative to the total amount of manufactured freight and (b) relative to the amount of manufactured freight suited to movement by each mode?
3. What is the potential amount of manufactured freight that could be transferred from one mode to the other?

SOURCE

The principal source of the data used in this study is the Commodity Transportation Survey, one of the reports generated from the 1972 Census of Transportation by the Census Bureau. The survey consists of 15 reports, each dealing with commodity-movement data grouped by major manufacturing industries. As examples, one report deals with food and kindred products and tobacco products, and another deals with lumber and wood products. Detailed commodity-movement data for the various specific commodities that comprise the major industry groupings are shown within these reports; i.e., the report that deals with food and kindred products and tobacco products details data separately for such categories as meat products, dairy products, canned foods, and cigarettes.

For each specific product, such as meats, the survey presents four tables. Tables 1 through 3 give origin and destination data, and Table 4 gives data relating to the weight and distance of shipments. The specific commodity data used here are taken from Table 4A of the survey, which is entitled Percent Distribution of Distance Shipped and Weight of Shipment, by Means of Transport, 1972: Tons of Shipment. It shows the total amount of cargo of the individual commodities shipped and the percentages of this cargo handled by each mode. The table is further detailed into seven distance brackets and, within each of these distance brackets, into six weight-of-shipment categories. The amount of cargo within each of these specific-weight-and-distance categories is distributed by percentage according to type of transport. The survey also gives a Table 4B for each commodity that measures megagram kilometers rather than megagrams. However, because the data of Table 4A were already cast in terms of weight-and-distance categories, these data were disregarded as they could be of only limited additional value.

Limitations Within Source Data

The Commodity Transportation Survey was designed to measure nonlocal shipments by large manufacturing concerns (those with 20 or more employees) of manufactured commodities only. The data in the survey in essence measure regular flows of intercity manufactured freight. Firms with 10 to 19 employees were surveyed by the Census Bureau, but these data could not be incorporated into this study because of the lack of comparable statistics in the weight and distance format. Firms with fewer than 10 employees were not surveyed.

Limitations Placed on Source Data

Only truck and rail will be considered in this analysis because the majority of freight transportation of manufactured goods involves movements by these two modes. For this purpose, truck movements handled by for-hire trucking companies (shown by the survey as motor carrier) were combined with those by private carriers (shown as private truck) to obtain a truck total. This combination involves some imperfection because these respective trucking groups, although generally having different characteristics in terms of length of haul and size of shipment, compete with each other to a degree. However, because both are substitutes for (or competitors with) rail service, their combination is consistent with the basic purpose of this analysis, which is to evaluate the relative advantages of truck and rail.

Air carriers, water carriers, and other and unknown, which were also covered in the Commodity Transportation Survey, will not be considered as to competitive impact. The exclusion of water carriers in this analysis required that Transportation Commodity Code (TCC) 29 (coal and petroleum products), one of the major industry categories, be disregarded. This group of products is handled predominately by water carriers (the Census Bureau does not measure pipeline movements) and, in fact, provided the overwhelming majority of total water-carrier cargo in all manufactured freight. This is the only group of products in which any mode other than truck and rail handle a significant share of the cargo. Another product group—TCC 27 (printed matter)—was also excluded because the data on this product were collected on the basis of value of shipments, rather than cargo amount.

The 13 remaining survey reports contain data on 89 separate products. These products represent 18 major manufacturing divisions. The total amount of cargo accounted for by the weight-and-distance data on these 89 products is 827 839 051 Mg (912 521 000 tons). Of this aggregate amount, the railroads carried 38.1 percent and trucks carried 58.1 percent (for-hire motor carriers, 36.2 percent and private trucks, 21.9 percent). Water carriers, air carriers, and other excluded modes carried only 3.8 percent of the total amount so that their exclusion does not materially affect the results of this study.

ANALYSIS

Limitations

The results presented in this paper add, surprisingly enough, a new dimension to traditional modal-share analysis, a view of a division of cargo to mode by shipments of a specific weight and distance. Older analyses have tended to rely on either

1. The overall division of the cargo of a product to mode (without considering weight or distance brackets),
2. The division of the cargo of a product to mode by distance bracket, or
3. The division of the cargo of a product to mode by weight bracket,

However, none have presented a combined picture of weight and distance. These older analyses have served to identify general trends in modal use, but their results are hardly definitive.

In this paper, only aggregate data are presented, i.e., combined data for the 89 commodities. Presenting results on an aggregate basis has its limitations as individual product peculiarities and small trends are lost in the processes of aggregating and averaging. Yet, because of the necessity to analyze the modal split of each of the 89 individual commodities to determine the aggregate figures, this analysis provides a methodology to recognize products that have peculiar shipping patterns.

Another limitation of this analysis concerns the level of detail presented. The product detail shown in the survey is only at the three-digit TCC level (equivalent to the standard industrial classification). Because the statistics were collected at as fine a level as seven-digit codes but summarized at the three-digit level, there is potentially an aggregation problem inherent in them. To the extent that the three-digit statistics fail to accurately reflect the finer level of data, this analysis could be misleading.

In addition, this analysis investigated cargo allocations only in specific weight-and-distance brackets, which are only two of the many factors affecting a shipper's choice of mode. Not considered independently were such factors as time-in-transit differentials between modes, loss and damage experiences, reliability of service, availability of specialized equipment, product density, product value, locational concentration of production or consumption, and many others that affect a shipper's choice of mode. However, if it is assumed that shippers have made a rational choice of mode by taking such factors into consideration, then these omissions have little, if any, impact on this particular analysis.

The inherent limitations of analyses based only on physical amounts of cargo (weight) should also be recognized. The amount of cargo handled by any mode should not be construed to be a measure of the relative economic position of that mode, nor should it be used as a proxy for the economic value that the cargo handled provides to the economy. Although no direct revenue comparison could be made between modes on the basis of manufactured freight only, the railroads carried twice as much total cargo (bulk, manufactured, scrap, and miscellaneous commodities) as did Interstate Commerce Commission regulated trucks in 1972—1389 million Mg (1531 million tons) versus 699 million Mg (771 million tons) respectively—but generated only about 70 percent of the total gross freight revenue ($12.8 billion versus $18.5 billion respectively). In terms of shipment values in manufactured commodities, the 38.1 percent of the manufactured cargo carried by rail accounted for only 29.9 percent of the total value of manufactured shipments, and the 58.1 percent of cargo carried by trucks accounted for 67.5 percent of the total value. Hence, amounts of cargo are a less than perfect measure of the economic roles of the modes of transportation. However, aside from this and the other previously mentioned limitations, this analysis will serve to identify some areas where the transportation roles of the modes tend to be uncertain (overlap) and, hence, where policy changes might have some impact.

Format

The data contained in Table 4A of the survey, while presented in a rather cumbersome tabular form, can be clearly and usefully displayed in matrix format. Figure 1 gives a partial example using the data of TCC 321 (Flat Glass).

This percentage matrix is one of two matrices constructed for each of the 89 products. (Matrix 1, which is not shown, contains the actual cargo amounts). The top left cell of the example matrix, indicated by TOTAL, shows the overall modal percentage distribution of the cargo of the product (represented as 100 percent) to rail, truck (motor carrier and private truck), and other modes. Horizontally displayed from the TOTAL cell are the percentages of total cargo (the top left figure of each cell) and the modal percentages thereof for each of the six weight-of-shipment divisions. Vertically displayed from the TOTAL cell are the percentage data in the seven distance categories. The body of the matrix consists of 42 cells giving cargo data within the specific weight-and-distance categories.

MEASUREMENT OF MODAL ADVANTAGE

Procedure

Each of the 42 specific weight-and-distance categories (cells) of each of the 89 products was analyzed as to the extent of modal involvement in handling the cargo in that cell. Each cell was classified as either truck dominated, competitive, or rail dominated on the basis of the following assumptions:

1. A cell is classified as competitive if the mode of secondary importance carried 10 percent or more of the total cargo of the cell (the mode of secondary importance is the mode carrying the second highest percentage of cargo), which is not to say that intermodal competition (competition between truck and rail) exists for each and every shipment in that cell, but indicates that shipper action favors both truck and rail in carrying cargo of that shipment weight and distance; and
2. If the mode of secondary importance did not carry at least 10 percent of the cargo of a cell, that cargo was considered to be noncompetitive or dominated by the mode of primary importance, which simply says that shipper preference favored one mode almost exclusively, for whatever reason(s).

This classification procedure can create distortions by an attribution of cargo to modes that differs from actual results. When a cell is classified as truck-dominated, this in essence attributes all of the cargo of that cell to trucks when, in actuality, rails may have carried up to 9 percent. However, while a number of

Figure 1. Illustration of basic matrix distributing cargo percentages by mode by shipment weight and distance for TCC 321 (flat glass), 1972.

TCC 321
FLAT GLASS
PERCENTAGE DISTRIBUTION
T. T. DETAILED=1,504,000

Distance		Mode	TOTAL	UNDER 453.6 kg (UNDER 1,000 LB.)	453.6 - 4,536.0 kg (1,000 - 9,999 LB.)	4,536.0 - 13,608.0 kg (10,000 - 29,999 LB.)	13,608.0 - 27,216.0 kg (30,000 - 59,999 LB.)	27,216.0 - 40,824.0 kg (60,000 - 89,999 LB.)	40,824.0 kg AND OVER (90,000 LB. AND OVER)
	TOTAL	Total Tons	100	0.86	4.19	7.65	19.27	10.51	27.53
		Railroads	30.9	—	—	7.8	6.3	41.1	84.1
		Truck	69.0	100	100	92.1	93.7	50.9	15.9
		Other	0.1	—	—	—	—	—	—
UNDER 160.9 km	UNDER 100 MILES	Total Tons	7.71	0.15	0.47	1.06	3.66	1.00	1.40
		Railroads	16.1	—	—	—	1.9	63.7	37.4
		Truck	83.8	99.2	98.8	100	98.1	36.3	62.6
		Other	0.1	0.9	1.2	—	—	—	—
160.9 km 321.8 km	100-199 MILES	Total Tons	13.56	0.18	0.53	0.86	6.65	1.46	3.92
		Railroads	26.7	—	—	—	1.7	13.9	85.4
		Truck	73.2	96.0	100	100	98.3	86.0	14.6
		Other	0.1	4.0	—	—	—	—	—
321.8 km 482.7 km	200-299 MILES	Total Tons	21.28	0.13	0.66	1.33	10.51	2.46	6.58
		Railroads	27.5	—	—	—	1.7	5.2	85.1
		Truck	72.5	98.8	98.7	100	98.3	94.8	14.9
		Other	—	1.2	1.3	—	—	—	—
482.7 km 804.5 km	300-499 Miles	Total Tons	36.05	0.20	1.20	1.99	13.34	3.06	10.24
		Railroads	33.1	0.2	—	8.7	4.1	49.0	93.7
		Truck	66.9	97.0	99.6	91.3	95.8	51.0	6.3
		Other	—	2.7	0.4	—	0.1	—	—
804.5 km 1,609.0 km	500-999 MILES	Total Tons	19.75	0.27	1.66	1.60	11.50	1.78	3.99
		Railroads	30.3	—	0.5	23.8	14.4	56.9	82.6
		Truck	69.7	95.6	99.1	76.6	85.6	43.1	17.3
		Other	0.1	4.3	0.4	—	—	—	—
1,609.0 km 2,413.5 km	1,000-1,499 MILES	Total Tons	3.39	—	0.07	0.13	2.33	0.78	0.18
		Railroads	21.1	—	—	25.4	9.2	44.2	100
		Truck	77.6	—	87.5	74.6	90.8	52.0	—
		Other	1.2	—	12.5	—	—	3.8	—
2,413.5 km AND OVER	1,500 MILES AND OVER	Total Tons	4.26	—	0.20	0.66	1.26	0.97	1.66
		Railroads	40.8	—	—	4.7	25.2	100	58.4
		Truck	59.1	—	100	95.4	74.8	—	41.6
		Other	0.1	—	—	—	—	—	—

cells of a product will be classified truck dominated, a number also might be classified rail dominated. This creates a trade-off of cells in which the distortions tend to be neutralized when the totals of individual cells are compared to the actual overall modal split of the product. The only time that a significant distortion might occur is when, e.g., a great majority or all of a product's cells are shown to be truck dominated. In this case, the amount of cargo considered to be truck dominated will be proportionally greater than the actual amount of the product cargo carried by truck. Such a distortion might occur for a commodity in which the rail mode carried 9 percent of the cargo of each cell. The classification procedure would attribute 100 percent of the product cargo to truck as truck dominated when the rails would have actually carried 9 percent of the overall cargo. However, this analysis determined that a significant distortion occurred in only a few products, all of which shipped very small amounts of freight for which the actual rail percentage was very low. Thus, the distortion was not significant.

The 10 percent level used to classify cells is arbitrary. If a level of 20 percent had been used, the number of cells classified as competitive would be less, and if a 5 percent level had been used, the number of cells classified as competitive would be greater. A 10 percent level was felt, however, to be a relevant delineation for this analysis.

Truck Versus Rail

When the cells of the individual products were classified by the 10 percent competitive definition, the modal division of cargo showed extremely consistent patterns. Low-weight shipments were consistently truck dominated and heavy shipments were rail dominated, as expected. Extensive intermodal involvement was generally evident only in the middle shipment weight-and-distance brackets. At first, the modal division of the cargo of products might appear to indicate extreme competition; yet, closer observation shows that this is not so. For example, the overall division of the cargo of TCC 243 (millwork, plywood, and such) was 50.5 percent rail and 48.7 percent truck. When classified by the 10 percent definition, however, 42.9 percent of the cargo of this product was truck dominated and 26.2 percent was rail dominated, leaving 30.9 percent showing extensive intermodal involvement. Similarly, although the overall division of the cargo of TCC 282 (plastic materials) was 44.9 percent rail and 51.5 percent truck, these modes actually competed for only 31.0 percent of the freight, the other 69.0 percent being either truck or rail dominated.

In fact, of the total manufactured cargo in this study [827 839 051 Mg (912 521 000 tons)], which in the aggregate divided 38.1 percent to rail and 58.1 percent to truck, the amount found to be truck dominated (and noncompetitive with rail) was 44.60 percent as shown below.

Cargo	Percentage of Cargo
Rail dominated	28.83
Truck dominated	44.60
Competitive	26.57

All rail-dominated cells within individual products accounted for 28.83 percent of the total, and the cargo of cells for which truck and rail involvement is evident was 26.57 percent.

These data can be expressed in matrix format as shown in Table 1, which was constructed by analyzing the classification of each specific cell for each of the 89 products. Each cell in Table 1 that is labeled T showed truck-domination in a majority of the 89 products.

R and C cells were classified by the same procedure. This matrix construction, therefore, gives equal weight to the shipment pattern of each product and represents the propensity for modal choice. (When this distribution patterns is referred to in the remainder of this analysis, it will be identified as typical.)

This typical modal-use pattern shows that truck domination generally extends to shipments of less than 4536 kg (10 000 lb), independent of the distance over which the shipment must be carried. The railroads generally dominate shipments of 40 824 kg (90 000 lb) or more, again virtually independent of the distance to be shipped. For shipments ranging from 4536 to 40 824 kg (10 000 to 90 000 lb), distance becomes a more important factor in modal choice. Shipments of 4536 to 13 608 kg (10 000 to 30 000 lb) generally move by truck up to 805 km (500 miles). Beyond 805 km (500 miles), there is a significant degree of competition from rail. For shipments of 13 608 to 27 216 kg (30 000 to 60 000 pounds), trucks generally dominated up to 483 km (300 miles) with longer hauls being competitive with rail. Truck competition with rails extends to 805 km (500 miles) in the 27 216 to 40 824 kg (60 000 to 90 000 lb) shipment-mass range, but beyond 805 km (500 miles), shipments tend to be rail-dominated.

