TRANSPORTATION RESEARCH RECORD 987

Travel Measurement Issues

TRB
TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
WASHINGTON, D.C. 1984

Transportation Research Record 987
Price $12.20
Editor: Scott C. Herman
Compositor: Joan G. Zubal
Layout: Joan G. Zubal

modes
1 highway transportation
2 public transit

subject areas
12 planning
13 forecasting

Transportation Research Board publications are available by ordering directly from TRB. They may also be obtained on a regular basis through organizational or individual affiliation with TRB; affiliates or library subscribers are eligible for substantial discounts. For further information, write to the Transportation Research Board, National Research Council, 2101 Constitution Avenue, N.W., Washington, D.C. 20418.

Printed in the United States of America

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Travel measurement issues.

(Transportation research record; 987)
1. Trip generation–United States–Congresses. 2. Transportation–United States–Planning–Congresses. I. National Research Council (U.S.). Transportation Research Board. II. Series.
TE7.H5 no. 987 380.5s 85-7116 [HE371.A3] [338.3'14'0973]
ISBN 0-309-03771-9 ISSN 0361-1981

Sponsorship of Transportation Research Record 987

GROUP 1–TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Kenneth W. Heathington, University of Tennessee, chairman

Transportation Forecasting Section
George V. Wickstrom, Metropolitan Washington Council of Governments, chairman

Committee on Traveler Behavior and Values
Mary Lynn Tischer, Federal Highway Administration, chairman
Peter M. Allaman, Julian M. Benjamin, Katherine Patricia Burnett, Melvyn Cheslow, David Damm, Michael A. Florian, Thomas F. Golob, David T. Hartgen, David A. Hensher, Joel L. Horowitz, Frank S. Koppelman, Steven Richard Lerman, Jordan J. Louviere, Arnim H. Meyburg, Richard M. Michaels, Alfred J. Neveu, Robert E. Paaswell, Earl R. Ruiter, Yosef Sheffi, Peter R. Stopher, Antti Talvitie

James A. Scott, Transportation Research Board staff

The organizational units, officers, and members are as of December 31, 1983.

Contents

Addresses of Authors

Boyle, Daniel K., Transportation Statistics and Analysis Section, Planning Division, New York State Department of Transportation, State Campus, Albany, N.Y. 12232

Brunso, Joanna M., Transportation Statistics and Analysis Section, Planning Division, New York State Department of Transportation, State Campus, Albany, N.Y. 12232

Chicoine, James E., Transportation Statistics and Analysis Section, Planning Division, New York State Department of Transportation, State Campus, Albany, N.Y. 12232

Clouse, Donald L., Schimpeler-Corradino Associates, 1429 South 3rd, Louisville, Ky. 40208

Greene, David L., Transportation Energy Group, Energy Division, Oak Ridge National Laboratory, Post Office Box X, Oak Ridge, Tenn. 37831

Hartgen, David T., Transportation Statistics and Analysis Section, Planning Division, New York State Department of Transportation, State Campus, Albany, N.Y. 12232

Hu, Patricia S., Transportation Energy Group, Energy Division, Oak Ridge National Laboratory, Post Office Box X, Oak Ridge, Tenn. 37831

Kaltenbach, Kenneth D., Schimpeler-Corradino Associates, 1429 South 3rd, Louisville, Ky. 40208

Kermanshah, Mohammad, Department of Civil Engineering, University of California-Davis, Davis, Calif. 95616

Kitamura, Ryuichi, Department of Civil Engineering, University of California-Davis, Davis, Calif. 95616

Kollo, Hanna P. H., Metropolitan Transportation Commission, Metrocenter, 101 8th Street, Oakland, Calif. 94607

Koppelman, Frank S., Department of Civil Engineering, Northwestern University, Evanston, Ill. 60201

Kostyniuk, Lidia P., Department of Civil Engineering, University of Michigan, 304 West Engineering Building, Ann Arbor, Mich. 48109

Ohstrom, Eric G., Humana Inc., River Front Plaza, Louisville, Ky. 40202

Ohstrom, Joan B., Dial-a-Dietician Board, Commonwealth Building, 680 South 4th Avenue, Louisville, Ky. 40202

Pas, Eric I., Department of Civil Engineering, Duke University, Durham, N.C. 27706

Purvis, Charles L., Metropolitan Transportation Commission, Metrocenter, 101 8th Street, Oakland, Calif. 94607

Roberts, Glenn F., Transportation Energy Group, Energy Division, Oak Ridge National Laboratory, Post Office Box X, Oak Ridge, Tenn. 37831

Stopher, Peter R., Schimpeler-Corradino Associates, c/o Bus Planning, Southern California Rapid Transit District, 425 South Main Street, Los Angeles, Calif. 90013

Supernak, Janusz, Department of Civil Engineering, San Diego State University, San Diego, Calif. 92182; formerly with Drexel University

Life-Cycle Concept: A Practical Application to Transportation Planning

JAMES E. CHICOINE and DANIEL K. BOYLE

ABSTRACT

The usefulness of the family life-cycle concept in trip-generation procedures is examined. A life-cycle classification scheme is constructed after consideration of important components and data availability. The Automatic Interaction Detector program is used to determine which variables are important in affecting the number of trips taken by a household. These variables are then calculated in light of published census tract information. The stages in the classification scheme are designed to be compatible with census categories, thus ensuring the usefulness of the scheme. Trip-generation tables based on stage in the life cycle and vehicle ownership are developed by using data from the 1973 Niagara Frontier Transportation Committee home-interview survey. These tables are compared with trip-generation tables based on household size and vehicle ownership. Analysis of variance is used to compare the life-cycle-based scheme and the household-size-based scheme. The applicability and replicability of the life-cycle-based trip-generation tables are also tested by using data from the 1974 Rochester, New York, home-interview survey. Results indicate that the life-cycle-based trip-generation procedure produces accurate results and has several advantages over other procedures. An example of an application at the town level in Albany County is briefly described.

