TRANSPORTATION RESEARCH RECORD 610

Passenger Travel Demand Forecasting

TRANSPORTATION RESEARCH BOARD

COMMISSION ON SOCIOTECHNICAL SYSTEMS
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C. 1976

The Transportation Research Board is an agency of the National Research Council, which serves the National Academy of Sciences and the National Academy of Engineering. The Board's purpose is to stimulate research concerning the nature and performance of transportation systems, to disseminate information that the research produces, and to encourage the application of appropriate research findings. The Board's program is carried out by more than 150 committees and task forces composed of more than 1800 administrators, engineers, social scientists, and educators who serve without compensation. The program is supported by state transportation and highway departments, the U.S. Department of Transportation, and other organizations interested in the development of transportation.

The Transportation Research Board operates within the Commission on Sociotechnical Systems of the National Research Council. The Council was organized in 1916 at the request of President Woodrow Wilson as an agency of the National Academy of Sciences to enable the broad community of scientists and engineers to associate their efforts with those of the Academy membership. Members of the Council are appointed by the president of the Academy and are drawn from academic, industrial, and governmental organizations throughout the United States.

The National Academy of Sciences was established by a congressional act of incorporation signed by President Abraham Lincoln on March 3, 1863, to further science and its use for the general welfare by bringing together the most qualified individuals to deal with scientific and technological problems of broad significance. It is a private, honorary organization of more than 1000 scientists elected on the basis of outstanding contributions to knowledge and is supported by private and public funds. Under the terms of its congressional charter, the Academy is called upon to act as an official–yet independent–advisor to the federal government in any matter of science and technology, although it is not a government agency and its activities are not limited to those on behalf of the government.

To share in the task of furthering science and engineering and of advising the federal government, the National Academy of Engineering was established on December 5, 1964, under the authority of the act of incorporation of the National Academy of Sciences. Its advisory activities are closely coordinated with those of the National Academy of Sciences, but it is independent and autonomous in its organization and election of members.

Transportation Research Record 610
Price $2.80
Edited for TRB by Frances R. Zwanzig

subject areas
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.

Notice
The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competence and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The views expressed in this report are those of the authors and do not necessarily reflect the view of the committee, the Transportation Research Board, the National Academy of Sciences, or the sponsors of the project.

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Passenger travel demand forecasting.

(Transportation research record; 610)
1. Traffic estimation–Congresses. 2. Choice of transportation–Mathematical models–Congresses. 3. Transportation–Passenger traffic–Congresses.
I. Title. II. Series.
TE7.H5 No. 610 [HE370] 380.5'08s [388.4] 77-23421
ISBN 0-309-02585-0

Sponsorship of the Papers in This Transportation Research Record

GROUP 1–TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
E. Wilson Campbell, New York State Department of Transportation, chairman

Transportation Forecasting Section
Charles V. Wootan, Texas A&M University, chairman

Committee on Passenger Travel Demand Forecasting
Daniel Brand, Massachusetts Executive Office of Transportation and Construction, chairman
Moshe E. Ben-Akiva, Robert B. Dial, Raymond H. Ellis, David S. Gendell, Thomas F. Golob, John R. Hamburg, Lawrence V. Hammel, Walter G. Hansen, David T. Hartgen, Kevin E. Heanue, Frank E. Horton, Louis E. Keefer, C. E. Klamm, Gerald Kraft, Eugene J. Lessieu, Herbert S. Levinson, William McGrath, James M. McLynn, Sydney R. Robertson, Peter R. Stopher, Vergil G. Stover, Darwin G. Stuart, Anthony R. Tomazinis, Edward Weiner, Charles C. Whitmarsh, George V. Wickstrom, Alan G. Wilson

James A. Scott, Transportation Research Board staff

The organizational units and the officers and members are as of December 31, 1975.

Contents

Tests of the Temporal Stability of Travel Simulation Models in Southeastern Wisconsin

Kenneth R. Yunker, Department of Civil Engineering, Northwestern University

The assumption of the stability of travel simulation models over time is an essential element of the urban transportation planning process. This assumption was tested using travel simulation models developed with data from an origin and destination survey conducted in 1963 and travel inventory data from a similar study conducted in 1972. Both surveys were conducted by the Southeastern Wisconsin Regional Planning Commission; the travel models tested were those that had been used in the preparation of a regional land use and transportation plan for southeastern Wisconsin that was completed in 1966. The testing performed as a part of the reappraisal of the land use and transportation recommendations of 1966, which was of the temporal stability of the three major travel simulation models—trip generation, modal split, and trip distribution—indicated that 1972 trip generation, transit use, and trip length characteristics within southeastern Wisconsin were predicted with adequate accuracy through the application of the original 1963 models.

A basic assumption of most urban transportation studies is that travel simulation models defined through analysis of base-year origin and destination survey data will remain stable over time, thus allowing the evaluation of alternative transportation plans for the future. In recent years considerable interest has been directed toward the validity of this assumption. The doubts of the validity of the assumption have been a result of assertions that the travel simulation models employed in most transportation planning efforts, which are based on a system of spatial aggregation and do not consider all the variables known to affect travel, have been developed on a descriptive, rather than a causative, basis (1). In consequence, the ability of such models to accurately predict future travel under conditions substantially different from those of the base year has been questioned. This assumption of temporal stability has never been adequately tested, as comparable data for the same area for two points in time have been available in only a limited number of instances (2). However, as a result of major origin and destination (O-D) surveys now being conducted in areas in which similar surveys were completed in the 1960s, the testing of this assumption—over short periods of time (10 years)—is now possible.

One of the areas in which two compatible O-D studies have been completed is the seven counties of southeastern Wisconsin. Major travel studies have been conducted by the Southeastern Wisconsin Regional Planning Commission (SEWRPC), which was established in 1960 to assist in solving areawide problems and in planning the physical development of the region. The first O-D study, performed in 1963, was part of the basis for the preparation of a regional land use and transportation plan for the area that was completed in late 1966. The second O-D study was conducted in 1972 and used in the reevaluation of the original land use and transportation recommendations. One of the factors prompting this reappraisal of the original planning effort was the recognized need to update the plans in light of changing conditions within the region, particularly the changes in those factors that would influence transportation system development.

A significant part of the analysis of changing conditions was the review of the ability of the travel simulation models used in the initial planning effort to predict 1972 travel, i.e., a test of the temporal stability of the relationships defined in the travel simulation models. The 1972 travel was predicted by applying the models developed in the original planning effort and comparing the results with the results of the O-D survey. The following sections summarize the approaches used by the three major travel simulation models—trip generation, model split, and trip distribution—and evaluate their continual validity.

TRIP GENERATION

In the SEWRPC 1963 regional land use and transportation study, trip generation was analyzed and simulated through the development of nine equations, four of which related total trip production by trip purpose to the land use within each traffic analysis zone and five of which related total trip attractions by trip purpose to such land use. The nine equations were developed with multiple regression analysis applied in a stepwise manner for the trip purposes of home-based work, home-based shopping, home-based other (a combination of personal business, medical-dental, social, and recreation), and non-home-based. Home-based school trips were analyzed and forecast using a growth factor technique. A balancing procedure was used for trip generation forecasts, which adjusted zonal totals of home-based shopping, home-based other, and non-home-based trip attractions so that the total trip attractions were equivalent to the total trip productions for these three purposes. For the home-based work trip purpose, zonal trip productions were adjusted so that regional home-based work trip productions were equivalent to total home-based work trip attractions.

The ability of the trip generation equations developed in the original land use and transportation planning program to simulate 1972 trip making was investigated by comparing the predictions of the equations to the actual 1972 travel survey data. Travel surveys conducted by the commission indicated that trip generation within southeastern Wisconsin increased by about 25 percent from 1963 to 1972, with work trips having the smallest increase (19 percent) and shopping trips the largest (30 percent). The ability of the trip generation equations developed and applied in 1963 to accurately predict these changes in regional trip generation is shown in the comparison of estimated and observed number of trips in 1972 given below.

Trip Purpose	Estimated No. of Trips	(Estimated/Observed) × 100%
Home-based work	1 151 800	108.2
Home-based shopping	770 600	113.6
Home-based other	1 552 700	99.8
Non-home-based	749 100	90.0
Total	4 435 100	102.5

The equations predicted regional trip generation with a high degree of accuracy considering the nature of the phenomena involved. That is, the actual 1972 regional trip generation data used as the basis for comparison, the 1963 trip generation data used to calibrate the original equations, and much of the data necessary to prepare predictions of 1972 trip generation—household socioeconomic characteristics—are all estimates derived from travel surveys. Thus, considering the limitations inherent in the data, the total trips generated in 1972 were predicted by the equations with a high degree of accuracy, although with some divergence with respect to trip purpose.

The ability of the original equations to estimate 1972 trip generation on the level of a small geographic area or traffic analysis zone is shown in Figures 1 and 2. These figures indicate the correspondence between observed and estimated 1972 zonal trip productions for the trip purposes of home-based work and home-based other; similar results were obtained for other trip purposes. Although there are considerable differences between actual and predicted trip generation by zone, there is no consistent bias of overestimation or underestimation. Moreover, much of the variance can be attributed to the random variations expected in any survey data, as well as to zonal characteristics not considered in the trip generation equations—both of which may cause deviations between observed and estimated values from regression procedures in a base year—rather than to possible changes over the past decade in the relationship between trip generation and the variables used to explain trip making in the equations developed in the initial planning effort (3). Again, considering the nature of the data used to develop the equations and to compare observed and estimated trip generation, and the detailed level at which this analysis and comparison were conducted, the 1972 trip generation was predicted with an adequate degree of accuracy on a zonal level.

MODAL SPLIT

Modal split was determined prior to trip distribution in the initial regional land use and transportation study for southeastern Wisconsin. The trip end models used were based on the relationships between the percent transit use in a traffic analysis zone, the average household automobile availability in the zone, and the relative availability and quality of highway and transit service as measured by an accessibility ratio (4). Two separate sets of modal split models were calibrated for the three urban areas within the region in which there was transit service in 1963: one set for the Milwaukee area and the other for the Racine and Kenosha areas combined. In the Milwaukee urban area the modal split relationships were developed for four trip purposes: home-based work, home-based shopping, home-based other, and non-home-based. The Racine and Kenosha urban area models were developed for three trip purposes: home-based work, home-based other and shopping, and non-home-based. The modal split relationships were defined mathematically by developing by hand three-dimensional surfaces whose axes were: automobile availability expressed in terms of the number of automobiles per household in a zone, the accessibility ratio of a zone for the trip purpose considered, and the percent transit use.

The modal split modeling procedure used in the initial land use and transportation study for the Milwaukee area was reviewed and modified slightly as a part of a Mass Transit Technical Planning Study in Milwaukee County begun in 1968 and completed in 1971. This modification included the consideration of home-based shopping, home-based other, and non-home-based trips in a single combined model as opposed to the three separate models of the initial study, and a redefinition of the accessibility ratio as used in the original model formulation (5).

The ability of the modal split models developed and used in the initial land use and transportation study and the Milwaukee County Mass Transit Technical Planning Study to predict 1972 transit use within the region was evaluated using actual 1972 O-D survey data. From 1963 to 1972 transit use in southeastern Wisconsin decreased significantly, in both the total number of transit trips and the percentage of the total market that used transit for trip making. The reduction in transit trip making was about 50 percent in the Milwaukee urban area and almost 80 percent in the Racine and Kenosha urban areas. The ability of the modal split models formulated and calibrated in the initial transportation study to estimate this change in regional transit use over the past 9 years is illustrated in Table 1. The model from the original land use and transportation study overestimated 1972 transit use within the region by approximately 10 percent; the modified model underestimated transit use within the Milwaukee area by about six percent. However, since the data—such as transit travel and total person travel by zone, zonal automobile availability, trip attractions, and transit and highway zonal interchange travel times—used in the application of the model are estimates, and since substantial changes in automobile availability and transit service and use have occurred over the past decade, the 1972 regional transit use predicted through application of the original and modified modal split models has a high degree of accuracy.

The ability of the modal split model to estimate transit use on a traffic analysis zone level in the Milwaukee urban area is shown in Figure 3, which displays the correspondence between predicted and observed 1972 zonal total transit trip productions. Similar results were obtained with the modified modal split model and, in the Racine-Kenosha areas, with the original predic-

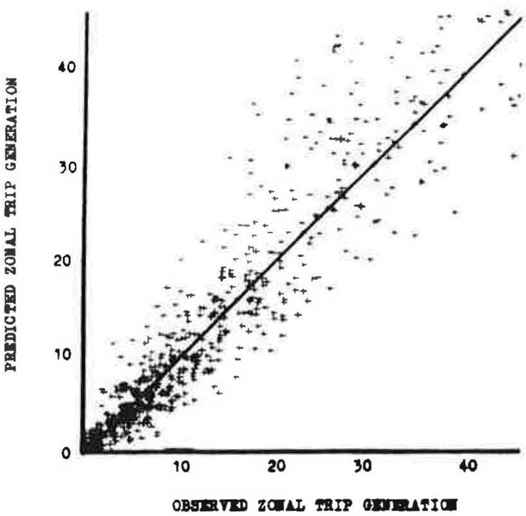


Figure 1. Comparison of predicted and observed 1972 total person home-based work trip generation by zone.

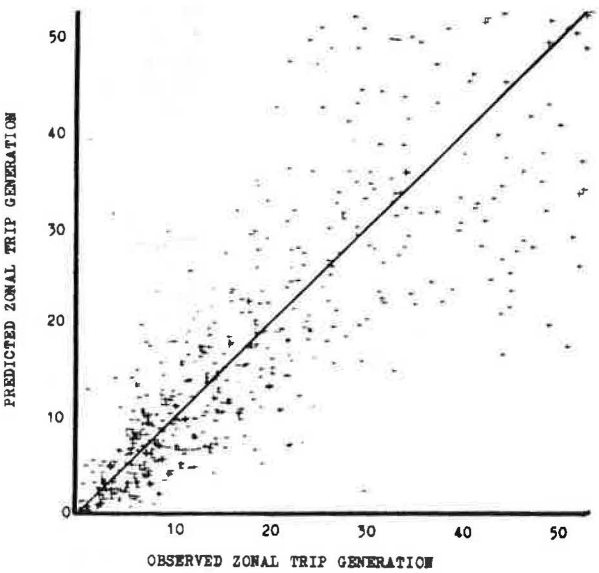


Figure 2. Comparison of predicted and observed 1972 total person home-based other trip generation by zone.

Table 1. Comparison of observed and estimated 1972 transit use within the southeastern Wisconsin region.

Trip Purpose	Estimated No. of Trips: 1963 Model	Estimated No. of Trips: Modified 1963 Model	(Estimated/Observed)ª: 1963 Model	(Estimated/Observed)ª: Modified 1963 Model
Milwaukee				
Home-based work	78 810	72 070	111.9	102.3
Home-based nonwork[b]	63 860	48 850	108.7	83.1
Home-based shopping	21 310		117.0	
Home-based other	25 070		90.5	
Non-home-based	17 370		136.6	
Subtotal trips[b]	142 670	120 920	110.4	93.6
Racine-Kenosha				
Home-based work	1 040		100.2	
Home-based shopping/other	1 510		125.8	
Non-home-based	240		82.0	
Subtotal	2 790		110.2	
Total[b]	145 460		110.4	

ªNo. of Trips × 100%. [b]Also includes school trips.

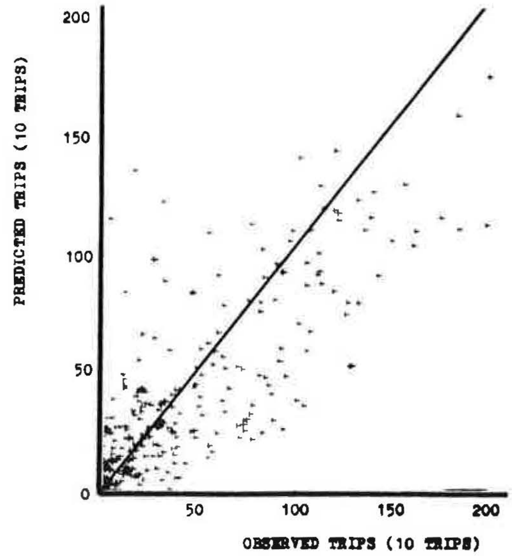


Figure 3. Comparison by zone of predicted and observed 1972 total transit trips in the Milwaukee urban area.

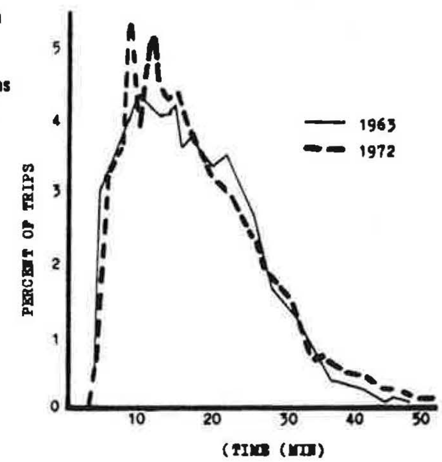


Figure 4. Comparison of observed total person trip length frequency distributions for home-based work travel within the region: 1963 and 1972.

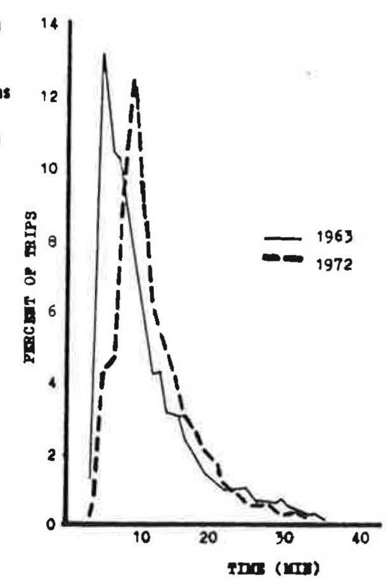


Figure 5. Comparison of observed total person trip length frequency distributions for home-based shopping travel within the region: 1963 and 1972.

tion procedures. Although there are considerable differences between the observed and predicted zonal transit trips, there is no consistent bias of substantial overestimation or underestimation. Much of the variance can again be attributed to random variation in survey data, rather than to changes in the relationship between modal split and automobile availability and the relative quality and quantity of highway and transit service over the past decade.

TRIP DISTRIBUTION

In the initial study internal person trip distribution was simulated by mode, following modal split through the uses of automobile driver and transit person gravity models, for travel with the trip purposes of home-based work, home-based shopping, home-based other, and non-home-based. For each of these gravity models, the calibrated friction factors were assumed to remain valid for the future. Zonal adjustment factors, although investigated, were not used for forecasting future travel patterns.

The stability of the 1963 trip distribution procedure was tested through a comparison of predicted and observed 1972 trip length characteristics. Trip length characteristics within southeastern Wisconsin have remained fairly stable from 1963 to 1972. As shown below, the average trip length for automobile and transit travel increased only slightly over the past 9 years for both modes for all trip purposes except automobile travel with the trip purpose of home-based shopping, which

Trip Purpose	Average Trip Length 1963	Average Trip Length 1972	Change (%)
Transit			
Home-based work	35.9	37.2	+3.7
Home-based shopping	28.5	31.9	+11.9
Home-based other	32.5	36.3	+11.7
Non-home-based	28.4	31.2	+9.9
Automobile driver			
Home-based work	17.9	17.9	0.0
Home-based shopping	9.2	11.6	+26.5
Home-based other	12.4	13.5	+9.4
Non-home-based	12.6	14.0	+11.9

increased significantly. Another measure of the trip length characteristics simulated in trip distribution is the trip length frequency distribution, a determination of the percentage of total trips that occur in 1-min time increments. From 1963 to 1972 the trip length frequency distribution for combined automobile and transit travel for the trip purpose of home-based work remained stable, as shown in Figure 4, but the frequency distributions for combined automobile and transit travel for all other trip purposes changed slightly. This change consisted of a shift in the peak trip length, as shown in Figure 5, which compares 1963 and 1972 trip length frequency distribution for aggregated automobile and transit travel for the trip purpose of home-based shopping.

The ability of the automobile driver and transit person gravity models, as calibrated in 1963, to estimate this change, measured in terms of average trip length and trip length frequency distributions, is shown below and in Figures 6 and 7. The average trip lengths for trips with a purpose of home-based work for both automobile driver and transit person were accurately predicted. The average trip lengths for home-based shopping, home-based other, and non-home-based trips for both automobile driver and transit person were predicted with reasonable accuracy, considering that the data used to establish both the actual and estimated trip distributions were estimates derived from travel surveys. The predicted trip length frequency distributions generally corresponded with the observed distributions; however, for all modes and all trip purposes except home-based work, the peak percentage of trips within a single time increment had been predicted to occur in a time increment shorter than that observed in the 1972 travel survey data. Figure 6 displays the accuracy with which the transit trip length frequency distributions of 1972 for the trip purpose of home-based work were predicted. The differences between predicted and observed 1972 automobile and transit trip length distributions for other trip purposes are shown by the example of Figure 7, which compares observed and predicted 1972 distributions for home-based other travel by the automobile. However, although trip length characteristics were predicted with reasonable accuracy with the 1963 models, a better test of the time stability of the trip distribution procedure would have been a test of its ability to predict zone-to-zone trip interchanges over time.

Figure 6. Comparison of 1972 predicted and observed transit trip length frequency distributions for home-based work travel within the region.

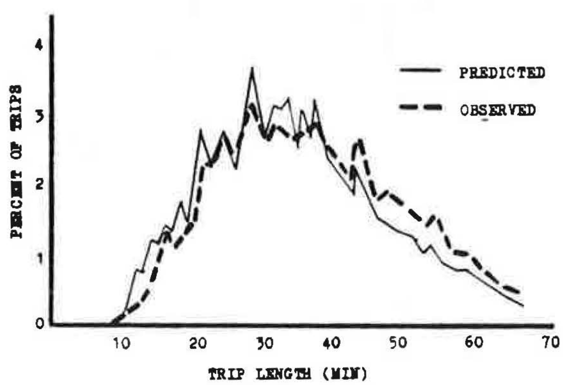
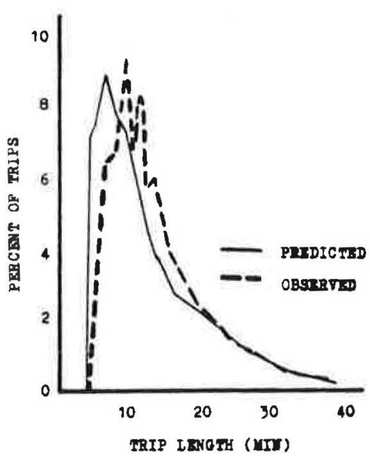


Figure 7. Comparison of 1972 predicted and observed automobile driver trip length frequency distribution for home-based other travel within the region.



Trip Purpose	1972 Average Trip Length Predicted	1972 Average Trip Length Actual	Difference (%)
Automobile driver			
Home-based work	17.9	17.9	0.0
Home-based shopping	9.3	11.6	-19.8
Home-based other	12.7	13.5	-5.9
Non-home-based	12.4	14.0	-11.4
Transit			
Home-based work	36.9	37.2	-0.8
Home-based shopping	29.6	31.9	-7.2
Home-based other	32.3	36.3	-11.0
Non-home-based	27.2	31.2	-12.8

SUMMARY AND CONCLUSIONS

The assumption of the stability over time of travel simulation models calibrated with base-year data is an essential element of present urban transportation planning. This assumption was tested in southeastern Wisconsin by using travel simulation models calibrated with data from an O-D survey conducted in 1963 and travel inventory data from a second survey completed in 1972. This was accomplished by comparing observed 1972 trip generation, transit use, and trip length characteristics with estimates derived by applying the original 1963 trip generation, modal split, and trip distribution models individually to 1972 observed independent variable data; the testing indicated that the predictions of the models on both regional and zonal levels were reasonably accurate. Even though the model relationships for travel forecasting purposes had remained stable over time, changes were made in the travel simulation modeling framework used in the reevaluation of the original land use and transportation plan for southeastern Wisconsin. These changes were made primarily as a result of advances in the state of the art in travel simulation and included the use of cross-classification analysis for trip generation as opposed to the aggregate technique used in the original study.

ACKNOWLEDGMENTS

The author wishes to acknowledge the assistance and advice given by Kurt W. Bauer, Executive Director of the Southeastern Wisconsin Regional Planning Commission, in the interpretation and presentation of the results of this study.

REFERENCES

1. P. R. Stopher and A. H. Meyburg. Travel Demand Estimation: A New Prescription. Traffic Engineering and Control, Nov. 1974, pp. 879-884.
2. N. Ashford and F. Holloway. Time Stability of Zonal Trip Production Models. Transportation Engineering Journal, ASCE, Vol. 98, No. TE4, Nov. 1972, pp. 799-806.
3. Forecasts and Alternative Plans. Southeastern Wisconsin Regional Planning Commission, Waukesha, Planning Rept. 7, Vol. 2, 1966.
4. E. Weiner. A Modal Split Model for Southeastern Wisconsin. SEWRPC Technical Record, Waukesha, Wisc., Vol. 2, No. 6.
5. Documentation of the Travel Simulation Models Used in the Mass Transit Technical Planning Study in Milwaukee County. Southeastern Wisconsin Regional Planning Commission, Waukesha, 1971.

Location, Housing, Automobile Ownership, and Mode to Work: A Joint Choice Model

Steven R. Lerman, Department of Civil Engineering, Massachusetts Institute of Technology

Household location decisions are closely related to other choices of housing, automobile ownership, and mode to work. This paper describes a model that considers these long-run decisions, termed the mobility bundle, as jointly determined, thus eliminating the need for an arbitrary set of assumptions about a sequence of choices. The model developed is based on disaggregate choice theory. Each potential location-housing-automobile ownership-mode to work combination is a distinct alternative, of which only one is selected by each household. The basic methodology used is the multinomial logit model. A sample of skilled, single worker households working and residing in the Washington, D.C., metropolitan area in 1968 was used to estimate the model. The variables used to describe each alternative included locational attributes, housing attributes, transportation level of service to work, spatial opportunities for shopping trips, automobile ownership attributes, and the socioeconomic characteristics of the household. Even with the relatively small sample used, a wide range of behavioral effects were measured. It is concluded that models such as the one described here could replace existing model systems used to forecast residential location patterns. The increase in behavioral content such models permit would allow credible work-trip forecasts to be made as a part of land use forecasting.

Transportation planners have long recognized that the facilities they implement today will have a great influence on the future location patterns of the activities those facilities will serve. These long-term interactions between transportation and location, termed activity shifts, have been the focal point of much of the debate about the desirability of various fixed transportation investments such as highways or rail rapid transit systems. The locational impacts of such facilities are not only important in and of themselves; they also have an enormous influence on the future transportation demand in that the size and characteristics of the served population may be greatly altered over the useful life of the facility.

These locational effects may be most significant in urban transportation planning. For example, a new transit system in a city will in the short run attract current highway users; this mode choice effect has been the principal focus of the vast bulk of travel demand modeling efforts and is probably the best understood aspect of travel behavior. However, in the longer run, the same system may have profound effects on residential and employment location patterns. Sites near transit stations will be more desirable (1), and those farther away may experience reduced levels of activity. Furthermore, as a consequence of the new location and work trip patterns, households may alter their automobile ownership and housing decisions. Thus, over its useful life, a major transportation facility may have the potential to completely alter the spatial structure of a city. Until the nature of the interactions among transportation services, location decisions, and travel demand is understood, there is little hope that existing forecasting models will provide reliable tools for long-run policy analysis (2).

This paper describes a model that applies disaggregate choice models to represent household location and the related choices of housing type, automobile ownership level, and mode to work. The particular choice theory methodology used is the multinomial logit model (3), which is analytically tractable and widely used (4, 5, 6).

The first section of the paper describes the set of available alternatives, i.e., the various location-housing-automobile ownership-mode to work combination or mobility bundles, that are feasible choices for representation in the model. The next section is a brief description of the way in which the various factors that affect the choice of mobility bundle are combined to form the variables entering into the household utility function. [This description is somewhat cursory; further descriptions of the model structure, the motivation for the variables used, and the final functional form can be found in Lerman (7)]. This is followed by a presentation of the parameter estimation results and, finally, by a revised forecasting framework that uses the model structure developed and is consistent with a behavioral theory of how household decisions are reached.

DEFINITION OF MOBILITY BUNDLES

Location, housing, automobile ownership, and mode of travel to work can be almost infinitely subdivided. Locations can be taken to be cities, towns, census tracts, blocks, zones, or any other geographical unit. Alternatively, location can be defined simply in terms of distance from the CBD or in terms of whether a site is in the central city, urban ring, suburbia, or rural fringe. Housing can be defined along a broad spectrum of dimensions, including age of the structure, lot size, architectural style, number of rooms of various types, garage space, quality and condition of unit, and type of tenure. Automobile ownership can consist of the number of automobiles as well as their make, age, gas mileage, horsepower, or operating cost. Mode to work can be classified as transit or car, or further described as bus, trolley, rail rapid transit, taxi, shared ride, paid car pool, or drive alone.