LIMITATIONS TO MATRIX REPRESENTING TYPICAL USE PATTERNS

Since the typical use pattern, as represented in Table 1, reflects averaging, it is not valid for the shipment pattern of each and every product. In some commodities, truck domination is found in cells that are typically represented as competitive or rail dominated. Conversely, there are products in which rail domination extends into the typical competitive or truck domains. However, an important value of the typical use pattern is that of providing a frame of reference by which to identify and measure the extent of deviation in the pattern by individual commodity. These deviations, if not reflective of weaknesses in the data, could provide valuable information on modal demand.

Despite this qualification, there is a strong relation between the typical distribution pattern and the patterns exhibited within individual commodities. Considering that shippers are free to use any mode to transport any shipment, and that the 10 percent definition of competition tends to make the occurrence of competition high in individual commodity cells, there is a remarkable consistency between the modal-preference patterns for individual commodities and the typical use pattern. In only 6.7 percent of the products (6 of 89) do the modal-use patterns deviate by more than 33.3 percent from the typical.

The vast majority of deviations were minor, involving truck domination or rail domination in typical competitive cells or competition in the typical truck or rail domains. There were only 49 instances, involving 30 products, of major deviation out of the 3299 cells analyzed in this study. A major deviation has been defined as an instance in which truck domination exists in a cell that would typically be considered rail dominated, or vice versa. In 16 of these 30 products, only one cell showed major deviation. In 11 products, only two cells showed this deviation. One product showed three major deviations, and two products showed four.

The distinction between minor and major cellular deviations is important because of the trend in modal-use patterns. That is, modal use generally favors truck for light-weight shipments and rail for heavy shipments. Instances of major deviations were in direct contrast to this pattern, where truck domination was evident in heavy shipments or rail domination in light shipments. The matrices for two commodities, TCC 202 (dairy products) (Table 2) and TCC 355 (special industry machinery) (Table 3), illustrate the distinction between minor and major deviations. While major deviations may represent weaknesses in the data, it is much more probable that they accurately reflect the actual peculiarities of some products for transportation service.

Of the 49 instances of major deviation, 45 involved truck domination in shipment weights and distances that typically involve rail movements. Some specific instances of major deviations were concentrated as shown below.

TCC	Industry Description	Number of Instances
36	Electrical machinery	9
35	Nonelectrical machinery	8
34	Fabricated metal products	6
22	Basic textiles	5
Total		28

RELATION BETWEEN MATRIX REPRESENTATION AND NUMERIC TABULATION OF DATA

While the percentages of cargo found to be truck dominated, competitive, or rail dominated pertain to the count of cargo performed on each cell of each product and thus represent absolute figures, the representation of these data in summary matrices requires the use of averages. The question to be addressed at this point concerns whether the actual instances of deviation invalidate the use of the matrix. Or, can Table 1 be reconciled with the absolute figures?

Table 4 shows the typical use pattern, giving the percentage of total manufactured cargo accounted for by each individual cell of the total matrix. This table, as summarized below, shows that the total of cargo within the truck domain of this aggregate matrix is 45.6 percent.

Percentage of Cargo	Truck Dominated (%)	Competitive (%)	Rail Dominated (%)
Total in Table 4	45.6	25.5	28.9
Total by count of individual commodity cells	44.6	26.6	28.8
Difference between Table 4 and cell-count amounts	1.0	1.1	0.1

This figure is greater by only 1.0 percent than the cargo amount determined to be truck dominated by the count of individual product cells (44.6 percent). The difference between the actual count of competitive cargo and the matrix-represented amount of competition is 1.1 percent. Hence, the matrix representation appears to accurately reflect the modal division of cargo at this aggregate level. The data given in Table 4 thus show that essentially all of the cargo actually carried by rail in the portion of the matrix that is typically truck dominated was offset by an equal amount of another product in the rail-dominated segment that is actually handled by truck.

COMPETITION: TRUCK AND RAIL

The evidence given above and the matrix as developed by Tables 1 and 4 show that only about a quarter of the total manufactured cargo involves intermodal choice. This competitive freight lies between light-weight shipments, for which the truck is inherently suited, and heavy-weight shipments, which are generally suited to rail. Specifically, this competitive cargo involves (a) shipment sizes that are related to the marginal use of the carrying capacity of each mode (the upper limits of truck capacity and the lower limits of rail capacity) and (b) shipment distances that are related to multiday delivery within the size-of-shipment constraints.

While this competitive cargo is about one-fourth of the total manufactured freight, it is a more significant amount relative to each mode individually. That is, most of the lighter weight shipments are unsuited to rail movements, and most of the heavy-weight shipments are unsuited to truck transport. As such, it is realistic to exclude the cargo accounted for by these respective shipment types from the total to determine the universe of cargo applicable to each mode. The cargo applicable to rail movement is, hence, that which is rail dominated or intermodally competitive [55.40 percent of the total (28.83 percent plus 26.57 percent)]. The cargo applicable to truck transport is 71.17 percent of the total (44.60 percent plus 26.57 percent). Relatively, truck dominance extends to 62.67 percent of the cargo applicable to this mode (or 44.60 percent/71.17 percent) and rail dominance extends to 52.03 percent of its applicable cargo (28.83 percent/55.40 percent).

While the above data indicate that the presence and importance of each mode are assured in the future virtually irrespective of competition, it is of significant importance to each mode to capture as much of this competitive cargo as possible. The question then becomes: What is the potential division of the competitive freight?

There are many different techniques available to the imaginative analyst by which to allocate this competitive traffic to mode. Within the framework of this analysis, the following procedure was used: The mode with the higher percentage of cargo within the individual competitive cell was awarded the total of the cargo of that cell. There are certainly weaknesses in using such a simple approach to allocate the competitive traffic, but it is supported by the data itself. Also, the data could not be analyzed in this detailed format over time, and there are no other divisional techniques that are totally devoid of subjective weaknesses. At the extreme, the most optimistic procedure would award the entire 26.57 percent to one mode or the other, but this would entail assuming stagnation and deterioration in the entire technology of the other mode. This approach would be totally unrealistic. Both trucks and rail will seek to increase their amount of cargo, and some of the competitive cargo will tend to truck and some to rail.

The results of this analysis, in which the competitive manufactured cargo was divided so that the mode that handled the larger percentage of freight in each individual product cell was awarded the total cargo of the cell are shown on the following page.

By actual count, the competitive cargo tended to divide in half, 13.33 percent to truck versus 13.24 percent to rail. As can be seen in Table 5, the traffic tending to rail was the heavier weight shipments, and

Table 1. Distribution of shipment sizes and distances, modally dominated and intermodally competitive for all products (1972).

Distance (km)	Weight (kg) ≤454	454 to 4536	4536 to 13 608	13 608 to 27 216	27 216 to 40 824	≥40 824
≤161	T	T	T	T	C	C
161 to 322	T	T	T	T	C	R
322 to 483	T	T	T	T	C	R
483 to 805	T	T	T	C	C	R
805 to 1609	T	T	C	C	R	R
1609 to 2414	T	T	C	C	R	R
≥2414	T	T	C	C	R	R

Notes: 1 km = 0.62 mile, 1 kg = 2.2 lb.
T = truck, C = competitive, R = rail.

Table 2. Distribution of shipment sizes and distances, modally dominated and intermodally competitive for TCC 202 (1972).

Distance (km)	Weight (kg) ≤454	454 to 4536	4536 to 13 608	13 608 to 27 216	27 216 to 40 824	≥40 824
≤161	T	T	T	T	C	T
161 to 322	T	T	T	T	C	C
322 to 483	T	T	T	T	R	R
483 to 805	T	T	T	T	R	R
805 to 1609	T	T	T	C	R	R
1609 to 2414	–	T	T	T	R	R
≥2414	–	T	T	T	R	R

Notes: 1 km = 0.62 mile, 1 kg = 2.2 lb.
T = truck, C = competitive, R = rail.

Table 3. Distribution of shipment sizes and distances, modally dominated and intermodally competitive for TCC 355 (1972).

Distance (km)	Weight (kg) ≤454	454 to 4536	4536 to 13 608	13 608 to 27 216	27 216 to 40 824	≥40 824
≤161	T	T	T	T	T	–
161 to 322	T	T	T	T	R	T
322 to 483	T	T	T	T	T	C
483 to 805	T	T	T	C	C	T
805 to 1609	T	T	C	C	R	C
1609 to 2414	T	T	C	C	T	R
≥2414	T	T	C	C	C	T

Notes: 1 km = 0.62 mile, 1 kg = 2.2 lb.
T = truck, C = competitive, R = rail.

Table 4. Percentage of cargo accounted for by applying Table 1 modal divisions to totaled amounts of cargo for all products (1972).

Distance (km)	Percentage of Cargo, Weight (kg) ≤454	454 to 4536	4536 to 13 608	13 608 to 27 216	27 216 to 40 824	≥40 824
≤161	0.51	2.28	6.05	11.84	1.29	6.93
161 to 322	0.27	1.24	2.51	7.22	1.02	4.57
322 to 483	0.20	0.93	1.60	4.88	1.05	4.09
483 to 805	0.30	1.22	1.90	4.84	1.37	5.46
805 to 1609	0.44	1.34	1.85	4.55	2.12	6.63
1609 to 2414	0.11	0.29	0.38	1.12	0.80	2.27
≥2414	0.12	0.37	0.33	0.77	1.17	1.76

Note: 1 km = 0.62 mile, 1 kg = 2.2 lb.

Cargo	Percentage of Cargo
Rail dominated	28.83
Truck dominated	44.60
Competitive	26.57
Tending to rail	13.24
Tending to truck	13.33

to truck the lighter weight shipments.

The portion of Table 5 tending to truck accounts for 14.36 percent of the aggregated cargo of all products. This figure can be derived from the data given in Table 4 and differs by only 1.03 percent from the actual amount of competitive cargo found tending to truck as shown above. The amount of cargo shown in Table 5 as tending to rail is 11.14 percent, which differs by 2.10 percent from the actual amount tending to rail. Again, the matrix representation appears to validly reflect the actual cargo-count data.

Based on the actual count of cargo, the potential distribution of freight to mode can be seen from the following data:

Mode	Modally Dominated (%)	Competitive Tending to a Mode (%)	Total Potential
Rail	28.83	13.24	42.07
Truck	44.60	13.33	57.93

That is, if the rail mode handled all of the cargo (a) designed as dominated and (b) currently competitive and tending to rail, the total rail percentage of cargo would be 42.07 percent. The truck percentage of cargo would be 57.93 percent.

Table 6 shows the composite matrix of all products and the potential division of all shipment size-and-distance categories. Based on this aggregate matrix representation, the division of total manufactured cargo between the modes reflects primarily the weight of the shipment being transported. Almost all cargo of less than 27 216 kg (60 000 lb) tends to trucks, and that of more than 27 216 kg (60 000) tends to rail. While this matrix presents clearly defined areas of modal advantage, not all shipments will, nor can they, move in compliance with these broad modally related weight-and-distance delineations. In fact, in aggregating the potential amounts of freight by mode as shown above, no attempt was made to allocate the cargo amounts accounted for by the instances of (major) deviation to what Table 6 would indicate to be the correct mode. The factor or factors that caused the deviant modal choices were obviously not weight and distance related in the first place and will not be so in the future unless there is a significant change in modal technology. In general, however, the shipment size versus mode of transportation relation is based on weight.

For comparative purposes, the three sets of modal percentage of cargo applicable to this analysis are listed below.

Category	Rail (%)	Truck (%)
Actual overall percentages of total cargo handled by mode in 1972	38.10	58.10
Overall percentages of total cargo derived through the allocation of individual product cells to mode in 1972	42.07	57.93
Typical (Table 2-represented) percentages of total cargo to mode in 1972	40.05	59.95

The actual overall percentages given above fail to account for the 3.80 percent of freight actually carried by the excluded modes; i.e., air, water, and other.

Concerning the amount of cargo potentially tending to each mode as shown above, one can readily see that the modal splits as classified by this analysis are proportionally not significantly different from the actual percentages of cargo in 1972. The actual rail percentage was 38.10, and the potential rail cargo was 42.07 percent. Trucks actually carried 58.10 percent and potentially carried 57.93 percent. When the actual percentages of cargo carried by the excluded modes are considered, the potential rail and truck cargo amounts will decrease minimally.

Thus, regardless of the way in which freight is allocated to mode, either by the real world, presently existing means (as represented by the actual figure) or by drawing an arbitrary line of comparative advantage, which reflects the typical shipper preference existing in 1972 (as represented by the individual cell count and table divisions), there would have been no significant change in the amount of cargo handled. A relatively insignificant amount would have been transferred from truck to rail for some products, and from rail to truck for others, but the net change in cargo carried would be inconsequential.

FUTURE ROLES OF TRUCK AND RAIL

Throughout this analysis, attention has been focused on the modal division of shipments for a single time period,

Table 5. Distribution of competitive shipment-size-and-distance categories by means of transport for all products (1972).

	Weight (kg)					
Distance (km)	≤454	454 to 4536	4536 to 13 608	13 608 to 27 216	27 216 to 40 824	≥40 824
≤161	T	T	T	T	C[a]	C[b]
161 to 322	T	T	T	T	C[b]	R
322 to 483	T	T	T	T	C[b]	R
483 to 805	T	T	T	C[a]	C[b]	R
805 to 1609	T	T	C[a]	C[a]	R	R
1609 to 2414	T	T	C[a]	C[a]	R	R
≥2414	T	T	C[a]	C[b]	R	R

Notes: 1 km = 0.62 mile, 1 kg = 2.2 lb.
T = truck, C = competitive, R = rail.
[a] Tending to truck. [b] Tending to rail.

Table 6. Potential distribution of shipment-size-and-distance categories by means of transport for all products (1972).

	Weight (kg)					
Distance (km)	≤454	454 to 4536	4536 to 13 608	13 608 to 27 216	27 216 to 40 824	≥40 824
≥161	T	T	T	T	T	R
161 to 322	T	T	T	T	R	R
322 to 483	T	T	T	T	R	R
483 to 805	T	T	T	T	R	R
805 to 1609	T	T	T	T	R	R
1609 to 2414	T	T	T	T	R	R
≥2414	T	T	T	R	R	R

Notes: 1 km = 0.62 mile, 1 kg = 2.2 lb.
T = truck, R = rail.

Table 7. Analysis of changes in modal percentage of cargo based on changes in relative amounts of cargo in lighter versus heavier shipments for all comparable commodities (1967 and 1972).

No. of Products	Amount of Cargo Decreased	Amount of Cargo Increased
Affected	43	42
In which rail percentage of cargo decreased	36	19
In which truck percentage of cargo increased	36	19
In which rail percentage of cargo increased	6	23
In which truck percentage of cargo decreased	6	23
In which no change occurred in rail percentage of cargo	1	–
In which no change occurred in truck percentage of cargo	1	–

Table 8. Analysis of changes in modal percentage of cargo based on changes of 5 percent or more in relative amounts of cargo in lighter versus heavier shipments for all comparable commodities (1967 and 1972).

