# CORRELATION OF SOCIOECONOMIC FACTORS WITH CORRIDOR TRAVEL DEMAND

Gerald R. Brown, Department of Civil Engineering,
University of British Columbia

Current interest in public transit to alleviate the urban transportation problem requires more research about the effects on urban structure of new modes. The purpose of this paper is to examine one aspect of this: the internal relationships between the socioeconomic characteristics of commuters and the transportation service characteristics they value in their mode choice to work. Specifically, the study is concerned with (a) the correlation of selected socioeconomic factors of individuals with their choice of travel attributes and (b) the effect of socioeconomic factors in causing car drivers to shift mode. Canonical correlation analysis illustrates that each modal group using a transportation corridor has unique socioeconomic characteristics and that these are related to the attributes of the transportation system. Car drivers are less dependent on system attributes than bus riders, and therefore planned changes in the system (such as the introduction of rapid transit) will have less effect on this group than on bus users in terms of use of the system. There is also some indication that changes in bus frequency would have an effect on car ownership in a given corridor. Statistical tests, using discriminant analysis, on a subsample of automobile drivers indicates that the socioeconomic characteristics of an individual probably exert an influence on his tendency to shift mode and on his sensitivity to specific transportation service level changes in the system. In particular it was found that as income levels increase the tendency to shift mode decreases. The findings of the study have implications for mode split planning, travel demand modeling, and urban structure.

•CURRENT interest in public transit as a means to rationalize urban transportation requires a better understanding of the effects of new modes on both travel generation and modal split. The purpose of this paper is to examine one aspect of this problem: the interdependent relationship between the socioeconomic characteristics of commuters in a travel corridor and the attributes of the transportation mode serving the corridor. The findings have implications for mode split, travel demand, and—indirectly—corridor development.

Research into mode split has shown that there are two major influences on the mode choice: (a) the socioeconomic makeup of the traveler and (b) the transportation system characteristics such as travel time, travel cost, and convenience. Early mode split methods can be grouped into those that were based primarily on the first of these influences and those that were based on the second (1). In the evolution of later mode split methods, the influence of income and car ownership has been considered by inference in the interpretation of model results (2). But by and large the variables used in recent models are primarily measures of transportation system characteristics, which permit the prediction of mode split when the magnitude of these variables is changed in a way that would simulate new modes. This process appears sound, given the assumption implicitly held in mode split analysis: that the introduction of a new mode has no effect on the absolute number of tips demanded but only on the mode split.

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

But what happens if this assumption is released? Indications from land use modeling studies are that a change of accessibility in a travel corridor will have some effect on land use activity, usually defined by the changes in population and employment structure in the urban region. This implies a change in travel demand—and a change in the socioeconomic structure of a transportation corridor—with the introduction of a new mode. Several authors have addressed this problem in the form of travel demand modeling, including Kraft and Wohl (3), Domencich, Kraft, and Valette (4), Brand (5), and Manheim (6). The thrust of the work of these authors is that travel demand is a derived demand based on the desire for activities at trip destination. Therefore, travel demand is a function of land use activities and transportation service. The essential structure of the early demand models (3, 4) included the classic urban transportation planning submodels of trip generation, trip distribution, and mode split. But the effect of transportation service on trip generation and trip distribution was not included in these early models. More recent studies have attempted to consider this feedback effect by general equilibrium models that model both travel flow and land use effects as a set of equations to be solved either simultaneously (5) or sequentially (6).

As a by-product of the mode split and travel demand modeling effects, inferences have been made about the effects of socioeconomic variables as well as the system variables in the demand function. However, for future operational travel demand models, including the feedback loop, our level of understanding of the internal structure of travel decisions and travel demand dimensions needs to be improved. This paper is an attempt to explore the internal structure of demand as it relates to urban structure changes and the mode split. The objectives of the research are (a) to assess the explicit relationship between socioeconomic characteristics of commuters and choice of system characteristics in the journey to work and (b) to determine what potential influence socioeconomic characteristics have in causing an individual to shift to a new mode.

The study method is an empirical analysis of a sample of automobile drivers and bus passengers of a commuter travel corridor with destinations in the central business district of Vancouver, Canada. The commuter-shed used in the analysis is in most important respects a typical high-income suburban area with some high-density residential districts. Incomes, car ownership rates, and occupational status are higher than for the region as a whole. The mode split in the corridor is about two-thirds automobile and one-thirds bus commuting. The mean travel time for automobile drivers is 31 minutes, and for bus commuters it is 37 minutes.