Clearly, at some level of detail the number of possible alternative mobility bundles is enormous. Even if suitable data were available, a model developed with such detailed alternatives would be almost impossible to estimate and apply. Some level of abstraction in defining alternatives is required.

Reducing the number of alternatives of the location dimension of the mobility choice presents some basic methodological difficulties. Ultimately, each household selects a particular dwelling unit, of which there are thousands or millions in any one metropolitan area. Any method of grouping alternatives is by its very nature somewhat arbitrary. Fortunately, it is possible under a set of fairly rigorous approximations to use the multinomial logit model in a way that allows one to rely on data about arbitrarily defined groups of dwelling units, but still obtain consistent estimates of parameters describing how households perceive the dwelling units themselves (7). This approximation makes it possible to consider census tracts as the basic locational alternative and permits the use of a large, relatively reliable data base.

The dimensions of housing alternatives used in this study are structure type and tenure. Four feasible choices, owner-occupied single family house, rented single family house, rented walk-up or garden apartment, and high rise dwelling, were used. This choice was determined primarily by the data available from the home interview survey. For the purposes of transportation planning, where the primary focus is on the spatial aspect of mobility rather than on the housing itself, this choice set should be adequate. However, the use of structure type and tenure only does restrict the applicability of this study to the analysis of housing policies, where issues of structure size and quality are very significant.

Automobile ownership choices have the dimensions of number owned, make, type, horsepower, and options. However, since the transportation planner is mainly interested in the ways in which people will alter their travel behavior in response to various policies, a consideration of the number of automobiles should be sufficient.

The final choice, mode of travel, is restricted to two modes of vehicular travel, car and transit. These two together constitute 93 percent of all work trips in the study city, Washington, D.C. (This figure does not include some very short work trips.) In order to further limit the scope of the empirical study, all forms of ride sharing were eliminated.

Not all possible residential location-housing-automobile ownership-mode to work combinations are feasible alternatives. For example, some locations are not served by transit, some tracts have a limited range of housing options due to zoning ordinances, certain households may not have some mobility bundles open to them (households without drivers do not own automobiles, and low-income households can only afford a limited subset of all feasible mobility bundles). In addition to restrictions such as these, the tracts available to any particular household are limited to the set of tracts actually selected by all workers (in the sample used) in the employment zone. This method of defining choice availability (5, 8) provides an operational way to avoid including infeasible alternatives in household choice sets: The fact that some feasible alternatives may be eliminated does not affect the properties of the parameter estimates in the multinomial logit model.

SPECIFICATION OF THE JOINT UTILITY FUNCTION

The variables that affect the choice of mobility bundle can be divided into six general categories. They are as follows:

1. Transportation level of service to work—travel time (in-vehicle and excess time) and cost for the work trip;
2. Automobile ownership attributes—taxes, depreciation, registration costs, maintenance, and title costs;
3. Locational attributes—neighborhood quality, demographic composition, taxes, urban services, parking availability, and local insurance rates;
4. Housing attributes—age of the structure, quality, size of the unit, garages, driveways, and structure type;
5. Spatial opportunities—measures of accessibility to shopping and other nonwork destinations; and
6. Socioeconomic characteristics—income, race, household size, number of drivers, number of workers, education, and marital status.

Each of the above six categories might be represented with a wide range of variables and each of the variables can, in theory, be combined with the others to produce a virtually limitless number of possible independent variables for each utility function. The final set of variables selected draws heavily on previous empirical work and on a set of more fully developed model estimations (7).

In a joint model of mobility choice there are an extremely large number of possible location-housing-automobile ownership-mode to work combinations. Thus, the number of utility functions is correspondingly great. Rather than consider each utility function individually, every variable will be defined as pertaining to all alternatives but as taking zero value for those utilities where it is not included.

The first group of variables are constant terms in the utility function. These constants measure the pure alternative effect, i.e., the net effect of all attributes of an alternative that are not measured by the other variables. In theory, a constant could be introduced into all but one utility function that would act as a base against which the effects of the other variables are measured. This choice of the base utility function would be arbitrary and have no influence on the parameter estimates of the choice probabilities. In practice, however, alternatives such as locations that are unranked and very numerous do not have constants associated with them unless they have particular attributes that make them distinguishable, such as the CBD in models of destination choice.

Even when the location choice group is ignored, the number of possible options is quite large. A household has a maximum of two modes: car and transit; three automobile ownership levels: zero, one, and two or more; and four housing types: to own a single family house, to rent a single family house, to rent a garden or walk-up apartment, and to rent a high-rise apartment. After the logically inconsistent alternative of zero automobile ownership and car to work is eliminated, there are 20 possible options for any one household. To reduce the number of dummy variables to less than 19 (the number of possible options minus one for a base), a way to approximate each independent effect by a linear combination of a smaller number must be found. The approach adopted here is to give each choice group a constant term for all its members but one, and assign constants to those of the interactions among choice groups that an exploratory data analysis had indicated to be significant. The resulting set of eight constant terms are as follows:

$$\text{DRENT1} = \begin{cases} 1 \text{ in the rent single family dwelling alternative} \\ 0 \text{ otherwise} \end{cases}$$

$$\text{DRENTG} = \begin{cases} 1 \text{ in the rent garden or walk-up apartment alternative} \\ 0 \text{ otherwise} \end{cases}$$

$$\text{DRENTH} = \begin{cases} 1 \text{ in the rent high rise apartment alternative} \\ 0 \text{ otherwise} \end{cases}$$

$$\text{DA01} = \begin{cases} 1 \text{ in the one automobile alternative} \\ 0 \text{ otherwise} \end{cases}$$

$$\text{DA02} = \begin{cases} 1 \text{ in the two or more automobiles alternative} \\ 0 \text{ otherwise} \end{cases}$$

$$\text{DCAR} = \begin{cases} 1 \text{ in the car to work alternative} \\ 0 \text{ otherwise} \end{cases}$$

$$\text{DAPTSTYL} = \begin{cases} 1 \text{ in the rent garden, walk-up, or high rise apartment and own less than two automobiles alternatives} \\ 0 \text{ otherwise} \end{cases}$$

$$\text{DSUBSTYL} = \begin{cases} 1 \text{ in the own single family dwelling and own two or more automobiles alternatives} \\ 0 \text{ otherwise} \end{cases}$$

The next two variables represent the travel time aspects of level of service to work. These variables have been expressed as the in-vehicle and out-of-vehicle time in most mode choice studies. More recent work (9, 10) has indicated that the disutility of out-of-vehicle time may be perceived as a function of the total trip length, which can be measured by travel distance. After experimentation with a number of alternative functional forms, the following specification (9) was chosen:

TOTIME = total two-way travel time (min) and

OVTT/DIST = two-way out-of-vehicle time (min) ÷ two-way travel distance (km)

In addition to these variables, a dummy variable was defined to reflect the added disutility associated with the use of a car in the downtown area as follows:

$$\text{DCITY} = \begin{cases} 1 \text{ for households with downtown workplaces in the car to work alternatives} \\ 0 \text{ otherwise} \end{cases}$$

The next variable arises from the fact that there are a large number of monetary measures such as household income; federal, state, and local taxes; housing costs; automobile ownership costs; and out-of-pocket travel costs for the work trip in the model. To avoid introducing a separate variable for each of these cost factors, they were combined into a single one termed for reference the Z variable, representing the money that would be available to the household if it selected each alternative. The value of Z (in dollars per year) is thus an estimate of the amount of money a household has after the following expenses: (a) federal taxes, (b) state taxes, (c) property taxes (if applicable), (d) housing cost, (e) direct automobile ownership costs, (f) automobile insurance, tags, and taxes, and (g) commuting cost to work. The coefficient of the Z variable in the utility function should always be positive, reflecting the fact that all else being equal households would rather have more money than less left for other things. The Z variable should not enter the utility functions linearly; the utility a poor family derives from extra money is much greater than that derived by a wealthy family. Thus, the marginal utility of money should decrease as the value of Z increases. This hypothesis can be reflected by using the natural log of Z rather than simply the value of Z as the independent variable.

The next variable used commonly appears in simple mode choice models in which automobile ownership is assumed fixed. It is defined as

$$\text{AALD} = \begin{cases} \text{number of automobiles in alternative} \div \text{number of licensed drivers in the household in the car to work alternatives} \\ 0 \text{ otherwise} \end{cases}$$

and represents the level of automobile availability that would be obtained if the household chose a given alternative. Alternatives with high automobile availability should be associated with high car to work utilities relative to those for transit to work alternatives; hence, the expected sign of its coefficient is positive.

The next variable was designed to reflect another effect of the number of licensed drivers within a household. While the number of licensed drivers impacts on choice of mode to work through the AALD variable, it also affects the level of automobile ownership directly: the more licensed drivers in a household, the more likely it will select a high automobile ownership level, independent of the mode to work selected. This effect was measured by introducing into each utility function a variable that reflects the number of licensed drivers with a different coefficient for each automobile ownership level, except one selected arbitrarily as a base. These variables are defined as follows:

$$\text{ILD1} = \begin{cases} 1/\text{number of licensed drivers in the household for one automobile alternative} \\ 0 \text{ otherwise} \end{cases}$$

$$\text{ILD2} = \begin{cases} 1/\text{number of licensed drivers in the household for two automobiles alternative} \\ 0 \text{ otherwise} \end{cases}$$

When these variables were originally introduced into the model, it was hypothesized that the effect for the two car alternative (as measured by the coefficient value) would be twice as great as the effect for the one car alternative. Statistical tests by Lerman and Ben-Akiva have indicated that this is indeed the case and that ILD1 and ILD2 can be combined into a single variable defined as follows:

$$\text{ILD} = \begin{cases} 0 \text{ for the zero automobile alternatives} \\ \text{ILD1 for the one automobile alternatives} \\ 2\text{ILD1 for the two automobiles alternatives} \end{cases}$$

The use of the inverse of the number of drivers rather than the number itself reflects the hypothesis that, as the number of drivers increases, the marginal effect of an additional driver on the need for automobiles decreases. Clearly, the coefficient of ILD should be less than zero.

Spatial opportunities influence the mobility decision in at least two ways. First, the absolute level of accessibility to shopping by either car or transit is probably important in a household choice of location. Second, the level of shopping accessibility by car relative to that of transit affects the mode with which the household will travel to shop, which in turn influences their desired level of automobile ownership.

The first of these effects was represented by

GPTINV = 1/expected generalized shopping price by transit

The generalized shopping price by transit is a weighted sum of estimates of the average in-vehicle time, out-of-vehicle time, and out-of-pocket cost of a shopping trip by transit that is derived from a disaggregate shopping trip choice model (10). The variable GPTINV is zero when transit is completely unavailable since the transit generalized price in such areas is, for practical purposes, infinite. The coefficient of this variable should be positive, since decreased travel costs resulting from improved transit service should increase the household utility.

Attempts to use a corresponding variable for the absolute level of car accessibility produced statistically insignificant coefficient estimates having an unexpected sign. This problem (15) may be the result of the high levels of externalities (such as noise and traffic congestion) often associated with locations with good highway accessibility (12). Thus, the car accessibility coefficient may also be measuring the effects of some omitted variables. For this reason, it was not included in the final specification.

The effect of the relative accessibility was measured by a variable defined as follows:

R = expected generalized car costs for shopping ÷ expected generalized transit cost for shopping

This variable does not change value for different automobile ownership alternatives, and therefore must be introduced into the utility function as alternative specific. Thus, the following two variables appear in the model:

$$R1 = \begin{cases} R \text{ in one automobile alternatives} \\ 0 \text{ otherwise} \end{cases}$$

and

$$R2 = \begin{cases} R \text{ in two automobiles alternatives} \\ 0 \text{ otherwise} \end{cases}$$

As the generalized shopping cost by car increases, the value of R increases. This increase in car cost should result in greater use of transit for shopping trips and, consequently, the likelihood of high automobile ownership should decrease. The coefficients of both R1 and R2 should therefore be negative, since they both measure the effect of shopping accessibility relative to the zero automobile-transit to work alternatives. Furthermore, the effect should be greater for the two automobiles alternatives than for the one automobile options, and the magnitude of the coefficient of R2 should be greater that that of R1.

In order to reflect the effect of household size on the desire for living in a single family dwelling, the following variable was defined:

$$\text{HHSIZE1} = \begin{cases} \text{household size in single family dwelling} \\ \quad \text{alternative (own or rent)} \\ 0 \text{ otherwise} \end{cases}$$

The coefficient estimate of this variable should be greater than zero, since, other things being equal, larger households probably have a stronger perference for single family dwellings than do smaller households.

The next group of variables are all locational attributes and are defined as follows:

$$\text{INCDIFF} = \begin{cases} (Y - \overline{Y})^2 \text{ for } Y \geq \overline{Y}, \text{ where } Y \text{ and } \overline{Y} \text{ are the} \\ \quad \text{household and average annual tract in-} \\ \quad \text{come, in thousands of dollars} \\ 0 \text{ otherwise} \end{cases}$$

$$\text{FBFORW} = \begin{cases} \text{fraction of nonwhite households in tract} \\ \quad \text{for white households} \\ 0 \text{ for non-whites} \end{cases}$$

$$\text{FBFORB} = \begin{cases} \text{fraction of nonwhite households in tract} \\ \quad \text{for nonwhites} \\ 0 \text{ for whites} \end{cases}$$

$$\text{DENSITY} = \begin{cases} \text{net residential density in households per} \\ \quad \text{acre} \end{cases}$$

$$\text{SCHOOL} = \begin{cases} \text{per pupil school expenditures (in dollars} \\ \quad \text{per year) except in District of Columbia} \\ 0 \text{ in District of Columbia} \end{cases}$$

$$\text{DOC} = \begin{cases} 1 \text{ in District of Columbia} \\ 0 \text{ otherwise} \end{cases}$$

The first variable is a measure of the neighborhood quality in a tract. The income differential is squared, reflecting the hypothesis that large differences are proportionately much more important than small ones. This variable should have a negative coefficient. The opposite variable, which is defined as non-zero when the household income is less than the average, consistently gave very small, statistically insignificant estimates having the wrong sign, and was omitted in the final specifications.

The two racial composition variables reflect the hypothesis that whites and nonwhites perceive the racial composition quite differently. The coefficients of FBFORW and FBFORB should be negative and positive respectively.

The density variable is self-explanatory; a negative coefficient should be expected. The DOC dummy variable was defined to correct for the setting of the annual per pupil school expenditure variable to zero in the SCHOOL variable, the coefficient of which should have a positive sign. The SCHOOL variable is defined to be zero for households without children even though the DOC variable is not. This was done to explore the possibility that the District has certain attributes that make it distinct from other locations regardless of whether or not a household has children.

The final variable in the model is the measure of tract size required to correct for the fact that a census tract is actually a group of housing units. Other conditions being equal, a very large tract (i.e., one with a large number of housing units) would have a higher probability of being selected than a very small one, since the number of disaggregate opportunities is greater in the former than the latter. If all units of a particular type in a given zone are relatively homogeneous and the logit model applies to each individual unit, then the appropriate term to correct for tract size is the natural logarithm of the number of units (7). This variable, denoted as lnN, should, under the previously cited assumptions, have a coefficient of one. However, if the assumptions of the logit model are violated, the coefficient may differ from one; for this reason, parameter estimates both with and without the coefficient of lnN constrained to unity are reported.

In order to derive the structure of the utility function for any particular location-housing-automobile ownership-mode to work combination, the variables that are set to zero for that alternative can be omitted.

ESTIMATION RESULTS

The variables in the model described above were estimated using the maximum likelihood method (13) with data from a 1968 Washington, D.C., home interview sur-

Table 1. Parameter estimates for the model.

No.	Variable	Unconstrained Estimates	Constrained Estimates
1	DRENT1	-0.361	0 393
		(-1.03)	(1.18)
2	DRENTG	2.31	2 93
		(2.87)	(3.58)
3	DRENTH	0.828	0.809
		(1.02)	(0.973)
4	DAO1	7 86	7.98
		(2.57)	(2.60)
5	DAO2	12.0	12.1
		(2.71)	(2.78)
6	DCAR	0.433	0.483
		(0.500)	(0.501)
7	DAPTSTYL	0.542	0.524
		(0.966)	(0.927)
8	DSUBSTYL	0 336	0.261
		(0.764)	(0.591)
9	TOTIME	-0.008 31	-0 008 18
		(-2.13)	(-2.05)
10	OVTT/DIST	-0.057 0	-0.052 6
		(-0.787)	(-0.708)
11	DCITY	-0.437	-0.415
		(-0.932)	(-0.879)
12	ln Z	1.07	1.20
		(2.64)	(2.81)
13	AALD	0.964	0.975
		(1.01)	(1.02)
14	ILD	-6.57	-6.56
		(-2.17)	(-2.16)
15	GPTINV	2.92	3.14
		(1.38)	(1.47)
16	R1	-1.35	-1.54
		(-1.08)	(-1.21)
17	R2	-4.05	-4.11
		(-3.01)	(-3.03)
18	HHSIZE1	0.850	0.875
		(5.21)	(5 16)
19	INCDIFF	-0.012 3	-0.012 1
		(-2.89)	(-2.80)
20	FBFORW	-2.18	-2.21
		(-3.79)	(-3.78)
21	FBFORB	1.95	1.85
		(2.23)	(2.12)
22	DENSITY	-0.005 57	-0 008 10
		(-1.25)	(-1.75)
23	SCHOOL	0.000 442	0.000 342
		(0.685)	(0.523)
24	DOC	-0.009 93	-0 100
		(-2.06)	(-0.204)
25	ln N	0.492	1
		(5.25)	—[a]
	$L^*(0)$	-824.4	-824.4
	$L^*(\hat{\beta})$	-645.9	-658.4
	NOBS	177	177
	NCASES	25 601	25 601
	Percent right	8.5	10.2

[a] Constraint imposed, hence t statistic not relevant

vey for a small sample of single worker households in which the worker was at least minimally skilled. This data set was augmented by 1970 census housing data (appropriately deflated) and transportation level of service from highway and transit networks. Two different estimations were made. The first allowed the coefficient of the tract size variable, lnN, to attain its highest probable value: These estimates are shown in column 3 of Table 1. The second set of estimates are based on the constraint that the coefficient of the tract size variable is unity: These estimates are shown in column 4 of Table 1. For each model, the asymptotic t-statistics are given in parentheses below their corresponding parameter estimates. In addition, five summary statistics are given.

1. $L^*(0)$ is the value of the log probability function when all of the parameters are zero (i.e., when every alternative has the same probability);
2. $L^*(\hat{\beta})$ is the value of the log probability function at the maximum probability coefficient values;
3. NOBS is the number of households in the sample;
4. NCASES is the number of available alternatives (in excess of one per household) used in the estimation; and
5. Percent right is the percentage of households for which the alternative with the highest systematic component of utility was actually selected. This value is maximized when the maximum score estimation technique is used (6).

All of the coefficient estimates in both the unconstrained and constrained models for variables about which hypotheses were formulated have the expected sign. However, the statistical significance of some coefficients such as the estimate for OVTT/DIST is marginal. This probably results from the very small sample used, since mode choice models with larger samples of the Washington data result in estimates significantly different from zero at fairly high levels of confidence.

The constrained estimates are similar to the unconstrained ones, with the exception of the coefficient of DRENT1. This suggests the possibility of some measurement error in the value of N, the number of units of rented single family dwellings.

As might be expected with an average of over 145 alternatives available for each household, the percent right result is low in absolute terms. A useful way of viewing this is in terms of the probability of a model classifying a given percent correctly if all the alternatives were actually equally likely. Suppose that all households have exactly 145 available alternatives, and that each is equally likely. In this case, the probability of a model classifying none of the 177 observations correctly is

$$(144/145)^{177} = 0.2938$$

and the probability of classifying k of 177 correctly is distributed as binomial, i.e.,

$$\Pr(k \text{ correct}) = \binom{177}{k} (1/145)^k (144/145)^{177-k}$$

According to this formula, the probability of classifying nine or more households correctly (about 5 percent right) is less than 0.0001, but the percent right found for the unconstrained and constrained estimates are 8.5 percent and 10.2 percent respectively.

IMPLICATIONS OF THE RESEARCH FOR REVISED ANALYSIS FRAMEWORK

Forecasts of urban land use have traditionally played an important role in the transportation planning process. However, the models used to forecast residential location patterns have usually been logically separable from those used to forecast both automobile ownership and trip-making patterns. Land use models have provided forecasts of zonal population and employment before a separate automobile ownership model is applied. These forecasts are then used as inputs to the four-step process, consisting of trip generation, distribution, mode split, and assignment.

A critical implication of the theory upon which this study is based is that such a forecasting approach fails to behaviorally represent the true process it seeks to model. In reality, it is more reasonable to assume that both automobile ownership and work-trip travel patterns arise as a logical consequence of a long-term choice process and should therefore be forecast within what has traditionally been termed urban land use forecasting. For work trips, trip generation actually represents labor force participation in a decision process that probably depends more on the household structure and life-style and the state of the regional economy than on the transportation system. Work-trip distribution is simply the outcome of urban location patterns rather than a distinct behavioral phenomenon that can be modeled separately.

Nonwork trips can then be modeled as conditional on the outcome of the work-trip pattern. Feedback between the longer run land use component and the nonwork trip patterns can readily be incorporated by extending the shopping generalized price variables to include other relevant trip purposes.

The choice models developed in this paper are prototypes for one part of such a system of models. Many of the other required model system components are the object of a great deal of research (8, 13). New models of the land use supply sector are being studied by the National Bureau of Economic Research. Thus, the proposed forecasting process represents a synthesis of the results of research in a variety of areas, and could possibly be implemented in a fairly short time.

This study represents one step in the process of improving land use forecasts by introducing joint behavioral choice models into the representation of the household mobility choice. Obviously, if the models described here are to be used effectively, they must be a portion of a much larger research effort to restructure the entire travel forecasting process to better reflect a behavioral understanding of the true causal mechanisms that determine supply and demand. Only by developing more behaviorally structured models can transportation analysts and urban planners hope to provide reliable forecasts of the impact of alternative policies on which an informed decision-making process can act.

ACKNOWLEDGMENTS

The author wishes to acknowledge the assistance of Moshe Ben-Akiva, Wayne Pecknold, Marvin Manheim, and Gregory Ingram, all of whom contributed their advice in the course of the study. In addition, the financial support provided to the author by the Motor Vehicle Manufacturers Association is gratefully acknowledged.

REFERENCES

1. R. E. Mudge. The Impact of Transportation Savings on Suburban Residential Property Values. Rand Corp., New York, June 1974.
2. D. B. Lee, Jr. Requiem for Large Scale Models. Journal of the American Institute of Planners, Vol. 39, No. 3, May 1973.
3. D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers in Econometrics (Paul Zarembka, ed.), Academic Press, New York, 1973.
4. Charles River Associates. A Disaggregate Behavioral Model of Urban Travel Demand. Federal Highway Administration, U.S. Department of Transportation.
5. M. Ben-Akiva. Structure of Passenger Travel Demand Models. Massachusetts Institute of Technology, PhD thesis, June 1973.
6. C. F. Manski. The Structure of Random Utility Models. School of Urban and Public Affairs, Carnegie-Mellon Univ., April 1975.
7. S. R. Lerman. A Disaggregate Behavioral Model of Urban Mobility Decisions. Center for Transportation Studies, Massachusetts Institute of Technology, Rept. 75-5, June 1975.
8. M. G. Richards and M. E. Ben-Akiva. A Disaggregate Travel Demand Model. Saxon House, D.C. Heath, Westmead, England, 1975.
9. F. Koppelman. Travel Prediction With Disaggregate Choice Models. Massachusetts Institute of Technology, PhD thesis, June 1975.
10. S. R. Lerman and M. E. Ben-Akiva. A Disaggregate Behavioral Model of Automobile Ownership. TRB, Transportation Research Record 569, 1976, pp. 34-55.
11. W. C. Wheaton. A Bid Rent Approach to Housing Demand. Department of Economics, Massachusetts Institute of Technology, Rept. 135, July 1974.
12. H. Theil. Principles of Econometrics. John Wiley, New York, 1971.
13. T. Adler and M. Ben-Akiva. A Joint Choice Model for Frequency, Destination, and Travel Mode for Shopping Trips. TRB, Transportation Research Record 569, 1976, pp. 136-150.

Transferability and Updating of Disaggregate Travel Demand Models

Terry J. Atherton, Cambridge Systematics, Inc., Cambridge, Massachusetts
Moshe E. Ben-Akiva, Department of Civil Engineering, Massachusetts Institute of Technology

In recent years much work has gone into the development of disaggregate travel demand models. However, little has been done to evaluate the ability of these models to predict travel behavior in locations other than the area for which the model was estimated. Unlike aggregate models, the parameters of disaggregate models are not dependent on a particular zonal system and therefore have the potential for transferability. The motivation behind transferring is clear—if a model estimated in one area can be transferred to another, the cost of conducting transportation studies could be greatly reduced. Several possible approaches for transferring are developed and discussed from a theoretical perspective. For an empirical evaluation, a work-trip modal-split model estimated on Washington, D.C., data is transferred to New Bedford, Massachusetts, using each of the proposed approaches. The results of estimating the original model on Los Angeles data are also represented. The most significant result is the exceptional performance of the original Washington work mode choice model on both New Bedford and Los Angeles data. This is noteworthy in view of the extreme differences of the means for several variables between these cities. Of the several approaches for transferring that were developed, Bayesian updating based on combining the existing model coefficients with the estimation results from a new sample gave the best overall performance. The results of this study indicate that the potential transferability of disaggregate travel demand models can be realized.

Traditional aggregate models of travel demand, which are based on existing relationships between aggregate variables, tend to be correlative rather than causal, and often are insensitive to proposed changes in transportation policy. Recently, travel demand models based on disaggregate data (i.e., individual observations of travel behavior) have been developed. These models can include the causal relationships between transportation level of service, household socioeconomic characteristics, and travel behavior and, therefore, provide a more meaningful analysis of various transportation policy options.

Often, particularly in small urban areas, there is neither the time nor the money to develop a travel demand model. This makes desirable the development of a travel demand model that could be transferred from one area to another. Disaggregate models are most likely to be transferable because they represent the average behavior of the individual traveler, and it is reasonable to expect individual travel behavior to be essentially the same in one area as in another. Moreover, the estimation of disaggregate models does not rely on a particular zonal aggregation so that a correctly specified disaggregate model that properly explains travel behavior in one area should be valid (or at least more valid than a comparable aggregate model) for predictions of travel behavior in other areas.

This paper discusses the theoretical justification for the transferability of disaggregate models. The results of transferring an existing disaggregate mode choice model for work trips, developed using 1968 data from Washington, D.C., to data sets representative of New Bedford, Massachusetts, in 1963 and Los Angeles, California, in 1967 are presented. Several possible approaches for updating are developed, compared from a theoretical standpoint, and evaluated empirically using the New Bedford data base. The most useful product of the research is a procedure for travel demand model development suitable for low-budget or short-duration transportation planning studies.

PROPERTIES OF DISAGGREGATE MODELS

Before empirical updating procedures are developed, the theoretical justification for the transferability of these models should be established by identifying the attributes that affect any model's ability to be transferred from one area to another. Clearly, all those factors that affect the reliability of predictions will also affect the transferability. If a model cannot successfully predict travel behavior in the area for which it was estimated, there is no reason to expect it to function better in any other area. To be transferable, then, it is not enough that the model merely fit existing data; it must also explain why travel behavior changes as conditions change. Rather than simply correlating existing travel behavior with socioeconomic characteristics and transportation level of service, the model specification must represent the causal relationships between these variables. Thus, the causal specification of a model is a precondition to its consideration for transferability.

From a practical point of view, no model is ever perfectly specified. Some variables that should be included in the model often must be excluded (e.g., when the estimation data set does not contain sufficient variability of these variables). In particular, when data for model development are taken from one urban area and applied to another, there may be cultural differences between the two areas that are not explicitly represented in the model. These peculiarities of the data will be implicitly hidden in the model coefficients and so the coefficients estimated in one area will not be valid for the other. For a model to be perfectly transferable its coefficients must be free from contextual factors.

What are the differences between aggregate and disaggregate models that affect their potential transferability? If we assume that the model specification is given, what effect does the use of aggregate or disaggregate data in its estimation have on its potential transferability? An implicit assumption in using aggregated data is that the characteristics of households within zones are relatively homogeneous as compared to the differences between zones. However, several studies have shown the opposite to be true—there is more variation within zones than between them (1, 2). Because of this, problems such as the loss of variability in the data, collinearity between variables, and the risk of an ecological fallacy (3) can arise in the estimation of aggregate models and adversely affect their predictive ability and, hence, their transferability.