No. of Products	Amount of Cargo Decreased	Amount of Cargo Increased
Affected	16	15
In which rail percentage of cargo decreased	15	5
In which truck percentage of cargo increased	16	4
In which rail percentage of cargo increased	–	10
In which truck percentage of cargo decreased	–	11
In which no change occurred in rail percentage of cargo	1	–
In which no change occurred in truck percentage of cargo	–	–

1972. If this 1972 division of cargo were constant over time, the future roles of truck and rail, in the absence of any artificial or arbitrary obstacles to modal use or interferences in market forces, would be dependent simply on the amount of cargo or the number of shipments in the weight-and-distance categories in any future time period. However, it is not at all certain that the 1972 modal division is constant. Certainly, if a long-term comparison could be made in the matrix format, perhaps over 20 or 30 years, material changes would have occurred, primarily reflecting the growth of trucking. Concerning the future roles of truck and rail, the question to be addressed is whether shippers' perceptions of modal comparative advantages will change and, if so, in what direction. Will trucking in the future dominate any current rail cells, or vice versa, or are the 1972-depicted limits stable? If changes are found to be occurring (assuming the rates of one mode have not changed drastically in relation to the rates of the other), this would indicate that shippers' perceptions of service differences are developing between the modes in one direction or another.

To determine trends in shippers' preferences, a time-series analysis is helpful. However, because of data-incompatability limitations, i.e., the lack of detailed data that would be necessary to create the matrix for a period other than 1972, this time-series evaluation can only be made in rough form and is only attempted because the 1972 potential modal split centered around 27 216 kg (60 000 lb) shipments, a shipment-size delineation that is available in the 1967 Census of Transportation.

This time-series comparison involved evaluating the relation between changes in the relative amounts of cargo suited to movement by a mode and changes in the actual overall percentages of freight handled by it. It was conducted under the assumption that modal relative advantage was stable at the 1972 limits, i.e., that there would be a strong and direct relation in the following measurements:

1. If the relative amount of cargo heavier than 27 216 kg (60 000 lb) decreased from 1967 to 1972, the overall truck percentage of freight would increase and
2. If the relative amount of cargo heavier than 27 216 kg (60 000 lb) increased from 1967 to 1972, the overall rail percentage of freight would increase.

This test could be performed on 85 products. Of the 89 products listed in the 1972 census, 2 were not contained in the 1967 source, namely, TCC 205 (bakery products) and TCC 225 (knit fabrics). An additional 2 products were eliminated because they failed to ship any cargo heavier than 27 216 kg (60 000 lb) in either 1967 or 1972. Those commodities were TCC 381 (engineering, laboratory, and scientific instruments) and TCC 387 (watches, clocks, and allied products). The results of this test are summarized in Table 7. These data pertain to all commodities in which there was any change in the percentage of cargo heavier than 27 216 kg (60 000 lb), irrespective of the magnitude of the change. There was a decrease in the percentage of heavy-weight shipments in 43 products. The overall rail percentage of cargo decreased in 36 products, increased in 6, and exhibited no change for 1. Conversely, the overall truck percentage of cargo increased in 36 products decreased in 6, and did not change for 1. These results for decreases in the relative amounts of heavy-weight cargo indicate that rail is not making serious inroads into the lighter weight shipment market, the truck domain. However, consider the relationship concerning the 42 commodities in which there was an increase in the percentage of heavy shipments. Contrary to the results expected by assuming stabilized modal use, in almost half of the commodities (19 of 42) the rail percentage of cargo decreased!

To further test the hypothesis, the commodities with only minor changes in the percentage of heavy-weight shipments were eliminated from the tabulation. The results for all commodities in which there was change of at least 5 percent of total cargo between heavier [over 27 216 kg (60 000 lb)] and lighter [under 27 216 kg (60 000 lb)] weight shipments are shown in Table 8. Sixteen commodities exhibited significant decreases in heavy shipments. The overall truck percentage of cargo increased in all 16. The actual overall rail percentage of cargo decreased in 15 and remained unchanged in 1. There were 15 commodities in which there were significant increases in heavier shipments. The overall rail percentage of cargo increased in 10 and decreased in 5, and the overall truck percentage of cargo decreased in 11 and increased in 4. Hence, it appears that modal-use changes have occurred from 1967 to 1972 to the detriment of the rail mode and, extrapolating from these results, could very well occur into the future. Since there is good reason to believe that no significant rate disparity occurred between 1967 and 1972, trucks apparently offer a service advantage that is tending to absorb a portion of the rail freight market.

Although service advantages are almost impossible to quantify, a recent study by the U.S. Department of Transportation helps to explain the attitude of shippers toward this performance factor. This study indicated that shippers felt that rail performance relative to truck performance was inferior in the following areas, even with the inherent modal differences (such as average time in transit) acknowledged: (a) on-time pickup; (b) on-time delivery; (c) arrivals without loss, short, or damage; and (d) availability of specified equipment. If the railroads cannot improve their service performance on inherently suited traffic, they cannot be expected to retain their present freight base, nor to gain traffic of other sizes or distances. However, if the rail time factors and damage performances improve and afford better reliability of service, then a significant increase in demand for rail service, especially long-distance line-haul service, could be expected. Improvements in these areas would also allow service innovations such as trailer on flatcar and container on flatcar to come up to their potential for the benefit of all modes, shippers, and consumers. Improving and expediting interchange arrangements and improving control of freight cars also could be expected to increase the amount of freight moving by rail.

At the same time, however, attention should be focused on productivity improvements that are available to the truck mode. An area with significant potential for such improvements is weight-limitation and vehicle-length and configuration laws, where there are differences among the various states. Changes in these laws can not only help trucking to better fulfill its obligation to provide service to the public, but can also have secondary impacts on resources allocated to trucking. They are a means to improve the freight-moved-to-fuel-consumed ratio, for example. They can impact on the number of trucks needed to move a fixed amount of freight with resultant fuel use and highway congestion and safety implications. Such changes can also impact on the cost of products to consumers if one truck can do the job that now might require two, or two trucks the job currently performed by three. In any event, there are reasonable means available to improve the performance of all modes, and these means must be seriously evaluated in the context of the continuing development of a national transportation policy.

CONCLUSION

The purpose of this paper has been to recast, or better display, the numbers used in the evaluation or analyses of the traffic split between modes. It is concerned principally with modal choices made in 1972 on the basis of shipment weight and distance moved. No attempt was made to identify or measure factors of modal choice other than shipment weight and distance, nor was any attempt made to allocate traffic to modes in any way that is inconsistent with actual shipper preferences as shown by the data. There was no attempt to quantify service or rate differentials between modes or to calculate social costs or benefits created by the modal split. This analysis simply investigates and depicts what shippers in 1972 felt were the roles or the relative advantages of truck and rail in their distribution systems. However, these data can be used as valuable aids in further research concerning modal choice. If one assumes that the shipper is oriented toward minimizing total costs, i.e., that the shipper has considered all relevant factors in the equation of costs before making a modal choice, then these data could help identify the level-of-service differential, given a wide spread in rates, or help dispel the premise that there are significant nonjustified economic costs created by the modal split.

This study provides a procedural frame in which to analyze the roles of truck and rail in the distribution of manufactured freight. The analysis, which uses uncomplicated statistical techniques, was conducted under the assumption that there are relative advantages, which are adhered to, in transportation. The data of the 1972 Commodity Transportation Survey not only supported this assumption as concerns the basic areas of advantage, but moreover, showed that shipment weight is a more significant factor in shipper choice of mode than is length of haul, although length of haul [centering around 483 to 805 km (300 to 500 miles)] does have some impact on modal choice for freight in the middle size-of-shipment brackets [4536 to 40 824 kg (10 000 to 90 000 lb)]. However, the advantages of trucks, especially that of for-hire motor carriers, generally pertain to shipments up to 27 216 kg (60 000 lb). The advantages of rail pertain to shipments of 27 216 kg (60 000 lb) and more.

In relation to the total amount of manufactured freight (with minor exceptions), 44.60 percent involved virtually only truck movement. The traffic moved solely by rail is approximately 28.83 percent of manufactured freight; intermodal competitive involvement is evident for the remaining 26.57 percent. Although these figures apply to a composite of all products, the basic data of this analysis show that there are noncompetitive areas within the shipment distribution patterns of any product.

This analysis has demonstrated that the amount of cargo subject to modal substitution is virtually inconsequential. Within individual commodities, some freight of some products could shift from truck to rail. But within other commodities, freight shifts will be from rail to truck. When all products are considered, the individual commodity freight shifts will tend to cancel each other.

The future roles of truck and rail will depend primarily on the amount of freight in the modally oriented shipment-size lots, although there will probably be shifts in modal use for freight in certain weight-and-distance categories. Factors such as service innovations, technological change, and transportation-oriented legislation can have significant impacts on modal roles; however, the primary factor should remain the size of the shipment.

The current split of cargo reflects the modal preferences of shippers who are themselves oriented toward minimizing the total cost of manufacturing and distribution of their product. The modal choice has already been made and the areas of advantage are relatively solid. The orientation of the carriers, the regulatory agencies, and policy economists should center on improving the efficiency and service offered by each mode within their respective freight bases. For only when efficient and satisfactory service can be offered in shipments for which a mode is inherently or relatively suited can the real cost of transportation decrease.

Publication of this paper sponsored by Committee on Passenger and Freight Transportation Characteristics.

Approximation Equations for Costs of Rail, Trailer-on-Flatcar, and Truck Intercity Freight Systems

Edward K. Morlok and John A. Warner, Department of Civil and Urban Engineering, University of Pennsylvania

This paper presents equations that approximate the fully allocated and variable costs contained in the Interstate Commerce Commission cost tables for rail-carload, trailer-on-flatcar, and truck intercity freight movements. These equations were developed to enable the user to approximate the costs quickly and easily. They should be useful in initial studies of costs where the exact values are not needed, such as in consideration of rate changes, studies of profitability, and general intermodal comparisons. The equations were used to develop estimates of cost for complete shipper to receiver shipments via the three carriers to illustrate general properties of the carriers, individually and with respect to one another.

The cost characteristics of rail-carload, trailer-on-flatcar (TOFC), and common-carrier-truck intercity freight systems, as estimated by the Interstate Commerce Commission (ICC), are discussed for two purposes: (a) to make available approximation equations for the ICC costs, which are in tabular form, and (b) to compare these costs, individually and with one another.

There are several reasons for using the ICC cost tables. Because they are in the public domain, costs estimated by using them have none of the problems associated with costs estimated by using proprietary methods. The ICC cost tables give estimates that are useful for many transportation-analysis purposes, their value in large measure being derived from the fact that they are used by regulators as a lower bound on the prices that carriers can charge. The fully allocated costs are somewhat analogous to average total costs, and thus also provide a useful measure of cost. The ICC costs are also useful to a shipper engaged in a rate negotiation with a carrier, as a means of estimating the cost to that carrier of providing the service in negotiation. These cost estimates may also be useful to carriers who wish to obtain estimates of the ICC-based costs for particular movements. Such estimates would be useful in studies of the profitability of various movements, as indicators of the commission's potential reaction to a rate change request, and as a basis for comparison of a carrier's true (i.e., internally developed) costs with those of the average carrier.

Balancing the advantages of the ICC cost tables is that the tables themselves are cumbersome to use and that it is difficult to obtain any general picture of cost characteristics from them. Therefore, analytical relations that approximate these tabular costs can be used advantageously, not only to simplify the computations, but also to gain a general understanding of the basic functional relations among the many variables that affect a carrier's costs. The relative costs of rail, TOFC, and road transport depend on a number of characteristics of the shipment and the carrier, such as the distance of the movement, the density of the material being shipped, the total weight of the shipment, the circuity of the routes, and the extent of empty versus loaded distance traveled.

In the following sections of this paper, we will discuss the cost characteristics of each of the freight systems individually, compare these characteristics, and make several comments pertinent to freight systems in general. The details of the cost-estimating equations are given in detail.

RAIL CARLOAD

Rail cost and performance will be illustrated for the two most common carriers of general-merchandise freight: unequipped and equipped (sometimes called damage-free or cushion-underframe) general-service boxcars. Although such cars have a wide range of sizes, capacities, and equipment configurations, we will model a typical car having a 59-Mg (approximately 65-ton) weight capacity and a volume capacity of about 139 m^3 (4900 ft^3). The costs are derived from the ICC statement, Rail Carload Cost Scales 1973 (1), one of a series published annually, which contains, by region, scales showing variable and fully allocated costs as a function of short-line rail distance and weight of load. The costs are given in three forms: (a) a summary table of average regional costs; (b) a breakdown of these unit costs into terminal and line-haul components, with the line-haul component further broken down into way trains and through trains; and (c) a detailed table that contains, for average trains, way trains, and through trains, by car type, variable and constant terminal and line-haul costs on a car-mile and hundredweight-mile basis. In addition to these three basic tables, provision is made to allow adjustment of average costs for situations in which it is known that operational procedures, such as circuity or empty-return ratio, are different from regional average procedures.

In this paper we illustrate costs for ICC region 3, the official territory of which includes those states east of Wisconsin and Illinois (including a portion of Illinois) and north of Kentucky and North Carolina (excluding a portion of Virginia). The costs of providing basic rail boxcar service in the region, as computed from the ICC cost data, are summarized in Figure 1 for equipped and unequipped cars for several distances as a function of shipment weight. The cost of moving a car is the largest portion of any shipment cost, and marginal increases in the net load of the car have a small effect on total shipment costs.

Cost-Estimating Procedure

The basic rail-boxcar cost-estimating procedures are as follows.

1. Determine shipment characteristics:

S = shipment weight (cwt);
L = shipment distance (actual miles)—(if this is not available, reasonable estimates are 1.25 times great-circle distance or 1.18 times rail short-line miles for boxcar movements, 1.09 times rail short-line miles for TOFC movements,

and 1.20 times great-circle distance or 1.06 times highway rate-making miles for truck movements);
D = commodity density (lb/ft^3);
t = type of rail car used; and
a = highway-access coefficient [(a) 0, indicating highway access at neither origin nor destination; (b) 1, indicating highway access at either origin or destination, but not both; and (c) 2, indicating highway access at both origin and destination].

(SI units are not given for the variables in these equations, inasmuch as they were derived for U.S. customary units.)

2. Compute number of rail cars required for shipment:

$$n_d = (100S/0.9V) \tag{1}$$

where n_d = volume requirement and V = volume capacity of equipment used (a typical value is 4900 ft^3).

$$n_s = (S/W) \tag{2}$$

where W = weight capacity of equipment used (a typical value is 1300 cwt). If n_s or n_d are not integers, they are rounded to the next higher integer.