## INTERDEPENDENCE OF SOCIOECONOMIC AND TRANSPORTATION SYSTEM ATTRIBUTES

If a travel corridor undergoes population structure changes with a change of accessibility, a knowledge of the cause-effect linkages is important to assess the specific impact of a new mode. The assumption is that equilibrium exists in a travel corridor between the type of people who commute and the characteristics of the transportation system. For example, if a corridor is served by a toll freeway that gets commuters to work quickly, the freeway will over the long run attract people to its influence area who are highly sensitive to travel time savings in the journey to work vis-à-vis the population as a whole. On the other hand, a corridor served by public transit, if travel costs are low, may attract people to its influence area who are cost-sensitive. In practice a travel corridor would likely have at least two influence areas, each of which would be defined as the group of persons who use a specific mode. It is probable that persons of any specific mode group are similar types and that this group will differ from other mode groups in important ways. It is also probable that members of each mode group will value alike the attributes of the transportation alternatives available and choose that combination of attributes (i.e., mode) that minimizes his level of dissatisfaction. In contrast, the members of a different mode group will select a somewhat different set of attributes. If there is an interdependence between the socioeconomic factors of a mode group and the travel attributes of the mode, and if the nature of this interdependence changes from mode group to mode group, we can begin to predict the socioeconomic composition of a transportation corridor if a new mode (i.e., a new set of transportation attributes) is introduced. The primary implication of this approach is that urban development models can be formulated on a disaggregated behavioral basis to follow sequentially the behavioral transportation flow models now in use.

## Preliminary Data Analysis

The data matrix consisted of 1,244 car drivers and 967 bus passengers. Twelve variables representing the socioeconomic and transportation characteristics of the sample were available for analysis. The variables were divided into two sets of six: $X^{(S)}$, representing the socioeconomic characteristics, and $X^{(T)}$, representing the transportation system characteristics. Set $X^{(S)}$ consists of age (AGE), occupation (OCC), the number of times per week a car is necessary at work (CNEC), car ownership (COWN), family income (INC), and the availability of a car for the work journey (CARA). Set $X^{(T)}$ consists of travel time at origin of trip (TTO), travel time at destination (TFRM), total travel time (OTT), out-of-pocket expenses for either parking charge or transit fare (OPE), OPE divided by income (E/INC), and bus frequency at origin (FREQ). A full description of the variables used in the analysis is in the Appendix.

Chi-square tests of comparable variables (Table 1) indicate there are two distinct socioeconomic groups associated with the two modes. For each mode, group socioeconomic characteristics are relatively homogeneous compared to the difference in characteristics between modes. Commuters in the prime working age groups of 40 to 60 are car-oriented while younger and older age groups are bus-oriented. Indications are that substantial commuters in the 20 to 40 age group prefer the bus. Occupational structure is related to mode choice somewhat differently than expected in that a high proportion of managers and professional employees use the bus. Also, a substantial proportion of unskilled workmen and clerical employees—groups that are often thought to be transit "captives"—are car-oriented. Secretaries and sales workers are transit-oriented as expected. Car ownership is high in this case and may be expected to moderate the usually high correlation between car ownership and mode split. Income shows the expected trend, with high-income households associated with car mode and low-income ones with bus use.

## Statistical Analysis

Canonical correlation analysis was used to investigate the relationship between the socioeconomic variables of each mode group and transportation system variables. This technique gives an optimal weighting and combination of the socioeconomic variables so that the combination will maximally correlate with the best combination of the transportation system variables. It tells us how the two sets of variables are related to each other and how the variables within each set contribute to this relationship. The conceptual implication is that an individual of a particular socioeconomic makeup will choose that combination of transportation service attributes that, in combination, minimizes his travel dissatisfactions. By looking at each mode group in turn we are able to interpret those characteristics of the group that are dominant in the mode choice. In this way socioeconomic-transportation system interdependence is analyzed as a system of attributes and not as a series of suboptional relationships inherent in a multiple-regression approach.

Canonical correlation brings out the nature of the interdependence of the two sets of variates when the linear combination of the two sets is maximally correlated. Consider an arbitrary linear combination $U = \alpha' X^{(S)}$ of the socioeconomic variates and an arbitrary linear combination $V = \gamma' X^{(T)}$ of the system variates. The maximum correlation is found by rotating the reference axes for each set of variates in the test space so that the axes of the socioeconomic variate set and those of the transportation service variate set form a new axes system. If the parameters $\alpha$ and $\gamma$ are normalized such that U and V have unit variance, then

$$\Sigma U^2 = 1 = \Sigma \alpha' X^{(S)} X^{(S)\prime} \alpha$$

and

$$\Sigma V^2 = 1 = \Sigma\gamma' X^{(T)} X^{(T)'} \gamma$$

The correlation between U and V is, therefore,

$$\Sigma UV = \Sigma\alpha' X^{(S)} X^{(T)'} \gamma \tag{1}$$

because

$$\Sigma X^{(S)'} \alpha = 0 \text{ and } \Sigma\gamma' X^{(T)} = 0$$

Thereby, substituting in Eq. 1, the correlation between the two sets is

$$\Sigma UV = \alpha' R_{12}\gamma \tag{2}$$

The canonical correlation problem is to find the values of $\alpha$ and $\gamma$ when $\Sigma UV$ is maximized, i.e., when the derivative of $\Sigma UV$ with respect to $\alpha$ and $\gamma$ is zero. Anderson (7) has shown that $\Sigma UV$ is maximized when

$$\begin{bmatrix} -\lambda R_{11} & R_{12} \\ R_{21} & -\lambda R_{22} \end{bmatrix} \begin{bmatrix} \alpha \\ \gamma \end{bmatrix} = 0 \tag{3}$$

in which

$R_{11}$ = intercorrelations among the socioeconomic variates,
$R_{22}$ = intercorrelations among the system variates,
$R_{12}$ = intercorrelations of socioeconomic with system variates,
$R_{21}$ = transpore of $R_{12}$, and
$\lambda$ = latent root, or eigenvalue.