One of the most profound recent changes in American society has been the rapid evolution of alternative living styles and family types. The proportion of single-head and single-person households has nearly doubled in the past decade alone, and the average size of the family has fallen sharply. These trends, well established in the literature of demographics and confirmed in the 1980 census, are likely to have widespread and far-reaching effects on family activity patterns and travel, and therefore it is incumbent on transportation planners to quantify and understand them.

In this paper the usefulness of the family life-cycle concept in the trip-generation phase of transportation planning is evaluated. The concept of life cycle as used in this paper refers to household structure or composition. Different structures are reflected in life-cycle stages, and a household passes through various stages as it evolves. Although not all households take the same path through these various stages, the concept has the ability to take into account structural changes in families and households more accurately than traditional variables (i.e., number of persons in a household, income), and this ability could possibly lead to better trip-generation models. Many researchers have examined the usefulness of the family life-cycle concept and have generally found it to be an important factor in explaining travel behavior (1-9). However, recent papers have cast doubts on its usefulness (10-12), and the issue deserves further examination.

The practical applications of the life-cycle concept to trip-generation procedures are stressed in this paper. The primary purpose here is to demonstrate that a useful life-cycle classification scheme can be developed and applied in trip-generation tables, where only readily available tract-level census data are required as input. A streamlined life-cycle classification scheme using readily available data is desirable for its practicality and usefulness. Because of the wide availability of published census information, development of a classification scheme is focused on the identification of stages that are compatible with census household categories. In this way trip-generation tables based on these life-cycle stages are easy to use, because of the ready availability of published tract-level census data.

Rather than establish stages of a life-cycle classification scheme based on a priori notions, the data in this paper rely on a computerized explanatory data analysis program known as the Automatic Interaction Detector (AID) to determine which life-cycle variables influence the number of household trips and how these variables should be arranged in a classification scheme. An examination of AID results can indicate which variables are important in explaining the variation of the dependent variable, and thus can provide insight into which variables should be considered as components of a life-cycle classification scheme. Once these ideal components of a classification scheme are identified, they are evaluated in light of available census tract information.

Data from the 1973 Niagara Frontier Transportation Committee (NFTC) home-interview survey in the Buffalo, New York, region are used in developing the life-cycle classification scheme and the trip-generation tables. Trip rates are developed for home-based work, home-based nonwork, non-home-based, and total trips; the primary focus of this paper is on total trips. The 1974 Genesee Transportation Council (GTC) home-interview survey in the Rochester, New York, region is used as a check on the life-cycle classification and trip rates developed from the NFTC data. Although use of GTC data is not a final test of replicability of the results, it provides a preliminary screening process to help judge the accuracy of the life-cycle-based procedure. The trip-generation tables based on life-cycle classification are tested for significance by using analysis of variance (ANOVA). Significance levels are then compared with those of trip-generation tables based on household size.

It should be noted that cross-classification tables based on income and automobile ownership are currently in favor for use in trip generation (13). Although automobile ownership is considered in the trip-generation tables (as described later in the paper), two problems preclude consideration of income here. The first is that trip-generation tables based on income require constant updating to account for inflation. The consumer price index is often used for this purpose, but an index more sensitive to changes in transportation costs may be more appropriate. The second problem is that this paper is based on data gathered in home-interview surveys, which have high nonresponse rates for income questions (more than 35 percent in both surveys used here). Consequently, no comparisons of results from life-cycle-based and income-based classifications are possible.

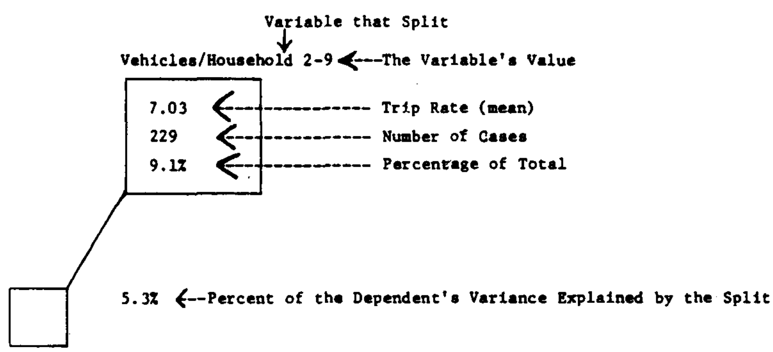
AID AND IDEAL COMPONENTS

As mentioned previously, there has been a considerable amount of research addressing the family life-cycle concept, and most researchers have found it to be an important factor in explaining travel behavior (1-9). A consensus has not yet emerged concerning the components of a family life-cycle classification scheme. In this paper potential components of a classification scheme are examined along with other demographic variables by using the AID program. AID is a sequential search procedure that divides the data set into subgroups through a number of binary splits based on the ability of the independent variables to account for the variation of a dependent variable (14). From the series of binary splits, a "tree" with various branches can be developed. In contrast to statistical methods such as multiple regression, the use of AID does not require assumptions concerning such factors as linearity.

The 1973 NFTC (Buffalo) and 1974 GTC (Rochester) travel surveys were used in the AID analysis. The analysis was done at the household level, and four dependent variables were used: total number of trips, home-based work trips, home-based nonwork trips, and non-home-based trips. Independent variables include all demographic and structural variables available or readily synthesized from the existing data. Figure 1 shows how to read an AID tree and also lists the independent variables.