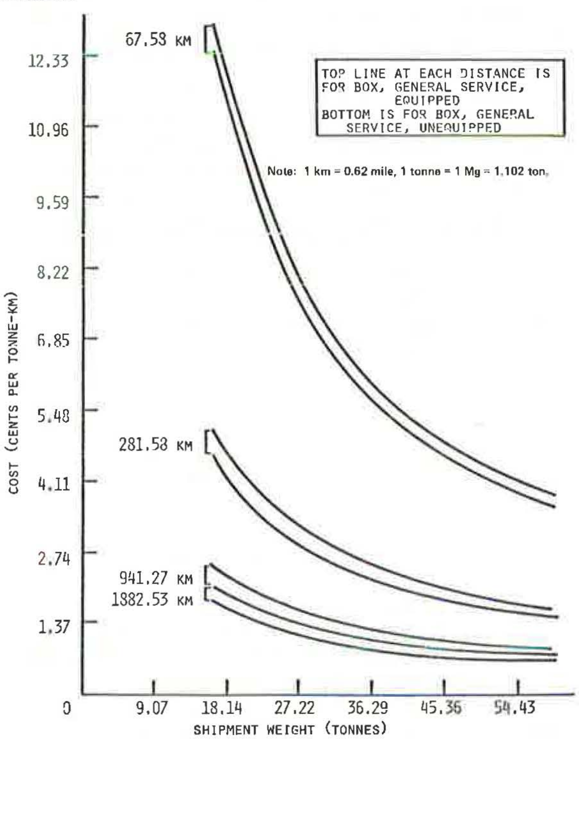
$$n = \mathrm{Max}\ (n_d, n_s) \tag{3}$$

3. Select applicable cost formula and compute cost: The basic cost equation has the form

$$C_i = B_i n + 0.000\ 36S + L\ (E_i n + 0.000\ 148\ 2S) \tag{4}$$

where B_i and E_i are parameters that vary with car type. The variable line-haul cost (c_t) is given by

Figure 1. Rail carload cost as a function of shipment weight and distance.



$$c_1 = 116n + 0.000\ 36S + L(0.317\ 35n + 0.000\ 148\ 2S) \tag{4a}$$

and

$$c_2 = 116n + 0.000\ 36S + L(0.422\ 49n + 0.000\ 148\ 2S) \tag{4b}$$

for unequipped and equipped general-service boxcars respectively. The fully allocated line-haul cost (c_t') is given by

$$c_1' = 116n + 0.024\ 46S + L(0.317\ 35n + 0.000\ 286\ 1S) \tag{5a}$$

and

$$c_2' = 116n + 0.024\ 46S + L(0.422\ 49n + 0.000\ 286\ 1S) \tag{5b}$$

for unequipped and equipped general-service boxcars respectively. The variable highway-access cost (h_s) for rail boxcar is given by

$$h_s = a(4.628S^{0.465} + 0.135\ 8S) \quad (S < 100) \tag{6a}$$

$$h_s = a(0.633\ 8S - 0.001\ 454\ S^2) \quad (100 \leqslant S < 200) \tag{6b}$$

$$h_s = a(0.474\ 8S - 0.000\ 659S^2) \quad (200 \leqslant S < 300) \tag{6c}$$

$$h_s = a(0.363S - 0.000\ 286S^2) \quad (300 \leqslant S < 437) \tag{6d}$$

$$h_s = a(0.238\ 1S) \quad (S \leqslant 437) \tag{6e}$$

The fully allocated highway-access cost (h_s') for rail boxcar is given by

$$h_s' = 1.11\ (h_s) \tag{7}$$

The total rail-boxcar system cost (c) is the sum of the line-haul and highway access costs (dollars per shipment).

Adjustments to Basic Procedure

Intuitively, we would expect the many different sizes, weights, and configurations of rail cars to have wide variations in cost characteristics. The two rail cars for which cost data are given above, unequipped and equipped general-service boxcars, are in the middle range of ranking by car costs. The least expensive to operate are large liquid tank cars, and the most expensive (depending on distance of shipment) are several types of special-service cars, such as refrigerated cars, special-service boxcars, and special-service gondolas. It is important to keep these variations due to car type in mind when analyzing specific commodity movements.

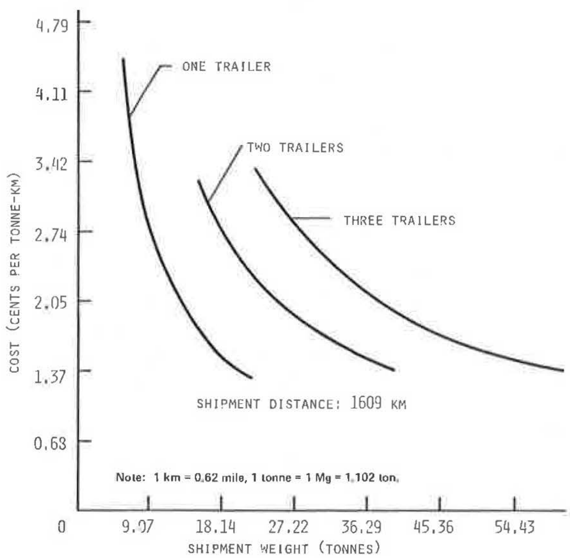
The cost characteristics of different types of rail cars can be taken into account by substituting for the basic coefficient values of Equation 4. Table 1 presents coefficient values for B_i and E_i, the two coefficients that vary with car type. The wide variation among car types in weight and volume capacity requires that these characteristics be determined for each individual situation.

Another major variation in cost for which provision can be made through use of the ICC cost data is the effect on total shipment cost of variations in frequency of intermediate yardings. These may be of three types: (a) an interchange movement between two railroads, (b) an intertrain switching between trains of the same railroad company, or (c) an intratrain switching of cars of the same train. The ICC assumes that an intertrain or intratrain switching occurs, on the average, about every 322 km (200 miles) (2). For example, there would be five intertrain or intratrain switchings in a

Table 1. Coefficient values of common railcars.

Car Types	B_i	E_i	Car Types	B_i	E_i
Unequipped general box	116	0.317 35	Mechanical meat refrigerator	83	0.564 99
Equipped general box	116	0.422 49	Mechanical other than meat refrigerator	83	0.535 08
Special box	125	0.482 18	Nonmechanical meat refrigerator	83	0.532 85
General gondola	125	0.368 45	Nonmechanical other than meat refrigerator	83	0.517 25
Special gondola	125	0.428 81	38 to 72 kL (10 032 000 to 19 008 000 gal) tank	83	0.538 61
Open general hopper	125	0.369 58	106 to 121 kL (27 984 000 to 31 944 000 gal) tank	83	0.615 66
Open special hopper	125	0.394 23			
Covered hopper	125	0.429 81			
Stock	125	0.384 86			
General flat	125	0.378 57			

Figure 2. TOFC cost as a function of shipment weight and number of trailers required.



1600-km (1000-mile) haul, resulting in an average distance between such switchings of about 269 km (167 miles). There is no explicit statement of the average frequency of interchanges. Calculations based on the average cost of each interchange in official territory, however, indicate an average frequency of one such interchange about every 965 km (600 miles). The cost difference among different intermediate yarding frequencies is substantial, especially for longer hauls. For example, over a 1600-km (1000-mile) haul, the unit cost difference for a shipment of about 23 Mg (25 tons) would be more than \$5.50/Mg (approximately \$5/ton).

The calculation of shipment costs for other than average intermediate switching conditions is straightforward. The cost estimated by the basic rail formulas includes the cost of average interchange and intermediate switching conditions. If the analyst knows that the movement being costed has other than average interchange switching, he or she may proceed as follows. (This example is for unequipped general-service boxcars.) The cost for x interchanges is given by

$$c_x = (36.54x - 0.066\,69L)n \tag{8a}$$

and the cost for y intermediate switchings is given by

$$c_y = (12.04y - 0.060\,17L)n \tag{8b}$$

The appropriate value (c_x or c_y) should be added to the cost computed by using the basic procedure.

TRAILER ON FLATCAR

TOFC costs are developed based on information in the Rail Carload Cost Scales 1973 (1). Cost scales are not presented, however, and the method of computation is quite different. Costs are computed on the basis of an assumed cost per ton mile carried, and other operational characteristics of TOFC are given on a regional average basis. This method allows the analyst to explicitly vary such operational characteristics as the number of trailers required for the shipment being costed, the number of trailers assumed to be riding on each rail flatcar, and the weight of the shipment. If information on these characteristics of the shipments is not available the analyst can use regional average data.

There are several TOFC plans, with variations in the degree to which responsibility for a shipment is divided between the railroad and the shipper. We will illustrate costs for plan 2, in which the railroads perform the entire service from consignor to consignee. Figure 2 shows cost per shipment for various shipment weights for distances of about 1600 km (1000 miles). The cost curves of Figure 2 illustrate three interesting characteristics of TOFC costs: (a) the significant contribution to total cost of terminal-related costs, (b) the influence on total cost of the number of trailers required to carry any given load, and (c) the relatively minor influence on total cost of the load carried in any given number of trailers. The second of these, that the number of trailers required for a movement, rather than the absolute size of the net load, is the primary determinant of cost, can be confirmed by inspection of single and double-trailer cost curves for movements of the same weight and distance. For example, a shipment of about 22 Mg (488 cwt) moving about 1600 km (1000 miles) in a single trailer has total shipment cost of approximately \$465. In a double trailer, the same shipment costs approximately \$820. Finally, the third point, that (as for rail-boxcar movements) the marginal effect of increasing shipment weight is minimal, can be observed by inspection of any single cost curve. For example, a three-trailer movement of about 22 Mg (488 cwt) moving about 1600 km (1000 miles) has a total shipment cost of approximately \$1200 dollars, but increasing the load to about 60 Mg (1317 cwt) increases the total shipment cost to only approximately \$1304, an increase in cost of slightly more than 11 percent.

Cost-Estimating Procedure

The basic TOFC cost-estimating procedure is as follows.

1. Determine shipment characteristics in the same way as for rail car.
2. Compute number of TOFC trailers required for shipment by using Equations 1, 2, and 3. (For TOFC trailers, typical values are $V = 2550$ ft^3 and $W = 490$ cwt.)

3. Select applicable cost formula and compute cost: The variable cost (c_n) is given by

$$c_1 = 193 + 0.001\ 18S + (0.206 + 0.000\ 136\ 7S)L \quad (n = 1) \tag{9a}$$

$$c_2 = 378 + 0.001\ 18S + (0.379 + 0.000\ 136\ 7S)L \quad (n = 2) \tag{9b}$$

$$c_3 = 566 + 0.001\ 18S + (0.585 + 0.000\ 136\ 7S)L \quad (n = 3) \tag{9c}$$

$$c_4 = 750 + 0.001\ 18S + (0.758 + 0.000\ 136\ 7S)L \quad (n = 4) \tag{9d}$$

and the fully allocated cost (c_n') is given by

$$c_1' = 193 + 0.025\ 28S + (0.206 + 0.000\ 262\ 2S)L \quad (n = 1) \tag{10a}$$

$$c_2' = 378 + 0.025\ 28S + (0.379 + 0.000\ 262\ 2S)L \quad (n = 2) \tag{10b}$$

$$c_3' = 566 + 0.025\ 28S + (0.585 + 0.000\ 262\ 2S)L \quad (n = 3) \tag{10c}$$

$$c_4' = 750 + 0.025\ 28S + (0.758 + 0.000\ 262\ 2S)L \quad (n = 4) \tag{10d}$$

both in dollars per shipment.

Adjustments to Basic Procedure

Interchange and intermediate switching adjustments are similar to those used for rail-carload costs. The cost for x interchange switchings is given by

$$c_x = (30.79x - 0.042\ 66L)\,(n/2) \tag{11a}$$

and the cost for y intermediate switchings is given by

$$c_y = (10.14y - 0.033\ 610L)\,(n/2) \tag{11b}$$

where, if (n/2) is not an integer, it is rounded to the next higher integer.

HIGHWAY COMMON CARRIER

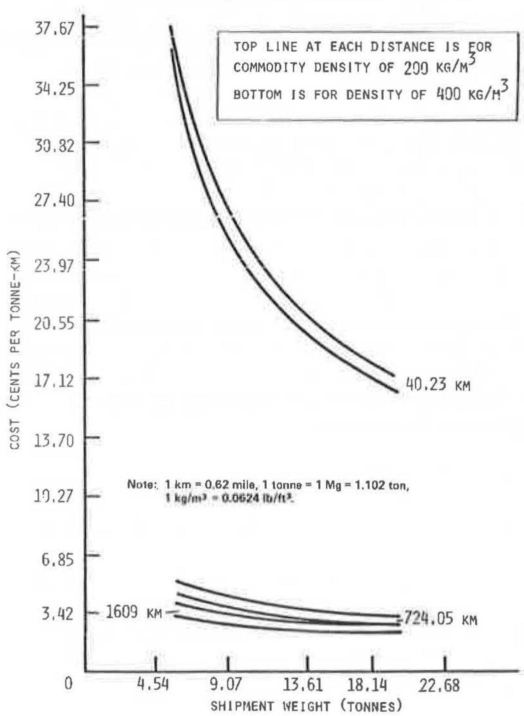
Highway, intercity freight-system costs are estimated by using the ICC statement, Cost of Transporting Freight by Class 1 and 2 Motor Common Carriers of General Commodities 1973 (3), which gives several tables of updated unit costs and operational characteristics that allow the analyst to develop cost scales for various shipment weights or weight brackets. As is the case for rail carload and TOFC, cost estimates are given for various regions or territories of the United States. For consistency, highway costs are here estimated for the eastern and central territory, which is similar to the rail industry's official territory, although it includes more of Illinois and none of Virginia.

There are two significant differences between the highway system cost-estimating procedures and those described above for rail boxcar and TOFC. First, it is not necessary that the analyst have explicit knowledge of the physical and operational characteristics of the highway service being costed. Data are given for a wide range of shipment weights and distances, and the assumption implicit in the associated unit costs is that the shipment moves in a service having average characteristics. This might mean, typically, that a small shipment would be first handled in local pickup-and-delivery service by a small vehicle suited to such operations and then be consolidated with other shipments with common destinations in a larger over-the-road vehicle for the intercity line-haul portion of the total movement.

The second difference, which is related to the first, is that in highway costing the density of the shipment being costed is explicitly taken into account. Since the capacity of a vehicle is limited by not only weight but also by volume characteristics, the highway costing technique, which does not include explicit determination of the number and type of vehicles required for a given shipment, must include some other method for taking variations in spatial occupancy of different commodities into account.

Figure 3 shows the cost characteristics of the highway mode. Cost is given by a band, rather than by a single curve, that shows the effect of different densities of commodities being shipped. The lower line at each distance is for the higher density commodity, and the higher line is for the lower density commodity. Although, intuitively, we would expect density to have effects on platform operations, on the pick-up and delivery portions of terminal costs, and on line-haul cost, it is taken into account in the source publication (3) by a weighted factor adjustment to the line-haul unit costs. As might be expected, cost decreases with increasing shipment weight, in part because of efficiencies associated with larger vehicles and in part because of a reduction in shipment platform handling by the carrier. (The probability that a shipment will be picked up at the consignor's dock and transported in a single truck to the consignee's dock without intermediate terminal handling by the carrier increases as the shipment size increases.) The rate of decrease in unit cost, however, becomes much lower above a shipment weight of approximately 18 Mg (399 cwt). One possible explanation is that this is the point beyond which vehicle size cannot be further increased so that the economies previously realized by increasing vehicle size are no longer available.

Figure 3. Highway common-carrier cost as a function of shipment weight and distance.



Cost-Estimating Procedures

The highway common-carrier cost-estimating procedure is as follows.