The determinantal equation of the first term

$$| R_{22}^{-1} F_{21} R_{11}^{-1} R_{12} - \lambda I | = 0 \tag{4}$$

is solved for all possible values of $\lambda$. For each characteristics root, the vectors of the coefficients $\alpha$ and $\gamma$ are found for the set U and V from the canonical equations

$$(R_{22}^{-1} R_{21} R_{11}^{-1} R_{12} - \lambda_1 I)\gamma = 0 \tag{5}$$

and

$$\alpha = (R_{11}^{-1} R_{12}\gamma)/\lambda_1 \tag{6}$$

where $\alpha$ gives the weighting of each of the socioeconomic variables in the interdependent relationship and $\gamma$ gives the weighting of each of the transportation system variables in the relationship.

A measure of the statistical significance of any canonical correlation is given by $R_c$, which is the correlation between any weighted linear combination of one set of variables and any weighted linear combination of the second set of variables. In geometrical terms it is the cosine of the angle between the vectors representing each set of variables. If the two vectors are coincident, $R_c = 1.0$, indicating that the two sets of variables are perfectly correlated. Significance level of any $R_c$ is a $\chi^2$ test of its significance in extracting the relationship between any two sets of variables.

Since the sets of variables may be correlated in several ways, several canonical vectors are possible, each correlation being given by $R_{c_1}$.

## Results of Correlation Tests

Table 2 shows one significant way in which the two sets of variables are interrelated. The result of the test shows reasonably strong association ($R_{c1} = 0.71$ and $p < 0.001$) between income, out-of-pocket expenses, and CBD walking distances, as demonstrated by the value of the coefficients. The test implies that high-income automobile commuters pay relatively high parking rates. However, in terms of their ability to pay, high-income workers pay a smaller proportion of their income to park than low-income workers. It is also evident that the relatively high parking cost is offset by a shorter walking distance. Therefore, as we would expect, commuters who can afford it park closer to their destination, at higher cost, to avoid walking.

It is useful also to examine a second way in which the variables are correlated as shown by Table 3. This table gives the vectors of coefficients for a second important dimension that is statistically independent, or orthogonal to the first set of vectors. This variate extracts the correlation between the other characteristics besides income and shows the effect of bus frequency. It shows that younger people of low occupational and car ownership status, for whom a car is seldom necessary for their work or not easily available, are concerned primarily with bus frequency. This implies that people of this general type locate where there is good bus service, even though they are, in this instance, car drivers. People of this group also park further from their job location, indicating their inability to command close-in parking spaces. In essence, these are low socioeconomic status car commuters who nevertheless need good bus service as a viable alternative, either because they cannot afford a second car and want good bus service for other members of the family or because they perhaps need to commute by bus periodically.

Tests on the bus group show results similar to the car group with some important differences. Table 4 gives a canonical correlation between income and cost factors. However, the linkage is not as strong as with the car group, particularly with transit fare. The ability to pay variable (E/INC) shows a fairly strong negative correlation, but since fares are relatively constant for all users, the correlation is probably increased because the test is showing the correlation between income and the increase of income in the E/INC variable. The simple correlation between these two variables (i.e., INC and E/INC) separate from all other variables is $r = -0.86$. It also appears from the table that older bus commuters seek a relatively short walking distance at trip destination.

The second important correlation between sets of variables among the bus commuters as given by Table 5 also shows the concern of those of low car ownership and availability to locate where there is good bus service. The test also shows a concern with transit fare. It also shows that older people place greater emphasis on efficient transit service than do younger bus users.

One inexplicable result of the canonical correlation tests is the absence of a significant correlation of income and total travel time. A separate analysis of variance test was carried out to find an explanation for this.

## Income and Travel Time Correlation

Table 6 gives the variation of incomes and travel times by mode. There is some indication that as income increases travel time by car increases and travel time by bus decreases. However, variance ratios show no significant variation in travel time either by mode or by income. While substantial differences appear to exist in modal travel times, when income is considered the differences in travel time are due mostly to the income factor. This finding is supported by the work of Zupan (8) and Domencich, Kraft, and Valette (4), who conclude that socioeconomic variables are more important than the system travel time savings between modes in determining mode choice.

In the case under study this explanation appears reasonable since the travel time variation across the whole sample was relatively narrow, with most trips taking between 20 and 45 minutes. Further tests with a broader range of trip lengths are needed to show the relative importance of travel time and travel cost factors.

**Table 1. Significance test of statistical difference in socioeconomic characteristics between car drivers and bus riders.**

| Socioeconomic Characteristic | d.f. | $\chi^2$ | p[a] |
|---|---|---|---|
| Age | 3 | 87.5 | p < 0.001 |
| Occupation | 5 | 407.4 | p < 0.001 |
| Car ownership | 3 | 385.4 | p < 0.001 |
| Income | 3 | 482.1 | p < 0.001 |

[a]p = probability of there being no difference in mode group for variable shown.