Figure 2 shows the AID tree for overall trips in the NFTC region. The box in the far left is the starting point (level 0) for the AID analysis; it contains all 1,963 households that average 7.9 trips per day. The first splitting variable is vehicle ownership. The top box on level 1 represents multiple-vehicle households, and these 774 households average 11.56 trips per day. In the bottom box on level 1 are the 1,189 households with zero or one vehicle; they average 5.56 trips per day. This partitioning of the data set into two groups according to level of vehicle ownership accounts for 17.5 percent of the total variation in household trips. An additional 1 percent is accounted for by splitting the multivehicle households into two groups based on occupation of the household head. The coefficient of determination (R^2) for the entire tree is 0.401. The uppermost box in the right-hand side contains eight white collar multivehicle households with six or seven children; these households average nearly 29 daily trips. The lowest box in the tree contains 310 households with no vehicle; these households average fewer than two daily trips.

Interpreting an AID tree is more an art than a science. It certainly appears that vehicle ownership has a strong effect on travel behavior. Household size, vehicle availability, and age of oldest child each accounts for at least 2 percent of the total variation in household trips. Occupation and number of children appear less important. A complete set of



INDEPENDENT VARIABLES USED IN AID:

Age of household head	free	10 year groupings (under 25, 35-34, etc., to over 65
Age of oldest child	free	None, 1-5, 6-10, 11-15, 16-20, 21+
Employment status of head	free	Various
Presence of spouse	free	Present, not
Employment status of spouse	free	various
Number of children	Monotone	Actual number
Number in household	Monotone	Actual number
Occupation of head	free	various
Number of vehicles	Monotone	Actual number
Income	Monotone	Various groupings in data set
Vehicles per licensed driver	Monotone	Actual number
Education	Free	various
Race	free	White, black, other
Presence of relatives (other than spouse or child)	free	Present, not
Presence of non-relatives	free	Present, not
Location	free	Urban, suburban, rural

Note: Free variables may break in any fashion; monotone variables are ordered and must break following that order (i.e., a split of two children and 0-1 or ≥ 2 children is not possible). See report by Ugolik and McDesmott (14) for more details.

FIGURE 1 Directory to AID trees.

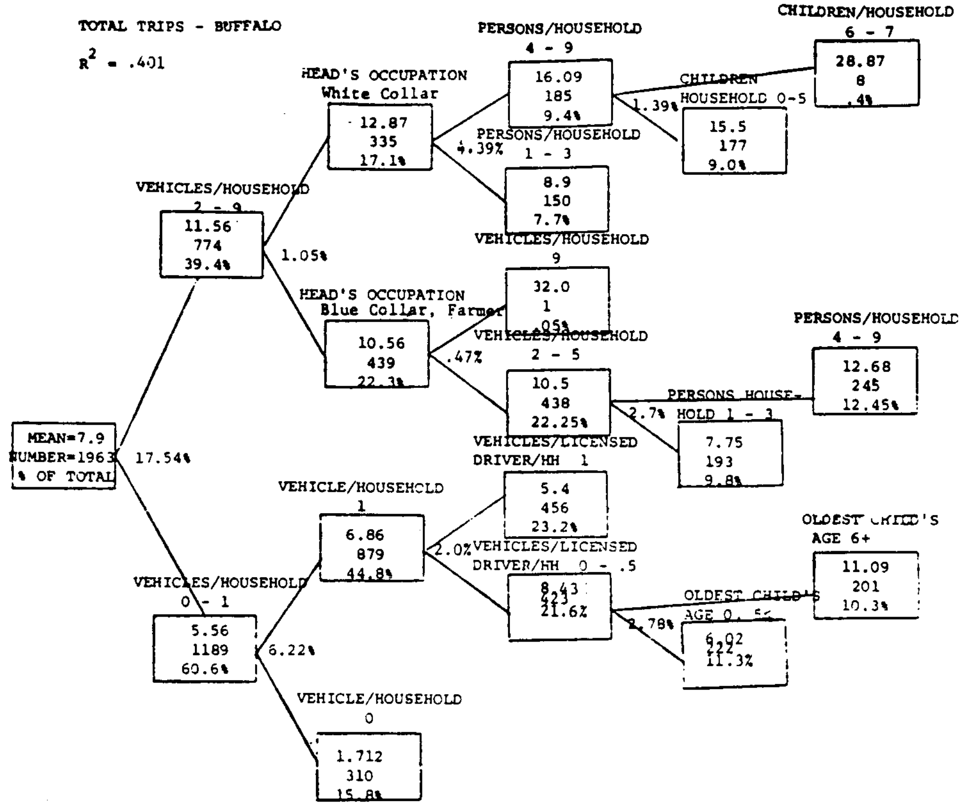


FIGURE 2 AID tree.

AID trees for both regions and for all trip purposes is contained in a report by Boyle and Chicoine (15); the results in terms of important variables are summarized in Table 1. Vehicle ownership, household size, and presence and age of children emerge from the AID analysis as important factors that affect the number of household trips. The importance of vehicle ownership indicates that it should be taken into account in developing trip-generation tables. Consequently, these will be cross-classification tables based on (a) stage in the life cycle and vehicle ownership and (b) household size and vehicle ownership. In terms of ideal components of a family life-cycle classification scheme, consideration should be given to the presence and ages of children.