Suppose for a moment that these problems have been considered and do not affect the estimation of an aggregate model. One serious problem that still remains is the linkage of the coefficients of the aggregate model to the zonal structure of the area for which it was estimated. This linkage is directly observed from the definition of an aggregate demand model. If the disaggregate model is denoted as $f(X, \theta)$ where X is a vector of independent variables and θ is a vector of the coefficients of the disaggregate model, the aggregate demand is the sum of the disaggregate demands and therefore the aggregate model is

$$\int_X f(X,\theta)\, h(X,\bar{X},\alpha)\, dx \qquad (1)$$

where $h(X, \bar{X}, \alpha)$ is the distribution function of the independent variables for the group on which the aggregation is performed, $\bar{X}$ is the vector of means of the independent variables, and α denotes other parameters (or higher moments) of $h(X, \bar{X}, \alpha)$. The result of this integral is an aggregate demand model that could be expressed as $F(\bar{X}, \alpha, \theta)$, where the function $F(\bar{X}, \alpha, \theta)$ does not necessarily have the same analytical form as $f(X, \theta)$. Traditional aggregate models do not explicitly include all the parameters of the within-zone distributions , $h(X, \bar{X}, \alpha)$, and therefore these parameters are implicit in the resulting coefficents of the model. Since these distributions would certainly differ from one area to another (4) they would have to be reflected in the model in order for it to be transferred successfully. However, existing aggregate models are not capable of this and so are less likely to be transferable than disaggregate models that are estimated on observations of individual behavior and have model coefficients that are not bound to any particular zonal structure. Thus, a disaggregate model is always more transferable than a comparable (i.e., same set of variables) aggregate model.

TEST OF TRANSFERABILITY

As an initial test of the transferability of disaggregate demand models, the specification of an existing mode choice model developed on 1968 Washington, D.C., data was reestimated on data sets representative of New Bedford, Massachusetts, in 1963 and Los Angeles, California, in 1967. The model coefficients from this reestimation and the statistical significance of the differences between these and those of the original model are discussed below.

Existing Model

The model selected for this test (and the subsequent empirical evaluation) is a multinomial logit mode choice model (1) that has been modified and extensively tested (5, 6, 7) [the multinomial logit formulation itself is described in many places (8, 9, 10, 11).] This model predicts the probability of a commuter driving alone, sharing a ride (i.e., two or more persons in a car), or using transit for the home-to-work trip. The model specification is given in Table 1.

The model contains all the normally expected variables—in-vehicle travel time, out-of-vehicle travel time, out-of-pocket costs, income, and automobile availability—plus some special variables to differentiate between alternative modes. A primary worker dummy variable is included for the drive alone mode under the hypothesis that the head of household has some priority in using any available automobile. The CBD dummy variables for the drive alone and shared ride modes express the added inconvenience of driving an automobile into the Washington, D.C., CBD above that reflected in the level-of-service variables. Three additional variables are included to account for the choice of the shared ride mode. These variables are a government worker variable (GW) that serves as a proxy for employer provided incentives for forming car pools, the destination employment density times one-way trip distance variable (DTECA), and the number of workers in the household variable (NWORK).

In transferring this model to the New Bedford and Los Angeles data sets, the specifications of the independent variables are identical to those of the original model with the exceptions that both CBD variables and the government worker variable are excluded. In the case of the CBD variables, the congestion and inconvenience associated with driving into the CBD of a large, dense city such as Washington are real factors in choosing between automobile modes and transit. In a small city such as New Bedford or in a very diffuse city such as Los Angeles, however, the distinction between CBD and non-CBD trips would probably have little effect on this choice. Therefore, the DCITY variables are assumed to have a value of zero. Similarly, the effects of large organizations offering incentives to car pool do not exist in either New Bedford or Los Angeles and the government worker variable also has a value of zero.

Estimation Results

The coefficients and statistics of the models estimated on the Washington, New Bedford, and Los Angeles data sets are given in Table 2. (The data base is given below.)

City	No. of Observations	No. of Alternatives	Log Likelihood at Zero	Log Likelihood at Convergence
Washington	1114	2924	-1054.0	727.4
New Bedford	453	1208	-436.4	-256.5
Los Angeles	879	2549	-930.0	-391.2

The coefficients of the original Washington model all have the correct signs, and, for the most part, are highly significant (i.e., having large t-statistics). The coefficients of the New Bedford model also have the correct signs.

The t-statistics, however, are not nearly as large as those of the original model, although for only three coefficients (in-vehicle travel time, shared ride automobile availability, and employment density times distance) are they seriously low. This relatively poor statistical performance may be related to the smaller sample size (453 versus 1114 observations) and the much lower variability observed for several of the level-of-service variables. (Data were available only for those trips having both origin and destination within the city of New Bedford itself.) The coefficients for the Los Angeles model also have the correct signs. In this case the overall statistical performance is much better than in the New Bedford case due in part to the larger sample size (879 observations).

Comparison of Coefficients

The three sets of coefficients are remarkably similar. The significance of the differences in coefficient values that do exist can be evaluated from two viewpoints: the practical policy analysis and the statistical. From the point of view of transportation policy evaluation the concern is with the consequences of the differences for transportation planning decisions, i.e., differences between coefficient values for level-of-service variables. As given in Table 2, with the exception of the travel cost coefficient for the New Bedford model, all level-of-service coefficients are sufficiently similar to warrant the conclusion that, even if the model as a whole may not be transferable, the level-of-service coefficients of the Washington model are.

The difference between two sets of coefficients can be tested by using the likelihood ratio test (12) where the null hypothesis is that the two sets of coefficients are equal. To perform this test it would be necessary to estimate the model with the two data sets pooled together in addition to the two separate estimations presented here. This was not done primarily because in an actual planning situation access to raw data cannot be assumed. Therefore, the original Washington coefficients were taken as constants (rather than random variables), and the likelihood ratio test was performed with the new data set only. The test statistic is given by

$$-2[L^*(\theta_{WASH}) - L^*(\theta_{NB})] \qquad (2)$$

where

$L^*(\theta_{NB})$ = the log likelihood of the New Bedford coefficients on the New Bedford data (= -256.5), and

$L^*(\theta_{WASH})$ = the log likelihood of the Washington coefficients on the New Bedford data (= -262.4).

From this, the value of the statistic is 11.8; it is chi-square distributed with 11 degrees of freedom. The probability of this statistic exceeding 11.8 is 38.3 percent. Therefore, the null hypothesis cannot be rejected, and the two sets of coefficients are not significantly different for the New Bedford data.

Rather than comparing sets of coefficients, the differences between individual coefficients can be evaluated by expressing the significance of the difference between the New Bedford or Los Angeles coefficients and the Washington coefficients as the t-statistic for the absolute difference. The test statistic used is the difference of the two coefficients divided by the square root of the sum of the variances of the two coefficients, and for large samples is normally distributed. Only for two of these coefficients (AALD. and AALD,) are the differences significant at the 90 percent level.

The facts that the original specification gave a reasonable model in other areas and that the sets of coefficients taken together and key level-of-service coefficients are not significantly different are encouraging. The differences between several of the coefficients indicate areas in which more research on improved specification could be fruitful, and show that the comparison of coefficients

Table 1. Work mode choice model: definition of variables.

Variable	Definition
D.	1, for drive alone 0, otherwise
D,	1, for shared ride 0, otherwise
OPTC/INC	Round trip out-of-pocket travel cost (¢) household annual income ($)
IVTT	Round trip in-vehicle travel time (min)
OVTT/DIST	Round trip out-of-vehicle travel time (min) one-way distance (miles)
AALD.	Number of automobiles licensed drivers, for drive alone 0, otherwise
AALD,	Number of automobiles licensed drivers, for shared ride 0, otherwise
BW.	1, if worker is head of household, for driver alone 0, otherwise
GW,	1, if worker is a civilian employee of the federal government, for shared ride 0, otherwise
DCITY.	1, if work place is in the CBD, for drive alone 0, otherwise
DCITY,	1, if work place is in the CBD, for shared ride 0, otherwise
DINC. ,	Household annual income—800 × number of persons in the household ($), for drive alone and shared ride 0, otherwise
NWORK,	Number of workers in the household, for shared ride 0, otherwise
DTECA,	Employment density at the work zone (employees per commercial acre) × one-way distance (miles), for shared ride 0, otherwise

Table 2. Transferability of work mode choice model to different cities.

	Washington		New Bedford		Los Angeles	
Variable	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
D.	-3.24	-6.86	-2.198	-2.648	-2.746	-4.85
D,	-2.24	-5.60	-1.535	-1.535	-1.830	-3.95
OPTC INC	-28.8	-2.26	-87.33	-1.576	-24.37	-2.07
IVTT	-0.015 4	-2.67	-0.019 9	-0.4849	-0.014 65	-2.25
OVTT DIST	-0.160	-4.08	-0.101 3	-2.903	-0.186 0	-4.02
AALD.	3.99	10.08	2.541	3.674	3.741	7.19
AALD,	1.62	5.31	0.449 9	0.847 8	0.609 3	1.58
BW.	0.890	4.79	1.026	3.769	0.810 1	3.28
GW,	0.287	1.78	—	—	—	—
DCITY.	-0.854	-2.75	—	—	—	—
DCITY,	-0.404	-1.36	—	—	—	—
DINC. ,	0.000 07	3.46	0.000 072	1.279	0.000 083	2.31
NWORK,	0.098 3	1.03	0.187 4	1.249	0.081 0	0.46
DTECA,	0.000 63	1.34	0.000 60	0.766 5	0.000 27	2.23

estimated for two different data sets is a powerful method of detecting specification errors. But, whatever improvements are implemented, no model will be perfectly specified and therefore perfectly transferable, hence the motivation for the application of updating procedures for the model coefficients.

PROCEDURES FOR UPDATING

This section develops several approaches for transferring a model from one area to another. Since the motive for transferring is to provide a reasonable travel demand model while meeting stringent resource constraints, the level of effort required for each of these approaches will be an important factor in evaluating their effectiveness. In terms of level of effort required, these approaches can be divided into two broad categories: those that require a disaggregate sample from the area in question and those that do not.

Transferring With No Disaggregate Sample

The simplest approach requiring the minimum level of effort is to use the existing model with its original coefficients. This assumes that all factors relevant to the choice process are embodied in the model, an assumption that will never be fully justified. For example, the specification of most models contains constant terms to account for factors not explicitly explained by the model. The presence of these constants indicates that in fact the model has not captured all aspects of the choice process and, because these other factors can vary between areas, the value of such a constant estimated in one area may or may not be appropriate for another. Therefore, although there is a theoretical basis for transferring the relationships estimated between time, cost, income, automobile availability, and such, there is no such basis for transferring these constant terms. Fortunately, in most applications data on existing conditions are available and the model will be used to predict changes in travel behavior that result from changes in the independent variables. For incremental predictions, therefore, the constant terms have no effect on the results. In some situations, however, data on existing conditions are not uniformly available at the required level of detail and the constants must be modified.

A suitable approach to this might be to use the existing model with adjustments of these constants. In this approach, the coefficients other than the constants are accepted and aggregate data on travel patterns in the new area are used to adjust the constants to better reflect the existing situation. The adjustment is performed by applying the model to the new area in the way in which it will be applied for forecasting. The results are aggregated to the level for which data (e.g., aggregate mode splits for work trips for the model tested in this paper) are available and the constants then adjusted until the model replicates the existing aggregate data. This improves the goodness of fit to the existing data, but the use of areawide averages for the independent variables could result in poorer estimates of the constants because of an aggregation bias. The primary criticism of this approach is that, in practice, other coefficients are not perfectly transferable and adjusting only the constant terms will compensate for these errors.

Transferring With a Disaggregate Sample

In this category, it is assumed that at least a small sample of observations on individual trip-making behavior representative of the study area will be available for use in updating the original model. The sample should be selected such that it could be used to reestimate the original model. (The effect of sample size on the performance of each approach is discussed later.) The most straightforward approach is to use the small disaggregate sample to reestimate all the coefficients of the original specification, reasoning that, since the model specification was successful in one area, it should work in another, and that, by using the coefficients estimated on data from the area where the model is to be used, none of the original coefficients need be accepted. However, because a model specification results in good statistical performance on one particular data set does not guarantee that estimating it in another area would result in reasonable coefficients. Even if the specification were correct, the use of a small sample for estimation is a potential source of problems. The maximum likelihood estimation technique used for these models gives coefficient estimates that have asymptotically optimal properties. For the small samples used in this approach, it is possible that the resulting biases and standard deviations will be large.

Another approach is to reestimate only the constant terms. In this approach, a single coefficient that modifies the scale of the other coefficients could also be estimated. This would retain the original trade-offs among the independent variables and should give a better goodness of fit on existing data. Forecasting accuracy for changes of individual variables should increase for those coefficients that benefit from the single scale coefficient and decrease for the others.

A better approach is to combine the original coefficients with those estimated on the small sample. Ideally, this should be done in such a way that all of the original coefficients are modified and at the same time any adverse effects resulting from the small sample available for the new area are minimized. Updating the original coefficients by using sample information should result in a model that better reflects travel behavior in the new area.

Bayesian Updating

The methodology used for combining sample information with prior information was that of Bayesian statistics (13), which relates the posterior distribution in an unknown parameter, θ, to the prior distribution in θ and the sample likelihood function by

$$\begin{pmatrix}\text{Posterior}\\ \text{probability}\\ \text{of } \theta \text{ given}\\ \text{the sample}\end{pmatrix} = C \times \begin{pmatrix}\text{likelihood}\\ \text{of the}\\ \text{sample}\\ \text{given } \theta\end{pmatrix} \times \begin{pmatrix}\text{prior probability}\\ \text{of } \theta\end{pmatrix} \qquad (3)$$

(The normalizing constant, C, is to ensure that the resulting posterior distribution is a proper set of probabilities.) The estimated coefficients of the original model are random variables that, for large samples, are normally distributed: This is the prior distribution. The data for the small sample for the new area are next used to reestimate the model to obtain a different distribution of the model coefficients: This is the sample distribution. These two distributions are then combined to obtain the posterior, or updated, distribution of the coefficients. This is shown in Figure 1 for the single coefficient case.

Since both the prior and the sample distributions are normal and the variance is assumed to be known, the mean and standard deviations for the posterior distribution in the single coefficient case are

$$\theta_2 = [(\theta_1/\sigma_1^2) + (\theta_s/\sigma_s^2)]/[(1/\sigma_1^2) + (1/\sigma_s^2)] \qquad (4)$$

and

$$\sigma_2 = [(1/\sigma_1^2) + (1/\sigma_s^2)]^{-1/2} \quad (5)$$

where

θ_1 = original coefficient,
θ_s = sample coefficient,
θ_2 = updated coefficient,
σ_1 = standard deviation of the original coefficient,
σ_s = standard deviation of the sample coefficient, and
σ_2 = standard deviation of the updated coefficient.

Thus, θ_2, the updated coefficient, is a weighted average of the original coefficient, θ_1, and the coefficient estimated from the new sample, θ_s, the weights being the inverse of their respective variances. The extension to the multivariate normal case is given by Raiffa and Schlaifer (13, p. 310). For the case shown in Figure 1, in which the prior information is reliable and a relatively small sample is used, the posterior distribution of θ will be based primarily on the prior information. For the case shown in Figure 2 in which the prior information is not very reliable and a relatively large sample is used, the posterior distribution of θ will be based primarily on the sample information. In both cases, the variance of the posterior distribution will always be less than that of both the prior distribution and the sample likelihood distribution.

This procedure also offers an opportunity to introduce subjective judgments into the model estimation process. Consider the case in which the data used to estimate the original model are thought to be inaccurate: The variance of the original coefficients can be increased to reduce the weight of these estimates in the updating process. Similarly, if the original estimation has been done a long time earlier, the relative weight placed on the prior distribution can be reduced.

It should be stressed that the key advantage of the Bayesian updating procedure is economic. By combining new sample information with prior information, it permits the use of small sample surveys that, by themselves, would not be statistically adequate for updating models. This procedure approximates the pooling of the sample used to estimate the existing model with the new sample but obviates the need to go back to the original sample data. It also has the advantage of being able to subjectively alter the weight of the prior sample with relative ease.

EMPIRICAL EVALUATION OF UPDATING PROCEDURES

This section describes an empirical evaluation of these approaches using New Bedford data. Two measures of effectiveness, goodness of fit on existing data and forecasting ability, are used to compare the models resulting from the various procedures of transferring. Each procedure is evaluated in terms of both; the results of this empirical evaluation indicate which method dominates at a given level of effort.

Goodness of Fit

One way of measuring how well a particular model fits the existing data is to use the log likelihood of the coefficients of that model for the New Bedford data to determine a value for a goodness of fit measure such as ρ^2, which equals the fraction of the log likelihood explained by the model, and is defined as

$$\rho^2 = 1 - [L^*(\hat{\theta})]/[L^*(0)] \quad (6)$$

where

$L^*(\theta)$ = the log likelihood of the sample for $\theta = \hat{\theta}$ and
$L^*(0)$ = the log likelihood of the sample for $\theta = 0$.

This approach has the advantage that it provides a single measure by which to rank the various models. However, ρ^2 is an abstract measure and it is difficult to grasp what differences in ρ^2 actually mean in terms of model performance.

Another approach is to compare the observed mode split of the data set with that predicted by the model by using values for the independent variables given in the data set. This is done by calculating individual probabilities of choosing available modes for each observation using the particular model being evaluated. These individual probabilities are then summed and compared with the observed mode split for the entire data set. These differences between observed and predicted mode splits provide a more specific measure of the goodness of fit of a model.

Forecasting Ability

To assess the forecasting ability of the models resulting from the transferring methods, the predicted changes in mode split resulting from policy changes are compared with the true changes predicted by the New Bedford model. The true changes are obviously unknown, but the model estimated with the entire New Bedford data set provides the best estimate available of the New Bedford conditions and therefore the best estimate of the true changes. The responses to policy changes are determined by recalculating the choice probabilities for each observation in the data set to account for the changed variables. These probabilities are then summed for each mode to find the forecasted mode shares. The policy selected to evaluate forecasting ability was that of assigning preferential lanes for multiple occupancy vehicles, resulting in a 15 percent decrease in shared ride and transit in-vehicle travel time.

EVALUATION OF APPROACHES FOR TRANSFERRING

The following models are used in this empirical evaluation:

1. True New Bedford model—the model estimated on the entire New Bedford data set (453 observations),
2. Washington model—the original model estimated on Washington data,
3. Washington model with updated (aggregate) constants—the original Washington model with the constant terms adjusted by use of aggregate mode split data,
4. New Bedford small sample models—models estimated on small random samples taken from the New Bedford data set [the size of the sample is indicated by the number of observations (44, 89, or 177)],
5. Washington model with updated (disaggregate) constants—models resulting from using disaggregate samples to reestimate the constant terms and a scale factor for all other coefficients, and
6. Model resulting from Bayesian updating—models resulting from Bayesian updating with the inverse of the variance-covariance matrix as the weighting factor.

Before discussing the evaluation results, a point should be made concerning the bias introduced into the

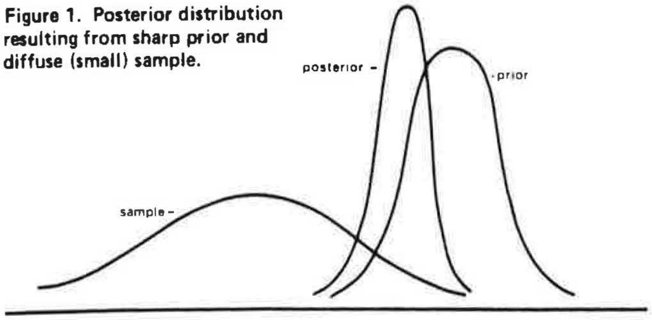


Figure 1. Posterior distribution resulting from sharp prior and diffuse (small) sample.

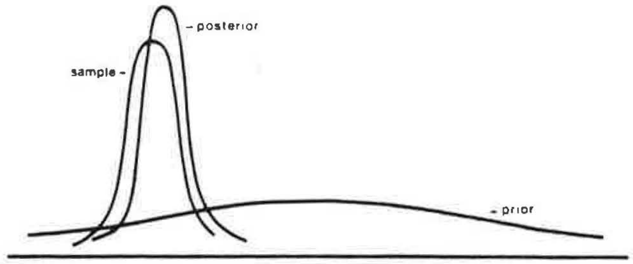


Figure 2. Posterior distribution resulting from diffuse prior and sharp (large) sample.

Table 3. Evaluation of approaches for transferring.

Updating Procedure	No. of Observations	ρ^2	Predicted - Observed Mode Shares[a] (%)			Predicted - True Changes in Mode Shares[b] (%)		
			Drive Alone	Shared Ride	Transit	Drive Alone	Shared Ride	Transit
True New Bedford model		0.412	0	0	0	0	0	0
Washington model		0.399	-0.39	-0.7	1.09	0.25	-0.24	0.01
Washington model with updated (aggregate) constants		0.398	-1.90	0.96	0.94	0.22	-0.21	-0.01
New Bedford small sample models	44	0.291	-4.91	3.55	1.36	-1.04	1.09	-0.05
	89	0.382	-1.69	0.31	1.37	-2.30	2.28	0.02
	177	0.397	-1.20	-0.58	1.78	0.82	-0.78	-0.04
Washington model with updated (disaggregate) constants	44	0.396	-3.51	4.22	-0.71	0.37	-0.34	-0.03
	89	0.401	-1.33	0.83	0.50	0.36	-0.34	+0.02
	177	0.398	-1.95	0.35	1.59	0.26	-0.26	0.0
Models resulting from Bayesian updating	44	0.399	-2.20	1.26	0.91	0.27	-0.26	-0.01
	89	0.400	-0.39	-1.16	1.56	0.24	-0.25	0.01
	177	0.400	-0.51	1.03	1.54	0.20	0.20	0.0

[a]Observed mode shares are drive alone = 55.12, shared ride = 37.76, transit = 7.12%.
[b]True mode share changes are drive alone = 1.01, shared ride = 0.96, transit = 0.05%.

evaluation measures. Because the observations for the small samples used in some of the transferring methods are taken from the data set that was used as the standard for New Bedford, the resulting measures of effectiveness for these approaches are biased toward indicating better performance, especially for the models estimated directly on these small samples. The magnitude of this bias varies with the sample size: For example, the bias for the model using 44 observations is relatively small since 409 observations (90 percent of the full data set) are different from those used in estimating the model. For the model using 177 observations, however, the bias may not be negligible since only 276 observations (61 percent of the full data set) are different.

The ρ^2 values for the different models are listed in Table 3. For sample sizes of less than 89 observations, performance is very poor; the small sample approach requires a sample size of at least 180 observations. In general, the Bayesian updating approach is best.

The comparison of predicted versus observed mode shares for the existing data is given in Table 3 for the entire data set. As observed with the ρ^2 values, below 89 observations the performance decreases for all approaches to transferring that require a disaggregate sample. The pattern of errors is not the same for all three modes and no general pattern of dominance emerges. Overall, the Washington model and the Bayesian updating models performed better than other approaches, particularly for the drive alone mode. The performance of the different approaches in predicting changes in mode shares due to a preferential lanes policy is also given in Table 3. The Bayesian updating approach in general performs better than the other approaches; the Washington model is superior to small sample models.

CONCLUSIONS

The most interesting result of this empirical study is the surprisingly good performance of the original Washington model on both the New Bedford and Los Angeles data sets. This is remarkable in view of the fact that the New Bedford and Los Angeles data sets represent very different conditions than those existing in the Washington data set. Although differences in individual coefficients between the models were observed, only three of these differences can be considered significant. Of all the approaches to updating the original model, Bayesian updating gives consistently better results. The small sample approach resulted in models that were inferior to the original Washington model for every measure of effectiveness and is clearly unreliable for the purposes of transferring.

Two approaches were taken to updating the constants: one using aggregate mode split data and the other using a small sample. For the first approach, because the original model had fit the data so well, the resulting model performed more poorly than the Washington model. Although slight improvements in some measures were observed for reestimating the constants, these improvements were insignificant when compared with those resulting from Bayesian updating. Therefore, the approach of reestimating the constants, like the small sample approach, is an inefficient use of the disaggregate sample for updating.

However, the superior performance of the Bayesian updating approach to that of the small sample models can be attributed to the good performance of the Washington model by itself for the New Bedford data. If the Washington model had had serious specification errors the small sample models would probably have performed much better relative to the Bayesian updating models.

Thus, it is clear that a credible specification is a precondition to any attempt to transfer a model.

In summary, three important conclusions are indicated from the empirical results:

1. A well-specified disaggregate mode choice model is transferable.
2. It is useful to update the model coefficients when transferring.
3. The Bayesian updating procedure using a small disaggregate sample is the most effective procedure for transferring well-specified models.

The empirical results reported in this paper are based on a model for the conditional probability of mode choice, which is only one component of the entire travel demand model system. These results are indicative but further work is needed in other aspects of travel demand for which model development effort has been significantly lower.

REFERENCES

1. M. E. Ben-Akiva and S. W. Haws. Estimation of a Work Mode-Split Model Which Includes the Car Pool Alternative. Department of Civil Engineering, M.I.T., Cambridge, 1974.
2. G. M. McCarthy. Multiple Regression Analysis of Household Trip Generation—A Critique. HRB, Highway Research Record 297, 1969, pp. 31-43.
3. R. de Neufville and J. S. Stafford. Systems Analysis for Engineering and Managers. McGraw-Hill, New York, 1971.
4. F. Koppelman. Travel Prediction with Disaggregate Choice Models. Department of Civil Engineering, M.I.T., Cambridge, PhD thesis, 1975.
5. Cambridge Systematics. A Behavioral Model of Automobile Ownership and Mode of Travel. Federal Highway Administration, U.S. Department of Transportation, 1975.
6. T. J. Atherton, J. H. Suhrbier, and W. A. Jessiman. Use of Disaggregate Travel Demand Models to Analyze Car Pooling Policy Incentives. TRB, Transportation Research Record 599, 1976, pp. 35-40.
7. Cambridge Systematics. A Study Design for Dual Mode Transit Planning Case Studies. Urban Mass Transportation Administration, U.S. Department of Transportation, 1976.
8. M. E. Ben-Akiva. Structure of Passenger Travel Demand Models. Department of Civil Engineering, M.I.T., Cambridge, PhD thesis, 1973.
9. T. Domencich and D. McFadden. Urban Travel Demand: A Behavioral Analysis. North-Holland, Amsterdam, 1975.
10. D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers in Econometrics (Paul Zarembka, ed.), Academic Press, New York, 1973.
11. M. G. Richards and M. E. Ben-Akiva. A Disaggregate Travel Demand Model. Saxon House, D.C. Heath, Westmead, England, 1975.
12. H. Theil. Principles of Econometrics. John Wiley, New York, 1971.
13. H. Raiffa and R. Schlaifer. Applied Statistical Decision Theory. Graduate School of Business Administration, Harvard Univ., Cambridge, 1961.

Guidelines for Aggregate Travel Prediction Using Disaggregate Choice Models

Frank S. Koppelman, The Transportation Center, Northwestern University

This paper describes procedures for aggregating disaggregate choice models after estimation of the choice model parameters to obtain an aggregated model structure. This structure consists of (a) a disaggregate choice model, (b) a representation of the distribution of explanatory variables, and (c) an aggregation procedure. A taxonomy of aggregation procedures that classifies models according to their structural characteristics is developed. Errors in prediction by use of alternative aggregation procedures are empirically estimated. Analysis of these errors leads to conclusions about the performance of different aggregation procedures. These conclusions suggest the following guidelines for aggregate travel prediction using disaggregate choice models: (a) Disaggregate choice models may be most effectively used for prediction at high levels of aggregation appropriate to policy analysis; (b) enumeration procedures should be used whenever adequate sample data are available, especially at high levels of aggregation; (c) when sample data are not available, classification procedures should be based on the most important class distinction, which will be differences in choice set when such differences exist; (d) when data are not available to predict class specific variable values, predictions by the naive procedure should be adjusted for differences in choice set when such differences exist; (e) the specification of the underlying disaggregate choice model should be developed and evaluated with particular care in the grouping of individuals with structurally different choice sets; and (f) incremental prediction should be used for prediction of the expected impacts of policy changes whenever an existing set of choice shares is available to use as a basis for adjustment.

A central component of transportation planning and policy analysis is the prediction of the future performances and impacts on the transportation system for each of the available plan or policy alternatives. These predictions should provide adequate information about the effects of each alternative so that an informed choice can be made among them and should distinguish among the effects of different policies with respect to travel flows, system performance, and external impacts. The evaluation and selection process requires that these predictions be at a level of aggregation that is relevant to the alternatives under study and the precision with which they have been formulated.