**Table 2. First canonical correlation test of the interdependence between socioeconomic and transportation system characteristics for car drivers.**

| Variables of Set 1 | Coefficients ($\alpha$) | Variables of Set 2 | Coefficients ($\gamma$) |
|---|---|---|---|
| AGE | 0.0214 | TTO | -0.0543 |
| OCC | 0.0065 | TFRM | -0.1007 |
| CNEC | 0.0758 | OTT | 0.0367 |
| COWN | -0.0363 | OPE | 1.6829 |
| INC | 0.9880 | E/INC | -1.7666 |
| CARA | 0.1617 | FREQ | 0.0063 |

Significance test: $R_{c_1} = 0.71$; $\chi^2 = 925.88$; d.f. = 36; $p < 0.001$.

**Table 3. Second canonical correlation test of the interdependence of socioeconomic and transportation system characteristics for car drivers.**

| Variables of Set 1 | Coefficients ($\alpha$) | Variables of Set 2 | Coefficients ($\gamma$) |
|---|---|---|---|
| AGE | -0.5591 | TTO | -0.1770 |
| OCC | -0.2481 | TFRM | 0.2499 |
| CNEC | -0.4512 | OTT | -0.1206 |
| COWN | -0.5526 | OPE | 0.0835 |
| INC | 0.1893 | E/INC | 0.1893 |
| CARA | -1.2391 | FREQ | 0.9740 |

Significance test: $R_{c_2} = 0.17$; $\chi^2 = 45.61$; d.f. = 25; $p < 0.01$.

**Table 4. First canonical correlation test of the interdependence between socioeconomic and transportation system characteristics for transit riders.**

| Variables of Set 1 | Coefficients ($\alpha$) | Variables of Set 2 | Coefficients ($\gamma$) |
|---|---|---|---|
| AGE | 0.0933 | TTO | 0.0168 |
| OCC | -0.0009 | TFRM | -0.0344 |
| CNEC | -0.0000 | OTT | 0.0184 |
| COWN | -0.0442 | OPE | 0.2974 |
| INC | 0.9500 | E/INC | -1.0485 |
| CARA | 0.0341 | FREQ | 0.0200 |

Significance test: $R_{c_1} = 0.91$; $\chi^2 = 1{,}940.0$; d.f. = 36; $p < 0.001$.

**Table 5. Second canonical correlation test of the interdependence between socioeconomic and transportation service characteristics for bus riders.**

| Variables of Set 1 | Coefficients ($\alpha$) | Variables of Set 2 | Coefficients ($\gamma$) |
|---|---|---|---|
| AGE | 0.0892 | TTO | 0.0083 |
| OCC | 0.0279 | TFRM | -0.0662 |
| CNEC | -0.0000 | OTT | -0.0998 |
| COWN | -1.8093 | OPE | -1.2693 |
| INC | 0.3940 | E/INC | 0.0762 |
| CARA | -2.1340 | FREQ | 0.9828 |

Significance test: $R_{c_2} = 0.49$; $\chi^2 = 278.0$; d.f. = 25; $p < 0.001$.

**Table 6. Mean total travel time, by income category and mode.**

| Income Category (dollars) | Mean Travel Time (minutes) | |
|---|---|---|
| | By Car | By Bus |
| <4,000 | 25.4 | 41.0 |
| 4,000-8,000 | 30.8 | 37.5 |
| 8,000-12,000 | 30.8 | 36.4 |
| >12,000 | 32.3 | 38.0 |

Significance test: Variance ratio[a]

| Source of Variation | d.f. | Sum of Squares | Mean Square | F | P |
|---|---|---|---|---|---|
| Mode | 1 | 54.50 | 54.50 | 2.32 | n.s. |
| Income | 3 | 55.44 | 18.50 | 0.79 | n.s. |
| Error | 3 | 70.44 | 23.50 | | |
| Total | 7 | 180.38 | | | |

[a]Variance ratio F is the estimate of $S^2$ based on the variation in travel time by mode ÷ the estimate of $S^2$ based on the variation in income.

## Comparison of Correlation for Mode Groups

Comparison of the two mode groups tells us something about the strength of the relationship between socioeconomic and transportation system factors for each group. First, the sensitivity of bus riders to transportation system characteristics is greater than in the car driver group, as shown by the relative values of the correlation coefficient ($R_{c_{bus}} = 0.91$ and $R_{c_{auto}} = 0.71$). The implication is that bus riders have the type of socioeconomic constraints that make them dependent on the attributes of the transportation system whereas car drivers have more freedom to choose different combinations of service attributes and therefore are less dependent on the attributes of the system. The correlation between the two sets of variables denotes a state in which decisions of car ownership, residential and/or job location, and mode choice are interdependent. The results support the idea that those who drive cars, presumably of a high-car-ownership category, have a better choice of residence and job location than those who are more dependent on transportation services. It also implies that pricing control policies will have a smaller locational effect on this group than on the group who currently use transit. That is, if the disutility of travel becomes unacceptable to this group they may tend to change job or home location to relieve the situation rather than adjust to a controlled change in the existing system.

Second, in both groups, income and system cost factors are the prime components describing the interrelationship of socioeconomic structure and behavior. It is evident, however, that parking charges (OPE for car drivers) have more influence relative to the other system components for the car drivers than fares do for transit passengers (OPE for transit group). Thus, high incomes are associated with high out-of-pocket expenses for both groups, but for transit riders this is more or less fixed, modifying its effect. Income and cost modified by the ability to pay (OPE/INC) shows a high interdependence in both groups.