1. Determine shipment characteristics in the same way as for rail car.
2. Select applicable cost formula and compute cost:

The variable unit cost (C_s) is given by

$$c_s = 924S^{-0.537} + (8.828\,6 + 0.169\,01L)\,[0.855 + 1.32 \exp -0.144\,7\,(D - 2.5)] + (0.292\,93S^{-0.736})L \quad (S < 100) \tag{12a}$$

$$c_s = 70.51 - 0.290\,7(S - 100) + (12.107 + 0.180\,98L)\,[0.855 + 0.68 \exp -0.161\,76\,(D - 7.5)] + [0.009\,88 - 0.000\,063\,8\,(S - 100)]L \quad (100 \leq S < 200) \tag{12b}$$

$$c_s = 41.44 - 0.131\,7\,(S - 200) + (6.25 + 0.172\,75L)\,[0.855 + 0.68 \exp -0.161\,76\,(D - 7.5)] + [0.003\,5 - 0.000\,029\,9\,(S - 200)] \quad (200 \leq S < 300) \tag{12c}$$

$$c_s = 28.27 - 0.057\,1\,(S - 300) + (2.706 + 0.150\,9L)\,[0.855 + 0.68 \exp -0.161\,76\,(D - 7.5)] + 0.000\,51 \quad (300 \leq S < 437) \tag{12d}$$

$$c_s = 23.153 + 0.134\,06L + [0.102\,61 \exp -0.161\,76\,(D - 7.5)]L \quad (S \leq 437) \tag{12e}$$

The fully allocated unit cost (C_s') is given by

$$C_s' = 1.11C_s \tag{13}$$

The total highway common-carrier variable and fully allocated costs (C_s and C_s' respectively) are given by

$$C_s = c_sS/100 \tag{14a}$$

and

$$C_s' = c_s'S/100 \tag{14b}$$

all in dollars per shipment.

Accuracy

The truck costs estimated by equations are not as good a fit as those for the rail modes. The errors, at representative shipment weights and distances, are given in Table 2. It is necessary, therefore, that an analyst using these equations be aware of the potential errors involved, although for the most common shipment distances and weights, the total cost error is less than 5 percent.

COST UPDATING

Although the costs given above are useful in providing information on general cost behavior in the study year (1973), most applications would require more current estimates.

One way of updating is based on the Bureau of Labor Statistics (BLS) indexes of rail freight costs. (The BLS term cost corresponds to our term rate—the price charged to shippers by transportation firms—rather than to our cost—the cost to transportation firms of providing that service.) For example, in 1973 the index stood at 129.3 (4); its last reported level was 198 (February 1977) (5). If we assume that rate increases mirror increases in carriers' costs, then we can update the 1973 cost estimates presented here by increasing them by the ratio of the current to the 1973 rate index. In the example above, this would mean multiplying the 1973 rail-carload and TOFC cost estimates by 1.53. A similar set of indexes for highway common-carrier costs (rates) is scheduled for release by BLS in 1977.

COMPARISON OF SYSTEMS

In the previous sections, we have presented cost models for the three intercity freight carriers. In this section, we will present a comparison of their characteristics for a few sample origin-to-destination intercity movements, including several levels of shipment size and distance. In all cases, the variable cost will be used.

The TOFC and highway systems as they are described above are capable of providing a direct dock-to-dock service. However, some shippers or consignees do not have direct rail access and, therefore, it is necessary to include in the rail-system cost estimates provision for highway access at either origin or destination or both, if required.

The highway portions of these access segments are assumed to be similar to the pick-up-and-delivery operations included in the highway cost estimation, and their cost is estimated on this basis. The physical transfer of cargo is assumed to be a rail operation. On this basis, access cost at both origin and destination is composed of the origin-and-destination terminal cost included in the highway cost estimates: (a) pick-up and delivery, (b) highway platform handling, (c) billing and collecting, and (d) a rail platform-handling charge estimated at about \$3/Mg (13.6 ¢/cwt)/handling (i.e., at each end) (1, p. 149).

Another factor that must be taken into account before we can make direct comparison among the three intercity freight modes is circuity, the deviation of the path actually followed by a shipment from some common or reference distance between its origin and destination. In the absence of actual knowledge of this information, we will adopt the results of a recent study (6). Since the typical circuity values for rail and truck are so similar (1.25 for rail and 1.20 for truck) (and we know that cost estimates may well have variations of the magnitude of the difference between these two circuity

Table 2. Absolute and percentage deviations of costs estimated by equations according to shipping distance and value.

	40.2 km				644 km				1609 km			
	\$10 to 15		\$20 to \$30		\$10 to \$15		\$20 to \$30		\$10 to \$15		\$20 to \$30	
Shipping Weight (Mg)	Absolute	Percentage	Absolute	Percentage	Absolute	Percentage	Absolute	Percentage	Absolute	Percentage	Absolute	Percentage
0.169	1.3	8.0	1.3	8.0	1.6	7.6	1.7	8.5	2.3	2.3	8.7	9.3
0.292	0.3	1.3	0.2	0.9	0.2	0.7	0.3	1.0	0.6	0.6	1.5	1.6
0.581	0.6	1.9	0.4	1.3	-0.1	-0.2	-0.1	-0.2	-0.5	-0.2	-0.1	-0.4
1.261	-1.2	-2.5	-1.5	-3.1	-3.2	-4.2	-3.2	-4.5	-3.8	-3.9	-3.4	-3.9
2.921	3.1	4.4	2.4	3.5	-0.2	-0.2	-0.2	-0.2	0.2	—	0.1	—
5.797	17.4	19.8	14.7	17.3	7.7	3.5	7.4	3.8	11.07	10.2	3.0	3.3
11.06	13.1	10.7	12.3	10.6	7.8	2.4	3.2	1.2	28.5	12.2	4.9	2.5
15.39	-7.5	-6.1	-8.2	-7.0	-2.0	-0.1	-7.5	-2.4	13.5	-6.1	1.9	-1.1
19.82	-13.1	-9.6	-13.4	-10.4	-5.7	-1.2	-13.1	-3.4	14.0	-11.4	1.6	-1.6

Note: 1 km = 0.62 mile; 1 Mg = 1.102 tons.

values), we will assume that these systems may be compared on an equal-distance basis.

These assumptions concerning highway access to rail carload and the circuity of the three systems are included in the cost-estimating procedures given above.

Effect of Distance and Shipment Weight

We are now able to compare the costs per megagram kilometer of the three systems for the same transport service. Such comparisons are illustrated in Figures 4 and 5 for 1609 and 280-km (1000 and 175-mile) line-haul movements respectively.

Figure 4. Comparison of shipper-to-receiver costs for 1609-km shipment distance.

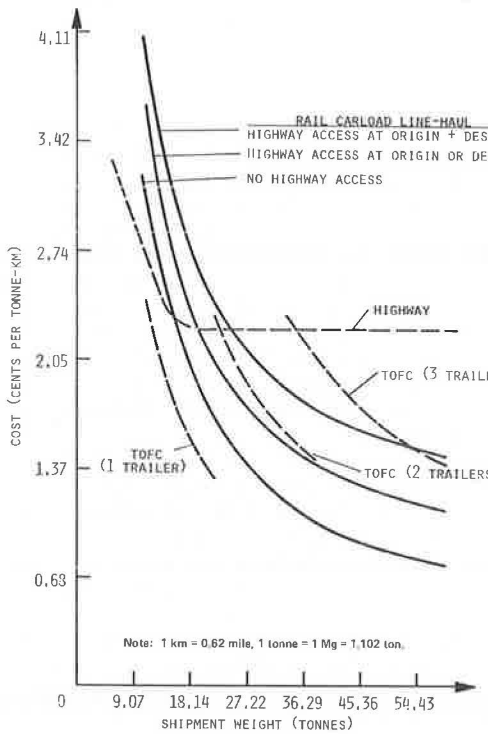
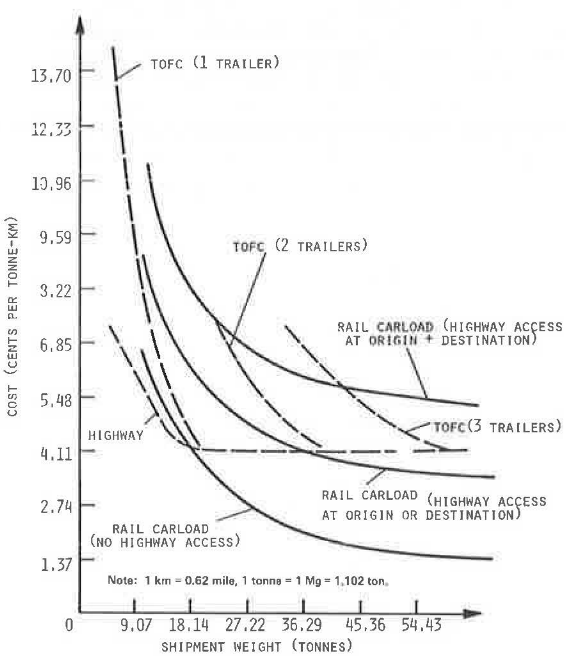


Figure 5. Comparison of shipper-to-receiver costs for 280-km shipment distance.



Analyzing Figure 4, we observe that truck is the least costly mode for shipments up to about 10 Mg (220 cwt), that TOFC is the least costly from that point up to a full (by weight) trailer load of about 22 Mg (488 cwt), and that relative ranking above that weight depends on whether or not highway access to the rail boxcar system is necessary. If access at either one or neither end is required, then conventional rail is the least costly for loads above about 22 Mg (488 cwt). If, however, highway access service to rail is required at both origin and destination, then TOFC is the low-cost mode for loads up to about 40 Mg (874 cwt) and above about 55 Mg (1220 cwt).

Figure 5 shows that the effect of distance on relative cost ranking is significant. For the shorter distance, the low-cost system is either highway or conventional rail, with the exception of a small region around 22 Mg (488 cwt), within which TOFC is less costly than highway. The relative rank of conventional rail and highway depends again on the extent to which highway feeder service to the rail line-haul is required. If it is not required, then rail is the low-cost system above about 16 Mg (360 cwt), and if it is required at both origin and destination, highway is least costly for all loads.

As we discussed in the section on TOFC costs above, drayage cost has an important effect on total TOFC cost levels. The results cited in the example will, of course, vary with changes in drayage, which was assumed here to be at average nonurban area (relatively low) levels.

Effect of Density and Shipment Weight

The assumption that weight, rather than volume, is the limiting characteristic in determining the number of vehicles or containers necessary to carry any given load is implicit in the system comparisons. Thus, a density of about 320 kg/m^3 (20 lb/ft^3) was assumed in the cost calculations. This explains the breaks in the TOFC cost curves: once a single trailer is loaded to its weight capacity, about 23 Mg (about 25 tons), any increase in load requires a second trailer. Similarly, a load of about 59 Mg (65 tons) in a typical boxcar requires a minimum density of about 430 kg/m^3 (27 lb/ft^3). These constraints were not addressed in the examples above, except as they are taken into account in the highway cost scales.

We can perform an analysis that illustrates the extent to which cargo density affects volume occupancy—the number of vehicles required for a given movement of specified weight and density—and, therefore, the density effect on relative cost ranking. The difference between this and the previous cost estimates is that these costs for highway and conventional rail systems were engineered on the basis of costs per basic unit of capacity and then used to determine costs per vehicles required, taking into account both weight and volume limitations of the vehicle used. Assume a load of about 22 Mg (488 cwt)—slightly less than the maximum highway trailer load by weight, trailers in both highway and TOFC service of about 65-m^3 (2300-ft^3) usable capacity (3, p. 23), and a typical rail boxcar of about 139-m^3 (4900-ft^3) usable capacity. For a cargo density of about 178 to 340 kg/m^3 (11.1 to 21.2 lb/ft^3), conventional rail boxcar is least costly for all distances, and TOFC is second least costly above distances of about 640 km (400 miles). For more dense commodities, highway is least costly at distances below about 200 km (125 miles), conventional rail between 200 and 1125 km (125 and 700 miles), and TOFC above

1125 km (700 miles). The higher density, of course, allows a single trailer (either highway or TOFC) to carry the entire 22-Mg (488-cwt) load. We see, then, that the effect of changes in cargo density may (as in our example) change the cost ranking of the three systems for some distance movements.

CONCLUSIONS

In this paper, we have presented cost-estimating equations for the variable and fully allocated costs of providing conventional rail boxcar, TOFC, and highway intercity freight services, based on ICC cost data and models. The paper covers three topics:

1. Presentation of the estimating equations and a comparison with the values obtained from the original ICC cost tables,
2. The basic costs of each mode and some discussion of the sensitivity of each to several different operating policies, and
3. An example comparison of the costs of each mode in the provision of total dock-to-dock freight transport service under various market conditions.

Some important conclusions about the characteristics of these three modes are that

1. The high cost of local access to line-haul facilities (shown here by highway access to rail line-haul) has significance for any multimodal system that includes existing technology and operating policies and specifically, potential improvements in line-haul cost and performance may be more than offset by the costs of local drayage and the transfer between the modes and
2. Variations in cargo density cause substantial variations in system costs, to the extent that the ranking by cost of the three systems may be changed.

Some conclusions about cost estimation in general are that

1. Although the costs given here are reasonable estimates of average costs in 1973 in the Northeast and Midwest regions, they are deficient in the following respects: (a) because of the regional average basis of both unit costs and service units, these costs would not be a sound basis on which to evaluate specific services in any but the most cursory analysis; (b) the current unsettled status of the rail system in the region might be resolved in such a way that institutional, operational, and managerial changes would cause the costs to become obsolete; and (c) the passage of time lessens the accuracy of any estimates based on historical data and
2. To correct these deficiencies will require research on the basic nature of transport system costs that goes far beyond the original ICC cost studies, which despite their widespread use for determining costs and rates, rely on many assumptions.

ACKNOWLEDGMENT

The research reported in this paper was partially supported by the U.S. Department of Transportation Program of University Research and the 1907 Foundation Endowment for Advanced Research in Transportation. This support in no way implies endorsement of the findings by the sponsors. This paper is an abridgment of a report (7) prepared for the U.S. Department of Transportation.

REFERENCES

1. Rail Carload Cost Scales 1973. Interstate Commerce Commission, Statement No. 1C1-73, Aug. 1975.
2. Explanation of Cost-Finding Procedures and Principles Relating to the Use of Costs. Interstate Commerce Commission, Statement No. 7-63, Nov. 1963, p. 119.
3. Cost of Transporting Freight by Class 1 and 2 Motor Common Carriers of General Commodities 1973. Interstate Commerce Commission, Statement No. 2C1-73, Nov. 1974.
4. Monthly Labor Review. Bureau of Labor Statistics, June 1975, p. 20.
5. Wholesale Prices and Price Indexes, February 1977. Bureau of Labor Statistics, April 1977, Table 14.
6. B. Hannon. A Railway Trust Fund. Transportation Research, Oct. 1974, p. 364.
7. Transport of Solid Commodities via Freight Pipeline. U.S. Department of Transportation, Vol. 3, Pt. A, July 1976.

Publication of this paper sponsored by Committee on Passenger and Freight Transportation Characteristics.

Abridgment

Evaluation of Potential Policies for Intercity Passenger Transportation in Canada

J. C. Rea, M. J. Wills, and J. B. Platts, Strategic Planning Group, Transport Canada

A quantitative, internally consistent assessment of the impacts of various policy options on national, multimodal, intercity passenger transportation has been studied. The results of the study are summarized in two reports (1, 2).

A multimodal travel demand model was developed to forecast travel by mode between pairs of cities. The model is responsive to (a) modal travel time and service frequency—reflecting the configurations and quality of the modal transport systems, (b) modal fares—reflecting fare structures, and (c) city population and linguistic characteristics—representing the demographic system. Each policy option was interpreted and analyzed to deduce its effect on modal level of service or fare structure. The resulting amended modal travel times, service frequencies, and fares were then used in the demand model to estimate the effect of the policy on modal demand.

The 1975 intercity air, bus, rail, and highway systems and subsequent variations were defined on the basis of 94 communities that represent two-thirds of the Canadian population.