In contrast with the need for aggregate predictions based on aggregate descriptions of transportation plan and policy options, travel behavior theory is postulated at the level of the behavioral unit—usually an individual or household. Group behavior, which is the object of prediction, is the aggregation of numerous travel choices of individual behavioral units. This aggregation is implicit in the use of models that are estimated on mean value aggregate data. When disaggregate choice models are used the aggregation is accomplished by use of an explicit aggregation procedure. The aggregate prediction will be sensitive to both the structure of the disaggregate choice model and the distribution of independent variables in the population or prediction group (1).

There are two approaches to the development of aggregate prediction models that are consistent with underlying disaggregate behavior. The first approach is to aggregate the model to obtain a consistent structure that can be estimated with aggregate data (Figure 1, part A) and then to use this model to make aggregate predictions. When the underlying choice model is nonlinear and the aggregate groups are not homogeneous, a consistent aggregate function will include parameters of both the choice model and the distribution of independent variables. The requirements on the structure of the choice model and the distribution of variables necessary to develop a consistent estimable aggregate relationship are extremely restrictive. The only successful example of this approach (2) is limited to use with the binary probit choice function and requires the assumption of multivariate normally distributed variables. The estimation is based on information about the distribution of variables in aggregate groups as well as on their mean values. The second approach is to estimate a disaggregate choice model using disaggregate data and then aggregate the estimated choice model when it is used for prediction (Figure 1, part B). The advantage of this approach is that it makes no assumptions about the future distribution of independent variables prior to the model estimation: Assumptions about the distribution of independent variables can be deferred until the time of prediction and may be varied for each prediction situation.

It is conceptually simple to make aggregate predictions of travel behavior when the characteristics of individual trip makers and the alternatives available to them are known or can be predicted. In this case, predictions of the expected choice behavior can be made for each individual and summed or averaged to obtain the aggregate travel predictions (3). Generally, however, this disaggregate information is not available and aggregate predictions must be based on a less than completely detailed set of independent variables.

This paper reports the results of a study of the accuracy of predictions based on disaggregate choice models using less than completely detailed data. It describes the structure of aggregated prediction models, develops a taxonomy of aggregation procedures, describes the error in prediction by alternative aggregation procedures in different prediction situations, and develops tentative conclusions about the performances of the different aggregation methods in different prediction situations. It then proposes a set of guidelines based on these conclusions for prediction with disaggregate models.

STRUCTURE OF AGGREGATED PREDICTION MODELS

Any model that predicts the travel behavior of groups of behavioral units is an aggregate prediction model. Such models predict travel demand at some level of aggregation based on input variables that are also aggregate.

An aggregated prediction model explicitly incorporates disaggregate behavioral relationships in a structure that describes the causal relationship between socioeconomic and travel service characteristics on the one hand and aggregate travel behavior on the other. A model structure for aggregate prediction based on disaggregate travel choice relationships has three components. These are (a) a disaggregate choice model, (b) a representation of the distribution of explanatory variables, and (c) an aggregation procedure that operates on the two other components to obtain the required aggregate prediction. This structure of an aggregated prediction model makes explicit two important advantages of aggregated prediction models over aggregate models based on correlative analysis of aggregate data. These are sensitivity to changes in individual behavior due to changes in travel service or other environmental attributes, including policy control variables, and sensitivity to changes in the distribution of the characteristics of the population that makes up the prediction group.

The disaggregate choice model relates the probability of choosing one alternative out of a set of available alternatives to the relative use of each alternative to the individual decision maker. The utility of an alternative is defined as a function of the characteristics of the individual and the attributes of the alternatives available to him. The choice model may have a wide range of functional forms that are derived from the underlying assumptions about the choice process of the individual (4, 5).

The distribution of independent variables describes the presence in the prediction group of individuals with different socioeconomic characteristics or facing different transportation service attributes. That is, the distribution represents the frequency of occurrence in the prediction group of different values of the socioeconomic and travel service variables that influence individual travel choice decisions. These distributions may be represented in a variety of ways and with different degrees of detail. The various methods for representing the distribution of variables are given below.

1. Enumeration represents the distribution of variables by actual or estimated values of the variables for individuals. Complete enumeration provides variable values for every member of the aggregate prediction group. Partial enumeration provides variable values for a subset of the prediction group.
2. Density functions represent the distribution of variables by the frequency of different variable values in the prediction group. These distributions are based on theoretical or empirical analyses or both, which describe the structure of the distributions and their parameters.
3. Distribution moments represent the distribution of variables in terms of moments and cross moments, which provide information about the spread and shape of individual distributions and their interactions.
4. Classification represents the distribution of variable values in terms of the proportion that is assigned to each of several relatively homogeneous subgroups.

The different methods of representing the distribution of independent variables provide similar information about the actual distribution in different forms. The representations may, to some degree, be transformed to alternative representations. Some transformations imply a loss of information while others imply an increase in information by the use of externally available information or simplifying assumptions. Enumeration, if based on a large enough sample, can be used to estimate parameters of a density function or distribution moment and can be used as a basis for classification. Density functions can be used to determine distribution moments or as a basis for classification. Distribution moments may be used to identify density functions directly or by augmenting them with an assumed distributional form. Classifications, unless they are extremely fine, cannot be used to generate any of the other distributional representations. The process of transformation may be used to modify an initial representation to an alternative representation that is required as input to a selected aggregation procedure.

The aggregation procedure operates on the disaggregate choice model and the distribution of independent variables to produce aggregate predictions. The theoretically consistent aggregation procedure is to estimate the choice probabilities for each individual and then average these choice probabilities to obtain the expected share choosing each alternative. The extreme data requirement of this procedure, prediction of all variable values for every member of the prediction group, motivates a search for aggregation methods that have less extensive input data requirements. A variety of alternative aggregation procedures have been proposed for this purpose. These procedures can be grouped in five categories:

1. Procedures of enumeration are procedures that represent the explicit theoretical relationship between aggregate and disaggregate demand. The expected share choosing an alternative is the average of the individual choice probabilities for that alternative. Complete enumeration averages the choice probabilities for all individuals in the prediction group; sample enumeration averages the choice probabilities for a sample or subset of individuals in the prediction group.
2. Procedures of summation or integration weight conditional disaggregate choice probability estimates by the probability density function for the independent variables. This is done by integration or summation over the multivariate distribution of explanatory variables in the prediction group. The computational requirements of procedures in this group are high if the integration or summation must be applied over a large number of variables. This group of procedures may use distributions (density functions) determined from theoretical and empirical analyses, or assumed distributions that offer computational or other advantages. The most advantageous distributional assumption is that the variables are multivariate normally distributed (2, 6).
3. Procedures of statistical differentials express aggregate shares as a function of the moments of the utility distribution. The aggregate function is obtained by lin-

Figure 1. Alternative procedures to obtain aggregate predictions based on disaggregate choice models.

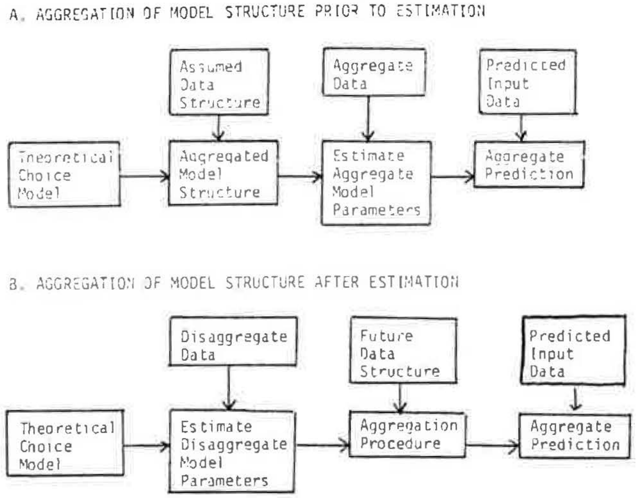
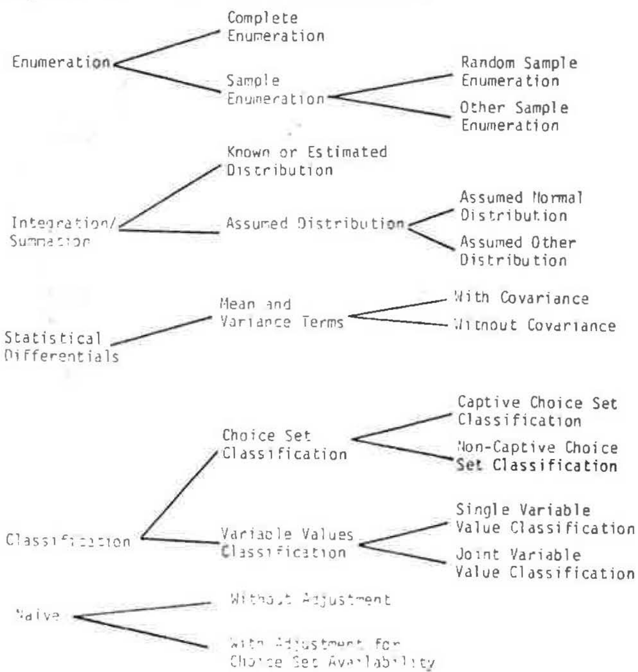


Figure 2. Taxonomy of aggregation procedures.



earizing the disaggregate choice function by the use of a Taylor series expansion and then taking expectation across the aggregate prediction group (7). If the utility function of the alternatives is a linear function of the independent variables, the share estimate is given by a series of terms that includes the nth order derivative multiplied by the nth distributional moments and divided by n factorial. The practical issues associated with estimating higher order moments and the instability of the series when the distribution is highly dispersed (8) suggest that the series be terminated after the second, or variance, term.

4. Procedures of classification assign members of the aggregate group to identifiable classes, use the average variable values for each class to predict aggregate choice shares for each class, and compute overall aggregate share as the weighted average of the class shares. Procedures in this group are differentiated by the basis of classification and the number of classes used. Alternative possibilities include classification by differences in choice set and classification by differences in variable values.

5. Naive procedures use the mean value of the choice influencing variables in the disaggregate choice function. These procedures implicitly assume that each individual acts as if he is described by the average values of the prediction group. The naive procedure is a special case of summation or integration procedures (the distribution is assumed to be concentrated at a point), statistical differentials procedures (truncating the series after the first term), or classification procedures (using one class only). It is useful to treat this procedure separately because (a) the data requirements are the same as those for models calibrated with aggregate data, (b) it is computationally and conceptually simple, and (c) it is the method most likely to be used in the absence of recognition of the aggregation problem. Predictions by naive procedures can be adjusted to account for differences in choice set availability when such differences exist.

The different types of aggregation procedures may be described in terms of the taxonomy represented in Figure 2. The major classes of aggregation procedures are divided into differentiable subgroups of procedures. These subgroups are defined by their important characteristics.

INFLUENCE OF PREDICTION ENVIRONMENT ON CHOICE OF AGGREGATION PROCEDURE

Choice of an aggregation procedure for use in a specific prediction environment depends on the expected magnitude of the aggregation error of alternative procedures and the contribution of the aggregation error to the overall error in prediction from all sources. The aggregation error of alternative procedures depends on the prediction situation, particularly the distribution of explanatory variables in the prediction group, the level of aggregation, the differences in choice set availability among group members, and the choice probabilities for the average member of the prediction group (9). The contribution of the aggregation error to the total error in prediction from all sources depends on the magnitude of the aggregation error relative to the magnitude of error from other sources. The prediction error from other sources is determined by decisions made in the model formulation and prediction process (3). These decisions include the specification of the disaggregate choice model including the functional form and variables to be included, the selection of data to be used in model estimation, the method selected to represent and predict the distribution of explanatory variables, and the selection of the aggregation procedures to be used. The characteristics of the prediction situation and the decisions made in the model development and prediction process influence both the aggregation error and the errors from other sources.

ANALYSIS OF ERRORS IN PREDICTION WITH DISAGGREGATE CHOICE MODELS

Errors in prediction with disaggregate choice models arise from each of the components of the aggregated model structure. For the purpose of the following discussion, these errors in prediction are separated into two categories:

1. Model and variable error, which includes errors in the specification of the choice model, errors in parameter estimation, and errors in the input variables; and
2. Aggregation error, which includes errors due

to the use of approximate aggregation procedures.

Empirical analysis of these errors was based on prediction of mode shares to work in the Washington metropolitan area for the drive alone, shared ride, and transit modes. The predictions were made for aggregate groups identified as 45 districts with an average sample size of 47, 10 super-districts with an average sample size of 213, and four rings with an average sample size of 533. The disaggregate mode choice model predicts the probability of drive alone, shared ride, and transit choices for the first work trip of breadwinners working in the CBD. The overall mode shares in the prediction sample were 38, 30, and 32 percent for the drive alone, shared ride, and transit alternatives respectively. The model specification and parameter estimates based on 824 work trips, of which 621 had all choices available and 253 did not have the drive alone alternative available due to lack of a car or driver's license, are given below.

Variable	Symbol	Estimated Coefficient	Standard Error
Drive alone dummy	D_d	-2.62	0.36
Shared ride dummy	D_s	-2.36	0.27
Automobiles per licensed driver (drive alone)	$AALD_d$	3.64	0.38
Automobiles per licensed driver (shared ride)	$AALD_s$	1.51	0.24
Out-of-vehicle cost per income	OPTC/INC	-0.028	0.012
Total travel time	TTT	-0.024	0.005
Out-of-vehicle time per distance	OVTT/DIST	-0.077	0.055
Government worker (shared ride)	GW_s	0.77	0.16
Number of workers in household (shared ride)	$NWORK_s$	0.24	0.10

The aggregation procedures used include the naive procedure, with adjustment for choice set availability, the statistical differentials procedure, with mean and variance terms, and classification by differences in choice set availability and automobile ownership.

Model and variable error in prediction was estimated by comparison of share predictions by the enumeration procedure against observed choice shares in the data set. Aggregation error was determined by comparison of share predictions by a selected aggregation procedure against share predictions by the enumeration procedure. Both types of error are reported as a percentage of the magnitude of prediction. Prediction errors were obtained for each prediction group (districts, super-districts, or rings) and each choice share. These errors are summarized in terms of average error, standard deviation of the error, and root mean square error (3). The combined error (model and variable error and aggregation error) is the square root of the squared model and variable error and the squared aggregation error. (This formulation implies independence between the errors from these two sources. This is a reasonable assumption in the absence of structural interdependence.) It is most sensitive to changes in the magnitude of error from the source that contributes the larger error.

COMPARISON OF AGGREGATION PROCEDURES

The naive procedure produced aggregation errors of approximately 10 percent of predicted values. Adjustment of the naive procedure by choice set availability reduced aggregation error to about 8.5 percent. Classification by choice set and automobile availability reduced the aggregation error to about 3 percent. The statistical differentials procedure resulted in higher aggregation error than the naive procedure. [Under certain conditions the statistical differentials produced can be expected to increase rather than decrease the aggregation error (9).] The errors are as follows:

Prediction Group	Error of Different Aggregation Procedures (%)			
	Naive	Naive With Adjustment	Statistical Differentials	Classification by Choice Set
Districts (45)	10.5	8.1	12.9	3.3
Super-districts (10)	9.8	8.5	11.8	2.9
Rings (4)	9.6	8.4	12.2	2.8

These results suggest the following conclusions:

1. When differences in choice set availability exist, these differences should be used as a basis for adjusting predictions by the naive procedure or as a basis for classification.
2. The statistical differentials procedure should be used only after verifying that it actually reduces aggregation error in prediction.

LEVEL OF AGGREGATION

Increasing levels of aggregation are expected, a priori, to have two effects on prediction error. First, model and variable error is expected to decline. This reduction of error results from averaging the expected choice probability over the larger sample of observations in the more aggregate prediction group. If the sample predictions are independent, the expected error in the estimate of choice shares would be inversely proportional to the square root of the number of observations in the prediction group. Predictions made with a common disaggregate choice model are not independent (3). Therefore, the effect of increasing the prediction group size should be less than proportional to the prediction group size. Second, aggregation error is expected to be larger for prediction groups with greater geographical dispersion as these groups are also expected to have greater dispersion of explanatory variable values.

The first expectation is supported by the results obtained. The model and variable error for each mode, and for all modes combined, declined with increasing level of aggregation.

Prediction Group	Model and Variable Error (%)			
	Drive Alone	Shared Ride	Transit Ride	All Modes
All districts (45)	19.8	35.4	28.7	27.8
Super-districts (10)	12.3	24.9	21.6	19.7
Rings (4)	5.4	18.0	20.1	15.4

This decline in errors was less than proportional to the square root of prediction group size. On the other hand, the second expectation is not supported by the results, as the aggregation error for the three different aggregation procedures was not strongly or consistently related to the level of aggregation. This is a consequence of the fact that increasing levels of aggregation do not lead to significant increases in intragroup variation in socioeconomic characteristics (10) or in intragroup differences in mode service characteristics. The net effect of these results, however, is that the combined error in prediction declines with increasing levels of aggregation as given below. As the size of the prediction group increases, the portion of the combined error attributable to aggregation error increases and the differences in prediction errors between aggregation procedures are amplified.

Prediction Group	Combined Error of Different Aggregation Procedures (%) Naive	Naive With Adjustment	Statistical Differentials	Classification	Enumeration
Districts (45)	29.7	29.0	30.7	28.0	27.8
Super-districts (10)	22.0	21.4	23.0	19.9	19.7
Rings (4)	18.1	17.5	19.7	15.7	15.4

These results lead to the following conclusions:

3. Predictions made with disaggregate choice models based on a fixed data set increase in accuracy as the level of aggregation increases.
4. The relative improvement in prediction error from the use of more precise aggregation procedures increases as the level of aggregation increases.

DIFFERENCES IN CLASSIFICATION STRUCTURE

The classification procedure used classifies the population according to both choice set availability and intra-household automobile availability. This classification was compared to classification solely by choice set availability and classification solely by intrahousehold automobile availability to identify the influence of different classification schemes on aggregation error. The aggregation errors for all three classification procedures at three levels of aggregation are given below.

Prediction Group	Aggregation Error of Different Classification Procedures (%) Choice Set and Automobile Availability	Automobile Availability Alone	Choice Set Alone
Districts	3.3	9.9	5.2
Super-districts	2.9	8.4	5.2
Rings	2.8	7.8	5.3

The most complete classification, that by choice set and automobile availability, has the least aggregation error. Classification by choice set alone has substantially lower aggregation error than classification by automobile availability alone. These results lead to two further conclusions:

5. Increasing refinement in classification leads to reduced aggregation error.
6. Classification by differences in choice set availability, when appropriate, results in lower levels of aggregation error than classification by variable values.

CHOICE MODEL SPECIFICATION

The large magnitude of model and variable error relative to aggregation error indicates the need to improve the choice model specification. The model and variable errors reported for the different levels of aggregation are substantially higher for the prediction of shared ride and transit shares than for the prediction of drive alone mode shares. The sources of these errors can be investigated by disaggregating them into average error and standard deviation of error for each mode as given below. The large errors in the prediction of shared ride and transit ride shares compared to drive alone shares are due to both higher average errors and greater standard deviation of errors. The opposite signs of the average error for shared ride and transit and the higher magnitude of average error and variability indicate that the model is deficient in predicting shares between these two alternatives. This deficiency may be due to inherent difficulties in specifying the utility function for these alternatives (due to large variations in excluded characteristics such as comfort and convenience or greater heterogeneity of preferences for different types of group ride alternatives). It may also be due to too great reliance on the independence of irrelevant alternatives axiom to estimate model parameters for individuals with different choice sets. The errors are disaggregated as follows:

Error Type	Model and Variable Error Disaggregated by Error Type Drive Alone	Shared Ride	Transit Ride	All Modes
Average error	+1.6	-7.1	+4.1	4.5
Standard deviation of error	19.7	35.3	28.4	27.3
Root mean square error	19.8	35.4	28.7	27.8

These results suggest two additional conclusions:

7. Analysis of error in prediction provides a basis for reevaluation and modification of the disaggregate choice model.
8. Individuals with structurally different choice sets should not be combined for estimation without testing the effect of this grouping on estimation and prediction errors.

PREDICTING CHANGED CHOICE SHARES

The discussion to this point has been concerned with the prediction of existing choice shares. We now turn our attention to prediction of choice shares after a change in travel service characteristics. The changes considered are (a) provide shared ride incentives to all trip makers (change 1), (b) reduce transit fares to zero (change 2), and (c) reduce transit times by one-half (change 3). The expected effect of these changes on mode share for the entire data sample is given below.

Prediction Situation	Mode Choice Shares Drive Alone	Shared Ride	Transit Ride
Change 1	0.35	0.36	0.30
Change 2	0.37	0.26	0.37
Change 3	0.35	0.23	0.42
Base	0.39	0.28	0.33

The new mode shares may be predicted directly by modifying the variables to reflect the policy changes. Alternatively, the new mode shares may be obtained by predicting the incremental change in shares resulting from a change in policy and using the predicted change to modify the observed choice shares. The aggregation error by the incremental prediction procedure is substantially lower than that by the direct procedure for all aggregation methods and for all of the policy changes. This result suggests the following conclusion:

9. Incremental prediction should be used to predict aggregate choice shares after policy change whenever predictions can be made for an observed set of choice shares to provide a base for adjustment.

SUMMARY

The preceding discussion describes a framework for the use of disaggregate choice models for the prediction of aggregate choice shares and proposes a taxonomy of procedures for making aggregate predictions based on disaggregate choice models. The aggregation error and the effect on the combined error of different aggregation procedures are empirically estimated for a variety of prediction situations. The results are summarized in a set of conclusions about the expected performances of selected aggregation procedures in specific situations. These conclusions suggest the following guidelines for aggregate prediction using disaggregate choice models.

1. Disaggregate choice models may be most effectively used for prediction at high levels of aggregation appropriate to policy analysis.
2. Enumeration procedures should be used whenever adequate sample data are available, especially at high levels of aggregation.
3. When sample data are not available, classification procedures used should be based on the most important class distribution (which will be differences in choice set when such differences exist).
4. When data are not available to predict class specific variable values, predictions by the naive procedure should be adjusted for differences in choice set when such differences exist.
5. The specification of the underlying disaggregate choice model should be developed and evaluated with particular care in the grouping of individuals having structurally different choice sets.
6. Incremental prediction should be used for prediction of the expected impacts of policy changes whenever an existing set of choice shares is available to use as a basis for adjustment.

ACKNOWLEDGMENTS

This research was supported by the Office of the Secretary, U.S. Department of Transportation. The empirical data used were provided by the Metropolitan Washington Council of Governments.

I wish to acknowledge the helpful comments and suggestions of Moshe Ben-Akiva, Marvin Manheim, Paul Roberts, and Steven Lerman of Massachusetts Institute of Technology, Daniel Brand of Harvard University, and Gerald Kraft of Charles River Associates.

REFERENCES

1. H. A. J. Green. Aggregation in Economic Analysis. Princeton Univ. Press, Princeton, N.J., 1964.
2. D. McFadden and F. Reid. Aggregate Travel Demand Forecasting From Disaggregated Behavioral Models. TRB, Transportation Research Record 534, 1975, pp. 24-37.
3. F. S. Koppelman. Prediction With Disaggregate Models: The Aggregation Issue. TRB, Transportation Research Record 527, 1975, pp. 73-80.
4. Charles River Associates. A Disaggregated Behavioral Model of Urban Travel Demand. Federal Highway Administration, U.S. Department of Transportation, 1972.
5. M. Ben-Akiva. Structure of Passenger Travel Demand Models. Department of Civil Engineering, M.I.T., Cambridge, PhD thesis, June 1973.
6. R. B. Westin. Prediction From Binary Choice Models. Journal of Econometrics, April 1974.
7. A. Talvitie. Aggregate Travel Demand Analysis With Disaggregate Travel Demand Models. Proc., Transportation Research Forum, 1973.
8. N. L. Johnson and S. Kotz. Distributions in Statistics, Vol. 1, Discrete Distributions. John Wiley, New York, 1969.
9. F. S. Koppelman. Travel Prediction With Models of Individual Choice Behavior. M.I.T. Center for Transportation Studies, Rept. No. 75-7, June 1975.
10. C. R. Fleet and S. R. Robertson. Trip Generation in the Transportation Planning Process. HRB, Highway Research Record 240, 1968, pp. 11-31.

A Disaggregate Modal-Split Model for Work Trips Involving Three Mode Choices

Jeffrey Ganek,* American Telephone and Telegraph Company, New York
Raymond Saulino,* Office of Budget and Management Systems, District of Columbia

This paper describes a disaggregate mode choice model with three travel modes: drive alone, car pool, and transit. A number of alternative model specifications were tested and the results analyzed. In general, the coefficients of in-vehicle time, out-of-vehicle time, and costs agree closely with the results of similar studies. Estimated coefficients of variables not included in previous logit model studies are also presented: Of these convenience, comfort, and flexibility influenced mode choice but mode unreliability and household income did not. Work location, cars per driver, and sex were the only socioeconomic variables for which statistically significant coefficients were found. Coefficients and models were also estimated for various subpopulations of commuters. The determinants of mode choice for CBD workers were different from those of non-CBD workers. Differences in the cost and time coefficients among travel corridors and income classes were also examined. The estimated models were validated by successfully predicting the mode choices of the commuters for whom the model was estimated, of other commuters, and, finally, of commuters for whom changes in the levels of service were made available.

This paper discusses a study that explored the determinants of commuter mode choice in the six-county region around Pittsburgh. The study was part of an evaluation of the car pool-public transit program administered by the Southwestern Pennsylvania Regional Planning Commission (SPRPC). Models of the multinomial logit form were estimated on disaggregate data to predict the short-run mode choices of commuters in the region, given the current work locations, residential locations, and automobile ownership patterns. The models predicted these choices from a set of alternative modes that include driving alone, car pooling, and riding on public transit as functions of the socioeconomic status (SES) of the commuters and the service attributes of the three modes. Variances in the influences of particular SES characteristics and mode attributes were determined by testing alternative specifications on subpopulations defined by work location and SES characteristics of the commuters. The objectives of this work were to learn why commuters choose the modes they do and to suggest how they could be enticed to choose shared ride modes.

The region around Pittsburgh has several concentrations of commercial and industrial employment. The largest of these is the CBD, which has five major transportation corridors leading to it from suburban areas. The quality of transportation services varies among these corridors. The southwest and east corridors have limited access highways that run to the CBD. The southern corridor is plagued by bottleneck problems as automobiles attempt to pass through Mt. Washington and over the Monongehela River. The northern corridors have a network of well-traveled streets where traffic flow is regulated by stop lights. Bus service to the CBD is available along all corridors, and streetcar lines, some of which have a right-of-way, run from the south. Commuter rail service is limited, and transit service to non-CBD locations is not extensive. The region served by these transportation systems has a population of over 2 million, with an average density of 300 people/km^2. The SES characteristics of the population are not different from the average characteristics of populations of other large standard metropolitan statistical areas. The region under study, then, has a diverse and well-established transportation system serving a dense population.

METHODOLOGY

The methodology used to estimate mode choice models was the multinomial logit form (1). This methodology was chosen because of its use for policy analysis due to its base in behavioral theory, its capacity to consider choice sets greater than two, and its successful application in other studies (2, 3).

DATA

The data used in this study were obtained from commuter surveys conducted as part of the study and from the Southwestern Pennsylvania Regional Planning Commission (SPRPC) travel time and cost networks. The complete data base included the SES characteristics, commuting preferences, and transportation services available to 740 commuters in the six-county region. This sample

*Both authors were at Carnegie-Mellon University when this research was performed.

population was generally representative of the working population of the region in terms of SES characteristics and available transportation services although there was a slight overrepresentation of commuters who work in the CBD. Additional data were obtained from a survey of commuters matched with potential car poolers by SPRPC.

The data used to estimate the model are discussed below.

In-vehicle time (IVT) is the time in minutes that commuters spend in their primary vehicle on a one-way work trip. These were obtained from the SPRPC network data.

Access time (ATIME) is the time in minutes that commuters spend between leaving home and reaching the primary vehicle for their work trip. Commuter estimates of the ATIME of their chosen modes were obtained from the surveys. Estimates of the ATIMEs of nonchosen modes were also derived from these survey results. ATIMEs for the transit alternative for nontransit riders were estimated to be 6.98 min, the mean ATIME report by transit riders. ATIME for the drive alone alternative was assumed to be 0 for all commuters. All car poolers were assumed to have a 4-min ATIME for the car-pool alternative, the mean time reported by car poolers. It was assumed that of all commuters those with lower car-pool ATIMEs would tend to car pool more often than those with higher; therefore, noncar poolers were assigned 6-min car-pool ATIMEs.