While the first canonical variate shows interdependence of income and cost, the second canonical variate brings out the positive relationships between the other socioeconomic characteristics and bus frequency. For both groups, car availability and bus frequency are related. This can be explained through the car-ownership factor, since low frequencies would precipitate higher car ownership, which is brought out by a high intercorrelation of car ownership with car availability ($r = -0.91$ for car group and $r = -0.86$ for the bus group).

The analysis supports the idea of a reciprocal relationship between car ownership and bus service. With the high car-ownership rate of the study area, the effect is probably due to the multiple-commuter households of many of the apartments in the area. Thus, the existence of a good bus service would doubtless delay a number of nonfamily households (single people living communally) from the purchase of their first car or the purchase of a second car.

## SOCIOECONOMIC STRUCTURE AND MODE SHIFT PROPENSITY

A second series of tests was carried out to determine if the socioeconomic structure of the corridor was likely to have a differential effect on causing automobile drivers to shift mode. A previous study shows that the propensity to shift from automobile commuting is dependent mostly on changes in the frequency of buses and the residential travel time, with secondary influences by overall travel time and parking costs (9). The question is whether these influences can logically be stratified by the socioeconomic profiles of the commuters. If so, the interdependence between the socioeconomic structure of the corridor and the particular system attributes that are important to each socioeconomic group adds a further dimension to be considered in the mode choice.

To eliminate some of the constraints on behavior patterns imposed by the existing system attributes, the tests were based on the stated preferences for an idealized system. The hypothetical system was a park-and-ride combination in the corridor using express buses for corridor line haul. Automobile drivers were asked to indicate the scale position at which they would shift to the system for each of the influences above. After editing, 465 commuters served as the data base for this test.

The ratio of the actual measurement of the system attribute experienced at the time of the trip compared to the measurement indicated on the preference scale was used as the relative measure of the variable. Variables derived, expressed as a ratio of actual system state to preferred state, were as follows:

$$X_1 = \frac{\text{actual overall travel time}}{\text{preferred overall travel time}}$$

$$X_2 = \frac{\text{actual out-of-pocket expenses (parking cost)}}{\text{preferred out-of-pocket expenses (parking cost and transit fare)}}$$

$$X_3 = \frac{\text{actual travel time from home to vehicle}}{\text{preferred transfer time at a park-and-ride terminal}}$$

$$X_4 = \frac{\text{actual frequency of available buses}}{\text{preferred frequency of buses at park-and-ride terminal}}$$

$$X_5 = \frac{\text{parking charge at CBD destination}}{\text{parking charge that would cause a shift to the new mode}}$$

The discrimination-classification statistical analysis was used to determine which of these relative system variables would be important in a mode shift for different socioeconomic characteristics. Discriminant analysis estimates vector coefficients of linear discriminant functions that, when solved, define the position of an individual on a line that best separates predetermined classes or groups (10). The estimated position on the line is a linear function of the travel characteristics of the individual. The original travel characteristics are transformed to discriminant "scores" by the statistical criterion of maximizing the ratio of the square of the differences between the mean of any group and the grand mean to the pooled within-group variance. Assumptions are that the variables are multivariate normal and that group variance-covariance matrices are equal. Classification into groups is accomplished by assignment of individuals to one of the predetermined groups based on the vector solution of the discriminant function for each group. The statistical validity of group assignment is made by posterior classification of the original group members.

Tests were made of each of four socioeconomic categories and discriminant functions used to determine the number of groups into which each category should be divided to achieve the best loading of the variables. This resulted in two age groups, those over 40 years and those under 40 years; two car-ownership groups, single- and multiple-car families; four income groups, very low (< $4,000 per annum), low ($4,000 to $8,000), medium ($8,000 to $12,000) and high (> $12,000); and six categories of occupation, managerial, professional, secretarial, clerical, sales, and other. The discriminant functions in all cases provided significant separation of the groups at $F_p < 0.05$.

Table 7 summarizes the results of the statistical tests. The summary shows the overall importance of the transportation service variables in discriminating between the socioeconomic classes tested. The underlined variable is the most important factor in giving the relative level of service change needed to cause a mode shift. For example, the frequency of bus service is the most important change that must be implemented for commuters of different age categories. In this case the influence of out-of-pocket expenses is also a significant factor.

The posterior classification is a measure of the ability of the linear function of scaled vector weights to assign an individual of unknown class with given relative transportation service preferences to a group category. In this case, although the discriminant functions are statistically significant, their ability to assign class is relatively poor, attesting to the fact that the test does not show conclusive results. The test does, however, point to some tentative conclusions about the interrelationship of socioeconomic characteristics and the tendency to shift mode.