DEMAND MODEL

A cross-sectional, quasi-abstract modal-travel demand model [Spatial Linkages Analysis Group (SLAG)] was estimated to assess modal demand response. This direct-demand econometric model was estimated in two steps: (a) a relative shares modal-split equation originally due to Warner (3) and later extended by Monsod (4), and (b) a total intercity-travel demand equation. Further details on the properties and empirical validity of the model have been given by Crow, Young, and Cooley (5) and by Crow and Savitt (6). The estimated form of the model is shown below.

$$T_{ijm} = [(\exp 4.12)(P_{ij}^{0.492} L_{ij}^{0.52})] \times \left[\left(\sum_m C_{ijm}^{-2.72} H_{ijm}^{-1.31} D_{ijm}^{0.128}\right)^{0.339}\right]$$
$$\times \left[(\exp K_m)(C_{ijm}^{-2.72} H_{ijm}^{-1.31} D_{ijm}^{0.128}) \Big/ \sum_m (\exp K_m) \times (C_{ijm}^{-2.72} H_{ijm}^{-1.31} D_{ijm}^{0.128})\right] \times F_{ijm} \quad (1)$$

where

- T_{ijm} = travel demand for city i to city j on mode m;
- P_{ij} = population cross products, cities i and j;
- L_{ij} = linguistic pairing index, cities i and j;
- C_{ij} = cost or fare (cents) of mode m from city i to j;
- H_{ijm} = travel time (h) of mode m from city i to j;
- D_{ijm} = departure frequency (per week) of mode m from city i to j;
- K_m = modal constants that may be interpreted as modal acceptability factors representing the unmeasured convenience involved in intercity travel (= -0.377, 0.979, 1.520, and 0 for air, rail, bus, and automobile respectively); and
- F_{ijm} = city-pair modal-specific adjustment factor.

In Equation 1, $(\exp 4.12)(P_{ij}^{0.492} L_{ij}^{0.52})$ is the travel-potential term, $[(\sum_m C_{ijm}^{-2.72} H_{ijm}^{-1.31} D_{ijm}^{0.128})^{0.339}]$ is the travel-impedance term, and $[(\exp K_m)(C_{im}^{-2.72} H_{ijm}^{-1.31} D_{ijm}^{0.128}) / \sum_m (\exp K_m)(C_{ijm}^{-2.72} H_{ijm}^{-1.31} D_{ijm}^{0.128})]$ is the modal-split term.

A calibration data set was developed for 1972, the latest year for which reliable data were available. The data set for each mode consisted of (a) modal-attribute variables—minimum travel time, weekly frequency of service, and economy-class fares between any origin-destination (O-D) pair; and (b) observed, annual aggregate travel demand by mode and O-D. Because automobile and bus O-D data were not available, these were estimated using link-flow data (7). The level-of-service attributes were seasonally weighted where appropriate.

Although the coefficients are measures of aggregate elasticity, it is also possible to obtain link and mode-specific elasticities by systematically varying the policy variables and computing the response from the model in terms of travel demand on a link and mode-specific basis. The results of such a procedure are shown in Table 1 for all city pairs in given distance intervals on a mode-specific basis. Thus, the effect of a constant-dollar 10 percent increase in long-haul air fares is a 13 percent decrease in long-haul air travel demand, other things remaining unchanged. In reality, such a change can be masked by an underlying structural trend toward more air travel. Service improvements may have two effects: (a) the diversion of demand from unimproved modes and (b) the generation of new intercity travel demand due to the substitution of this for local travel and other commodities. Both of these effects are represented in the elasticities presented in Table 1.

A comparison of SLAG-model elasticities with those obtained in other studies is given below.

Model or Data	Time	Cost	Departure Frequency
Monsod (4)	-2.15	-2.99	0.65
Crow and Savitt (6)	-1.94	-2.34	0.36
SLAG (2)	-1.31	-2.72	0.128
CTC (8)	-1.35	-2.59	–
Gupta and others (9)	-0.53	-2.17	–
Kraft and Kraft (10)	-1.84	-2.89	–

The model forecasts of 1975 domestic traffic at Toronto airport were within 6 percent of observed traffic volumes. In general, the model, when used for the comparative assessment of macropolicy implications at an aggregate level, appears to give a reasonable indication of modal demand response when policy variables are altered.

DEVELOPMENT OF INTEGRATED MODAL-SYSTEM OPTIONS

To construct integrated scenarios, 14 different modal systems were developed. All of them were described in terms of the same set of 94 community nodes although all nodes are not necessarily present in every system.

An integrated multimodal system can be created by grouping selected modal systems. By grouping together air, bus, rail, and highway systems of requisite characteristics, an integrated package can be created that reflects a desired option. Thirteen multimodal combinations were created as shown in Table 2, which describes general characteristics of each combination.

The O-D travel times and service-frequency attributes were computed by analyzing each modal system by using the Canada passenger system developed by Transport Canada.

DEVELOPMENT OF INTEGRATED MODAL-FARE OPTIONS

Ten different modal fare structures were initially constructed. From these elements, four integrated modal-fare options were assembled as shown in Table 3. The demand model does not include an income term because this proved not to be significant, which makes it necessary to deflate the above fare structures to 1972 values by using a disposable income index, so as to preserve the 1972 fare-and-income relation embedded in the demand model.

DEVELOPMENT OF INTEGRATED POLICY SCENARIOS

Combinations of the modal-system sets and modal-fare-

Table 1. Model elasticities.

Change In	Due to 10% Change In	Distance Range	Time	Cost	Frequency	Change In	Due to 10% Change In	Distance Range	Time	Cost	Frequency
			Change in Variables						Change in Variables		
Air	Air	1	-0.373	-2.193	0.128	Bus	Air	1	0.001	0.005	-0.000
		2	-0.421	-2.247	0.126			2	0.005	0.026	-0.001
		3	-0.536	-2.339	0.122			3	0.022	0.093	-0.005
		4	-0.360	-1.558	0.082			4	0.161	0.693	-0.037
		5	-0.370	-1.327	0.068			5	0.308	1.103	-0.056
Air	Rail	1	0.004	0.009	-0.001	Bus	Rail	1	0.005	0.011	-0.001
		2	0.010	0.023	-0.001			2	0.011	0.025	-0.001
		3	0.021	0.049	-0.002			3	0.022	0.050	-0.002
		4	0.047	0.099	-0.005			4	0.062	0.137	-0.007
		5	0.074	0.166	-0.008			5	0.101	0.219	-0.011
Air	Bus	1	-0.007	-0.016	0.001	Bus	Bus	1	-0.973	-2.082	0.129
		2	-0.010	-0.022	0.001			2	-1.077	-2.393	0.130
		3	-0.009	-0.020	0.001			3	-1.144	-2.571	0.129
		4	0.011	0.023	-0.001			4	-1.142	-2.603	0.128
		5	0.008	0.018	-0.001			5	-1.195	-2.659	0.127
Air	Automobile	1	0.860	1.784	0.000	Bus	Automobile	1	0.867	1.798	0.000
		2	0.846	1.753	0.000			2	0.853	1.769	0.000
		3	0.784	1.626	0.000			3	0.797	1.667	0.000
		4	0.193	0.684	0.000			4	0.286	0.859	0.000
		5	0.078	0.326	0.000			5	0.098	0.404	0.000
Rail	Air	1	0.001	0.006	-0.000	Automobile	Air	1	0.001	0.005	-0.000
		2	0.005	0.025	-0.001			2	0.006	0.028	-0.002
		3	0.025	0.095	-0.005			3	0.022	0.090	-0.005
		4	0.136	0.611	-0.032			4	0.131	0.542	-0.029
		5	0.328	1.147	-0.059			5	0.284	1.054	-0.054
Rail	Rail	1	-0.195	-2.160	0.128	Automobile	Rail	1	0.004	0.010	-0.001
		2	-0.027	-2.400	0.126			2	0.012	0.026	-0.001
		3	-1.023	-2.480	0.125			3	0.022	0.051	-0.003
		4	-1.134	-2.482	0.123			4	0.064	0.148	-0.007
		5	-1.058	-2.293	0.110			5	0.095	0.209	-0.010
Rail	Bus	1	-0.007	-0.015	0.001	Automobile	Bus	1	-0.007	-0.015	0.001
		2	-0.009	-0.020	0.001			2	-0.009	-0.020	0.001
		3	-0.007	-0.015	0.001			3	-0.008	-0.018	0.001
		4	0.005	0.011	-0.001			4	0.005	0.011	-0.000
		5	0.044	0.093	-0.004			5	0.012	0.027	-0.001
Rail	Automobile	1	0.862	1.788	0.000	Automobile	Automobile	1	-0.448	-0.935	0.000
		2	0.850	1.762	0.000			2	-0.463	-0.964	0.000
		3	0.777	1.643	0.000			3	-0.497	-1.059	0.000
		4	0.383	0.980	0.000			4	-0.481	-1.715	0.000
		5	0.092	0.402	0.000			5	-0.546	-2.323	0.000

Note: Distance ranges are 1 = 124 to 1232, 2 = 1232 to 2341, 3 = 2341 to 3450, 4 = 3450 to 4558, and 5 = 4558 to 5673 km respectively.

structure sets gave a total of 52 integrated system-fare policy scenarios. These were assigned alphanumeric designations from 1A to 13D by using the numerical system-set designations and alphabetic fare-set designations given in Tables 2 and 3. Obviously, many more scenarios could have been constructed from the elemental variables, but the time available constrained the scope of the study.

Each exploration of a policy scenario required that the demand model be presented with an integrated multimodal set of 12 variables (i.e., O-D fare, O-D travel time, and O-D service frequency for each of the air, bus, rail, and automobile modes) in addition to the demographic and linguistic index variables.

Forecast travel patterns for the year 1975, based on the 1A scenario, which closely approximates the real situation, serve as a basis for comparing the effects of the other scenarios. The forecast 1975 modal split, as function of trip length, is shown below.

Mode	All Trips	Short Trips	Medium Trips	Long Trips
Automobile	88.5	91.9	43.1	20.3
Air	5.4	2.7	39.3	58.2
Bus	4.5	4.2	8.7	9.7
Rail	1.7	1.2	8.9	11.8
Total	100.0	100.0	100.0	100.0

The market shares of intercity passenger kilometers are shown below.

Mode	All Trips	Short Trips	Medium Trips	Long Trips
Air	29.31	8.83	40.31	48.27
Bus	7.25	5.00	8.84	10.29
Rail	6.78	2.94	8.79	12.24
Automobile	56.66	83.23	42.07	19.20
Total	100.0	100.0	100.0	100.0

These differ considerably from the passenger-trip modal-split figures, as would be expected.

IMPACTS OF SCENARIOS

On Air Travel

The effect of imposing a cost-based air fare (B) is to decrease total air patronage by about 25 percent. Short-distance travel is reduced by one-third, and medium-length trips are reduced by about one-sixth.

The effect of an increase in air fares to reflect a cost of crude oil of $79/m^3 ($13/bbl) is to decrease air travel by about 5 percent. [Remember that, in this analysis, the other modes will suffer a simultaneous corresponding increasing in fares (i.e., fare set C).]

When gasoline costs are further increased to $0.33/L ($1.50/gal) in the above context (i.e., fare set D), total air travel increases by about 20 percent, and short-haul air travel increases by 50 percent.

Changes in the rail system do not have a great effect on air travel. However, the imposition of an 88-km/h (55-mph) automobile speed limit causes an approximately

Table 2. Multimodal system sets.

Designation of Integrated System Set	General Characteristics of Multimodal System Set
1	Existing 1975 systems
2	Existing 1975 systems amended to reduce service route kilometers by 25 percent
3	Amended 1975 systems with maximum automobile speed of 88 km/h (55 mph)
4	Amended 1975 systems with rail improved to give 129-km/h (80-mph) average speed on main lines and 80-km/h (50-mph) speed on branch lines
5	Amended 1975 systems with 129-km/h (80-mph) rail service on main lines and 80-km/h (50-mph) service on branch lines in the context of an 88-km/h (55-mph) automobile speed limit
6	Amended 1975 systems with 129-km/h (80-mph) rail service and 161-km/h (100-mph) service on Toronto to Montreal corridor
7	Amended 1975 systems with 129-km/h (80-mph) rail service and 161-km/h (100-mph) service on Windsor to Quebec corridor
8	As 6, but maximum automobile speed of 88 km/h (55 mph)
9	As 7, but maximum automobile speed of 88 km/h (55 mph)
10	Amended 1975 systems with 161-km/h (100-mph) rail service in Windsor to Quebec and Edmonton to Calgary corridors and 129-km/h (80-mph) regional daytime services, but all other rail services eliminated
11	As 10, but maximum automobile speed of 88 km/h (55 mph)
12	Amended 1975 systems with rail service offered only in 161-km/h (100-mph) Windsor to Quebec and Edmonton to Calgary corridors
13	As 12, but maximum automobile speed of 88 km/h (55 mph)

Table 3. Multimodal fare-structure combinations.

Designation of Integrated Fare Set	General Characteristics of Integrated Fare Set
A	Uniform intramodal fare structure based on 1975 fares
B	As A, but with cost-based air-fare structure; long-haul versus short-haul cross-subsidies for air partially removed
C	1975 fares increased to reflect a 100 percent increase in crude oil costs
D	As C, but automobile gasoline costs increased to $0.33/L ($1.50/gal)

5 percent increase in air travel, as compared to corresponding scenarios based on a 113-km/h (70-mph) freeway speed limit. As would be expected, short-haul air travel has a relatively larger increase (about 8 percent).

On Intercity Bus Travel

Air and bus demand are virtually independent of each other, and the imposition of cost-based air fares (B) causes only a slight increase in aggregate bus demand. An increase in the cost of crude oil to $79/m^3 ($13/bbl) (C) increases bus travel by 17 percent. When gasoline cost increases from $0.19 to $0.33/L ($0.86 to $1.50/gal) (D), bus travel increases dramatically and attracts twice as many patrons as in the 1975 (1A) situation.

Changes in the rail system have, in general, little impact on bus travel. Only when the spatial extent of the rail system is reduced (by 25 percent), does bus patronage respond to any noticeable extent (about 11 percent). The 88-km/h (55-mph) maximum automobile speed limit generally increases bus travel by about 11 percent. [It was assumed that bus travel times would not be changed by the imposition of an 88-km/h (55-mph) maximum automobile speed limit.]

On Intercity Rail Travel

The effects of fare sets A, B, and C are almost coincident. Cost-based air fares increase rail patronage by only 1.5 percent, indicating that rail does not substitute for air travel to any great extent. One might perhaps have expected rail to benefit when energy is expensive (C). In practice, however, rail-load factors in Canada are quite low. When the increased operating costs are passed on to the small number of patrons, the increase in rail fares (compared to the increase suffered by other modes) is great enough to actually deter rail patrons. Under fare option D, rail demand increases by 45 percent with almost all of the increase due to deterred automobile travelers switching to short-haul rail.

In the 1975 1A situation, 40 percent of the rail demand occurs on only 6 percent of the route kilometers. Consequently, a 25 percent reduction in route kilometers (principally by the elimination of underused branch lines and one of the two transcontinental services) causes a reduction of only 11 percent in patronage. This indicates that rationalization of the Canadian rail system could significantly improve its economic situation.

With a partially rationalized rail system, improving service to give 129-km/h (80-mph) average speeds and 80-km/h (50-mph) speeds on branch lines increases patronage to 40 percent of the 1975 patronage. The cost of achieving this 40 percent increase in demand would be billions of dollars. The addition to the above system of the 161-km/h (100-mph) average-speed service in the Windsor to Quebec corridor increases aggregate demand by only a small percentage.