CENSUS DATA AND A FAMILY LIFE-CYCLE CLASSIFICATION SCHEME

A major purpose of this paper is to develop a classification scheme using as input published tract-level census data. The availability of such data ensures the widest possible use of the scheme in trip-generation procedures. Thus published 1980 census information was examined (16) and appropriate household categories sought for use in constructing a family life-cycle classification. With the AID findings in mind, a breakdown of households by presence and ages of children was particularly sought, without particular success. Several alternate classification schemes were drawn up; details may be

TABLE 1 Important Variables by Trip Type

Trip Type	Variable	Categories Indicated by AID Analysis
All trips	Number of vehicles	0, 1, ≥ 2
	Vehicles per licensed driver	0-0.5, ≥ 1; or 0, ≥0.3
	Number of persons	0-3, ≥ 4
	Number of children	0-1, ≥ 2
	Age of oldest child	None or 1-5, ≥ 6
Home-based work trips	Employment status of spouse	Full-time or part-time, not employed, or no spouse
	Employment status of head	Full-time or part-time, not employed
	Age of oldest child	None or 1-20, ≥ 21; or none or 1-15, ≥ 16
	Number of persons	1-2, ≥ 3; or 1, ≥ 2
Home-based nonwork trips	Number of persons	1-3, ≥ 4
	Age of oldest child	11-20, none or 1-10 or ≥ 21; or none or 1-5, ≥ 6
	Vehicles per licensed driver	0, ≥0.3
	Location	Urban-rural, suburban; or urban, suburban–outer ring
Non-home-based trips	Number of vehicles	0, 1, ≥ 2
	Employment status of head	Full-time, part-time, or unemployed
	Employment status of spouse	Full-time, part-time, or unemployed
	Age of head	17-54, ≥ 55; or 17-44, ≥ 45
	Occupation of head	Categories unclear

found elsewhere (15). After some consideration, the following census-compatible family life-cycle classification scheme was selected:

1. Singe-person households,
2. Households of unrelated persons without children,
3. Families with children younger than 18 years old, and
4. Families without children or families with the youngest child older than 18 years old.

Census information did not include the age of the oldest child, and so the presence of children is a major component of this life-cycle classification scheme. It should be noted that this classification does not differentiate between single-parent and two-parent households. AID results indicated that the presence of a spouse is not a significant element in determining the number of household trips.

DEVELOPMENT AND ANALYSIS OF TRIP-GENERATION TABLES

Trip-generation tables are prepared by using data from the NFTC survey. Two sets of tables are developed: the first is based on stage in the life cycle and vehicle ownership, and the second is based on household size and vehicle ownership. Mean trip rate, standard deviation, and number of observations are presented in each cell. The trip-generation tables for overall trips in the NFTC region are given in Table 2. Detailed tables by trip purpose may be found elsewhere (15). Table 4, discussed later in this paper, gives trip rates for all trip purposes and for both classification schemes in both NFTC and GTC regions.

With number of household trips as the dependent variable, a two-way ANOVA was run by using vehicle ownership and either the life-cycle or the household-size classification as the two independent variables. Because examination of the data indicated unequal variances, the Welch statistic was used to determine F-values and tail probabilities. The Welch statistic was chosen because it is approximately distributed as an F-statistic and does not assume equality of variances (17,18). Although tail possibilities are directly comparable, F-values are not because their level of significance is based on the degrees of freedom. Therefore, F-values resulting from the ANOVAs are examined in general terms.

The data in Table 2 also give the results of the ANOVAs. For overall household trips, the F-values are comparable, although slightly higher for the family-size-based classification. Standard deviations are similar for both schemes. These findings also apply to other trip types [see Table 4 and the report by Ugolik and McDermott (15)]. There is no indication that a significant improvement is obtained by use of one scheme instead of the other. Thus either classification scheme may be considered valid as an analytical tool for use in examining differences in travel behavior.

APPLICATION OF TRIP RATES TO GTC REGION

Another method of comparing the two classification

TABLE 2 Trip Rates, All Trips, NFTC

Life Cycle Stage		VEHICLE OWNERSHIP 0	1	2	3+	Household Size	VEHICLE OWNERSHIP 0	1	2	3
1	$\bar{x}$	0.9	3.3	3.3	3.0	1	0.9	3.3	3.3	3.0
	s	1.4	2.5	2.5	—		1.4	2.5	2.5	—
	n	144	105	6	1		144	105	6	1
2	$\bar{x}$	2.2	5.2	8.1	18.6	2	1.4	4.9	6.9	8.2
	s	2.5	5.1	5.7	10.6		2.1	4.2	3.7	5.4
	n	17	16	16	5		82	315	136	13
3	$\bar{x}$	3.6	9.4	13.0	16.0	3	3.3	7.4	9.1	10.6
	s	4.1	6.9	7.8	7.8		2.6	5.4	5.4	5.5
	n	69	396	359	81		35	151	144	43
4	$\bar{x}$	1.4	5.2	8.3	10.8	4+	3.4	9.8	13.8	15.4
	s	2.1	4.1	5.1	5.3		4.6	7.1	7.8	7.5
	n	79	362	228	78		48	308	323	108

Welch Statistic:			Welch Statistic:		
	F-Value	137.46		F-Value	146.65
	Tail Probability	0.0000		Tail Probability	0.0000
	Degrees of Freedom	14,105		Degrees of Freedom	14,182

$\bar{x}$ = mean trip rate
s = the standard deviation of the mean rate
n = number of households

Life Cycle Stages
(1) Single person households
(2) Households of unrelated persons without children
(3) Families with children under 16 years old
(4) Families with no children or with youngest child at least 18 years old

schemes is to apply each to a different data set and compare the results. The data set from the GTC home-interview survey was used to do this. The number of GTC households in each cell is given in Table 3. Trip rates from Table 2 were applied to these household distributions, and the resulting numbers of trips in all cells were summed to obtain total calculated GTC trips for each classification scheme. These are compared with the actual number of GTC trips:

1. Actual number of GTC trips (sample only) = 18,920;
2. Calculated number of GTC trips using life-cycle-based method = 18,739; and
3. Calculated number of GTC trips using household-size-based method = 18,246.