## Influence of Age and Car Ownership in Mode Shift

Tables 8 and 9 give the group means on the significant variables for different age and car-ownership categories. The interpretation of the tables is through the means of the relative transportation service variables. The larger the mean value for each variable, the more resistant that group is to shifting mode. This is true because there is a greater difference between the measure of the existing service and that any individual in the group would prefer in those cases where a smaller value of the attribute

**Table 7. Statistical summary of interdependence of socioeconomic characteristics with mode shift propensity.**

| Socioeconomic Characteristic | Scaled Vector Weights for System Variable[a] | | | | | Test of Group Separation | | | Posterior Classification, Percent Misses |
|---|---|---|---|---|---|---|---|---|---|
| | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $\Lambda$[b] | $F_p$[c] | Percent of Trace[d] | |
| Age | 5.98 | 37.67 | -0.56 | 47.46 | 2.12 | 0.96 | <0.01 | 100 | 40.6 |
| Car ownership | 9.73 | 8.92 | 8.52 | -0.52 | -2.06 | 0.97 | <0.05 | 100 | 46.4 |
| Income | -3.50 | 104.71 | -14.38 | 12.14 | -3.80 | 0.90 | <0.001 | 63 | 62.2 |
| Occupation | -3.51 | -12.11 | 11.43 | -28.64 | 7.04 | 0.91 | <0.01 | 50 | 79.8 |

[a]Only the first discriminant function is given.
[b]Wilk's lambda criterion for discriminating power of the function: $\Lambda = |W|/|T|$ or equivalently $\Lambda = \pi_{i=1}^{r}\left[\frac{1}{1+\lambda_i}\right]$, r = number of roots.
[c]$F_p = F_{probability}$ = the chance of producing group differences this large or larger! This test is based on $F = \left(\frac{1-\Lambda}{\Lambda}\right)\left(\frac{N-g}{g-1}\right)$ N = number of observations, g = number of groups.
[d]Percent of trace = the percentage of the total discriminating power of the vector shown.

**Table 8. Group means on significant variables for those over 40 years and those under 40 years.**

| Significant Variable | Group Means | |
|---|---|---|
| | $U_{40}$ | $O_{40}$ |
| $X_2$(OPE) | 3.26 | 2.07 |
| $X_4$(FREQ) | 4.88 | 3.91 |

**Table 9. Group means on significant variables for single- and multiple-car families.**

| Significant Variable | Group Means | |
|---|---|---|
| | Single | Multiple |
| $X_1$(OTT) | 1.11 | 0.94 |
| $X_2$(OPE) | 3.01 | 2.30 |
| $X_3$(TTO) | 0.47 | 0.30 |

**Table 10. Group means on significant variables for very low, low, medium, and high income commuters.**

| Significant Variable | Group Means | | | |
|---|---|---|---|---|
| | VL | L | M | H |
| $X_1$(OTT) | 1.11 | 1.07 | 1.08 | 0.90 |
| $X_2$(OPE) | 1.31 | 1.79 | 3.14 | 2.94 |
| $X_3$(TTO) | 0.71 | 0.61 | 0.29 | 0.24 |
| $X_4$(FREQ) | 3.58 | 4.21 | 4.43 | 4.54 |
| $X_5$(PKCHG) | 1.04 | 0.55 | 0.53 | 0.62 |

**Table 11. Group means on significant variables for managerial, professional, secretarial, clerical, sales, and other occupation categories.**

| Significant Variable | Group Means | | | | | |
|---|---|---|---|---|---|---|
| | Managers | Professionals | Secretaries | Clerical | Sales | Other |
| $X_1$(OTT) | 1.01 | 1.04 | 1.18 | 1.11 | 0.90 | 0.88 |
| $X_2$(OPE) | 2.28 | 2.85 | 1.27 | 1.04 | 4.78 | 4.87 |
| $X_3$(TTO) | 0.35 | 0.32 | 0.30 | 0.52 | 0.20 | 0.84 |
| $X_4$(FREQ) | 4.10 | 4.79 | 4.18 | 4.02 | 5.24 | 3.64 |
| $X_5$(PKCHG) | 0.62 | 0.52 | 0.97 | 0.41 | 0.66 | 0.76 |

is logically preferred, that is, variables $X_1$ to $X_4$. For the variable measuring the parking charge required to shift, $X_5$, the opposite rule applies: The lower the mean value the more resistance to shift mode with a given service value.

The results for the different age groups indicate that those under 40 years are more reluctant to shift mode than those over 40, considering changed service levels for bus frequency and out-of-pocket costs. Since the ratio of actual to preferred frequency is about 5:1 for those under 40 and 4:1 for those over 40, if the existing frequency was 1 bus every 20 minutes the under-40 group would require that it be decreased to a 4-minute frequency before shifting mode while the over-40 group would shift with a 5-minute frequency. A greater decrease in out-of-pocket cost would also be demanded by the younger group.

Similar results for car ownership implies that those of single-car families would be more resistant to shift than those of multiple-car families.

These results are not in the direction expected, which was that reluctance to shift mode would increase with age and car ownership. There appears to be no good explanation for the apparently odd results here except that the statistical tests extracted a very low percentage of variation in the data, indicating that further tests on a different sample may show more conclusive results.

## Influence of Income in Mode Shift

The best association was between income categories and mode shift tendencies (Table 10). The results show that as income increases the resistance to a mode shift increases, as shown primarily through the service variable of cost and frequency. The higher income groups are less sensitive to cost than the middle range, a result we would expect. It is also evident that an increase in income creates more tolerance to travel time variables (OTT and TTO), a result consistent with the finding of the previous section. It appears that the variation across either the income groups or the travel times is not great enough to produce the suspected relationship between incomes and travel time. The conclusion from the results here is that some increase in walking times would be acceptable providing that bus frequencies were substantially increased (i.e., by 300 to 400 percent).