When the rail system is reduced to one that provides six 129-km/h (80-mph) coterminal regional services and 161-km/h (100-mph) service in the corridor [a 68 percent reduction in route kilometers to 8850 km (5500 miles)], forecast rail patronage approximates the 1975 level. The elimination of all rail service except the 161-km/h (100-mph) service in the corridor [a 94 percent reduction in route kilometers to 1600 km (1000 miles)] would attract about 50 percent of the 1975 patronage. The effect of an 88-km/h (55-mph) automobile speed limit depends on the extent of the rail system and ranges from a 10 percent increase in patronage with a national rail system to a 15 percent increase when only a corridor system exists.

On Automobile Travel

The scenario of crude oil costing $79/m^3 ($13/bbl) (C) reduces automobile demand by about 13 percent. Increasing the cost of gasoline to $0.33/L ($1.50/gal) is forecast to decrease automobile travel by about 40 percent. (Remember that the demand model was calibrated on 1972 data, and thus reflects the behavioral patterns of that time. The subsequent energy price increases, inflation, and more energy-efficient automobiles since then would probably imply a much weaker response to operating cost increases.)

POSTSCRIPT

The study reported in this paper was the first attempt in Canada to carry out a national-scale, multimodal passenger transportation policy analysis. Because of the limited time available, the need to use available secondary data, and the fact that the demand model reflects 1972 conditions, the results should be interpreted with care. However, the results give a reasonable indication of the directions and relative strengths of modal-demand shifts in response to various policy options.

Since the above study was completed, an extensive data base has been established, a comprehensive analytical capability has been developed, and a new demand model based on a 1976 data set has been calibrated. Preliminary results indicate that sensitivity to fare has decreased considerably since 1972, that service frequency is more significant, and that sensitivity to travel time has increased slightly.

ACKNOWLEDGMENT

This paper is based on a study carried out by the Research Branch of the Canadian Transport Commission in 1975 to support Transport Canada's Task Force on National Transport Policy. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policy of Transport Canada.

REFERENCES

1. Intercity Passenger Transport in Canada: A Review of Existing Intercity Passenger Transport Systems. Canadian Transport Commission, Res. Rept. No. 252.
2. Intercity Passenger Transport in Canada: Analysis of the Consequences of Alternative Pricing and Network Strategies. Canadian Transport Commission, Res. Rept. No. 254.
3. S. L. Warner. Stochastic Choice of Mode in Urban Travel. Northwestern Univ., 1962.
4. S. Monsod. Further Developments in the Cross-Sectional Rail-Passenger Demand Model. Mathematica, Studies in Travel Demand, Vol. 3, July 1967; NTIS, Springfield, Va., PB 177 610.
5. R. T. Crow, K. H. Young, and T. Cooley. Alternative Demand Functions for Abstract Transportation Modes. Transportation Research, Vol. 7, 1973, pp. 335-354.
6. R. T. Crow and J. Savitt. The Impact of Petroleum Shortages on Intercity Travel and Modal Choice. Transportation Research, Vol. 8, 1974, pp. 383-397.
7. M. J. Wills. Linear and Nonlinear Estimators of the O-D Matrix. PhD dissertation, Univ. of British Columbia, 1978.
8. Intercity Passenger Transport Study. Canadian Transport Commission, 1970.
9. M. Gupta, S. Monsod, and K. H. Young. Contrast of Elasticities in the Rail Model With the Standard Abstract Mode Model. Mathematica, Studies in Travel Demand, Vol. 3, July 1967; NTIS, Springfield, Va., PB 177 610.
10. J. Kraft and A. Kraft. Empirical Estimation of the Value of Travel Time Using Multimode Choice Models. Journal of Economics, Vol. 2, 1974, pp. 317-326.

Publication of this paper sponsored by Committee on Passenger and Freight Transportation Characteristics.

Analysis of Truck Deliveries in a Small Business District

Gerd A. Ahrens,*Vienna
Keith W. Forstall,* Multisystems, Inc., Cambridge, Massachusetts
Raymond U. Guthrie,* Carnegie-Mellon University
Byron J. Ryan,* Operations Department, Burlington Northern

The nature of goods deliveries in a small commercial district in Pittsburgh was analyzed. The considerations in choice of the site were the expected cooperation from merchants, the diversity of store types available for analysis, and a high level of business activity. The surveys collected data on several aspects of the delivery process and its relationship to related activities in two phases. The first phase involved interviews at 59 stores in two blocks of the study site. The second phase involved recording truck pickup-and-delivery movements for one week in each block. The 400 observations covered the 8 a.m. to 12 n. period each day. The distributions in the data identified such things as hourly and daily delivery patterns, delivery times, and total number of deliveries to particular stores. Multiple linear regression was performed on the business data to test for equations that could predict the average number of deliveries per week. Regression on the movement-survey data was performed to test equations with the handling time of deliveries as the dependent variable.

Only recently has the urban goods-movement problem received a significant amount of attention from transportation planners. However, since 1970 interest in the question has increased, and it is now considered an integral part of the urban picture.

The object of this study was to identify and formulate the character of urban goods movement in a typical small commercial district. Specifically, the analysis was to

1. Validate and calibrate or refute the existing models that are designed to forecast the demands for goods movement that are associated with various categories of business establishments;
2. Collect and form the data base that is needed to study goods movement in a small urban business district made up primarily of single private owners; and
3. Analyze truck delivery and pickup patterns, including arrival and departure times, handling and dwell times, at-vehicle times, means of transporting goods, parking situations, and internal handling methods.

METHODOLOGY

Two of the primary considerations in the selection of the site were (a) the amount of cooperation anticipated from the businessmen in the area and (b) the inclusion of a typical variety of stores in a suitable district.

The Squirrel Hill business area of Pittsburgh was chosen because one of its two major streets was scheduled for renovation. Many of the local businessmen were concerned that goods delivery service during this period could not be maintained. It was felt that one of the primary considerations in site selection, that of business cooperation, could be met because of this impending problem. In addition, the Squirrel Hill business district includes a typical variety of stores suitable for the study. At the beginning of the study, an agreement of cooperation in carrying out the study and a joint letter to the businessmen was prepared in conjunction with the local merchant's organization.

The data collection was accomplished in two steps. First, a business survey was developed to obtain information about a sample of at least 50 percent of the stores that were observed in the movement survey. The data in the first section of the business survey were determined without asking the store manager. The information included address, type of store, and products sold. It also included the location of the store in terms of distance to traffic signals and type of road, to obtain information about the impacts on traffic of goods movements.

The second section of the business survey required an interview. It was composed to obtain as much information as possible in less than 10 min; to make the answers easy to understand and reliable; and to be not too detailed or confidential, which might interfere with the cooperation of the businessmen. This section included information about floor space and number of employees and was used to correlate numbers of generated trips with the structures of the stores. As other studies have shown, it is difficult to obtain data on inventory turnovers (4). Therefore, the survey asked for the number of sales per day, which is equivalent to customers served.

All but one or two of the businessmen agreed to allow observation of the deliveries that they received. Overall, they were very cooperative and provided more information than was asked for in the survey.

The second step in the data collection was a movement survey that was developed to gather information from interviews of truck drivers and through observation of the pickup and delivery processes. As with the business survey, cooperation was generally good.

The data were collected between 8 a.m. and 12 noon. From the business-survey questionnaire, it was con cluded that about 80 percent of the deliveries and pickups in the study site could be observed during those hours. The information collected included parking locations, delivery times, goods delivered, delivery-process used, and general truck-route details.

ANALYSIS OF BUSINESS SURVEY

A sample of 59 business establishments was taken on two streets. This included 15 clothing stores, primarily on Forbes Avenue, and 14 retail food stores, primarily on Murray Avenue. Figure 1 shows the distribution of answers to six of the interview questions. The bar graphs compare the characteristics of all stores with those of clothing and retail food stores.

More than 50 percent of the establishments have only one entrance, which is also used for deliveries, and almost all of them (97 percent) have no loading docks. There are some side doors (28 percent of all stores), with retail stores having a higher percentage (43 percent), but few of them are used for deliveries.

The floor-space distribution shows that most of the establishments are small to medium sized, which means that they have less than 372 m^2 (4000 ft^2) of business space. Ninety-five percent of the retail food stores were this size, while clothing stores tended to be larger.

Quite a few of the clothing stores have more than 15 salespersons, while the retail food stores are generally small family-type establishments with fewer than five employees. Almost 30 percent have only the owner and one helper to run the store. On the whole, small stores are commonest. This is also confirmed by the distribution of sales per day. The fifth bar diagram, which indicates the number of enterprises ordered from, shows that most of the stores order from more than 20 different suppliers. This is particularly true for clothing stores. Only 13 percent of them order from fewer than five different suppliers. These are either highly specialized firms, such as a fur store, or branch-store boutiques, which obtain their goods from the main stores. The number of food stores in this low category is also significant. These stores are typically delicatessens, meat markets, and other specialized small places without an extensive variety of different products.

The distribution of deliveries per week was similar for the three categories of all stores, clothing stores, and retail food stores.

Because graphic illustration alone is inadequate to establish the store characteristics that determine the nature of related goods movements, regression analysis was performed on the complete set of business data and then on a data subset that included only businesses that sell at least two different types of products. This second set of regressions was performed to test the conclusions of a similar study conducted at the Polytechnic Institute of New York (PINY) (1). The dependent variable for all regressions was the average number of deliveries per week.

The table below shows the correlations of a number of variables for the complete data set.

Dependent Variable	Weekly Deliveries (NW)
Number of product types sold (PN)	-0.1634
Number of enterprises delivering (EN)	0.1420
Floor space (000 ft^2) (FL)	0.2883
Full-time employees (EMF)	0.5434
Part-time employees (EMP)	0.1509
Total employees (TEMP)	0.5525

The strongest correlation is that between the number of employees and the number of weekly deliveries. The only apparent anomaly is the negative correlation between the number of products (referred to in the PINY report as the specialization index) and the number of weekly deliveries. This may be representative of the high number of deliveries associated with food stores and restaurants, which deal in only one type of product.

The best regression equation developed for all businesses was the following:

$$NW = 0.187EN - 0.888PN + 1.04FL + 0.471EMF + 0.738EMP \quad (1)$$

The explained sum of squares for this equation is 0.64. (These equations were developed for U.S. customary units; therefore SI units are not given for the variables.) An equation with equivalent performance for the explained sum of squares is the following:

$$NW = 0.21EN + 0.58TEMP \quad (2)$$

This equation, however, did not perform as well on the chi-squared normality test.

The PINY equation for stores selling more than one type of product had the following form:

$$NW = 9.0PN - 16.6 \quad (PN \geq 2) \quad (3)$$

but an equation of this form failed to explain the Pittsburgh data. On the other hand, an equation similar to

Equation 1 performed even better than it did for the full data set. For the data subset, the best equation was as follows:

$$NW = 0.438EN + 1.89PN - 6.82FL + 1.11EMF + 0.534EMP \quad (PN \geq 2) \tag{4}$$

The explained sum of squares for this equation is 0.83 and the chi-squared value for nine degrees of freedom was 5.76.

ANALYSIS OF MOVEMENT SURVEY

The daily patterns of deliveries and vehicle stops are shown in Figure 2. On Forbes Avenue there was a rather regular distribution on Monday through Wednesday, a slight increase on Thursday, and a peak on Friday. From an average of 48 deliveries/d at the beginning of the week, there was a 35 percent increase on Thursday and a 55 percent increase on Friday. The distribution of vehicle stops was more homogeneous, with an increase of 37 percent on Thursday and of 31 percent on Friday. A large percentage of the additional deliveries were made by the United Parcel Service (UPS) truck.

Murray Avenue did not show such a significant number of multiple deliveries. Monday, Wednesday, and Thursday are rather similar days on Murray, while Tuesday is the least busy and Friday the most busy day. Here there was an increase of about 40 percent for Fridays, which is different from the 5 percent average increase observed in Brooklyn for Thursdays and Fridays (1) and the uniformly distributed deliveries in the German town of Braunschweig (2, 3).

There was a peak period for deliveries after 10 a.m., which is comparable to the 10 to 12 a.m. peak in Brooklyn. A distinct difference between food deliveries and other deliveries is that food deliveries are more frequently made in the early hours.

There are some obvious differences in the dwelling, delivery, at-vehicle, and delay times between food stores and clothing stores (Figure 3). The average dwelling time for food stores (19.8 min) is close to the Brooklyn value of 22 min. The 33.5-min average dwelling time for clothing stores, however, is misleading, because the UPS truck makes multiple deliveries from the same curb space.

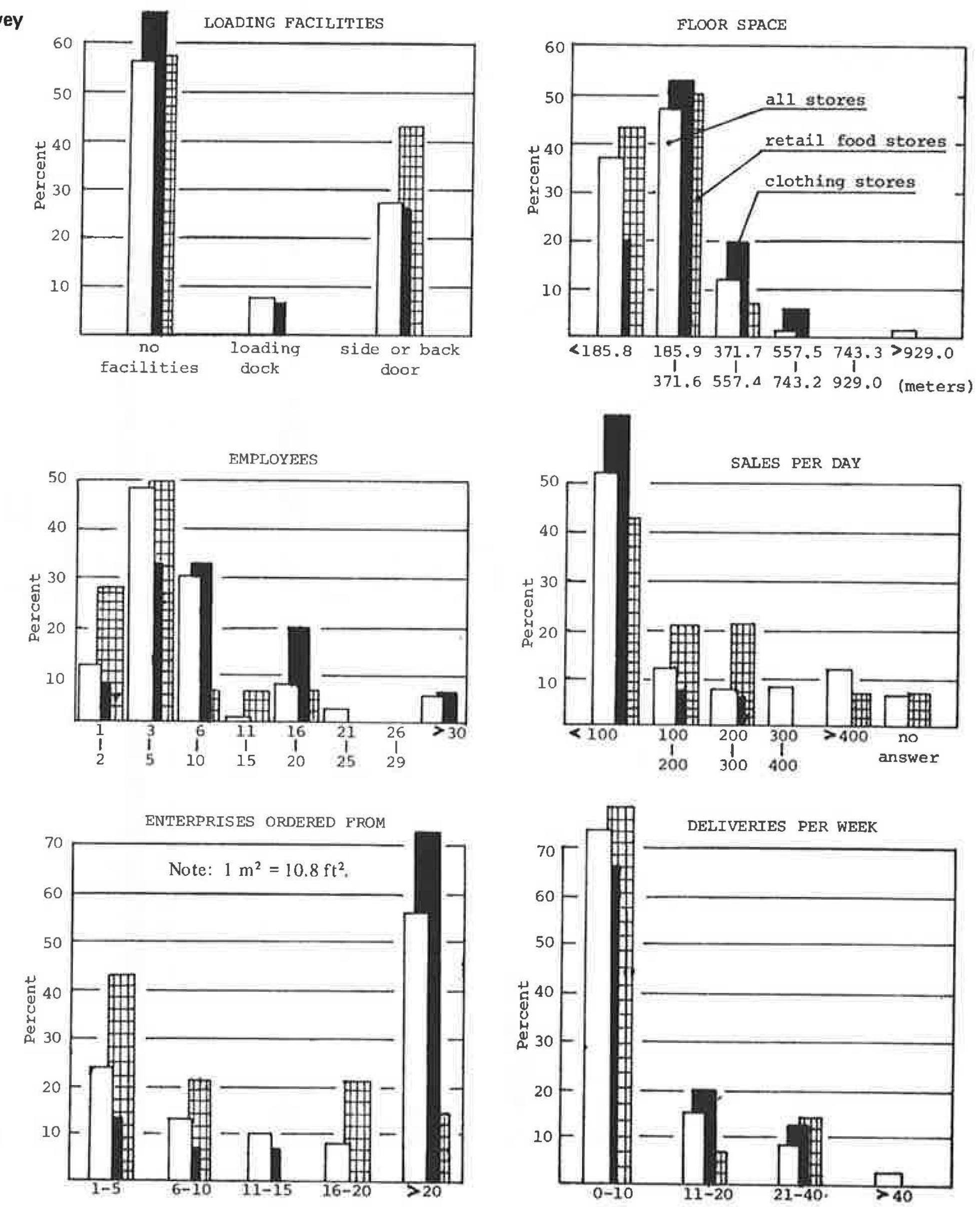


Figure 1. Distribution of business-survey responses.