Although one application certainly is not conclusive, it is interesting that use of the life-cycle-based trip table produced a total number of trips within 1 percent of the actual number, whereas use of the household-size-based trip table produced a total number of trips 3.5 percent less than the actual number.

TABLE 3 Distribution of GTC Households by Cell

	Vehicle Ownership			
	0	1	2	>3
Stage in life cycle				
1	281	259	18	2
2	32	36	38	12
3	126	365	474	119
4	81	341	245	84
Household size				
1	281	259	18	2
2	118	347	194	23
3	46	123	170	48
>4	75	272	393	144

Note: Life-cycle stages are 1 = single-person households, 2 = households of unrelated persons without children, 3 = families with children younger than 18 years old, and 4 = families with no children or with youngest child at least 18 years old. Total number of GTC households = 2,513.

COMPARISON OF NFTC AND GTC TRIP RATES

The final test of the life-cycle-based trip-generation tables also concerns their applicability to other areas. If the trip rates could be applied to several different data sets where the actual number of trips is known, this would indicate whether use of these trip rates produced consistently accurate results. Because only one other data set is used here, variations in household distribution among cells may mask differences in trip rates. A better way of testing the accuracy of the trip-generation tables is to derive a set of tables from the GTC data and compare the trip rates in each cell between the two regions. This can serve as a preliminary test of whether the life-cycle-based trip rates are replicable. For this test, all trip types are considered [see report by Ugolik and McDermott (15) for detailed data].

The data in Table 4 present the trip-generation tables by classification scheme, by region, and by type of trip. For the life-cycle-based tables, trip rates in each cell were examined for differences between the two regions. Those cells with greater than a 10 percent difference were tested to determine whether the difference was statistically significant. Only 6 cells (out of 52) were found to have trip rates different at a significance level of 0.05 in the two regions:

1. Total trips, stage 1 (single person), no vehicle;
2. Total trips, stage 4 (families without children), no vehicle;
3. Home-based nonwork trips, stage 1 (single person), one vehicle;
4. Home-based work trips, stage 3 (families with children), no vehicle;
5. Home-based work trips, stage 4 (families without children), no vehicle (trip rates also different at a significance level of 0.01); and
6. Non-home-based trips, stage 4 (families without children), three or more vehicles.

The NFTC trip rates are generally replicable using GTC data. Although the results cannot be used to proclaim the replicability of the life-cycle-based trip rates, these preliminary indications are promising.

Related to the concerns of accuracy and replicability is the issue of the stability of trip rates over time. One study of differences between the results of home-interview surveys conducted in the NFTC region (1962 and 1973), and the GTC region (1963 and 1974) indicates that trip rates tend to be stable over time, at least for an 11-year period (19). The question of the stability of trip rates in the post-energy-crises era remains to be answered.

TRAVEL PROJECTIONS: ALBANY COUNTY

An interesting application of the life-cycle-based trip-generation procedure was carried out by using town-level data in Albany County. Projections of 1990 town households, broken down by life-cycle stage, were made by using 1970 and 1980 data and previous New York State Department of Transportation forecasts (20). The life-cycle-based trip-generation procedure was then used to forecast the number of trips generated in 1990 in each town under two scenarios. The first scenario held the number of households in each town constant at the 1980 level, thus measuring solely the effects of changes in household structure. The second scenario allowed the number of households to grow to the levels forecast for each town, thus measuring the actual number of trips expected in 1990. Results indicate that the number of trips shows an 11 percent increase in 1990 over 1980, with a 13 percent growth in number of households. When the number of households is held constant, changes in household structure produce a 2.3 percent decrease in the number of trips in 1990 compared with 1980. These results suggest that, if present trends continue, changes in household structure will dampen the increase in travel expected with an increase in number of households.

CONCLUSIONS

The concept of family life cycle has been used to construct trip-generation tables based on a life-cycle classification scheme developed in this paper. The stages in the classification scheme are developed in such a way as to require only published tract-level census data as input. Important components of a life-cycle classification scheme were not assumed a priori, but were determined through use of the AID program. Results from AID were evaluated in light of available census information, leading to a scheme in which the presence of children is emphasized more than ages of children. By designing life-cycle stages to be compatible with census categories, the practical usefulness of these life-cycle-based trip-generation tables has been ensured.