Egress time (ETIME) is the time in minutes that commuters spend between leaving the primary vehicle and arriving at the workplace. The ETIMEs that commuters reported for their chosen modes were used in estimating the model. ETIMEs for nonchosen modes were derived from the same data. The ETIMEs were generally homogeneous within commuter subpopulations defined by chosen mode and location of workplace (i.e., CBD or non-CBD). Therefore, the mean ETIME for each mode ridership to CBD workplaces and the mean time for each mode ridership to non-CBD workplaces were used as estimates of the ETIME associated with nonchosen modes.

Travel cost data were obtained from SPRPC network data. Transit travel costs were taken directly from the SPRPC data. These data do not consider the possible purchases of monthly and yearly discount fare passes. Automobile travel costs were the zone-to-zone costs of driving alone, calculated by SPRPC, and include gasoline, oil, tires, and maintenance costs. Parking costs, which were not included in the SPRPC calculations, were obtained from survey results. The reported parking costs of commuters who drive alone to work were added to the SPRPC cost estimates to obtain daily commuting costs. Parking costs for other commuters were assigned according to the traffic zone in which the commuter worked from the average of the daily parking costs of commuters who drive alone to that zone. These costs were added to the SPRPC cost estimates to establish total drive alone costs for commuters who do not drive alone.

The same data were used to estimate car-pool costs. All car pools were assumed to share costs equally among all members. Car-pool costs, then, were obtained by dividing the drive alone costs for the same zone-to-zone trip by the mean size of car pools as reported by car-pooler respondents.

Commuter perceptions of the relative levels of comfort and convenience offered by various modes were obtained from their responses to questions that asked them to compare their nonchosen modes with their chosen mode. Their responses were coded on an integer scale of -2 to +2, indicating the service level of the nonchosen mode as compared to that of the chosen mode. A negative number indicated that the nonchosen mode had a lower level of service with respect to either comfort or convenience than the chosen mode with 0 taken as the service level of the chosen mode. Although the data are subject to uncertainties with respect to scaling and definition (4), they are the best available approximation of commuter perceptions, and were used as such.

Flexibility is a measure of the possible irregularity in commuter schedules. It was determined from the survey by asking commuters how often they arrive at or leave work early or late. Those who arrive at or leave work early or late between 5 and 15 times a month are assumed to have flexible schedule patterns. Others are assumed to have regular schedule patterns. All commuters who commute flexibly were assigned a value 1 for flexibility; others were assigned a value of 0. This measure of flexibility reflects the preference of a commuter for flexibility, although data from which the measure is constructed may also represent an interaction between the preferences of the commuter and the actual flexibility of the service offered by his chosen mode.

Mode unreliability is a measure of how often each mode is not available for the trip to work. Data describing this service attribute were gathered in the mode choice survey by asking transit riders how many times per month their bus or trolley was more than 0.5 h late, car poolers how many times per month their car pool was unavailable for the trip to work, and those who drive alone how often their car was not available for the trip.

The standard SES attributes (including sex, income, and automobile ownership) of the sample population were obtained from the mode choice survey. The mode constants account for the SES characteristics, commuter preferences, and service attributes that influence the choice decision but are not explicitly contained in the model. The mode choice process is highly complex, with many factors influencing the decision. If a model were estimated without mode constants, then the factors left unaccounted for could affect the values of the coefficients associated with the variables that are explicitly included in the model.

ESTIMATION RESULTS

A random subsample of 400 observations from the sample population was used to estimate the model parameters for a number of model specifications of which the most interesting are models 1 and 2 in Table 1. Table 1 also includes models 3 and 4, which represent separately estimated specifications for CBD and non-CBD workers respectively.

Some observations about the success of the modeling efforts can be made from Table 1.

1. The signs on all the significant coefficients are as expected.
2. The relative magnitudes of the coefficients are reasonable.
3. In general, the results agree with similar studies made elsewhere.
4. The variances in the magnitudes of the coefficients estimated on different subgroups of commuters were as expected.

The coefficients, then, appear to be stable estimates of the true parameters.

The coefficients for the IVT are consistently negative and statistically significant, reflecting the dislike of in-vehicle time. Further analysis of these coefficients suggests that different groups of commuters have different sensitivities to IVT. For example, analysis of a model that included two IVT variables, one for all commuters and one for commuters with annual household incomes greater than \$15 000, showed that all commuters have an IVT coefficient of -0.023 and that wealthier commuters have an additional IVT coefficient of -0.015. While the additional factor of -0.015 is not statistically significant, it does suggest that income level may have some influence on how commuters value IVT.

The hypothesis that the value placed on IVT depends on the quality of service available to the commuter was tested by the estimation of five corridor-specific IVT coefficients; these coefficients were not statistically different from one another.

Separate coefficients, which were different from each other, were estimated for ATIME and ETIME. This difference may explain the disparity between these estimates and the coefficients for out-of-vehicle time (OVT) estimated elsewhere, i.e., the higher ATIME here may reflect a division of the independent influences of ATIME and ETIME on mode choice that are aggregated in single OVT coefficients. These variables, however, do not discern whether commuters disvalue ATIME or ETIME differently for different modes but the data were insufficient to estimate mode-specific ATIME coefficients. However, the coefficient here has the appropriate sign and agrees with other results (2, 3, 5) that changes in OVT have a greater effect on mode choice than similar changes in IVT.

The ETIME coefficient for commuters who work in non-CBD locations is close in magnitude to the OVT coefficients reported (2). It is not significant for CBD workers for whom the ETIME differences for different modes were too small to significantly affect mode choice decisions.

Table 1 also shows that cost has a negative coefficient and that, although it is relatively small, its difference from zero is statistically significant. The magnitude of the coefficient is not incongruous with other cost coefficients. The cost elasticities derived from the coefficient shown below are similar to those reported elsewhere (5).

Percent Change in Ridership	Transit	Car Pool	Drive Alone
From a 1 percent change in transit costs	-0.17	+0.09	+0.09
From a 1 percent change in car-pool costs	+0.06	-0.14	+0.06
From a 1 percent change in drive alone costs	+0.18	+0.18	-0.48

Therefore, the relatively small magnitude of the coefficient may be accepted as a reasonable estimate.

Since it was not possible to estimate mode-specific differences in the cost coefficients, a generic cost variable was used in the model. Attempts to identify differences in the sensitivity to cost of commuters in different income classes gave a coefficient that was not statistically significant. This may be due to the small number of low-income commuters in the sample.

The comfort associated with a mode is important to commuters with household incomes ≥\$15 000/year, but is not important to those with smaller incomes. This is indicated by the statistically significant coefficient estimated for comfort (≥\$15 000 only). The coefficient for comfort (<\$15 000) was not statistically different from zero.

Commuters also highly value the mode characteristic identified as convenience. The statistically significant coefficient for convenience is very important in mode choice decisions, but, because of similarities in meaning, comfort and convenience were not included in the same equation. Qualitative factors like comfort and convenience can be included in choice models along with more quantified variables such as time and cost.

The preference for flexibility has been included in the model as a mode-specific variable because car pooling is regarded by the commuters in the sample as very inflexible relative to the other modes (6). Therefore, commuters are assumed to take special consideration of their preference for flexibility only when they assess their car-pool alternative. The statistically significant and negative value of the estimated coefficient confirms this assumption.

Car availability is included in the model twice. The two estimated coefficients represent the relation between the number of cars per driver in the household and the value that a commuter gives to the drive alone and car-pool modes relative to transit. These coefficients show that a strong factor in the mode choice is the number of cars per driver in the household. The more cars per driver, the higher the tendency to commute via automobile. High car availability especially favors the drive alone choice. The influence of car availability on mode choice is a short-run effect. In the long run, car availability becomes a function of mode choice, and, in the long run, the tendencies of commuters to change modes as a result of changes in the levels of service offered may be greater than those estimated here. In the short run, few drivers with high car availability can be enticed to change their mode of commuting.

Work location was a mode-specific variable for both the drive alone and car-pool modes. The resulting negative and highly significant coefficients indicate that for CBD workers the problems of driving and the availability of transit service to the CBD strongly discourage use of the automobile modes and encourage use of transit. The magnitudes of these coefficients suggest that commuting to the CBD is very different from commuting to non-CBD locations. Traffic congestion, the difficulty and cost of parking, higher worker density, and transit availability make the determinants of mode choice in the CBD different from those elsewhere. Therefore, two sets of coefficients, one for CBD workers (model 3) and one for all other commuters (model 4), were estimated. The results show the following influences of work location:

1. CBD workers find IVT more onerous than do non-CBD workers. This may result from the greater traffic congestion in the CBD.
2. Non-CBD workers are much more sensitive to ATIME than are CBD workers, and policies that affect ATIME should have greater impact on non-CBD commuters. The relative magnitudes of these coefficients, however, may be due to the method used to assign ATIMEs. However, non-CBD commuters, who do not have to consider the inconveniences of congestion and parking, undoubtedly discriminate between modes on other dimensions, and it is not surprising that ATIME is more important to them.
3. All commuters are relatively insensitive to costs.
4. CBD workers are very sensitive to the convenience of commuting, but the coefficient of convenience for non-CBD workers is not statistically significant. Thus, where the work trip does not require coping with heavy congestion, parking problems, and the like, commuters are not very concerned with convenience.
5. The preference of a commuter for flexibility detracts more from the value of the car-pool alternative for commuters bound for CBD destinations than for non-CBD workers. The greater and more frequent availability of transit to the CBD may make car pooling, with its rigid schedules, a less attractive alternative there.
6. Car availability (as measured by the number of cars per driver in a household) is more influential in the

Table 1. Estimated variable coefficients in four models.

Variable	Description	Model 1[a] Coefficient	Model 1[a] Standard Error	Model 2[a] Coefficient	Model 2[a] Standard Error	Model 3: CBD[b] Coefficient	Model 3: CBD[b] Standard Error	Model 4: non-CBD[c] Coefficient	Model 4: non-CBD[c] Standard Error
IVT	One-way (min)	-0.032	0.008	-0.028	0.008	-0.046	0.009	-0.029	0.011
ATIME	Home to primary vehicle (min)	-0.221	0.047	-0.212	0.046	-0.835	0.037	-0.835	0.105
ETIME (non-CBD only)	Primary vehicle to workplace (min)	-0.178	0.074	-0.170	0.075	–	–	–	–
COST	Commuting cost (cents per day)	0.002	0.000 9	-0.002	0.001	-0.002	0.001	-0.002	0.001
Relative comfort (≥$15 000 only)	Scaled from -2 to +2	0.443	0.119	–	–	–	–	0.001	0.137
Relative convenience	Scaled from -2 to +2	–	–	0.600	0.093	0.942	0.115	–	–
Flexibility$_{CP}$	1 if commuter prefers flexibility, 0 otherwise; car-pool function only	-0.802	0.357	-0.736	0.365	-1.45	0.456	-0.691	0.339
Mode unreliability	Times per week mode not available for work trip	0.552[d]	0.391	0.696[d]	0.439	0.360[d]	0.225	0.816[d]	0.510
Cars per driver$_{DA}$	Drive alone function only	2.30	0.578	2.26	0.595	3.482	0.681	1.213	0.931
Cars per driver$_{CP}$	Car-pool function only	0.698[d]	0.532	0.817	0.539	1.508	0.562	0.656[d]	0.939
CBD workplace$_{DA}$	1 if CBD, 0 otherwise; drive alone function only	-2.17	0.452	-2.21	0.473	–	–	–	–
CBD workplace$_{CP}$	1 if CBD, 0 otherwise; car-pool function only	-1.90	0.433	-2.00	0.451	–	–	–	–
Household income$_{DA}$	$ per year; drive alone function only	-0.000 005[d]	0.000 03	0.000 02[d]	0.000 03	0.000 02[d]	0.000 03	0.000 03[d]	0.000 04
Household income$_{CP}$	$ per year; car-pool function only	0.000 009[d]	0.000 03	0.000 02[d]	0.000 03	0.000 02[d]	0.000 03	0.000 05[d]	0.000 04
Sex$_{DA}$	1 if female, 0 otherwise; drive alone function only	0.943	0.378	-1.10	0.397	0.688[d]	0.424	-0.578[d]	0.546
Sex$_{CP}$	1 if female, 0 otherwise; car-pool function only	-0.980	0.375	-1.32	0.391	-0.451[d]	0.358	-0.889[d]	0.567
Mode constant$_{DA}$	1 if drive alone, 0 otherwise; drive alone function only	-2.49	0.825	-3.12	0.849	-6.316	0.935	-5.597	1.31
Mode constant$_{CP}$	1 if car pool, 0 otherwise; car-pool function only	-0.769[d]	0.755	-0.968[d]	0.777	-3.823	7.16	-1.534[d]	1.07

[a]Sample size = 400.
[b]Sample size = 347.
[c]Sample size = 361.
[d]The estimated coefficient is not significantly different from 0 at the 0.05 level or better.

Table 2. Validity of model 1: predicted and actual distributions over the modes for the estimation and validation samples (percentage for each mode).

Distributions	Estimation Sample Transit	Estimation Sample Car Pool	Estimation Sample Drive Alone	Validation Sample Transit	Validation Sample Car Pool	Validation Sample Drive Alone
Actual	38.5	30.2	31.3	32.7	43.3	24.0
Predicted (aggregate of predicted individual mode choices)[a]	43.8	25.2	31.0	40.3	29.3	30.3
Expected value (mean of individual probabilities)[b]	38.8	28.7	32.6	35.8	31.3	32.9

[a]Computed by assigning each commuter in the sample to highest probability mode and calculating the resulting proportion of commuters on each mode.
[b]Derived by computing the mean of the individual probabilities $[\Sigma p_t(i)]/n$ where n is the number of commuters in the sample and $p_t(i)$ is the probability that commuter t chooses mode i.

mode choices of CBD workers. This may be due to the general lack of transit service to non-CBD locations, which forces non-CBD workers to commute by automobile, regardless of their car availability.

7. A number of the estimated coefficients were not significantly different from zero. These include: relative comfort (≤$15 000 only), mode unreliability, sex, and household income. A priori, these variables had been considered important factors in the mode choices of commuters; their actual lack of significance, then, is important.

All of the coefficients in models 3 and 4 have the expected signs and are of reasonable magnitudes. Their significance is that they demonstrate the necessity to consider work location when modeling mode choice.

While the coefficients for sex in models 1 and 2 are highly significant they are not significant in the separate CBD and non-CBD models. This is due to the fact that almost twice as many female commuters work in the CBD than in the non-CBD. When variables are correlated, their estimated coefficients are a clue to the total effect of both variables on mode choice. Since sex became unimportant when the population was partitioned by work location, the total negative effect of sex and CBD must be primarily a CBD effect.

Several variables had insignificant coefficients; i.e., the commuters in the sample were insensitive to variations in these variables when discriminating among the alternative modes. Since these variables were previously considered to be important, their failure to be significant here is worthy of note.

The failure of household income to discriminate was particularly troubling. None of the attempts to include household income (e.g., cost per household income and household income adjusted for family size) gave coefficients that were significantly different from zero. Household income may be an important discriminator only in the lower ranges (e.g., under $8000) in which the sample was deficient.

Attempts to estimate a statistically significant coefficient for the SES of the job held by the commuter also failed. Blue collar workers did not have any greater or lesser tendencies to choose any one mode over the others. Nor did the number of workers employed by the employer of a commuter have any influence on mode choice.

MODEL VALIDATION

As a test of the validity of model 1, the probabilities of choosing each of the three modes were calculated for each commuter included in the sample of 400 commuters used to estimate the coefficients. In 68 percent of the cases, the mode with the highest probability of choice was the one actually chosen. This percentage is similar to the percentages reported elsewhere (2).

The predicted distribution over the modes (i.e., the mode split) was calculated in two different ways. First, the mode choice of each individual was predicted by the highest probability method and the percentage of individuals predicted to choose each mode then calculated. Second, the individual probabilities of each mode were summed and averaged to give the expected value estimate of the modal-split distribution. These two distributions are compared with the actual distribution for the estimation sample in Table 2. The proportion of commuters predicted to drive alone is almost equal to the actual proportion in the sample. At the worst, transit ridership is overpredicted by 5.3 percent and car poolers are underpredicted by 5 percent. The model, then, does very well at aggregate prediction of the distribution over the modes and reasonably well at prediction of the individual mode choices for the estimating sample.

To further test the predictive ability of the model, the three mode choice probabilities were calculated for 300 commuters who were not used in the estimation of the model. The model predicted their individual mode choices correctly in 59 percent of the cases.

CONCLUSION

This modeling effort has identified numerous characteristics of car pools and potential car poolers that should be considered in managing programs to encourage the use of car pooling. It identifies subpopulations of commuters who have high car-pool potential and service attributes of car pooling that would entice more commuters into car pools if they were improved.

The subgroups of commuters at whom car pool encouragement efforts should be directed are

1. Commuters who have regular (as opposed to flexible) commuting schedules,
2. Commuters who have moderate or low ratios of cars per driver in their households,
3. Commuters who have annual incomes of less than \$15 000 (suggested by the coefficient of comfort for commuters having incomes greater than \$15 000 and by the fact that commuters who drive alone generally view car pooling as a considerably less comfortable mode), and
4. Commuters who drive alone to the CBD who will not easily be lured into car pools (because CBD commuters are very sensitive to convenience and those who drive alone generally consider car pooling to be a less convenient mode).

The modeling effort also helped to identify improvements in the service attributes of car pools that would have a significant impact on mode choice. These include

1. Provision of a back-up service for car poolers—This would lessen the influence of its inflexibility on mode choices. Taxi and emergency car-pool matching operations are possible back-up services.
2. Policies that minimize the OVT of car pooling—Here again, preferential parking spots for car poolers would be effective. To minimize the OVT of car pooling, commuter computer matching operations should strive to match only commuters with very similar origins, destinations, and schedules. Also, since the 4-min average ATIME of car poolers in the sample is significantly lower than previous estimates, advertising its real magnitude could alter the perceptions of car pooling held by commuters. The magnitude of the OVT coefficient suggests that such a change could have a sizeable impact on mode choices.
3. Policies, such as exclusive car-pool lanes, that influence IVTs.
4. Policies that improve the convenience of car pooling—Such policies include preferential parking or flexible work hours for car poolers. Another effective policy would be an advertising campaign designed to dispel subjective notions about the inconvenience of car pooling.

In addition to these conclusions, the model casts doubts on some efforts that previously have been effective in drawing commuters to car pools. These include:

1. The economy of car pooling does not substantially encourage many commuters to car pool. Therefore, advertising campaigns that emphasize cost will not be as effective as the themes suggested above.
2. The estimated cost coefficients show that in the short run nondrastic cost incentives (e.g., taxes on driving alone) should not be expected to yield substantial changes in mode choices.

Finally, the modeling effort has defined some areas needing further research. Further work on the long- and short-run influences of costs, the significance of OVT, and the importance of comfort and convenience will be useful.

ACKNOWLEDGMENT

This paper was part of the systems synthesis project course taken by MS students in the School of Urban and Public Affairs at Carnegie-Mellon University. Other students participating in the course were Pearl Davis, Alan Standring, Ahmed Thomas, Douglas Gunwaldsen, Douglas Doxey, and Richard Kuzmyak. Alfred Blumstein was the faculty advisor supervising the project. Significant assistance was provided by Charles Manski and Martin Wohl of Carnegie-Mellon University and Moshe Ben-Akiva of Massachusetts Institute of Technology. Additional assistance was provided by Daniel Nagin and Jacqueline Cohen of the Urban Systems Institute of Carnegie-Mellon University. The effort was supported by the Southwest Pennsylvania Regional Planning Commission. Wade Fox of the Southwest Pennsylvania Regional Planning Commission was extremely helpful in providing guidance and suggestions. The support of all these individuals is very much appreciated.

REFERENCES

1. M. Ben-Akiva. Structure of Passenger Travel Demand Models. Department of Civil Engineering, M.I.T., Cambridge, PhD thesis, 1973.
2. M. Ben-Akiva and S. Haws. Estimation of a Work Mode Split Model Which Includes the Car Pool Mode. Department of Civil Engineering, M.I.T., Cambridge, 1974.
3. A Disaggregated Behavioral Model of Urban Travel Demand. Charles River Associates, Cambridge, Mass., 1972.
4. P. R. Stopher, B. D. Spear, and P. O. Sucher. Toward the Development of Measures of Convenience for Travel Modes. Department of Environmental Engineering, Cornell Univ., Ithaca.
5. Free Transit. T. Domencich and G. Kraft. Charles River Associates, Cambridge, Mass., 1972.
6. J. Ganek, R. Saulino, J. Cohen, and R. A. Stafford. Evaluation of the Car Pool-Public Transit Program of the Southwestern Pennsylvania Regional Planning Commission's Engery Conservation Plan. Southwestern Pennsylvania Regional Planning Commission, Pittsburgh, 1975.

An Application of Mode-Choice Methodologies to Infrequent Commuter-Rail Service

R. Ian Kingham, National Cooperative Highway Research Program

The feasibility of commuter-rail transit in the southwest Baltimore corridor was studied by a variety of passenger estimation methodologies. The methodologies selected were required to be applicable to the corridor scale, to be run manually, and to be capable of quick response. They were also required to be responsive to the addition of one or two trains per peak period, changes in station location and accessibility, and changes in costs such as parking charges and gasoline costs associated with the automobile. No one methodology met all of the above requirements. However, two methodologies were adapted to consideration of infrequent rail service (one or two trains per peak period) and applied to the corridor. The first methodology involved the application of a simple graphical technique that related mode split to station distance from the CBD; the second involved the application of a marginal utility model to corridor census tracts. The infrequent service capability was added, in the case of the graphical approach, by applying experience factors, and in the computational approach by relating automobile captivity to the number of trains per peak period. Both methodologies were transferable, without reestimation of coefficients, to the southwest Baltimore corridor. Both approaches could be applied manually in a person-week or less; the need for any greater sophistication than the graphical methodology is seriously questioned.

In many urban areas of the United States, development of commuter-rail transportation is being advocated in lieu of further highway development. This is the case in the Baltimore-Washington corridor where a policy to improve present train service on the Baltimore and Ohio Railroad right-of-way has been adopted. The pressure to improve this service has come from the northeast Washington, D.C., corridor, and little thought has been given to service improvements for the corresponding southwest Baltimore corridor. Although the Baltimore regional plan has recommended that commuter-rail service to downtown Baltimore, using existing rail lines, be given serious study and adequate experimentation, for the most part it has concentrated on two modes of public transportation: a rapid transit system and a mixture of express and local bus service. The question that initiated this study is whether resources should be directed to commuter-rail service in the southwest Baltimore corridor.

Good corridor planning should consider ridership potential. This requires a methodology that is responsive to a number of proposed transportation strategies, such as parking regulation, gasoline taxation, commuter-rail fare changes, and commuter-rail service improvements. Since the number of alternative strategies may be large and the time available for evaluation short (a month or less in a citizen participatory process), the methodology must be simple and easily applied. Models must consider travel time and cost within their structures. For quick response, a methodology should be sufficiently simple to allow for manual computations. Data requirements should be confined to the corridor of interest. Since census tract data are easily obtained, the methodology should be designed to use them directly. Home interview data (other than those obtained in the 1970 census) should not be required. For commuter-rail passenger estimation, the methodology should consider the exact number of peak-period trains under consideration.

DESCRIPTION OF THE CORRIDOR AND EXISTING COMMUTER-RAIL SERVICE

Corridor Description

The southwest Baltimore corridor served by the Baltimore and Ohio (B&O) Railroad is shown in Figure 1. The population of the corridor, from the greater Laurel area to the Baltimore beltway (Interstate 695), is approximately 68 000. Average family incomes in 1969 for corridor census tracts ranged from $10 281 to $12 672 except for one of $16 632. The number of cars per capita is 0.41, a figure almost as high as the 0.44 in Los Angeles (1). There are 21 000 housing units and an average of 1.32 cars/household. A length of approximately 33 km (20 miles) of the corridor has a significant Baltimore orientation (based on newspaper circulation observations). The residential density is generally low, and, except for the Laurel area, existing communities have remained approximately the same size since World War II. The area surrounding Laurel has been part of the fastest growing area of Prince Georges County.

The corridor routes are all very heavily traveled, primarily by automobile, truck, and rail freight. There are two freeways, the Baltimore-Washington Parkway and Interstate 95, and two primary highways, US-1 and US-29, parallel to the freeways. The parkway carries no truck traffic but provides access to the Baltimore CBD. Interstate 95, however, terminates inside the beltway and leads to an arterial street, requiring a 15 to 20-min drive to the CBD in the peak hour.

There are two railroads. The Penn Central approaches Baltimore almost directly from the south and then changes direction to skirt the west side of the Baltimore-Washington International (BWI) Airport and enter the city from the west. However, the Penn Central Station is north of the CBD and well out of walking range. The second railroad, the B&O, approaches Baltimore between the two expressways from a southwesterly direction and terminates at Camden Station, approximately 1.2 km (0.7 mile) from the key employment centers of the CBD.

Highway transportation perpendicular to the rail lines consists almost exclusively of primary state highways. There are no freeways at present although several are planned. Thus, there is no east-west accessibility of Interstate standards outside the Baltimore beltway, approximately 8 km (5 miles) from the CBD.

Existing Commuter-Rail Service and Patronage

Passenger service in the southwest Baltimore corridor is maintained by the Penn Central, which has one stop between Washington and Baltimore, and the Baltimore and Ohio Railroad, which has four stops (Riverdale, Laurel, Jessup, and St. Denis).

The location of the terminal stations serving the CBD is extremely important to any ridership estimate. A geographical plot of the origins and destinations of 10 000 Philadelphia area commuters has shown that these commuters used trains only if their CBD destinations were within 10 min by foot or transit from their arrival station (2). The Camden Station of the B&O Railroad is not well situated with respect to the Charles Center, the major employment area in the CBD, for those who presently take the train. A walk of 1.2 km (0.7 mile), approximately 15 min, is typical. The Penn Central Railroad Station, however, is well beyond the CBD, and without shuttle service to meet the trains, this system is not likely to attract potential riders. For this reason, the Penn Central Railroad was not considered an attractive alternative to bus and automobile in the corridor and is not further considered in this paper.

The B&O runs one train, which leaves Laurel at 7:22 a.m. and arrives in Baltimore at 7:48 a.m. In the evening, the train leaves Baltimore at 4:55 p.m. and arrives in Laurel at 5:22 p.m.

Passenger counts and an origin-destination (O-D) survey made in 1973 showed 63 people traveling into Baltimore in the morning and 62 returning in the evening in a train providing 130 seats. The largest number of passengers (23) boarded at the Laurel station, 30 km (18 miles) from the CBD. There were 43 passengers using the Laurel, Jessup, and St. Denis Stations. Almost all respondents were employed in the Charles Center.

MODE-CHOICE METHODOLOGY EVALUATION AND DEVELOPMENT

Evaluation

Mode-choice methodologies have been developed for both intercity and intracity passenger demand forecasting. The length of corridor and the results of the O-D survey suggest that the B&O commuter-rail service is an intracity rather than an intercity service. The majority of fare-paying passengers travel a distance of 33 km (20 miles) or less. Those traveling longer distances, particularly those traveling from Washington to Baltimore, were B&O employees traveling on passes. For these reasons, the intercity methodologies developed for the northeast corridor were considered to be inappropriate here and are not considered further in this paper.

Methodologies presently being used by departments of transportation and highways are limited in their usefulness for commuter-rail passenger forecasting. The principal packages available are those developed by the Federal Highway Administration and the Urban Mass Transportation Administration. These require input data for a network that comprises the entire urban area, whereas in this study only a corridor is of interest. It is difficult to break away from the sequence of trip generation, trip distribution, and mode split when, as in this case, a trip table can be obtained from census data. For the headways greater than 30 min that are common for commuter-rail service, the models are not appropriate. Both packages required a computer and provide exact numerical computations when only a rough estimate may be necessary. The use of the computer places the analysis at a distance from the analyst, which may hide good strategy alternatives that might appear during manual manipulation of the data.

Conceptual Development

The methodology of estimating commuter-rail patronage shown in Figure 2 (2, 3) appears to be a logical and simple approach for making quick estimates of travel demand for proposed transportation strategies. It is a deductive methodology that considers the potential market and its characteristics (such as residential density and length of corridor), the competing transportation modes in the corridor, and the frequency of train service, in order to determine the anticipated patronage. However, to use this methodology requires an experienced analyst. For a less experienced analyst, a modal-choice relationship that considers all those factors is required. To meet this need, the methodology was expanded in two ways. The first approach was to test the use of an existing graphical technique in place of the deduce ¢ rail shown in Figure 2. The second (computational) approach was to impose an existing modal-split model on the methodology as shown in Figure 3.