The delivery times represent a more realistic impression. The unloading of trucks with food usually took longer than that of any other commodity. The reasons for this are the larger size of the shipments, the larger number of smaller packages, and often the unpacking and setting up of displays inside the store.

The at-vehicle times were not always easy to estimate. The high average value for clothing stores is again caused by the UPS truck, because the at-vehicle times were not individually recorded, but were computed by subtracting the delivery times and delays from the dwelling times and then dividing the result by the number of multiple deliveries per stop. The at-vehicle time for single deliveries could not always be recorded.

Four subsets of the movement data were used in the regression analysis. Subset A, 378 observations, included all of the data for single and multiple deliveries and single and multiple pickups and deliveries. Subset B, 167 observations, included the data for single deliveries only. Subset C further restricted the single-delivery data to only those deliveries in which the means of transport was by one-man carry and included 87 observations. Subset D, 61 observations, restricted the single-delivery data to deliveries in which the means of transport was by dolly or hand truck.

The dependent variable selected for the most analysis was delivery time. Preliminary regression runs eliminated most of the potential independent variables from consideration. The variables that remained—number of packages, total weight, number of persons per truck, and means of transport—were used as explanatory variables in most of the regression analysis. The table below shows the coefficients of correlation between these variables and delivery time.

Figure 2. Daily vehicle stops and deliveries on Forbes and Murray avenues.

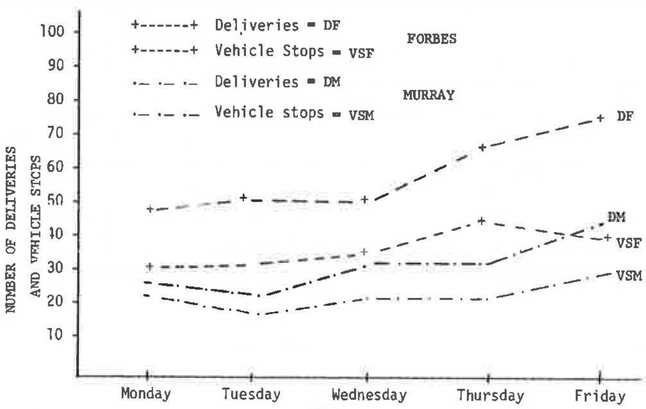
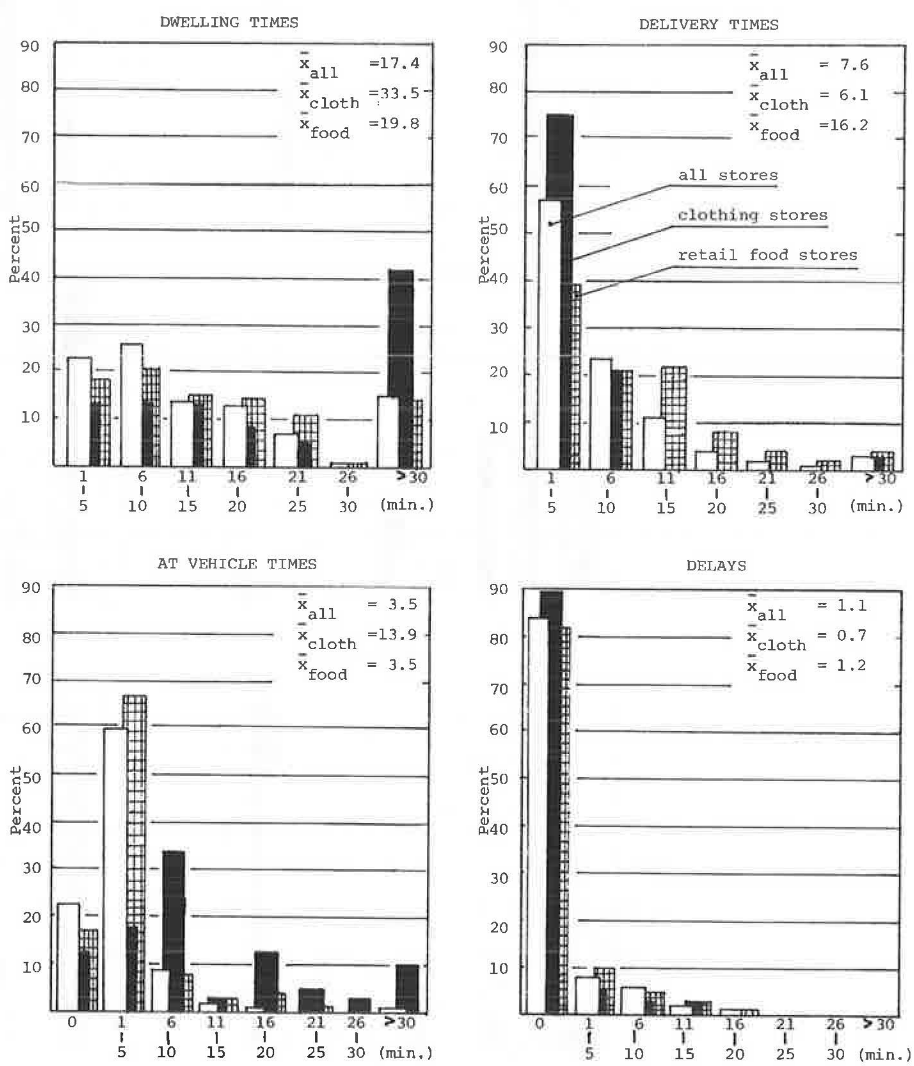


Figure 3. Comparison of delivery characteristics.



Dependent Variable	Delivery Time (T) (min)
Number of packages (P)	0.7215
Total weight (W)	0.7699
Number of persons in truck (N)	0.1413
Means of transport (MT)	0.2648

The PINY work obtained the best results by using delivery time as the dependent variable and number of packages and total weight as the independent variables. The basic equation was

$$T = 1.9 + 0.11P + 0.008W \quad (r^2 = 0.62) \tag{5}$$

where

T = delivery time,
P = number of packages, and
W = total weight.

The best results in testing Equation 5 were obtained by using the data subset that represented single-delivery data only. The equation was

$$T = 4.453 + 0.177P + 0.007W \quad (r^2 = 0.68) \tag{6}$$

The use of the other three data subsets in testing this equation caused little change in the constant term, some changes in the coefficients of the explanatory variables, and significantly lower r^2 values.

Better results were obtained with our data by dropping the constant term from the equation. The equation developed by using the data for single deliveries only was

$$T = 0.28P + 0.008W \quad (r^2 = 0.76) \tag{7}$$

Further improvement in the explanatory powers of the equation was obtained by including the variable for the number of persons per truck. Equations 8, 9, and 10 were developed by using the data for single deliveries only, one-man carry, and dolly or hand truck respectively.

$$T = 0.196P + 0.007W + 3.601N \quad (r^2 = 0.84) \tag{8}$$

$$T = 0.041P + 0.014W + 3.856N \quad (r^2 = 0.80) \tag{9}$$

$$T = 0.302P + 0.001W + 4.578N \quad (r^2 = 0.84) \tag{10}$$

where N = number of persons per truck.

Quantitatively, these results are in substantial agreement with those of the PINY research. The differences in the equations, which may be considered site specific, predict longer delivery times than do those in the PINY equation in both the constant and nonconstant forms. Clearly, there is a need for quantitative analysis of additional study sites having different characteristics to provide a better site basis for comparison.

QUALITATIVE FINDINGS

A number of behavioral characteristics were observed in the study area that could not be adequately quantified. These findings were a composite of the observers' intuitive notions, information supplied by vehicle operators, and opinions expressed by several of the merchants.

The issue of parking legality was of interest because both Forbes and Murray avenues are extensively metered throughout the study area. Not once during the 10 d of observation was a delivery-vehicle operator ever observed inserting coins in a meter. Illegal parking zones in front of fire hydrants, bus stops, driveways, and near intersections were extensively used by delivery vehicles with no evidence of reprimands from policemen. Because the periods of occupancy of these zones were generally quite brief, very little or no interference with their intended purposes was observed. This seems to suggest that many illegal parking zones could function as temporary loading zones with a minimum of interference to normal community operations.

Vehicle operators almost always turn off their engines while making deliveries. Exceptions were noted during cold weather when some operators kept their engines running to provide heat for the cabs of the vehicles. In addition, some refrigerated vehicles have engines left running to keep the perishable payload cooled, and some rubbish vehicles have engines left running to power the hydraulic compactors.

Delays experienced by vehicle operators were recorded in the movement survey, but causal factors were not formally recorded. A general consensus of operators attributes the longest delays to searching for establishment owners or the persons responsible for unlocking delivery-service doors. In one particular instance, a vehicle operator attempted to deliver to a restaurant three different times before successfully gaining access to the basement storage area, because the manager was late arriving at the establishment. Delays were also observed when goods were being picked up. Quite often, packages were not prepared for shipment until the vehicle arrived. Delays were observed from time to time in connection with signing the bill of lading and because of conversations carried on between operators and establishment owners or employees.

There was less correlation between vehicle capacity and shipment size than was expected. Several large straight trucks with payloads of less than five pieces made deliveries during the observation period, but small vans and station wagons with all available space occupied by their shipments were also observed. There was a great deal of variation in packing efficiency among vehicles. The UPS vehicles have shelves catalogued by the names of the establishments along the vehicle route. This system eliminated much of the en-route sorting required of vehicle operators. In contrast, U.S. postal-vehicle operators generally spend a great deal of time along the route sorting packages.

Most vehicle operators were quite concerned with operating efficiency. However, there did appear to be a relation between prevailing weather conditions and operating efficiency. No hard data were collected to support this contention, but observations were made through a wide range of conditions, including 15°F (-9.5°C) with snow showers and sunny days with mild temperatures. There was clearly a reduction in operating efficiency on those days when the weather was inclement.

One example of the general concern for operating efficiency was found in an operator's route plan that served establishments on the east side of a street, went on to the next community, and then served establishments on the west side of the street during the return trip. In cases of pickup and delivery, very few steps were wasted as operators used return trips to the vehicle to transport packages that were being picked up. Some operators expressed a fear that members of the observation team might be time-study personnel representing their employers. However, we believe that what we observed was an accurate portrayal of normal daily operations.

The experiences with the survey completion, data computerization, and data analysis of the business and movement surveys all led to the development of certain conclusions about desirable format and content improvements.

Deliveries that involved one man with a dolly demonstrated the importance of package volume on delivery time. The dolly made it possible for one man to transport a full load of even the heaviest goods, yet required two or more trips for relatively small numbers of packages of large size. Because the number and weight of the packages was less critical in these instances than their number and volume, an estimate of the volume of packages delivered should be included in the movement survey.

It would also be valuable to include a question about the number of store employees involved in the delivery process. The relevance of this was particularly evident in loading-dock deliveries, where packages were typically transferred by a bucket brigade consisting of the driver and one or two store employees. This was one of the fastest delivery systems observed.

At the conclusion of movement surveys in a location, a check should be made with the stores to establish whether the week of observation was a normal one for deliveries. This permits eliminating significantly atypical observations.

SUMMARY AND CONCLUSIONS

The survey format used had the advantage of collecting both empirical and interview data, which allows cross-checking of the results for consistency. The data-collection process was designed specifically to acquire as much data as possible with a very small staff. The movement data were collected in two concentrated areas chosen to provide information about a wide variety of retail establishments, with specific emphasis on food and clothing stores. The observation periods were chosen so that daily patterns would be observed. The hours of observation were chosen to coincide with high levels of delivery activity. The collection periods were chosen to allow comparisons with the results of the business survey.

The distribution analysis indicated similar daily patterns of deliveries. The peaks generally occurred on Friday, regardless of type of establishment. The average Friday delivery rate was about 30 percent higher than the delivery rate on other days. This peak was somewhat higher than that found in the PINY study. Hourly patterns of various retail types were not markedly different, although food stores generally received goods earlier than did other stores.

Analysis of the business-survey data indicated that delivery frequency can be generally correlated to employment, floor space, and product diversity. While the PINY analysis was proven to be too simplistic in this respect, this data base must also be considered too small to permit extensive generalization.

The movement survey was used to find indicators of delivery times. The variables that appeared to have the most correlation were the number and weight of packages. This is in agreement with the PINY findings although the coefficients of the equation developed for the Pittsburgh data were generally higher than those for the New York data. This appears to be an indication of site-specific differences. The volume of a shipment is also fundamental to the delivery time required.

Analysis of the present data base is not complete. Future work will include analysis of other data disaggregations to further determine the characteristics that are most critical in defining the delivery process. Cross-checking the results of the two surveys will aid in identifying consistent biases or errors in interview responses. The effects of the delivery process on traffic congestion are largely unresolved.

The present study leaves as many questions as answers, but it adds needed insight to the urban goods-movement problem. The specific equations found in the analysis are of limited applicability, but the importance of general relationships (and nonrelationships) cannot be overstated. In particular, the points of correspondence and contradiction vis-a-vis the PINY results indicate the types of generalizations and models that may be valid under more extensive testing. The problem of deriving general conclusions from site-specific data is an acute one in urban goods-movement research because there are so many variables distinguishing commercial districts from each other. To date, no clear evidence exists to distinguish the variables that are descriptive from those that are irrelevent. It is hoped that the present study has helped to narrow the focus toward some of the more important relationships.

There is a need for a coordinated series of studies to compile data that are significant to goods deliveries. These data should be collected on a common basis from a wide variety of sites to define valid general relationships and important site-specific variables. Ultimately, a better understanding of these topics can provide tools for planners to use to anticipate the nature of goods movements that any commercial area can be expected to generate.

REFERENCES

1. K. W. Crowley and others. Facilitation of Urban Goods Movements. Department of Transportation Planning and Engineering, Polytechnic Institute of New York, First Year Final Report, 1974.
2. W. Schwerdtfeger. Staedtischer Lieferverkehr: Bestimmungsgruende, Umfang, und Ablauf des Lieferverkehrs von Einzelhandels und Dienstleistungsbetrieben. In Veroeffentlichungen des Instituts fuer Stadtbauwesen (H. Habekost, ed.), Technische Universitaet Braunschweig, W. Germany, No. 20, 1976.
3. W. Schwerdtfeger, Wirtschaftwerkehrsaufkommen und Warenauslieferung grosser Einzelhandelsbetriebe: Zwei Beispiele. In Veroeffentlichungen des Instituts fuer Stadtbauwesen (H. Habekost, ed.), Technische Universitaet Braunschweig, W. Germany, No. 14, 1974.
4. Trucks at Work in Selected Square Miles of the Tri-State Region. Tri-State Regional Planning Commission, Internal Technical Rept., New York, July 1968.

Publication of this paper sponsored by Committee on Urban Goods Movement.

**The authors were students at Carnegie-Mellon University when this analysis was completed.*