The tests also imply that as income increases the charge for parking necessary to cause a mode shift also increases, except again for the highest income group. The very-low-income group is very sensitive to parking charge increases, indicating that any increase whatever would cause them to shift mode.

## Influence of Occupation and Mode Shift

Occupation does not appear to have much influence on the propensity to shift mode. Table 11 indicates some tendency for managers and professional employees to be sensitive to out-of-pocket expenses. However, this does not appear to be intuitively reasonable and bears further investigation. Sales people and craftsmen on the other hand are shown to be sensitive to cost factors, as expected. Managers, professional workers, and sales people are prepared to walk some distances while craftsmen and laborers are not. Secretaries are the most sensitive to parking fee increases, while clerks are least sensitive. Laborers, sales people, and managers are above average in sensitivity to parking fee increases.

# CONCLUSIONS

Correlation tests show an interdependency between level of income and travel expenses when the correlation between socioeconomic structure and system attributes is maximized. Those of higher incomes have an ability to select the more expensive alternative. This is as expected and has been seen in the increased ownership and use of automobiles as the general level of affluence has increased. Also of interest are the interdependencies of other structural variates with system performance. Car ownership and availability are related to excess travel times and the frequency of bus service. This follows intuitive reasoning and can be seen in current transit planning. The decision to extend bus lines into new areas usually follows development of these areas when there is a sufficient perceived demand to justify the extensions. If car ownership is high, demand for bus service is low, and therefore service is only provided when those who have no transportation choice become a significant group to warrant it. On the other hand, where transit service is already adequate, the need for a second or third car diminishes. The inference is that, if a high-quality transit service is provided to a developing area, the demand for such a service would cause a switch away from multicar ownership. This, however, would require a change in the "pay as you go" philosophy of transit decision-making and would require subsidization of capital and operating expenses until demand increased to economic levels.

One finding of the study is that socioeconomic characteristics of commuters exert some influence on sensitiveness to shift mode. In general, those under 40, from families with one car, with middle to higher incomes, in the nonprofessional occupations, would be more resistant to shift than the population as a whole. The age factor may influence mode shift propensity through out-of-pocket expenses and frequency of service. Car ownership is a potential factor because of overall travel time and residential walking time. There is probably a spatial effect here as well, since car ownership no doubt increases as distance from the CBD increases, thereby pointing up the trade-off between travel time and owning another car. Income influences are varied. The lower income groups are more time-sensitive than the higher income groups, although the latter groups are more sensitive to transit frequencies and out-of-pocket costs. On the other hand, lower income groups would tolerate less walking than the higher income groups and are more sensitive to parking charge increases.

The findings of this investigation are still tentative, but they do point to possible interrelationships between socioeconomic characteristics of a transportation corridor and the transportation service attributes serving the corridor. Further research is necessary using a data base collected for the purpose of examining some of the hypotheses expressed here in which both the range and extent of socioeconomic and transportation variables can be enlarged.

## ACKNOWLEDGMENT

This research was carried out under a grant from the National Research Council of Canada.

## REFERENCES

1. Fertal, M. J., et al. Modal Split: Documentation of Nine Methods for Estimating Transit Usage. U.S. Bureau of Public Roads, 1966.
2. McGillivray, R. G. Demand and Choice Models of Modal Split. Jour. of Transport Economics and Policy, May 1970.
3. Kraft, G., and Woho, M. New Directions for Passenger Demand Analysis and Forecasting. Transportation Research, Vol. 1, No. 3, 1967.
4. Domencich, T. A., Kraft, G., and Valette, J. P. Estimation of Urban Passenger Travel Behavior: An Economic Demand Model. Highway Research Record 238, 1968, pp. 64-78.
5. Brand, D. Theory and Method in Land Use and Travel Forecasting. Highway Research Record 422, 1973, pp. 10-20.
6. Manheim, M. L. Practical Implications of Some Fundamental Properties of Travel-Demand Models. Highway Research Record 422, 1973, pp. 21-38.
7. Anderson, T. W. An Introduction to Multivariate Statistical Analysis. John Wiley, New York, 1958.
8. Zupan, J. M. Mode Choice: Implications for Planning. Highway Research Record 251, 1968, pp. 6-25.
9. Brown, G. R. Analysis of User Preferences for System Characteristics to Cause a Modal Shift. Highway Research Record 417, 1973, pp. 25-36.
10. Cooley, W. W., and Lohnes, P. R. Multivariate Procedures for the Behavioral Sciences. John Wiley, New York, 1962.

## APPENDIX

### VARIABLES USED IN ANALYSIS

The variables selected for analysis were those collected in a survey on one bridge approach to the CBD of Vancouver, Canada. The survey was a hand-out-mail-back questionnaire that both car drivers and bus passengers were asked to fill out and return. The data matrix consisted of 1,244 car drivers and 967 bus passengers, after editing to remove car passengers, zero entries, and captive bus passengers. The variables used in the analysis were as follows.