Elements Common to the Two Approaches

Demand Area

The determination of a demand or geographic area over which patronage will be drawn is common to both methodologies. Experience must be the guide for this. A 6-km (4-mile) corridor width based on observations in Philadelphia that suggest most potential riders will not drive more than 3 km (2 miles) perpendicular to the corridor direction has been used (4). In a study of park-and-ride facilities (5), it was found that 50 percent of those using park-and-ride resided within a distance of 4 km (2 to 3 miles) from the lot with 90 percent residing 5 to 10 km (3 to 6.5 miles) from the lot. Since these distances include travel in all directions, their perpendicular component could be close to the 3-km (2-mile) criterion observed in Philadelphia. By analogy with these observations, the southwest Baltimore corridor was defined as being 6 to 11 km (4 to 7 miles) wide, except in Laurel where, to include the entire greater Laurel area, a width of approximately 15 km (9 miles) was used. These limits are compatible with census tract boundaries and are shown in Figure 1.

The total distance of the automobile journey as compared to that of the commuter-rail journey was also used in establishing corridor boundaries. The criterion here was that the distance traveled by automobile not be greater than the distance traveled by train. In all cases the ratio of automobile distance to train distance was actually considerably less than one.

Data

The 1970 census had enumerated for each census tract the employed persons who journeyed to the CBD. Since commuter-rail service is used for work trips almost exclusively, only such trips were considered in the methodology developed. The total market is 508 persons, of whom approximately 40 percent reside within the demand area for the St. Denis Station, approximately 11 km (7 miles) from the CBD.

The census data also included the number of households having one, two, and three or more automobiles. These data were used to determine the average car ownership per household and per capita for each census tract.

Figure 1. The southwest Baltimore corridor.

Competing Modes

Another step in the methodology common to both the graphical and computational approaches is the identification of modes competing with commuter-rail service. In the graphical approach, experience in other corridors must reflect similar competition, which in most instances includes the automobile and transit bus. In the computational approach, running times, access times, and costs must be determined for all modes.

Elements of the Graphical Approach

Mode-split relationships for the graphical approach are based on a correlation of the commuter ridership, expressed as the percentage of the total number of trips to the CBD, and the distance from the station to the CBD (4). Relationships based on experience in Philadelphia, Chicago, and San Francisco are shown in Figure 4. The selection of the appropriate relationship should be based on the similarity of residential densities and competing modes. The lower commuter-rail percentages represent combinations of low residential densities and strong competition from parallel freeways.

From the mode-split curve, the appropriate commuter-rail percentage is multiplied by the number of CBD employees for each census tract in the demand area to determine the potential patronage. Since patronage determined in this manner is representative of service frequencies of three or more trains in the peak period, it must be reduced for lower frequencies. The following reductions are suggested (2):

No. of Trains per Peak Period	Reduction in Patronage (%)
1	38 to 43
2	53 to 60

Elements of the Computational Approach

Mode-Split Models

There are several additional important characteristics that a commuter-rail mode-split model must have in order for it to be policy responsive. The first is that

Figure 2. The deductive methodology.

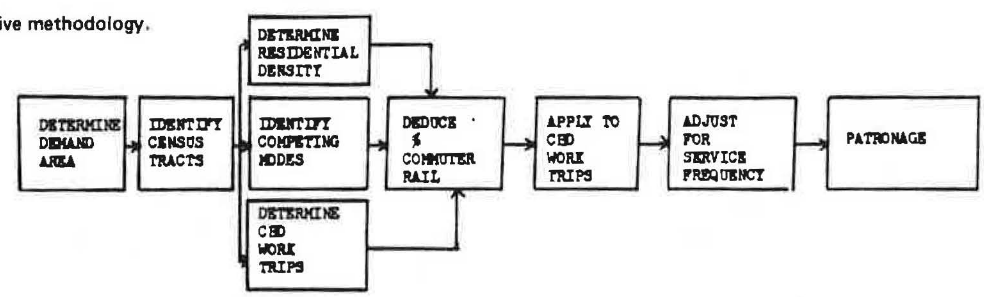


Figure 3. Computational approach for commuter-rail ridership estimation.

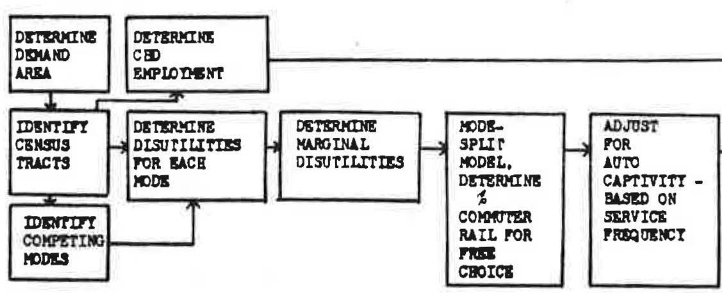
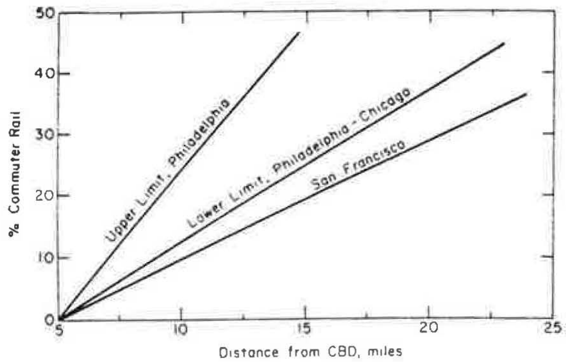


Figure 4. Graphical approach for commuter-rail ridership estimation.



the model must consider low service frequencies as characterized by one or two trains per peak period. The second is that the model should account for income or be calibrated using high-income patronage as commuter-rail studies have shown that commuter-rail patrons have access to a car and higher than average incomes.

The models reviewed are trip interchange or post distribution models. These synonymous terms describe a model requiring knowledge of the trips made between zones. This type of model is advantageous because CBD-employed persons are known for the corridor census tracts.

To meet the need to be policy responsive, the models reviewed were limited to those that considered both travel time and cost. A disutility or impedance is determined for each mode by a linear combination of running times, waiting and walking times, and costs, appropriately weighted by factors derived from previous studies. The greater the disutility, the greater the travel time and cost; hence, the likelihood that an individual will choose the mode is reduced. Thus, the assumption is made that a person rationally measures, for each travel mode, the disutility (such as time and cost) necessary to arrive at a destination, and chooses the mode that will minimize the disutility. There are many models fitting this description (1, 4, 6, 7, 8, 9, 10, 11, 12). Other models based on this premise (12, 13) were not evaluated in detail since their complex structures make manual computations more difficult.

Of the models evaluated, none were capable of considering service frequencies as low as one or two trains per peak period: All accounted for long headways through the waiting time variable. Some (4, 12) were not calibrated for commuter-rail, and one (12) did not consider income directly.

The model for the Washington Council of Governments (WASH COG) (11) considers income directly, but is calibrated with an areawide bus system. However, this model also considers automobile and transit captives as functions of transit accessibility and household income. Accessibility is defined as the proportion of jobs within a 45-min transit travel time. Automobile and transit captive rates were deduced from Washington data that considered accessibility at both origin and destination points. Hence, for a commuter-rail problem accessibility would be high for the destination and low for the origin, and household income would be high. For such a problem, automobile captivity is given as 38 percent and transit captivity as 1 percent.

Study Hypothesis

Since the WASH COG model was calibrated from data for a city having an areawide bus system, the frequency of service was much greater than that typical for a commuter-rail system. For this study, it was hypothesized that automobile captivity would be substantially greater for a commuter-rail operation having only one or two trains per peak period. Furthermore, it was assumed that transit captivity would be zero since an automobile is necessary for access to the rail stations.

Development of Automobile Captivity Per Number of Trains Relationship

It is possible to compute the free choice patronage for the commuter-rail and automobile modes for the Laurel, Jessup, and St. Denis Stations using the WASH COG free choice model for work trips by an equation that relates the percentage of patrons using transit to the marginal utility, which is defined as the difference between the automobile and transit disutilities for each census tract.

To determine commuter-rail and automobile disutility measures, the following assumptions were made:

1. Passengers have a 15-min walk from the B&O Station to their place of work.
2. The perceived automobile cost in 1970 was $0.04/km ($0.06/mile) for the automobile mode only trip.
3. For the rail trip, the average patron does not perceive the cost of driving his automobile to the station except for the time taken.
4. The average commuter-rail patron has an income of $15 000/year.
5. The average passenger perceives a wait time of 5 min if he expects the train to be on time. (It was further assumed that a passenger would allow 4 min to find a parking space and to reach the station platform. Hence, a 9-min wait time was assumed for dependable service. Since present patrons feel that the service is not dependable, it was assumed that this increased the perceived wait period from 9 to 12 min.)
6. The time required to drive to a station is assumed to be a component of excess time and is weighted by a factor of 2.5.

Disutilities were computed for the automobile and commuter-rail modes. In general, there were no competing public transit modes, except in the St. Denis Station area. In that area the Halethorpe bus provides service at 9-min headways during the peak hour. The disutilities determined for the bus were approximately the same as those determined for the commuter-rail service even though the commuter-rail trip has a much shorter running time. However, the bus discharges its passengers at the Charles Center with little or no walk and has short wait times due to the short headways. Because these disutilities were equal overall, the commuter-rail patronage determined from the automobile-rail mode-split computation was halved for the St. Denis Station demand area.

The only other public transportation in the corridor is provided by the Greyhound Company, which runs three buses in the peak period along US-1 to its Baltimore terminal. This service suffers as a commuter service because the terminal is considerably north of the CBD. It was assumed, then, that Greyhound does not serve CBD-oriented employment. From the percentage of free choice patrons using the commuter-rail service determined from the marginal disutilities for each census tract, passenger estimates can be generated as a function of automobile captivity as follows:

$$P = \left[\sum_{i=1}^{n} (CR\%_i)(E_i) \right] [1 - AC] \tag{1}$$

where

P = number of passengers,
$CR\%_i$ = percent commuter rail for census tract i,
E_i = CBD employees for census tract i, and
AC = automobile captivity expressed as a fraction.

To determine the automobile captivity corresponding to one, two, or three trains per peak period, it is necessary to independently determine the patronages for one, two, and three trains. For one train, the current patronage can be used. However, some reduction in patronage must be expected if all passengers are required to pay the regular fares. In a Maryland Department of Transportation on-board passenger survey it was estimated that, of the 43 persons boarding at Laurel, Jessup, and St. Denis, 38 were fare paying passengers. For two and three trains, a factoring process can be used to determine revenue passengers, and, for the revenue passengers determined, the automobile captives can be determined from equation 1. These results are shown in Figure 5. The shape of the curve is reasonable because it can be anticipated that reduction in automobile captivity will be less as the number of trains increases. With such an adjustment the WASH COG model becomes responsive to the policy action of increasing the number of trains.

If 66 percent automobile captivity is used, the WASH COG model gives the present patronage for the study corridor. However, the appropriateness of such a mode-split relationship for other corridors must be tested. Thus, the model was compared to relationships (1) that describe the variation in percent transit as a function of the disutility for various values of car ownership per capita. For three or more trains and dependable service, a patronage estimate from the Chicago model (9, 10) should compare closely to the patronage estimated using the WASH COG model with 26 percent automobile captivity.

PATRONAGE ESTIMATION

Commuter-Rail Strategies

The following strategies were considered for the improvement of commuter-rail patronage:

1. Increased number of trains,
2. Improved service dependability (to give a perceived reduction in wait time from 12 to 9 min),
3. Increased parking charges (from \$36 to \$72/month) in the CBD), and
4. Reduced transit access times (by improving station access roads, constructing new stations, or eliminating old ones).

The Graphical Approach

It is a simple matter to compute anticipated ridership patronage from the known market by using Figure 4. The results, given below, indicate the present service carrying 46 fare-paying passengers.

Strategy	No. of Fare-Paying Passengers
One train (present service)	46
Two trains in peak period	59
Three trains in peak period	99

The curve for San Francisco was chosen for use in the study corridor because of similarities in residential density and competing highways. The estimated increases in patronage for two and three trains per peak period were determined by using factors inferred previously (2). The graphical approach does not provide a means to determine changes in demand caused by changes in parking charges or in transit excess times.

The Computational Approach

Travel Forecast for Service and Parking Cost Changes

Travel forecasts were prepared using both the modified WASH COG and Chicago models (9, 10). Present ridership for the Laurel, Jessup, and St. Denis Stations is compared with ridership estimates below.

Patronage Condition	Chicago, 3 Trains	WASH COG, 1 Train	WASH COG, 2 Trains	WASH COG, 3 Trains
Present patronage	75	37	51	85
Patronage if dependability improved	101	42	56	94
Patronage if parking charges increased and dependability improved	90	55	74	123

The improved dependability estimates are comparable to those developed from the graphical approach. There were, however, large differences for the doubling of parking charges, due largely to differences in converting cost to time. (Variation in the value of time with income is considered in the WASH COG model and not in the Chicago model.)

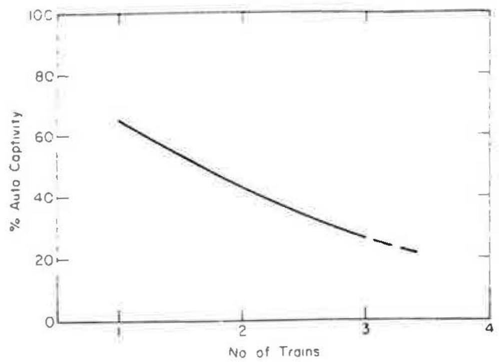
Whatever the strategy, estimates of patronage are low. If increasing the service frequency by adding trains adds only 20 to 30 passengers/train, such increases are not cost-effective. These figures, however, could be useful in estimating additional revenue if trains were deadheaded over the route during peak periods to meet service requirements in Washington.

These estimates for the Laurel, Jessup, and St. Denis stations should be expanded to obtain total patronage by including passengers boarding at Riverdale and Washington. At present, the 20 such passengers are approximately one-third of the patronage. It is not likely that this number will increase for fare-paying passengers, primarily because of the small number of persons who reside in the Washington suburbs and work in the Baltimore CBD.

Travel Forecast for Changes in Station Locations

A close study of population concentrations in the corridor suggests only two possible additional locations for stations. The first was Hanover Road, which would better serve Elkridge, Dorsey, and the area west of BWI Airport. However, forecasts indicate that there would be no overall increase in commuter-rail patronage from this. The second possible station location is the Baltimore beltway at Hollins Ferry Road. The accessibility provided by the beltway would increase the potential market from 508 to 957. However, most of this market is served by Metro transit buses, and, in order for this station to be feasible, transit bus service would have to be reduced. Aside from that consideration, moreover, the practicality of a station only 7 km (4.5 miles) from the CBD is questionable; most commuter railroads have abandoned stations less than 10 km (6 miles) from the CBD.

Figure 5. Automobile captivities and train service.



Travel Forecast for 1980 Travel Conditions

A limited study was conducted to determine the impact of more congested highways; 1980 travel speeds provided by the Maryland Department of Transportation were used with 1970 demographic data. The forecasts showed commuter-rail passenger counts increasing by only 15 passengers. Even though the peak-hour travel speed from Laurel to the Baltimore CBD is forecast to be 29 km/h (18.5 mph) on the Baltimore-Washington Parkway, the impact on commuter-rail patronage will be small because of the congested east-west arterials leading to the commuter-rail stations.

THE NEED FOR A MODEL

A comparison of patronage estimates made by the graphical and computational approaches shows the differences between the two approaches to be very small and provides evidence that the graphical approach is good enough to be used in an initial approach to determining passenger demand.

COMPARISON OF RESOURCE REQUIREMENTS

Both the deductive and modeling approaches were conducted manually using Maryland DOT highway maps and regional traffic assignments, census publications, published transit schedules, and a desk calculator. The state DOT traffic assignments were not necessary to the success of the study. Average speeds, and hence average times, could be judged from the road classification, speed limit, and density of development.

An estimate of the person-hours required to do two analyses is given below.

Task	Graphical Approach	Computational Approach
Determine demand area Identify census tracts Determine CBD work trips Identify competing modes	16	16
Determine residential density Select mode-split curve	4	
Determine disutilities for each of 11 census tracts for each mode		24
Determine mode split for each of 11 census tracts Correct for one or two trains per peak period, if necessary Determine patronage for each census tract and total ridership	4	4
Total	24	44

A generous amount of time is allowed to the first four tasks of assembling materials. The computational approach time requirements are almost double those of the graphical approach because of the time required to determine travel times and disutilities for each mode.

CONCLUSIONS

The question of allocating greater resources to commuter-rail transit in the southwest Baltimore corridor was answered by the very small market of 508 persons. Unless the demographic characteristics of the area were to change drastically, the potential of commuter-rail transit could never be exploited. However, the study did demonstrate the successful manual application of several mode-choice methodologies and the response to a number of policy options relating to infrequent service.

Commuter-rail corridors can be characterized by residential density, competing parallel freeways, and terminal station location relative to the CBD. Patronage experience in corridors defined by these characteristics can provide a valuable guide in planning rail service for other corridors. Such experience has been described (4) by a function that relates distance from the CBD to the percentage of the market using commuter-rail transit. Estimates of patronage determined from this graphical approach can be factored from past experience to determine the effects of low-frequency service. This methodology can eliminate the need for a more sophisticated model and be applied with approximately 24 person-hours of effort. However, the relationship is based on past experience and is tied to time, cost factors, comfort, and convenience experienced in the past. Also, there is no way to determine patronage changes resulting from changes in accessibility or parking charges. These limitations may require a more sophisticated approach such as disutility models.

Of existing post-distribution, mode-split models, those using utility models are responsive to policy and planning issues such as parking taxes, transit fares, station location, and road access to rail stations. One such model, the WASH COG model, can be adapted to low-frequency operations by adjustments to automobile captivities and the development of a relationship between automobile captivity and the number of trains in the peak period.

This model can be applied at the corridor scale by using CBD employment for corridor census tracts. Detailed transportation zone data or regional networks are not required. Because of the narrowness of a commuter-rail corridor and the resulting few census tracts involved, the entire computational process can be done by hand. The model can be applied to a 32-km (20-mile) corridor in a person-week of effort. The model exhibited transferability without recalibration in its application to a Baltimore corridor.

ACKNOWLEDGMENTS

The assistance of Mike West and Jay Hierholtzer of the Maryland Department of Transportation in providing access to data is gratefully acknowledged.

REFERENCES

1. General Approach to the Development of a Macro Choice Model: A System Sensitive Approach to Areawide Urban Travel Forecasting. Alan M. Voorhees and Associates, Inc., Technical Memorandum 4, May 1971.
2. C. R. Englund. Commuter-Railroad Feasibility Study on Selected Lines in the Los Angeles Metro-

politan Area. Southern California Association of Governments, Jan. 1974.
3. C. R. Englund. Washington Metropolitan Area Rail Commuter Feasibility Study. Urban Mass Transportation Administration, U.S. Department of Transportation, May 1971.
4. Traffic Forecast, Proposed Interim, R. F. and P Railroad Commuter Demonstration Project. Alan M. Voorhees and Associates, Inc., McLean, Va., Jan. 1968.
5. G. H. Tanner and R. Barba. Park-and-Ride Transit Service; Some Guidelines and Considerations for Service Implementation. New York State Department of Transportation, Rept. PRR44, April 1973.
6. A New Procedure for Urban Transportation Planning. De Leuw, Cather and Co. of Canada, Sept. 1969.
7. S. J. Bellomo, C. G. Turner, and D. K. Johnston. Modal-Choice Model for Relating Demand to Investment. HRB, Highway Research Record 392, 1972, pp. 1-12.
8. G. A. Shunk and R. J. Bouchard. An Application of Marginal Utility to Mode Choice. HRB, Highway Research Record 322, 1970, pp. 30-39.
9. G. W. Schultz and R. H. Pratt. Estimating Multi-Mode Transit Use in a Corridor Analysis. HRB, Highway Research Record 369, 1971, pp. 39-46.
10. R. H. Pratt. A Utilitarian Theory of Travel Mode Choice. HRB, Highway Research Record 322, 1970, pp. 40-53.
11. Development and Calibration of the Washington Mode Choice Models. R. H. Pratt and Associates, Inc., Technical Rept. 8, June 1973.
12. T. E. Lisco. The Value of Commuters' Travel Time—A Study in Urban Transportation. HRB, Highway Research Record 245, 1968, p. 36.
13. P. R. Rassam, R. H. Ellis, and J. C. Bennett. The Dimensional Logit Model: Development and Application. HRB, Highway Research Record 369, 1971, pp. 135-147.

The Subarea Focusing Concept for Trip Distribution in the Puget Sound Area

Donald G. Miller, Peat, Marwick, Mitchell and Company
Nancy L. Nihan, University of Washington

This paper explores one method of reducing the computational costs associated with urban travel demand modeling. The technique investigated is a data base approach called subarea focusing. The distinction between this technique and other data simplification methods is in the detailed analysis of only a portion of a study area and the simultaneous presence of several levels of areal detail in the data base. A computerized technique for subarea focusing was developed and used for a sample trip distribution application to data from the Puget Sound region. The procedure was not unduly expensive in either manpower or computer time. When aggregate data sets were used to obtain travel demand predictions, substantial savings in computer time were realized. The error analyses indicated some sacrifice in accuracy, but not a serious sacrifice. The results appear to justify continued refinement of the aggregation procedure and investigation of the effects of subarea focusing on other demand models.

Conventional travel demand forecasting procedures have received substantial criticism from professionals and policymakers since the late 1960s. A repeated complaint [e.g., Bouchard (1)] is that conventional forecasting procedures are too time-consuming and too expensive. In response to these and other criticisms, new methods are being discussed, developed, and tested.

One method of reducing the excessive costs of computer modeling is the simplification of model inputs. Dial and others (2) have proposed three types of data base simplification: regionwide abstraction, subarea focusing, and subarea windowing. Figure 1 illustrates these three concepts. Regionwide abstraction is the uniform aggregation of network and zone information across a study region to create a district system (areal scheme b in Figure 1). Subarea focusing is the extraction of a subarea of interest (called the window) from the original data base and the abstraction of zone and network information outside the window (areal scheme c in Figure 1): All data within the window boundary are kept at a zonal level of detail. Subarea windowing also extracts a subarea of interest from the original data base but then collapses trip ends outside the window onto the window boundary much like the treatment of external stations in the original network (areal scheme d in Figure 1).

For analyses in which only a small section of an entire region is of interest, subarea focusing appears to be the best method for reducing data requirements without adverse effects on investigations inside the windowed section. It is best suited to local planning, corridor analyses, and major updates of existing regional plans when only a small number of subareas are under investigation.

Although subarea focusing has played a role in transportation planning studies for many years (3, 4), it has been formally recognized only recently with the advent of sketch planning techniques. Thus, while the concept has been applied informally for some time, there is little documentation on how this was done, and no studies have been initiated on the effects of applying it. Thus, the prime objective of this investigation was the development and automation of an aggregation methodology for subarea focusing. A second objective was the demonstration of the effects of such simplification on predicted trip flows to, from, and within a subarea of interest for a realistically sized study region.

A few recent studies have considered means of alleviating many of the time and money problems associated with conventional modeling techniques. Peat, Marwick, Mitchell and Company (5) have proposed new simulation models, network and zone aggregation techniques, and an interactive computer environment in an early but detailed look at sketch planning techniques. When operational these concepts were to have been added to the UMTA Transportation Planning System, but they have since been superseded by others (2). These techniques included an integrated multimodal demand model, a time-sharing computer environment, and the three types of data simplification described above. Mann (6), on the other hand, has developed a composite model using simplified versions of existing models with a district-level data base that achieves significant savings in computer time. The most unconventional sketch planning technique for travel demand forecasting is that of Schleifer, Zimmerman, and Gendell (7), which does not require extensive network coding since arterials are considered ubiquitous and areal data are aggregated to the community level. The model is noniterative and requires only trip end data by mode for each community. Transit estimation is presently done manually.

Two other studies were of interest to this project. One aided in the study design; the other gave some insights about the effects of zonal aggregation.

Bovy and Jansen (8) define the problem and present an experimental design for investigating the effects of spatial abstraction in transportation planning. They propose investigating the effects of abstraction on travel predictions and costs for three types of trip assignment models, three degrees of network abstraction, and three degrees of zonal abstraction. The models investigated are conventional in nature and include all-or-nothing, capacity restraint, and multipath assignment techniques. Although not clearly stated, the data simplifications seem to be of the regionwide abstraction type.

Wilbur Smith and Associates (9) studied the effects of zonal data simplification on trip distribution and assignment models. The areal simplifications were of a regionwide abstraction type, manually created, and spanned a range of total centroids from 630 to 73. Three trip purposes in the trip distribution stage, two trip assignment models, and six different sized areal systems were analyzed for both trip assignment and trip distribution. It was concluded that major reductions in the number of centroids on a regionwide basis could be achieved before errors due to aggregation became larger than errors inherent in the trip assignment process.

PROJECT DEFINITION

Study Scope

The salient features of the subarea focusing study are presented here. A more detailed description of the experimental design, aggregation procedure, and error results is given in Miller (10).

In applying the subarea focusing concept the data simplification was limited to zonal information. Network modifications were avoided for three reasons: (a) Zonal aggregation appeared far easier to achieve, (b) although a network aggregation technique has been developed (11) it has not been successfully automated for large networks, and (c) significant regionwide reductions in areal units have been achieved without incurring unreasonable trip prediction errors (9).

In the initial demonstration of the subarea focusing procedure, one design variable, the degree of aggregation, was varied while others such as window location and size were held constant. [In an associated project (12) the sensitivity analysis is extended by examining two additional window locations for approximately the same degrees of aggregation. Both studies limit their investigations to trip distribution predictions using the gravity model.]

The study area was the three-county mainland portion of the Puget Sound region. The data base for this area consists of a base 635-zone system containing 597 traffic analysis zones and 38 external stations, a network with 9495 one-way links and 3186 nodes, projected 1990 home-based person work-trip productions and attractions, and projected impedance factors for work trips in 1990.

The Seattle CBD was the subarea of interest. There are 42 traffic analysis zones within this window. Aggregate districts were created from one or more original zones outside the window and approximately aggregated to the same level of detail regardless of distance from the window boundary. The degree of aggregation was measured by the percentage reduction in the number of centroids outside the window. Three degrees of aggregation, 55, 79, and 90 percent reductions in centroids outside the window, were analyzed. The total number of centroids present in each system was 302, 165, and 103 respectively.

A forecast trip matrix was obtained for the 635-zone system by applying the gravity model to the base data. This forecast trip matrix was considered correct, and any deviation from it was attributed to data aggregation. A computerized aggregation procedure was developed to generate aggregate district data from the base zone system. An aggregate trip matrix was obtained for each degree of aggregation by applying the gravity model to the district-level data. The fundamental comparison in the error analysis was between the forecast trip matrix (compressed into appropriate district-level dimensions) and each of the aggregate trip matrices.

Spatial Aggregation Procedure

There are several assumptions associated with the aggregation procedure developed. They are: that zonal trip and network data are available for a particular region and that some subarea of the study region is of special interest; that the network representing the transportation system need not be modified and that the zones cannot be split during aggregation; and that the conventional demand prediction methods are sufficiently accurate for subarea analysis.

Several means of achieving zonal aggregation were considered and three criteria used to select the procedure that was finally implemented. These were: simplicity, transferability, and reasonable accuracy of trip prediction with respect to the window. The three major steps of the aggregation procedure selected are given below:

Step 1. Construct aggregate district boundaries,
Step 2. Establish district centroids, and
Step 3. Assign terminal and intrazonal travel time for districts.

In step 1, outside zones were aggregated into districts based on a minimum travel time difference criterion. That is, zones that displayed the smallest differences in travel time to a common node on the window boundary were combined first. The process was then continued until the desired number of districts was reached.

In step 2, a district centroid was chosen from the set of zone centroids originally contained in the district. The chosen centroid was centrally located in the district and ranged nearest the median value with regard to travel time to the window boundary.

In step 3, a simple unweighted average approach was used to calculate terminal and intradistrict times for each district.

With the exception of minor hand adjustments, the procedure was fully automated. An example of an areal system obtained by this method is shown in Figure 2. [A more detailed discussion of the process is given by Miller (10).]

PRESENTATION OF RESULTS

Time Savings and Expenditures

The reduction in computer time for trip distribution was calculated for each level of aggregation. This involved three tasks: (a) the creation of a minimum path travel time matrix that included terminal and intrazonal travel times (SKIMTREE PROGRAM), (b) the production of a line printer list of a matrix (T PRINT PROGRAM), and (c) the production of a trip matrix by the gravity model (WILSON PROGRAM). The computer execution times, the gross reductions in computer times, and the percentage reductions for the three programs are shown in Table 1.

The preparation of the data inputs for the aggregation procedure required 2 person-days. Additional costs of

10, 7.5, and 4 person-days were incurred to obtain the district-level data sets for the 55, 79, and 90 percent degrees of aggregation respectively. These manpower requirements were for the existing aggregation procedure, which should be considered a prototype. Significant savings may be realized by further automation. (The manpower requirements included learning time. Thus, it should decrease as more experience is gained.)