TABLE 4 Trip Rates

	LIFE CYCLE CLASSIFICATION									FAMILY SIZE CLASSIFICATION								
	BUFFALO (1973)				ROCHESTER (1974)					BUFFALO (1973)				ROCHESTER (1974)				
	VEHICLES/HOUSEHOLD									VEHICLES/HOUSEHOLD								
LIFE CYCLE	0	1	2	3+	0	1	2	3+	FAMILY SIZE	0	1	2	3+	0	1	2	3+	
TOTAL TRIPS																		
1	.9	3.3	-	-	1.1	3.2	5.3*	-	1	.9	3.3	-	-	1.1	3.2	5.3*	-	
2	2.2*	5.2*	8.1*	-	3.0	7.6	9.1	14.8*	2	1.4	4.9	6.9	8.2*	2.0	5.5	6.8	8.9*	
3	3.6	9.4	13.0	16.0	2.8	9.2	13.0	16.5	3	3.3	7.4	9.1	10.6	3.0	7.2	9.6	10.6	
4	1.4	5.2	8.3	10.8	2.2	5.6	7.6	11.6	4+	3.4	9.8	13.8	15.4	3.4	10.1	13.8	16.7	
ANOVA WELCH STATISTIC																		
F-VALUE	137.46				144.97					146.65				156.63				
TAIL-PROBABILITY	0.0000				0.0000					0.0000				0.0000				
HOME-BASED NON-WORK TRIPS																		
1	.5	1.6	-	-	.6	1.3	2.8*	-	1	.5	1.6	-	-	.6	1.3	2.8*	-	
2	1.1*	2.1*	3.7*	-	1.5	3.7	4.7	6.4*	2	.9	2.8	2.9	3.8*	1.0	2.9	3.0	3.2*	
3	1.9	5.7	7.9	9.2	1.8	5.6	7.9	10.1	3	1.8	3.9	5.2	5.3	1.6	3.9	4.9	4.8	
4	.9	2.9	4.0	5.2	1.1	3.0	3.6	4.6	4+	1.6	6.1	8.3	8.6	2.4	6.3	8.5	[illegible]	
ANOVA WELCH STATISTIC																		
F-VALUE	90.05				96.33					95.12				101.95				
TAIL-PROBABILITY	0.0000				0.0000					0.0000				0.0000				
HOME-BASED WORK TRIPS																		
1	.3	.8	-	-	.4	.9	.8*	-	1	.3	.8	-	-	.4	.9	.8*	-	
2	.7*	1.9*	2.0*	-	.9	2.3	2.2	4.2*	2	.4	1.1	2.2	2.8*	.8	1.4	2.1	2.7*	
3	1.2	1.8	2.4	3.7	.8	1.9	2.5	3.3	3	1.4	1.8	2.0	2.7	1.1	1.7	2.6	3.3	
4	.5	1.3	2.4	3.3	.9	1.4	2.4	3.8	4+	1.1	1.9	2.6	3.9	.8	2.0	2.6	3.8	
ANOVA WELCH STATISTIC																		
F-VALUE	72.35				73.08					80.47				75.21				
TAIL-PROBABILITY	0.0000				0.0000					0.0000				0.0000				
NON-HOME BASED TRIPS																		
1	.2	.9	-	-	.2	1.0	1.8*	-	1	.2	.9	-	-	.2	1.0	1.8*	-	
2	.4*	1.4*	2.5*	-	.5	1.8	2.2	4.3*	2	.2	1.1	1.8	1.7*	.2	1.2	1.7	3.1*	
3	.6	2.0	2.7	3.3	.2	1.7	2.7	3.2	3	.1	1.8	2.0	2.7	.4	1.6	2.1	2.8	
4	.1	1.0	2.0	2.3	.2	1.2	1.7	3.3	4+	.8	1.9	2.9	3.0	.2	1.8	2.7	3.5	
ANOVA WELCH STATISTIC																		
F-VALUE	37.90				37.02					39.60				36.33				
TAIL-PROBABILITY	0.0000				0.0000					0.0000				0.0000				

- = Empty cells or cell count very small (under 10)

* = Cell Size ≤ 30

Life Cycle Stages

(1) Single person households
(2) Non-related person households without children
(3) Families with children under 18 years old
(4) Families with no children or families with youngest child over 18 years old

The explanatory power, accuracy, and replicability of the life-cycle-based trip-generation tables were tested by various means. ANOVA showed that the life-cycle-based scheme is comparable in terms of F-values to a scheme based on household size (with vehicle ownership being a second independent variable for both schemes). When applied to data from the GTC region, the life-cycle-based trip-generation table produced a more accurate number of total trips than did the household-size-based trip-generation table. Life-cycle-based trip rates were also shown to be replicable using GTC data.

The advantage of a life-cycle-based trip-generation procedure over regression models lies in its simplicity and its ability to handle non-numeric values. It is preferable to a procedure based on family size because it explicitly addresses family structure and thus takes intrahousehold interactions into account. Finally, a life-cycle-based procedure uses readily available data; an income-based procedure is vulnerable to high nonresponse rates if a noncensus data source is used, and such a scheme must be constantly adjusted to account for the effects of inflation.

It is anticipated that critiques of this paper will focus on the difficulty in forecasting household structure, the usefulness of the census tract as the basic areal unit for travel analysis, and the justification for changing established trip-generation procedures. Each of these points deserves to be addressed. First, the question of the pattern of family structure in the future needs further investigation and cooperation with demographers and sociologists so that accurate means to forecast household structure can be developed or put into more widespread use. Related to this, the sensitivity of the life-cycle-based procedure to the projections of future household and family structure needs to be investigated. Second, as noted previously, use of the census tract as the basic areal unit of analysis ensures the availability of the necessary data.

Finally, although it has been demonstrated in this paper that use of the family life-cycle concept in trip generation is practical and produces accurate results, the main justification for this procedure is based on theoretical considerations. The premise behind this investigation is that the family life-cycle concept holds the potential to improve the trip-generation process by increasing its sensitivity to household structure. Consequently, this analytical tool should improve the ability of the transportation analyst to account directly for underlying factors that influence travel behavior.

ACKNOWLEDGMENTS

The authors would like to thank David Hartgen, Alfred Neveu, and Wayne Ugolik for their invaluable assistance and guidance throughout the development of this paper. Thanks are also due to Catherine E. Meyers, who performed the trip-generation analysis for Albany County and authored the microcomputer program. Special thanks are due to Tedi Toca for preparing this manuscript.