1. Age. The age category of each respondent was coded between 1 and 4, designating 0-20, 21-40, 41-60, and over 60 years.

2. Occupation. The occupational categories in the survey were reorganized in a new hierarchy in accordance with Blishen's socioeconomic index for occupations. Blishen devised an interval-scaled index of occupations in Canada. The occupation variables in the multivariate analysis were rank-ordered according to Blishen's hierarchy. A code number between 1 and 6 was assigned to the six categories of the work-purpose trip in accordance with the rank of that occupation.

3. Car Necessary for Work. The question, "How often is it absolutely necessary to use your car during the day?" was coded from 1 to 4 to designate "never", "very seldom", "once a week" or "frequently". All bus passengers were coded "1".

4. Car Ownership. This variable was coded 1 to 4 designating none, one, two, and three or more cars in the household. Since the analysis is for "choice" trips, only those households that reported owning one or more cars were included in the data matrix.

5. Income. Income was coded 1 to 4 designating under $4,000, $4,000 to $8,000, $8,000 to $12,000, and over $12,000 categories.

6. Car Availability. Since household ownership of an automobile may not reflect the actual availability of a car for a given trip, a proxy based on the estimated availability of a car and the competition among members of the household for car use was included. This variable accounts for different levels of car ownership, persons per household (zonal average), and income:

$$\text{Car availability} = \frac{\text{Persons per household} \times \log \text{income}}{\text{Cars per household}}$$

This proxy variable defines the competition between members of the household in terms of the number of persons in the household and the number of cars owned, modified by the demand for travel. If the car was purchased for commuting only, the availability of a car for commuting is a simple relationship between the number of commuters and the number of cars owned. If there were one car per commuter there would be no competition for the use of the car, and the mode choice would be related to factors other than car ownership. Intuitively, the use of transit would increase as this simple ratio decreased from the value of one. However, the number of persons and cars owned per household for commuting must be related to competition from other purpose trips as well, such as shopping and social-recreation trips carried out by noncommuting members of the family. The income factor takes this into account. As income per person increases, the number of trips demanded increases. But there is a limit to the number of trips and there are also possible economies of scale. Therefore it is postulated that the desire for individual travel is related to income, but at a decreasing rate, and the log of income becomes the proper modifier for car availability. Income is the reported individual income, and persons per household is the average number of persons per household in each zone. Persons per household is taken from the 1966 population census tracts and converted to the survey zones. Cars per household is that reported by survey respondents.

7. Residential Travel Time. This variable measures the length of time to travel from the start of the work journey to the bus stop or to a car. Although the original survey times were categorized (0 minutes, 1-3 minutes, 3-5 minutes, 5-10 minutes, and

10 or more minutes), the intervals were unequal and a significant number of responses were in the open-ended category (10 or more minutes), and thus it was decided to treat this variable as an ordinal scaled response. Therefore the categories were coded from 1 to 5 inclusive.

8. CBD Travel Time. This variable is the walking time from the parking location or bus stop to final destination. It was also ordinalized in the same way as residential travel time, since all conditions were parallel.

9. Overall Travel Time. This measures door-to-door travel time for both auto and bus passengers in 5-minute increments. These increments are probably the minimum that can be perceived by users. That is, perceived and reported times are given to the nearest 5 minutes, and any finer breakdown is probably meaningless in terms of subjective responses.

10. Parking Charge. Parking charges levied are the manifestation of out-of-pocket expenses for automobile drivers, as are fares for transit riders. Most commuters pay parking fees by the month. Reported monthly fees were converted to daily rates to be compatible with the other components of daily travel. On the assumption there are 20.8 working days per month, less 10 statutory holidays per year (or 0.8 per month), the reported monthly rate is equivalent to the reported daily rate if divided by 20 (20.8 - 0.8). Occasionally researchers split parking fees, with half of the fee assigned to the journey to work and half to the journey from work. It was reasoned here, however, that the perceived parking charge would be the total charge per day, since this would be the manner in which the transfer of money would take place. This is particularly true of subscribers, who it is postulated would not attempt to cost a single trip.

11. Fare. Both transit agencies reported that most fares are in fact commuted fares, but no precise breakdown is available. Therefore it was assumed that all commuters would take advantage of commuted fares where available. The one-way fares were then doubled to conform with the nature of the perception of parking charges.

12. Parking Charge/Income. This variable is a proxy incorporating parking charge divided by income, on the premise that the parking rate as an influence in mode choice would be more meaningful if related to income.

13. Frequency. This was taken as the average frequency over the period between 7 and 9 a.m. Since the minimum difference in average frequency for commuters is 5 minutes, and since the original survey did not report exact frequencies facing individual travelers, it was decided that average frequency for each zone would adequately describe the perceived dimension of transit frequency.

The variables are measured in different scales of time, space, and dollar units. The score matrices were standardized and adjusted by Sheppard's correction factor by dividing each row score by the corrected standard deviation of the variate. Standardization, in which the mean of the variate over the population is zero and the standard deviation equal to unity, was necessary in some tests to reduce both sets of variates (the socioeconomic ones and the behavior ones) to the same metric. Standardization of the variates has no effect on the results of the analysis.