The computational costs of the aggregation procedure in terms of computer execution time were minimal and relatively inelastic with respect to aggregation level. The total costs for a single application of the algorithm for either 55, 79, or 90 percent degrees of aggregation were 81, 74, or 73 s respectively. When all three levels of aggregation were desired, the total CPU time was 164 s.

Errors Attributable to Degree of Aggregation

Analysis of the differences between the forecast and aggregate trip matrices was limited to the calculation of seven error measures. Three of these were regional measures. The remaining four measured matrix errors on an interchange-by-interchange basis. In calculating these four measures it was necessary to compress the forecast trip matrix to a size that comformed with each of the aggregate trip matrices. Thus, compression of the single 635-zone trip matrix created a distinct compressed trip matrix for each degree of aggregation investigated. (These will be termed compressed trip matrices.) Analysis of pairs of equally dimensioned compressed and aggregate trip matrices was performed for all matrix cells.

In addition to the comparisons of full matrix pairs, matrix quadrants were compared to isolate errors for trips to, from, and within the subarea of interest. Specifically, the three groups of interchanges of interest were: trips produced at centroids outside the window and attracted to centroids inside the window; trips produced at centroids inside the window and attracted to centroids outside the window; and trips made completely within the window. To facilitate this analysis the aggregate and compressed trip matrices were arranged so that the lowest numbered centroids were those contained within the window. The specially arranged matrices were separated into four quadrants based on the types of trip interchanges as illustrated in Figure 3.

Matrix Error Measures

Pairs of trip matrices of conformable sizes were compared on an interchange-by-interchange basis to provide a microscopic analysis of the differences between them. The error measures, absolute percent deviation and absolute deviation, are defined in equations 1 and 2 below.

$$\gamma_{ij} = |C_{ij} - A_{ij}|/C_{ij} \tag{1}$$

subject to

$C_{ij} > 0$ trips and
$C_{ij} > 3$ trips, except when $|C_{ij} - A_{ij}| > 10$ trips;

and

$$\beta_{ij} = |C_{ij} - A_{ij}| \tag{2}$$

where

C = compressed 635-zone trip matrix (dimensions $r \times r$),
A = aggregate trip matrix (dimensions $r \times r$),
C_{ij} = trips predicted from i to j for matrix C,
A_{ij} = trips predicted from i to j for matrix A,
r = number of total centroids (i.e., 302, 165, or 103),
γ_{ij} = absolute percent error for interchange i, j, and
β_{ij} = absolute deviation for interchange i, j.

The absolute percent error measure was constrained to consider only compressed trip interchanges with the number of trips greater than zero since division by zero is undefined. Compressed trip interchanges having magnitudes of less than three trips were not analyzed unless the absolute deviation was greater than ten trips because many small trips created large absolute percent errors (e.g., 2 trips ± 1 trip gives a 50 percent error). (These large percent errors were insignificant because of the small number of trips involved.)

The mean and standard deviations for the two error measures for each level of aggregation are given in Table 2. The computer program that calculates these measures also stratifies the error by row and column for the full matrix and for particular matrix quadrants. These results are not included here but are available for future, more detailed analyses and refinements. Such analyses might isolate centroids with unusually high error magnitudes and study their properties for further refinement of the aggregation procedure.

The maximum error for a single trip interchange in the full matrix and in each matrix quadrant was determined for the three degrees of aggregation. These errors are shown in Table 3. In all instances the maximum error found in quadrant IV is also the maximum error found in the entire matrix. Since the aggregation criteria were purposely biased against quadrant IV trips, this is an expected result. (Only a subset of quadrant IV trips, namely through trips, are of interest to the subarea planner. These are discussed later.) The number of trips listed for each error measure is the number of trips for the associated compressed trip interchange. For example, the full 302-by-302 matrix had a maximum percent interchange deviation of 304 based on six trips. [These percentages were calculated before trips were rounded off. In terms of whole trips this was (24 - 6)/6 = 300.]

Error Length Distributions

The error length distributions indicate the bias of deviations for trips of different lengths. Two error measures, the mean absolute percent deviation per time interval and the mean absolute deviation per time interval, were used to calculate these distributions. They are given by

$$\lambda_M = \frac{\sum_{i,j \ni t_{ij} \in M} [(|C_{ij} - A_{ij}|)/C_{ij}]}{N} \tag{3}$$

and

$$\rho_M = \frac{\sum_{i,j \ni t_{ij} \in M} |C_{ij} - A_{ij}|}{N^2} \tag{4}$$

where

C_{ij} = trips predicted from i to j for matrix C,
A_{ij} = trips predicted from i to j for matrix A,
M = time interval,
N = number of trip interchanges in time interval M,
t_{ij} = travel time from i to j in the aggregate system,
λ_M = mean absolute percent error for time interval M, and
ρ_M = mean absolute deviation for time interval M.

Figure 1. Types of data base simplifications.

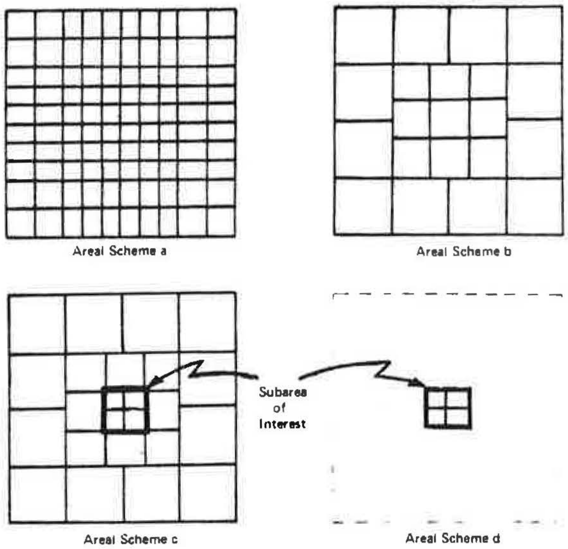


Figure 2. Aggregate traffic analysis districts.

Table 1. Comparison of computer execution times for trip distribution programs.

Degree of Aggregation (%)	No. of Centroids	Skimtree[a] Program	Wilson[b] Program	Skimtree and Wilson Programs	Tprint[c] Program
Computer Execution Times					
0	635	822	360	1182	218
55	302	370	55	425	49
79	165	213	19	232	15
90	103	137	9	146	6
Gross Reduction in Computer Time[d]					
55	302	452	305	757	169
79	165	609	341	950	203
90	103	685	351	1036	212
Percentage Reduction in Computer Time[d]					
55	302	55	85	64	77
79	165	74	95	80	93
90	103	83	98	88	97

Note: All time is measured in CPU sec for CDC 6400. These may be converted to dollar terms by the following conversion unit: 0.0583 (dollars/CPU sec).

[a]See (13).
[b]Doubly constrained gravity model carried through four interactions of balancing procedure.
[c]Print out program that prints 50 × 28 entries per page.
[d]Based on 635 zone system execution time.

Figure 3. Quadrant definition.

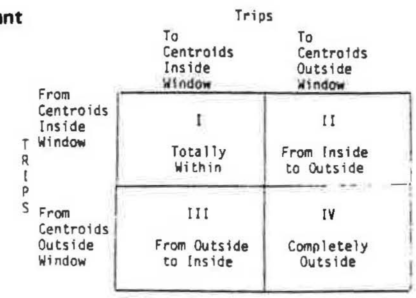


Table 2. Trip matrix error summary.

No. of Centroids	Mode of Analysis	Absolute Percent Error: Mean	Absolute Percent Error: Standard Deviation	Absolute Deviation: Mean	Absolute Deviation: Standard Deviation
302	Full matrix	9.93	11.57	1.94	20.61
	Quadrant I	1.02	0.70	0.80	0.38
	Quadrant II	6.26	6.47	0.15	1.12
	Quadrant III	5.82	6.24	1.00	3.52
	Quadrant IV	11.77	12.66	2.43	23.87
165	Full matrix	12.45	18.17	7.68	60.88
	Quadrant I	1.63	1.01	0.28	0.63
	Quadrant II	9.00	7.84	0.51	2.70
	Quadrant III	8.73	15.16	2.35	5.29
	Quadrant IV	15.77	20.28	12.81	81.22
103	Full matrix	11.84	15.63	21.75	266.09
	Quadrant I	1.90	1.09	0.31	0.60
	Quadrant II	11.91	11.20	1.46	6.77
	Quadrant III	8.95	11.17	5.27	10.44
	Quadrant IV	18.02	19.71	57.23	447.01

Table 3. Maximum trip matrix error magnitudes.

No. of Centroids	Mode of Analysis	Absolute Percent Error: Maximum Error	Absolute Percent Error: Number of Trips	Absolute Deviation: Maximum Error	Absolute Deviation: Number of Trips
302	Full matrix	304	6	2 836	3 573
	Quadrant I	4	11	4	216
	Quadrant II	37	25	57	188
	Quadrant III	57	6	96	553
	Quadrant IV	304	6	2 836	3 573
165	Full matrix	704	3	2 976	1 965
	Quadrant I	5	7	10	288
	Quadrant II	41	5	106	326
	Quadrant III	191	10	76	43
	Quadrant IV	704	3	2 976	1 965
103	Full matrix	214	317	15 195	29 485
	Quadrant I	5	4	7	275
	Quadrant II	62	3	124	607
	Quadrant III	101	11	146	821
	Quadrant IV	214	317	15 195	20 485

Figure 4. Error length distribution results for 302-zone system: Seattle CBD window.

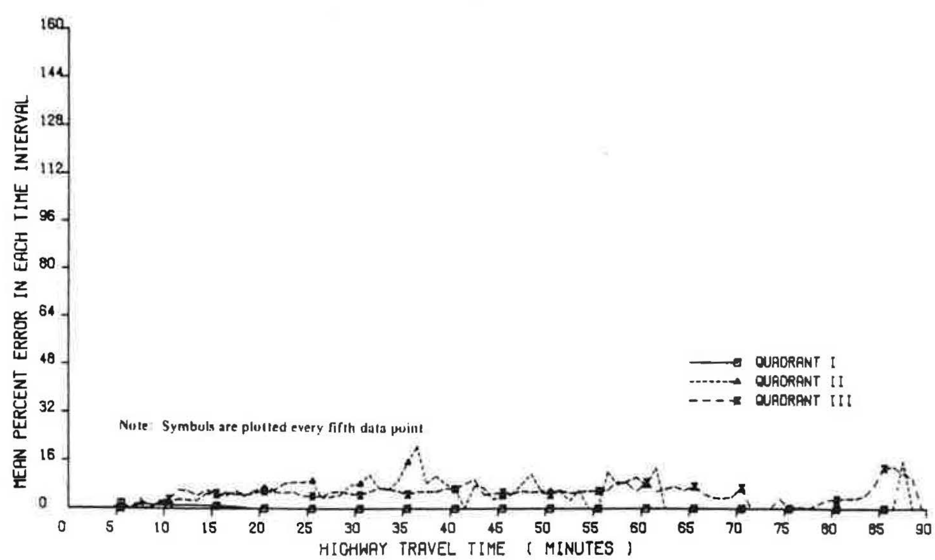


Figure 5. 1990 home-based person work-trip distribution gravity model output: Seattle CBD window.

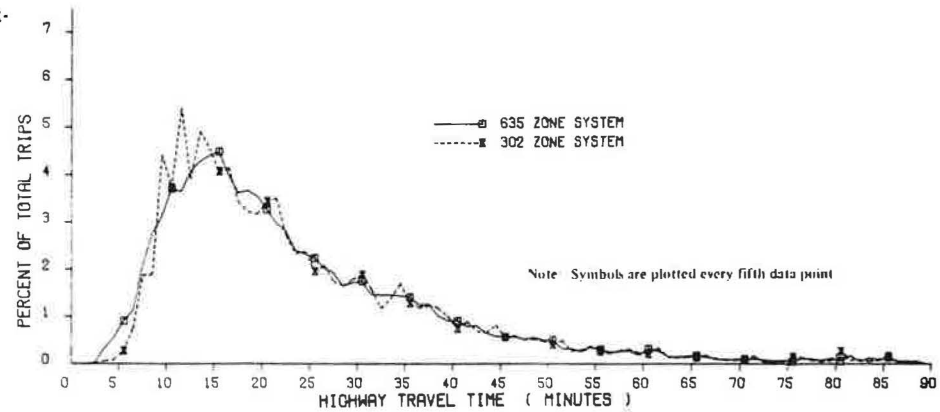
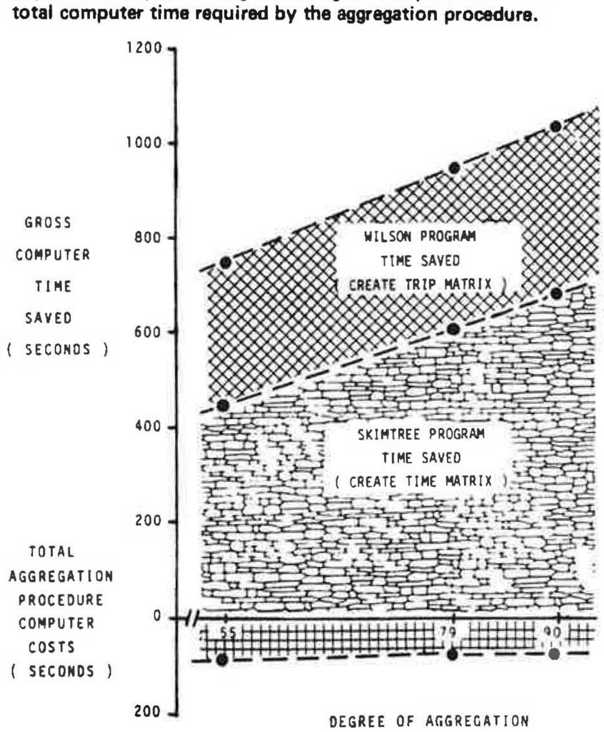


Figure 6. Comparison of gross savings in computer time with total computer time required by the aggregation procedure.



The error length distributions were calculated for the full matrix and all four quadrants for each degree of aggregation. An example of the mean percent error distribution is illustrated in Figure 4. In all instances, large percent errors were associated with small numbers of trips when compared with the absolute deviation distributions.

Trip Length Distributions

Trip length distributions were calculated for the forecast matrix and for each aggregate matrix. Figure 5 illustrates a sample comparison of trip length distributions for the 635 and 302-centroid systems. Since quadrant IV errors are of less relative importance than errors in the other three quadrants the trip length distribution for each quadrant was also calculated. This permitted the isolation of the types of trip interchanges that caused the largest deviations in the trip length distribution curve.

Regional Trip Length Measures

The final measures calculated were the mean and standard deviations of the trip length and person-kilometers traveled in the study region. The mean trip length for the 635-zone base system was 23.98 min with a standard deviation of 14.66 min. The mean trip lengths for the 55, 79, and 90 percent aggregation levels were 24.09, 23.71, and 25.01 min respectively, with corresponding standard deviations of 14.61, 13.84, and 13.59. The total person-hours of travel was 507 790 for the base system and 510 202, 502 193, and 529 598 for the respective aggregation levels.

INTERPRETATION OF RESULTS

The significant computer time savings achieved by subarea focusing were shown in Table 1. The data indicate that it is possible to perform six 103-centroid, four 165-centroid, or two 302-centroid skimtree runs in approximately the same computer time as required for one 635-centroid run. Similarly, either forty 103-centroid, nineteen 165-centroid, or six 302-centroid gravity model runs can be made for approximately the same cost as one 635-centroid run. Therefore, more alternatives can be examined by subarea focusing than by conventional methods for the same costs, at least with regard to computer time.

The costs of preparing the three degrees of aggregate district data appear to be modest in terms of both manpower and computer time. The total manpower expended in the creation of the aggregate district data sets ranged from approximately 6 to 12 person-days. A substantial portion of the manual tasks could be automated in further refinements to the aggregation procedure. The total computer time consumed in preparing the aggregate district data sets was small relative to the computer time saved during trip prediction. This is shown graphically in Figure 6. For example, it required 81 s to run the aggregation programs for the 302-centroid system, but the computer time saved in skimtree and gravity model runs for the same system was 757 s. (TPRINT was not included in this calculation since it was unlikely that a full trip matrix would be printed for each run.)

The loss of accuracy between each of the three aggregate trip matrices and the 635-zone trip matrix appears acceptable. The average percent errors for the quadrants of special interest (e.g., I, II, and III) for the 302, 165, and 103-centroid systems were 6, 9, and 12 respectively. These error magnitudes are within the acceptable range (9) and agree with other results (7). Similarly, the largest average absolute deviations among the key quadrants are one, two, and five trips respectively. The results for trip interchanges with the worst error magnitudes indicate that large percent errors exist but are associated with small numbers of trips. The largest percent error found among the three primary quadrants for all three degrees of aggregation is 190 based on 10 trips in the compressed trip matrix [i.e., (10 - 29)/10 = 190]. The largest absolute deviation found for these quadrants is 146 trips based on 675 trips (i.e., |675 - 821| = 146 trips).

The results of the error length distribution analyses do not indicate any systematic bias with regard to travel time, and the magnitude of errors in these results is not serious. [Serious error is defined here as any error over approximately 15 to 20 percent (7, 9).] Only a single 1-min time interval (70 to 71 min, quadrant II, 103-centroid system) showed a large percent error, but, since this time interval contained only five trips, the total error was insignificant. It appears, therefore, that a balance has been achieved for quadrants I, II, and III and that the simple choice of unweighted averages for calculating district-level terminal and intrazonal times is not overly damaging. Of course, more complicated weighted average computations could lead to incremental improvements.

The results indicate that the trip length curve becomes less dispersed about the mean as the degree of aggregation increases. This is to be expected since fewer trips of very short or very long duration will occur as the number of centroids is reduced. This shift of short and long trips occurs at the 103-centroid system. The large peaks in intermediate length time intervals appear to offset this shift so that average trip length varies only slightly.

The mean trip length and the person-hours of travel for the 302 and 165-centroid systems closely correspond to those measured for the 635-zone system. Even at the highest degree of aggregation (103 centroids) both the mean trip length and person-hours of travel are only 4 percent in error. Since trip distribution models are usually calibrated within a 5 percent error of the origin-destination matrix, these errors are quite small.

CONCLUSIONS

The subarea focusing concept as applied to trip distribution appears to be both economical and reasonably accurate:

1. The computer time cost of the aggregation procedure is very low in comparison to the computer time savings realized for trip prediction.
2. The procedure is not unduly expensive in terms of the manpower required to develop an aggregate data set.
3. Substantially more alternatives may be examined by subarea focusing than by conventional methods for the same cost.
4. At very high degrees of aggregation subarea focusing may be suitable for preliminary screening of subarea alternatives (i.e., sketch planning) even with conventional models.
5. Data base simplification can be limited to zonal information and still achieve significant savings in computer time.
6. Assuming subarea focusing is applied, the additional savings in computer time that might be realized by network aggregation appear to be small.

The last conclusion, the potential of network aggregation given subarea focusing, requires more substantiation. Both the gravity model and the printout process will not be affected by network aggregation. Thus, the only means of varying computer time in the trip distribution phase involves the minimum travel time path program. Consider the example of the highest degree of zonal aggregation. If it is assumed that the network aggregation is 80 percent successful in reducing computer time (the percent reduction realized for zonal aggregation), there will be a saving of approximately 134 CPU s. This saving will, of course, be diminished by the costs of applying the network aggregation technique and the additional uncertainty brought into the demand predictions. But even ignoring these costs, the results still compare poorly with the savings attributable to zonal aggregation alone. If the costs of creating the aggregate district data are included, the computer time saving for one run is 612 CPU s. Since subarea focusing by definition requires zonal aggregation, additional savings due to network aggregation do not appear to merit the probable added costs and uncertainties in the trip predictions.

The error investigation was initially limited to one urban region (the Puget Sound region), one particular window type and location (the centrally located Seattle CBD), one window size (42 zones), three degrees of aggregation (55, 79, and 90 percent reduction in the number of centroids outside the window), and one travel demand model (gravity model). Thus, only preliminary conclusions about accuracy can be made. The initial results indicate that subarea focusing is sufficiently accurate for conventional planning at the first two aggregation levels and for sketch planning at the highest aggregation level. Subsequent analyses of additional windows and through-trips (12) support these preliminary conclusions.

ACKNOWLEDGMENTS

The material in this paper is based on a study financed through a grant from the U.S. Department of Transportation, Urban Mass Transportation Administration.

REFERENCES

1. R. J. Bouchard. Relevance of Planning Techniques to Decision Making. HRB, Special Rept. 143, 1973, pp. 17-19.
2. R. B. Dial, P. S. Loubal, E. L. Luther, and G. J. H. Brown. New Systems Requirements Analysis Program: A Procedure for Long Range Transportation (SKETCH) Planning. U.S. Department of Transportation, Technical Rept., July 1973.
3. Bellevue Area Micro-Model Transportation Analysis. Redford Engineers, Inc., Nov. 1972; Puget Sound Governmental Conference, Seattle.
4. A Transit Plan for the Metropolitan Area, Seattle and King County: Technical Appendix. Puget Sound Governmental Conference, Seattle, May 1972.
5. New Systems Requirements Analysis Program, UMTA, Project Implementation Plan, Work Item 3: Development of Interactive Sketch Planning Techniques. Peat, Marwick, Mitchell and Co., July 28, 1972.
6. W. W. Mann. T.R.I.M.S.: A Procedure for Quick Response Transportation Planning. Alan M. Voorhees and Associates, McLean, Va., Jan. 1975.
7. H. Schleifer, S. L. Zimmerman, and D. S. Gendell. The Community Aggregate Planning Model. TRB, Transportation Research Record 582, Jan. 1975, pp. 14-27.
8. P. H. L. Bovy and G. R. M. Jansen. Spatial Abstraction in Transportation Planning: Problem Analysis and Experimental Design at a Research Project. Delft Univ. of Technology, The Netherlands, 1974.
9. The Effect of Zone Size on Traffic Assignment and Trip Distribution. Wilbur Smith and Associates, 1971; Australia Commonwealth Bureau of Roads.
10. D. G. Miller. Sensitivity of the Gravity Model to Spatial Aggregation: Development of the Subarea Focusing Concept for the Puget Sound Region. Urban Systems Research Center, Univ. of Washington, Research Rept. 75-3, Aug. 1975.
11. Y. Chan, K. G. Follansbee, M. L. Manheim, and J. R. Mumford. Aggregation in Transport Networks: An Application of Hierarchical Structure. Department of Civil Engineering, M.I.T., Cambridge, R68-47, June 1968.
12. N. Nihan and J. Lassabliere. Sensitivity of the Gravity Model to Subarea Focusing: Effects of Alternative Window Sizes and Locations in the Puget Sound Region. Paper presented at ASCE National Convention, Denver, Nov. 3-7, 1975.
13. R. B. Dial. Algorithm 360: Shortest Path Forest, With Topological Ordering. Communications of ACM, Nov. 1969.

Statewide Disaggregate Attitudinal Models for Primary Mode Choice

Stephen M. Howe and Gerald S. Cohen, Planning and Research Bureau, New York State Department of Transportation

Statewide disaggregate models for predicting the choice of primary mode were developed, using attitudinal and situational (demographic and trip-related) data collected in a recent home interview survey in New York (1). The modeling effort had two objectives: to build mode choice models having both predictive and explanatory value, and to assess the relative significance of situational and attitudinal variables in regional or statewide models. Hartgen (2) has found that situational variables possess greater strength in disaggregate models for urban mode split. Recker and Golob (3), however, have used situational, choice-constraint variables to define distinct market segments, and found attitudinal variables to be highly significant in models built for separate segments.

Disaggregate binary logit models for the choice of bus as primary mode were constructed for six area-purpose cases: NYC-work, NYC-nonwork, upstate-work, upstate-nonwork, statewide-work, and statewide-nonwork. Except for the NYC-work and statewide-work cases, the models had significant statistical strength and goodness of fit. The most significant explanatory variables were situational ones, in particular the number of automobiles available to the household and whether the respondent possessed a driver's license. Attitudinal variables were generally less significant than situational ones.

DATA SET AND VARIABLES

A sample of 1000 households selected at random throughout New York State was surveyed. Controls ensured that the sample matched statewide population characteristics (1). The demographic variables surveyed (one respondent per household) included age, sex, employment status (EMPSTA), possession of a driver's license (LICEN), distance to nearest bus stop (DISTOP) and bus fare, household income, number of automobiles owned (AUTOS), number of workers (WORKRS), and household size (HHSIZE). Several automobile availability indexes were also constructed from the basic variables.

The attitudinal variables included comfort, cleanliness, quietness, crowdedness, safety from crime, safety from accidents, reliability, convenience, speed, frequency, and cost. The survey contained two attitudinal responses for each attribute: general importance and bus-specific satisfaction, each measured on a seven-point scale. For each travel attribute, a third attitudinal variable was created from the product of the importance and satisfaction responses. The dependent variable for both work and nonwork models was a [0-1] dummy (1 = bus, 0 = otherwise).

MODELS

Correlation matrices and stepwise order-of-entry tables were prepared from linear regression runs to guide the selection of independent variables. These runs clearly demonstrated the significance of the AUTOS and LICEN variables as factors in mode choice. The automobile availability indexes were generally no stronger than the basic AUTOS variable, and were not used in the logit models. Although LICEN is correlated with AUTOS, the coefficients of both were relatively uniform across different area-purpose contexts and were also stable when other variables were introduced. Below AUTOS and LICEN, there was a second tier of situational variables having less consistent strength. The attitudinal variables formed a third tier that was generally of less significance than the tiers of situational variables. Among the attitudinal variables, the importance variables were stronger than either the satisfaction or the product variables. No logit models were developed for NYC-work or statewide-work as attempts to build models for these cases showed early entry of variables having incorrect signs and failed to adequately explain observed travel behavior.

The program PROLO was used to construct binomial logit models for choice of primary mode, for the area-purpose cases NYC-nonwork (II), upstate-work (III), upstate-nonwork (IV), and statewide-nonwork (VI). For each case, a series of models was built by incorporating the independent variables selected earlier. Two models for each case are presented in Tables 1 to 4: Model A contains only the AUTOS and LICEN variables; other situational and attitudinal variables are added to form model B. Three indicators of model strength are given in these tables. The pseudo-R^2 and the statistic ($-2 \log \lambda$) are both widely used goodness-of-fit measures, but, as pointed out by Stopher (4), appear to be weak in discriminatory power. Thus, a third indicator, the 0.50 classification, which consists of the four entries in a 2-by-2 classification table as shown below, is also given.

Group	$P_t < 0.50$	$P_t > 0.50$
Nonusers	—	—
Users	—	—

A strong model presumably should estimate $P_t \geq 0.50$ for most transit users in the data set, and $P_t < 0.50$ for most

Table 1. New York City-nonwork—bus versus other (sample size = 208).

Model	Independent Variables: Name	Coefficient	t-Statistic	-2 Log λ	Pseudo-R^2	0.50 Classification
IIA	CONST	0.93	1.47[a]	64.5[b]	0.41	142 22 16 28
	AUTOS	-2.20	-5.16			
	LICEN	-0.61	-1.31[a]			
IIB	CONST	-3.96	-1.40[a]	82.7[c]	0.51	151 13 19 25
	AUTOS	-2.26	-5.20			
	LICEN	-0.52	-1.06[d]			
	DISTOP	-0.06	-1.48[a]			
	AGE	0.04	2.81			
	QUIE-IMP	0.43	2.10			
	ACCI-IMP	0.52	1.15[d]			
	COMF-IMP	0.39	-1.65			

[a]Low (0.80) significance. [b]2 d.f. [c]7 d.f. [d]Insignificant.

Table 2. Upstate-work—bus versus other (sample size = 148).

Model	Independent Variables: Name	Coefficient	t-Statistic	-2 Log λ	Pseudo-R^2	0.50 Classification
IIIA	CONST	2.40	2.33	36.7[a]	0.43	125 6 9 8
	AUTOS	-1.70	-3.35			
	LICEN	-1.73	-2.55			
IIIB	CONST	-23.52	-1.68	60.8[b]	0.66	126 5 8 9
	AUTOS	-3.31	-3.53			
	LICEN	0.44	0.42[c]			
	SEX	-2.41	-2.60			
	HHSIZE	1.60	2.67			
	AGE	0.05	1.39[d]			
	INCOME	-0.17	-1.98			
	RELI-IMP	4.58	1.89			
	CONV-IMP	-1.12	-1.43			

[a]2 d.f. [b]8 d.f. [c]Insignificant with incorrect sign. [d]Log (0.80) significance.

Table 3. Upstate-nonwork—bus versus other (sample size = 317).