At the time when this report was originally prepared, James E. Chicoine was working for the New York State Department of Transportation on a fellowship funded by the Indiana University Foundation.

REFERENCES

1. P.M. Allaman, T.J. Tardiff, and F.C. Dunbar. New Approaches to Understanding Travel Behavior. NCHRP Report 250. TRB, National Research Council, Washington, D.C., Sept. 1982, 142 pp.
2. F.S. Chapin. Human Activity Patterns in the City. Wiley, New York, 1974.
3. P.M. Jones, M.C. Dix, M.J. Clarke, and J.G. Heggie. Understanding Travel Behavior. Gower Publishing Co., Aldershot, England, 1983.
4. E. Ampt. Exploring Choices. Presented at 1982 Australian Transport Research Forum, Hobart, Australia, March 1982.
5. C. Zimmerman. A Life Cycle of Travel by the American Family. Report 7, 1977 Nationwide Personal Transportation Study. FHWA, U.S. Department of Transportation, July 1981.
6. C. Zimmerman. The Life Cycle Concept as a Tool for Travel Research. Transportation, Vol. 11, No. 1, March 1982.
7. R.H. Knapp. Life Cycle Stages and National Travel Surveys. In Recent Advances in Travel Demand Analysis (S.M. Carpenter and P.M. Jones, eds.), Gower Publishing Co., Aldershot, England, 1983.
8. L.P. Kostyniuk and R. Kitamura. Life Cycle and Household Time-Space Paths: Empirical Investigation. In Transportation Research Record 879, TRB, National Research Council, Washington, D.C., 1982, pp. 28-37.
9. D. Damm. Theory and Empirical Results: A Comparison of Recent Activity-Based Research. In Recent Advances in Travel Demand Analysis (S.M. Carpenter and P.M. Jones, eds.), Gower Publishing Co., Aldershot, England, 1983.
10. Schimpeler-Corradino Associates. Oahu Model Update Study. Draft Final Report, Volume I. Oahu Metropolitan Planning Organization, Honolulu, Hawaii, April 1982.
11. K.G. McDonald and P.R. Stopher. Some Contrary Indications for the Use of Household Structure in Trip-Generation Analysis. In Transportation Research Record 944, TRB, National Research Council, Washington, D.C., 1983, pp. 92-100.
12. J. Supernak, A. Talvitie, and A. DeJohn. Person-Category Trip-Generation Model. In Transportation Research Record 944, TRB, National Research Council, Washington, D.C., 1983, pp. 74-83.
13. A.B. Sosslau, A.B. Hassan, M.M. Carter, and G.V. Wickstrom. Quick-Response Urban Travel Estimation Techniques and Transferable Parameters: User's Guide. NCHRP Report 187. TRB, National Research Council, Washington, D.C., 1978, 229 pp.
14. W. Ugolik and J. McDermott. User's Guide to AID, an Exploratory Data Analysis Program. Prelim. Res. Report 190. Planning Research Unit, New York State Department of Transportation, Albany, Oct. 1980.
15. D.K. Boyle and J.E. Chicoine. Application of the Life Cycle Concept to Trip Generation Procedures. Transportation Analysis Report 24. Planning Division, New York State Department of Transportation, Albany, Aug. 1983.
16. Census Tracts (100 Percent Tables) PHC 80-2. Bureau of the Census, U.S. Department of Commerce (undated).
17. W.J. Dixon and M.B. Brown, eds. BMDP-79: Biomedical Computer Programs, P-Series. University of California Press, Berkeley, 1979.
18. M.B. Brown and A.B. Forsythe. The Small Sample Behavior of Some Statistics Which Test the Equality of Several Means. Technometrics, Vol. 16, 1974.
19. G.S. Cohen and M.A. Kocis. Components of Change in Urban Travel. In Transportation Research Record 775, TRB, National Research Council, Washington, D.C., 1980, pp. 42-47.
20. A. Rapoli. New York State: A Profile of the 1980s. Transportation Analysis Report 8. Planning Division, New York State Department of Transportation, Albany, Dec. 1981.

Automobile Occupancy, Vehicle Trips, and Trip Purpose: Some Forecasting Problems

ERIC G. OHSTROM and PETER R. STOPHER

ABSTRACT

The problems with estimating automobile occupancy by trip purpose for use in travel forecasting and in the policy decisions that frequently follow from forecasts are described. Investigations of data and development of logit models of mode choice reveal that the occupants of multioccupant automobiles frequently have disparate trip purposes, even within the restricted trip-purpose definitions usually encountered in practical transportation planning. These disparate purposes mean that, although occupants can be classified by trip purpose, the automobile vehicle cannot be defined as being used for a single trip purpose, as is necessary to compute accurately the automobile occupancy for a purpose and to convert automobile-person trips by purpose to automobile-vehicle trips for assignment of automobile vehicles to the highway network. This has serious repercussions on a variety of contemporary policy decisions. The problems are discussed, and some alternative procedures that can be used as a compromise computation of vehicle occupancy by purpose are given. The problems and solutions are demonstrated in the context of a case study.

Automobile occupancy plays a number of roles in practical transportation planning. First, it is used as a statistic to verify the correctness of collected data and the validation of forecasting models. In both cases it is usually used as a purpose-specific measure. Second, it is used to convert automobile-person trips (the product of standard modeling procedures) to automobile-vehicle trips for assignment of vehicles to the highway network. This is again purpose specific, except in the case of estimating 24-hr assignments (1). A peak-hour or peak-period assignment uses purpose-specific occupancy in building a peak trip table from different proportions of trips by each of the purposes. Finally, automobile occupancy is an important component in policy decisions concerning high-occupancy vehicles (HOVs), where the forecasts of automobile trips in such vehicles is of critical importance. Again, occupancy is generally required to be purpose specific, particularly because most HOV facilities will operate only during peak periods (2,3).