Model	Independent Variables: Name	Coefficient	t-Statistic	-2 Log λ	Pseudo-R^2	0.50 Classification
IVA	CONST	1.71	2.50	86.4[a]	0.49	263 21 11 22
	AUTOS	-2.32	-5.17			
	LICEN	-1.50	-2.79			
IVB	CONST	1.68	0.94	92.0[b]	0.52	273 11 13 20
	AUTOS	-2.31	-5.11			
	LICEN	-1.52	-2.81			
	FREQ-IMP	0.35	1.44[c]			
	FARE	-0.53	-1.79			

[a]2 d.f. [b]4 d.f. [c]Low (0.80) significance

Table 4. Statewide-nonwork—bus versus other (sample size = 525).

Model	Independent Variables: Name	Coefficient	t-Statistic	-2 Log λ	Pseudo-R^2	0.50 Classification
VIA	CONST	1.24	2.68	158.8[a]	0.46	405 43 27 50
	AUTOS	-2.34	-7.31			
	LICEN	-0.97	-2.76			
VIB	CONST	-2.38	-1.78	169.7[b]	0.49	418 30 37 40
	AUTOS	-2.44	-7.26			
	LICEN	-0.76	-2.06			
	AGE	0.03	2.42			
	FREQ-IMP	0.40	1.98			

[a]2 d.f. [b]4 d.f.

nonusers. In each of the four cases, model A shows high overall significance in the log-likelihood statistic. Both AUTOS and LICEN show a negative relation to the choice of bus, and (except in case II) both AUTOS and LICEN have significant statistics in model A. In general, at least 85 percent of the nonusers and at least half of the observed bus users are correctly classified by using only the first-tier variables AUTOS and LICEN.

The inclusion of additional variables in model B appeared to provide only a marginal improvement in overall model strength. The LICEN variable was considerably affected by the addition of other variables, becoming insignificant in cases II and III, but the coefficient of the AUTOS variable remained significant and stable in all four cases.

A cross-comparison of the effects of the first-tier variables, for NYC versus upstate areas and for work versus nonwork trip purpose, can be made on the basis of the models IIA, IIIA, and IVA. AUTOS has a substantially smaller coefficient and T-statistic in the work-purpose model IIIA than in either of the nonwork models IIA and IVA, implying that automobile availability is of lesser importance in mode choice for work than for nonwork purposes. The coefficient of LICEN is strong in both upstate models (IIIA and IVA), but has low significance in the downstate model (IIA). In New York City, transit is a viable alternative even for licensed drivers.

CONCLUSIONS

These findings suggest that attitudinal variables are less significant than situational variables in predicting primary mode choice. This conclusion is consistent with findings reported in previous research at local and regional levels. However, attitudinal variables may be significant determinants of variation within market segments defined by situational context, as found by Recker and Golob (3).

An analysis of the data showed that automobile users consistently express more satisfaction with the attributes of transit than did those who actually used transit, possibly because many automobile users simply do not know what the local bus system is like. Relationships between mode choice and satisfaction variables were therefore largely spurious and often contradictory.

The principal conclusions are the following:

1. Relatively small disaggregate data sets appear to be quite adequate for the development of useful mode choice models at the statewide level.
2. The independent variables best explaining mode choice were consistently the number of automobiles available to the household and the possession of a driver's license. Cross-comparison and sensitivity analyses suggest, in addition, that AUTOS is of greater importance in mode choice to work, and that LICEN is more important in upstate than in NYC models.
3. Variables have different levels of influence on mode choice in different areas. This suggests that transferability is most appropriate for similar areas and trip purposes. Thus, the NYC model should be appropriate for very large cities that have extensive transit systems and the upstate models for other areas.
4. Attitudinal variables are generally of less significance than situational variables in explaining mode choice at the statewide level.

REFERENCES

1. E. P. Donnelly and others. Statewide Public Opinion Survey on Public Transportation: Technical Report. Planning and Research Bureau, New York State Department of Transportation, Preliminary Research Rept. 80, 1975.
2. D. T. Hartgen. Attitudinal and Situational Variables Influencing Mode Choice: Some Empirical Findings. Transportation 4, 1974, pp. 377-392.
3. W. W. Recker and T. F. Golob. An Attitudinal Modal Choice Model. Transportation and Urban Analysis Department, General Motors Research Laboratories, Warren, Mich., 1975.
4. P. Stopher. Goodness-of-Fit Measures for Probabilistic Travel Models. Transportation 4, 1974, pp. 67-83.

Quick Policy Evaluation With Behavioral Demand Models

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

Many of the important policy issues confronting present day urban planners involve regionwide transportation system changes that will have many effects. Conventional urban transportation planning models do not capture the full range of travel impacts, and are cumbersome and resource consuming for evaluation of these policy options. In response to this, new behavioral travel demand models have been developed; these are policy sensitive and can be generalized among urban areas. However, there are several unresolved questions about these disaggregate demand models that prevent their widespread application. These problems are:

1. Models estimated in one urban area have not been validated on other urban areas to test their generality.
2. Models estimated on small data sets have not been applied to other small data sets to predict regionwide travel behavior.
3. Disaggregate logit models give biased forecasts when applied to sketch plan or district sized zones.
4. Disaggregate demand models estimated on automobile drive alone and transit mode choices will not predict the full range of choices available to trip makers, which may include car pooling, chauffeuring, and walking, in response to a change in system performance.

This paper presents methods that apply disaggregate, probability choice demand models to a sample of sketch plan zones to evaluate various automotive pollution control strategies in the Los Angeles region. Each of the above problems is considered.

THE AGGREGATION PROBLEM

In the logit specification of probability choice models, the probability of an individual choosing any given mode has the following functional form:

$$P(A) = 1 \left[1 + \sum_{\substack{b=1 \\ b \neq a}}^{n} \exp(-Y_{ab}) \right] \quad (1)$$

where a is one among n alternatives and Y_{ab} is the relative costs and attributes between alternatives a and b. Each Y_{ab} represents the log of the ratio of the probability of a to the probability of b. The prediction of travel behavior in a zone of T individuals requires estimates of individual probability choice of a:

$$N_a = \sum_{t=1}^{T} P_t(a) \quad (2)$$

Typically, the only information available about the arguments of the Ys is their means for a zonal interchange and, possibly, the variances and covariances of the terms in Y. There is no analytical form to translate this information into an estimate of N_a.

A Taylor's series approximation of equation 1 evaluated about the zonal means of the data has been suggested to adjust for this problem (1). The expected value of the resulting expression, truncated after the third term, gives the following equation:

$$E[P(a|Y)] = P(a|\bar{Y}) \left\{ 1 + \sum_{\substack{b=1 \\ b \neq a}}^{n} \mathrm{var}[Y_{ab}] * P(b|\bar{Y}) - \tfrac{1}{2} \right\} \quad (3)$$

where

E [ ] = expected value operator and
var [ ] = variance.

The variables with bars over them are means. Equation 3 is somewhat different from those derived by Talvitie. The operational difference is that stochastic independence between the attributes of alternative a and other alternatives is not assumed here but is by Talvitie (1).

The expected value of choices other than $\bar{a}$ is

$$E[P(c|Y)] = P(c|\bar{Y})\left\{1 + \sum_{\substack{b=1 \\ b\neq a}}^{n} \mathrm{var}[Y_{ab}]\,[P(b|\bar{Y}) - \delta]\,[P(b|\bar{Y})] - \tfrac{1}{2}\right\} \quad (4)$$

where $\delta = 1$ if $b = c$ and 0 if $b \neq c$.

Consider the comparison function, Y_{ab}. For the mode choice, the functional form for Y (2) is

$$Y_{ab} = \alpha_0 + \alpha_c(C_a - C_b) + \alpha_T(T_a - T_b) + \alpha_s(S_a - S_b) + \beta 0 \quad (5)$$

where

- C = operating cost of the trip,
- T = waiting and line-haul time of the trip,
- S = walking time for the trip,
- 0 = availability of an automobile, and
- a, β = estimated constants.

There are 28 possible variance-covariance terms for this equation. About half of these can be presumed to be zero because of stochastic independence or constancy over a zone. Of the other half, there is a presumption that most are proportional to, or simple functions of, the variance of the distance traveled in a zone interchange.

With the information available, one cannot be very precise in measuring the variance in distance in zonal interchanges. The approach here is to assume that distances (or origin and destination points) are distributed over the area of a zone pair according to a well-defined probability density function. This approach, ultimately, allows estimates of the variance of distance as a function of the areas of the two zones in a zonal interchange.

In deriving the appropriate density functions, it is presumed that, for a given zone pair, trips are distributed over a range that reflects both the distance between the zone centroids (geographic centers) and the sizes of the zones. In symbols this is

$$D \in [D' - (Y_i + Y_j)/2,\ D' + (Y_i + Y_j)/2] \quad (6)$$

where

- D = a stochastic variable that represents the distance between zones i and j for person trips,
- D' = the distance between the geographic centers of zones i and j, and
- Y_i, Y_j = some measure of the size of zones i and j.

If another stochastic variable, X, which can have values in the range from 0 to $(Y_i + Y_j)$, is now introduced, the distance for any trip can be represented by the sum of two variables:

$$D = \{D' - [(Y_i + Y_j)/2]\} + X \quad (7)$$

The above relation indicates that trips must travel a nonstochastic minimum distance (the term in [ ]) and that the rest of the distance varies randomly between 0 and $Y_i + Y_j$. The variance of D is

$$\mathrm{var}[D] = E[X^2] - (E[X])^2 \quad (8)$$

The distribution function for X is assumed to be

$$f(X) = \{3/[2(Y_i + Y_j)]\} - [X/(Y_i + Y_j)^2] \qquad \text{for } 0 < X < Y_i + Y_j \quad (9)$$

The premise of the density function is that the distribution of trips can be approximated by a linear declining function over the range bounded by $Y_i + Y_j$.

The first two moments about zero of the distribution are

$$E(X) = [5(Y_i + Y_j)]/12 \quad (10)$$

$$E(X^2) = (Y_i + Y_j)^2/4 \quad (11)$$

The variance of distance can then be calculated from equation 9 and the above moments:

$$\mathrm{var}[D] = [11(Y_i + Y_j)]/144 \quad (12)$$

The measures of zone size over which trips are distributed should reflect the length of the zone. Use of an intuitive measure of length, the square root of area, leads to the following:

$$\mathrm{var}[D] = [11(\sqrt{A_i} - \sqrt{A_j})^2]/144 = [11(A_i + 2\sqrt{A_iA_j} + A_j)]/144 \quad \text{for } i \neq j \quad (13)$$

For intrazonal trips, the above equation must be modified to account for the fact that the stochastic part of the range is only half, on an average, of that of interzonal trips:

$$\mathrm{var}[D] = 11A_j/144 \qquad \text{for intrazonal trips} \quad (14)$$

The remaining terms in the Taylor expansion tend toward zero. However, the truncation of the series after the third term opens the possibility that, for values of Y that diverge rather far from $\bar{Y}$, equations 2 and 3 will not provide a measure of probability that increases monotonically with $P(a|Y)$. In symbols, $E[P(a|y)]$ must satisfy the following three conditions:

$$\sum_{a=1}^{n} E[P(a|Y)] = 1 \quad (15)$$

$$0 < E[P(a|Y)] \quad \text{for all a} \quad (16)$$

$$\partial E[P(a|Y)]/[\partial P(a|\bar{Y})] > 0 \quad \text{for all } P(a|\bar{Y}) \in [0, 1] \quad (17)$$

Conditions 15 through 17 ensure that $E[P(a|Y)]$ is a probability measure. Stronger conditions are required if $E[P(a|Y)]$ is to have plausible properties in terms of individual choice behavior. One of these is that the elasticity be greatest at $E[P(a|Y)] = 0.5$, i.e.,

$$\partial^2 E[P(a|Y)]/\partial Y^2 = 0 \quad \text{at } E[P(a|Y)] = 0.5 \quad (18)$$

$$\partial^3 E[P(a|Y)]/\partial Y^3 < 0 \quad \text{at } E[P(a|Y)] = 0.5 \quad (19)$$

Of the above five conditions, 15 and 18 hold for all values of the variances. The following constraints on the variances are sufficient to ensure that the other conditions are met:

$$\sum_{\substack{b=1 \\ b\neq a}}^{n} \mathrm{var}[Y_{ab}] < 16 \quad (20)$$

$$\mathrm{var}[Y_{ab}] < 1 \quad \text{for all } b \neq a \quad (21)$$

The variance and covariance of terms in equations 3 and 4 are computed as proportions of 13 and 14 subject to the above constraints.

Table 1. Estimated versus actual work trips.

Mode	Mode Shares for 172 Zonal Interchanges		Shares for LARTS Region (%)	
	Actual	Estimated	Actual	Estimated*
Automobile driver	1040	960	84	79
Transit	47	54	1	4
Automobile passenger	123	196	12	16
Driver serve-passenger	—	40	—	—
Walk	—	10	—	—

*Excludes driver serve-passenger and walk

THE NEW MODE PROBLEM

If mode choices are constrainted to be automobile drive alone and bus transit, the model will not predict the range of responses caused by a policy that significantly alters system performance. The consideration of new modes requires a heuristic approach that results in the construction of new comparison functions, Y_{ab}.

The new comparison functions are formed by attributing to each new mode a variable cost per mile, a time spent in-vehicle and waiting, and a walking time for person round trips between each i, j zone pair. Each of these trip system performance variables can then be substituted for their transit counterparts in the estimated mode split equations to derive the odds between automobile choice and the given new mode choice, and from equation 1, the probability of choosing any alternative among all modes—automobile alone, transit, car pool, serve passenger, and walk—can be derived.

The values for level of service for the new modes are largely the results of assumptions about extra time penalties involved with car pooling and chauffeuring. Such assumptions are required because of the paucity of data about these alternatives. In particular, for individuals who currently drive alone, virtually nothing is known about the availability and attributes of potential car pools.

TRAVEL DEMAND MODEL WITH LOS ANGELES REGIONAL TRANSPORTATION STUDY DATA

This section compares actual Los Angeles Regional Transportation Study (LARTS) data for 1967 against predictions from the demand model. The data given are the number of person round trips between zone pairs by travelers surveyed in the 1967 household survey (a 1 in 100 sample). The level of aggregation is sketch plan zones defined by LARTS in 1970; there are about 12 traffic analysis zones to each sketch plan zone and 69 sketch plan zones for the analysis area (Los Angeles and Orange Counties—the Los Angeles Air Quality Control Region).

The tests described below attempt to determine whether (a) application of the disaggregate demand model estimated on Pittsburgh data and adjusted for zonal variations and new modes can be generalized to Los Angeles, and (b) a small but representative sample of zonal interchanges can be used to predict regionwide effects.

The approach to applying the model is summarized in the following steps:

1. Odds functions for automobile versus other modes are estimated for each zonal interchange using the zonal averages for system performance;
2. Probabilities of each mode choice for each zonal interchange are calculated from application of equation 1;
3. The mode shares for each zonal interchange are calculated using the probabilities, the calculated variance-covariance terms from the formulas in the previous section, and equations 3 or 4, to adjust for aggregation; and
4. The estimated mode shares are multiplied by total trips in the zonal interchange to derive predicted trips by mode.

A random sample of 172 zonal interchanges was chosen for testing and applying the approach. Because the policies in the study were evaluated by their effects on vehicle-miles traveled (VMT) the model also was tested by placing the most emphasis on actual versus predicted VMT.

Table 1 compares the mode shares predicted from the sample to the work-trip mode shares for the entire LARTS region. The model gives reasonable predictions of mode split. Although the total vehicle trips were underpredicted by 7.69 percent, the VMT (estimated, 15 302; actual, 15 211) were predicted with virtually no error.

CONCLUSION

The results presented in this paper indicate that disaggregate demand models hold promise for quick evaluation of transportation-related policies. Although actual policies are not discussed (3, 4), the models were used to simulate the effects of various pollution control strategies by projecting 1974 base case trip behavior and the changes that would have been caused by gasoline taxes, emissions taxes, parking surcharges, and bus system improvements. The resulting predictions were comparable with other research efforts; for example, the implied elasticity of gasoline in Los Angeles in 1974 was between 0.19 and 0.24, which corresponds with many econometric estimates of short-run gasoline demand.

ACKNOWLEDGMENTS

The research reported in this paper was part of a project performed for the Washington Environmental Research Center, Environmental Protection Agency on Policies for Controlling Automotive Pollution in Los Angeles. The author wishes to thank Eleanor Maxwell for performing the calculations in this paper.

REFERENCES

1. A. Talvitie. Aggregate Travel Demand Analysis With Disaggregate Demand Models. Proc., Transportation Research Forum, Vol. XIV, No. 1, Oct. 1973, pp. 583-603.
2. A Disaggregated Behavior Model of Urban Travel Demand. Charles River Associates, 1972.
3. F. C. Dunbar. Evaluation of the Effectiveness of Pollution Control Strategies on Travel: An Application of Disaggregated Demand Models. Proc., Transportation Research Forum, Vol. XVI, No. 1, 1975.
4. F. C. Dunbar. Policies for Controlling Automotive Air Pollution in Los Angeles. Charles River Associates, in press.

A Sensitivity Evaluation of the Traffic Assignment Process

Vergil G. Stover and Jimmie D. Benson, Texas Transportation Institute, Texas A&M University
Jay Buechler, East Central Wisconsin Regional Planning Commission

The nature of the input and the nature of the output (computer printout that gives an impression of very precise and accurate traffic volumes for each link) lend a very deterministic appearance to the traffic assignment process. This paper reports an investigation of the sensitivity of the assignment results to the inputs from the preceding modeling phases (1). Additionally, analyses of the assignment results produced by different trip matrices provide a means of evaluating the sensitivity of various commonly used measures of assignment accuracy.

METHOD OF STUDY

A better-worse approach was used in developing data for analyzing the sensitivity of the measures of accuracy of traffic assignment results. Four different trip matrices were used to generate different traffic assignments to one network. The existing network for the Tyler, Texas, Urban Transportation Study was used for test and evaluation.

The better-worse gradient hypothesized that the least desirable assignment (i.e., the worst case) would result from a stochastic trip matrix constrained only to total trips. The fully modeled trip matrix developed in the urban transportation study was used as the standard for comparison in the analyses. The four matrices used in the analyses are defined as follows:

Stochastic matrix 1—a stochastic trip matrix constrained only to the total trips for the urban area,

Stochastic matrix 2—a stochastic trip matrix constrained to the total trips as well as to the desired trip length frequency for the urban area,

Stochastic matrix 3—a stochastic trip matrix constrained to the total trips, the desired trip length frequency, and the desired trip ends at each external station for the urban area, and

Existing trip matrix—the fully modeled trip matrix as developed and used in the urban transportation study.

Analyses

Comparison of the three stochastic matrices with the existing fully modeled matrix indicates that these matrices represent significant differences at both the zonal level (i.e., comparison of zonal trip ends) and at the zonal interchange level (i.e., cell by cell comparison of the trip matrices). An indication of the differences observed at the zonal level is shown below.

Matrix	Range of Trip Ends per Zone	Matrix	Range of Trip Ends per Zone
1	2000 to 3000	3	500 to 14 000
2	500 to 4500	Existing	0 to 15 500

The assignment analyses used a variety of common measures such as vehicle miles of travel (VMT), screenlines, cutlines, and travel routes. They also focused on various statistical measures of link differences (i.e., assigned volume minus counted volume) by counted volume group such as mean differences, standard deviation of the differences, percent standard deviation, RMS error, and percent RMS error.

The results of these analyses indicated that the fully modeled existing trip matrix gave consistently superior assignment results. The most important observations from these analyses, however, were that the assignment results from the stochastic matrices were not nearly as different as had been expected in view of the major differences at the zonal and zonal interchange levels reflected by the matrices.

Comparison of these assignments with those from various urban transportation studies indicated that the results obtained using stochastic matrices 2 and 3 were consistently well within the range of accuracy observed in other studies. Of 10 recent studies in Texas, only 3 had a smaller total percent RMS error than the assignment using stochastic matrix 3. Analysis of the percent RMS error by five volume groups (Figure 1) indicated that the stochastic assignments compare favorably at volumes greater than 4000.

Comparison of assigned to counted volumes for 17 cutlines shows that assignment 1 (i.e., the assignment using stochastic matrix 1) generally resulted in over assignment, but that assignments 2 and 3, as well as the existing trip assignment, tend to be underassigned. As shown below, assignment 3 was consistently better than the other stochastic matrix assignments in the percent of cutlines with assigned volumes within a stated percent difference from the ground counts, but not as good as the existing trip assignment.

Stated Percent Difference From Ground Count	Percent of Cutlines by Assignment 1	2	3	Existing Trip
±10	24	29	35	47
±15	41	41	47	71
±20	47	41	53	76
±25	47	47	71	100
±50	76	94	100	100

These results indicate that, while the fully modeled existing matrix gives the best assignment results, the stochastic matrices with the trip length frequency constraint give reasonable assignment results.

Evaluation

Measures of goodness such as percent RMS error, error range, and standard deviation (type II measures) showed the greatest improvement between assignment 3 and the existing trip assignment. VMT, travel routes, cutlines, and screenline (type I measures) all showed the greatest improvement between assignments 1 and 2. Thus, the type I measures appear to be relatively more sensitive to the trip length frequency than do the type II measures. However, the type II measures are more sensitive to the distribution of zonal trip ends. The sensitivity of the type I and type II measures (shown in italics) to the trip length frequency and the distribution of zonal trip ends appears to relate in the following manner:

Increasing Sensitivity to Trip Length Frequency ↑

VMT
Travel Route
Mean Difference
Cutline
Screenline

Error Ranges
Percent RMS Error
Standard Deviation

↓ Increasing Sensitivity to the Distribution of Zonal Trip Ends

This suggests that, as the measures are listed from top to bottom, there is a decreasing tendency to mask matrix inaccuracies. As a measure of the accuracy of an assignment, VMT is the least discriminating of the eight measures analyzed, while percent RMS error is the most discriminating. (Standard deviation probably is most sensitive to the distribution of trip ends, but it is difficult to know a reasonable value of standard deviation for any assignment because it is so dependent on network size.)

Since percent RMS error is calculated in terms of network size, it is the preferred measure of assignment accuracy. However, the single most important conclusion from these analyses is that several measures must be used in combination, with full awareness of the strengths and weaknesses of each.

Interpretation

As in virtually all urban transportation studies, the Tyler zonal structure tends to reflect the geographical distribution of activities in the urban area. This may be illustrated by subdividing the area into four concentric rings: Ring 1 consists of the CBD; rings 2 and 3 comprise the remainder of the developed urban area; and ring 4 contains those zones in the fringe area. As the intensity of activities within a ring (reflected in the trip ends per square mile) declines, the average number of zones per square mile tends to decline in a similar manner.

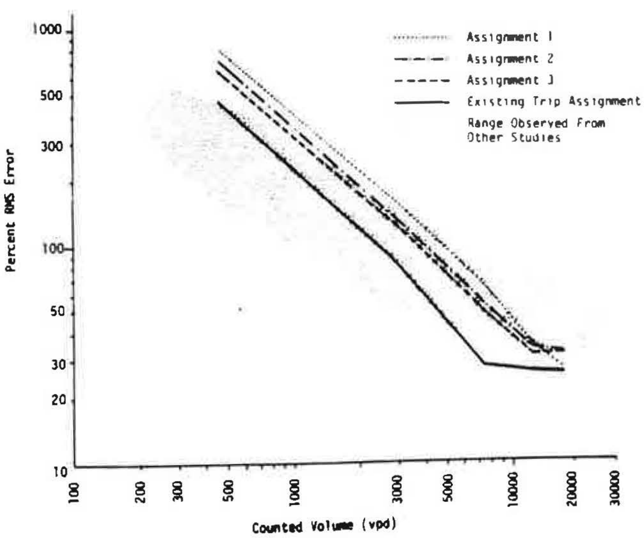
The application of the trip length frequency constraint tends to increase the trip ends in rings 1 and 2 (i.e., the CBD and the inner portion of a developed urban area) where the more intense activities are reflected in the average number of zones per square mile. This simply reflects the disproportionate opportunities to travel at the shorter separations (i.e., 1 to 5 min) within rings 1 and 2, which results from the smaller zone sizes in these rings. The zonal structure imposed on the urban area is, therefore, a major determinant of the trip end distribution resulting from the stochastic matrices. For example, if the zonal structure is redefined such that the CBD consists of only two zones, the resulting trip ends will substantially underestimate the desired trip ends within the CBD. In essence, the zone structure provides a crude tool for a distribution of activities in the urban area.

IMPLICATIONS RELATIVE TO ORIGIN-DESTINATION TRIP TABLES

Previous research, based on a 100 percent home interview survey of three selected zones (2), showed that the estimates of zonal trip ends, based on the expansion of home interview data from that zone, are subject to substantial error. For example, the observed expected error ranges at the 95 percent probability level varied from ±32 to ±66 percent, when using a 5 percent sampling rate for a zone containing 424 occupied dwelling units. Other research (3), using the same data base, demonstrated that estimates of interchange volumes from expanded survey data are subject to even greater variance of estimate than the zonal trip ends.

While expanded origin-destination trip tables are subject to substantial error, in terms of the resulting zonal trip ends and interzonal interchange volumes, these trip tables have generally given reasonable assignment results. This has led practitioners to feel confident of the accuracy of their survey data. In reality, the power of the assignment process masks inaccuracies at the zonal level (i.e., the trip end estimates) and at the zonal interchange level. The assignment results from the stochastic matrices demonstrate

Figure 1. Percent RMS error as a function of counted volume.



the power of the assignment process to overcome and mask most of the data inadequacies that are encountered in an origin-destination trip table.

In spite of inaccuracies, expanded origin-destination trip tables provide good estimates of total trips and trip length frequency for the urban area, and at least a crude estimate of the geographical distribution trip ends and travel patterns. From the perspective provided by the stochastic matrices, it is not surprising that the expanded origin-destination trip tables generally give reasonable assignment results and that mathematical modeling of urban travel patterns gives even better results.

While the comparisons of trip ends and travel patterns indicate that there are significant differences between assignment 3 and the existing trip matrices, the differences in the assignment results, due to their aggregative nature, are not nearly as significant. This suggests that the assignment results are not overly sensitive to the results of the preceding modeling phases (i.e., the trip generation and trip distribution phases). Therefore, a simplified or short-cut trip generation analysis procedure might be used in conjunction with traditional distribution and assignment models for preliminary system evaluations.

The land use patterns could be described by a map reflecting the desired land use categories. These categories should be kept reasonably simple but have sufficient detail to reasonably describe the urban area being studied. In addition, a number of special land use categories to handle unusual situations may be used. With a description of the land use categories in each zone (i.e., the number of acres or of units of each land use category within a zone), a set of vehicle trip generation rates consistent with the land use categories may be applied to determine the zonal productions and attractions. The resulting geographical distribution of trip ends will be adequate input to the subsequent trip distribution and assignment procedures to yield acceptable assigned volumes for preliminary system evaluation.

CONCLUSIONS

Due to the aggregative nature of the assignment procedure, many of the differences observed at the zonal and zonal interchange levels tend to disappear in the assignment results (1, 4). This implies that much of the precision in the preceding modeling phases (i.e., trip generation and trip distribution phases) can be sacrificed while still producing reasonably accurate assignment results. Therefore, abbreviated or sketch planning techniques should produce assignment results of sufficient accuracy for valid evaluation and comparison of system alternatives.

ACKNOWLEDGMENT

The research leading to the development of this paper was sponsored by the Texas State Department of Highways and Public Transportation in cooperation with the U.S. Department of Transportation, Federal Highway Administration, as part of Study 2-10-74-17.

The contents of this paper reflect the view of the authors who are responsible for the facts and the accuracy of the data presented. The contents do not necessarily reflect the official views or policies of the Texas State Department of Highways and Public Transportation or the Federal Highway Administration. The paper does not constitute a standard, specification, or regulation.

REFERENCES

1. J. Buechler, V. G. Stover, and J. D. Benson. A Sensitivity Evaluation of Traffic Assignment. Texas Transportation Institute, Research Rept. 17-2, Nov. 1974.
2. V. G. Stover, J. D. Benson, and L. J. Ringer. Accuracy of Trip End Estimates From the Home Interview Survey. Texas Transportation Institute, Research Rept. 167-7, Aug. 1973.
3. V. G. Stover, J. D. Benson, and D. F. Pearson. Accuracy of Travel Pattern Estimates From the Home Interview Survey. Texas Transportation Institute, Research Rept. 167-8, March 1974.
4. R. E. Foster, J. D. Benson, and V. G. Stover. Consequences of Small Sample Origin-Destination Data Collection in the Transportation Planning Process. Federal Highway Administration, U.S. Department of Transportation, Rept. FHWA-RD-76-43, Jan. 1976.