Before the general introduction of multimodal logit models of mode choice in practical transportation planning, occupancies by purpose were estimated outside the standard modeling stream and were introduced for the conversion of automobile-person trips to automobile-vehicle trips. HOV policies were not of much interest at that time, and automobile occupancy was not an issue in model or data validation. Usually, occupancy by purpose was obtained from roadside interviews, with the driver's trip purpose defining the vehicle trip purpose.

The introduction and expanding use of the logit mode-choice model with varying levels of automobile occupancy or the use of an automobile driver and automobile passenger split in the automobile alternatives has revealed hitherto unrecognized problems and issues in the use of purpose-specific automobile occupancy. Briefly, the issues explored by this paper are that

1. Automobile occupancy by purpose cannot be estimated from modal-split models that specify occupancy levels by purpose, and these models cannot be validated by use of automobile occupancy;
2. Standard measurement procedures for automobile occupancy do not estimate occupancy by purpose, and it is not clear if this can be estimated by any current methods; and
3. Use of automobile occupancy by purpose for any of the uses previously described must involve some approximation, for which currently there are neither empirical nor theoretical rules available to guide the practitioner.

In this paper these problems are described in more detail, the additional common problem of measurement of automobile occupancy is explored, and the problems with a case study from Honolulu, Hawaii (4), are discussed. Some suggested ad hoc procedures are outlined, although no final solutions to the problems are offered. It is hoped that the problems discussed in this paper will serve to alert practitioners to inherent problems in working with purpose-specific automobile occupancies, will assist in discouraging the practice of using automobile occupancy by purpose to validate data and models, and will encourage research to deal with this problem more effectively than is done by the ad hoc procedures outlined here.

OUTLINE OF PROBLEMS

The problems that arise can be defined most clearly by considering the two alternative automobile-occupancy model specifications most commonly used for logit mode-choice models. In the first model specification, the automobile mode is defined as the submode of drive-alone automobile, two-occupant automobile, and three-or-more-occupant automobile (5-7); person trips in each submode are divided by the average occupancy for the submode (1, 2, and about 3.3, respectively) to derive automobile-vehicle trips. The second specification defines the two submodes of automobile driver and automobile passenger (8,9), in which automobile-vehicle trips are set equal to the number of automobile drivers, and the number of automobile passengers plays no role in the assignment. Before developing these descriptions further, however, some discussion of trip purposes is necessary.

Trip Purpose

In most practical applications, trip generation and trip distribution use six to eight trip purposes, whereas modal-split and highway and transit assignments use three or four purposes. In the Honolulu case study (4), as in a number of other transportation studies, trip generation and trip distribution each use six trip purposes for resident travel:

1. Home-based work,
2. Home-based school,
3. Home-based shopping,
4. Home-based social-recreational,
5. Home-based other, and
6. Nonhome based.

After trip distribution, the six purposes are aggregated to four by forming a new home-based other category [sometimes referred to as modal-split other (MSO) to distinguish it from the category 5 trip purpose] by combining purposes 3-5.

Of particular concern in the issue of automobile occupancy and trip purpose is the treatment of serve-passenger trips. In common with conventional procedures, the 1982 modeling in Honolulu treated home-based serve-passenger trips as home-based other trips, whereas non-home-based trips with a serve-passenger origin or destination were classified as non-home-based trips.

Definition of Principal Issues

Bearing in mind the definitions of trip purpose, the problems associated with the automobile-occupancy models can be described.

Multioccupancy Reporting Error

The reporting of automobile occupancy for multioccupant automobiles may exhibit one or more of several systematic and random errors in the recording of the actual occupancy of the vehicle:

1. Sampling error, resulting in driver and passenger bias,
2. Automobile drivers differing from automobile passengers in reporting occupancy,
3. Occupants improperly include or exclude themselves (depending on the wording of the question) in determining the occupancy, and
4. Children younger than 5 years are generally included in the occupancy response, although no travel information is usually collected for this age group (e.g., this results in four person trips using a five-occupant automobile).

These errors are critical to the correct analysis and application of these data to automobile-occupancy models.

Automobile Occupancy by Trip Purpose

Automobile occupancy by trip purpose is frequently derived by cross-tabulating person trips by automobile occupancy and trip purpose. However, multioccupant vehicles with two or more trip purposes will necessarily include an unknown number of trips of other purposes in the occupancy response. In effect, this will lead to varying levels of double counting, as is discussed later in the case study.

Model Specification Mixtures

The two model specifications previously discussed may be used for different trip purposes. However, this leads to additional error in converting person trips to automobile-vehicle trips for multioccupant, multipurpose vehicles. Consider the common case of a two-passenger vehicle with a serve-passenger driver taking a student from home to school: a home-based other and a home-based school trip. If home-based other trips are modeled with a driver and passenger model, the driver yields 1.0 automobile-vehicle trips. If home-based school uses the occupancy model, the passenger converts to 0.5 automobile trips, yielding 1.5 automobile-vehicle trips where only 1.0 actually occurred.

SOLUTIONS

The following case study gives techniques to quantify the multioccupancy reporting errors and to adjust the data accordingly. As mentioned previously, these adjustment procedures are ad hoc and somewhat arbitrary, but they represent the state of the art for this problem.

For multioccupant, multipurpose automobile trips, it would appear that the first potential solution might be to restrict calibration data to those automobile trips where all occupants are traveling for the same purpose. Two problems arise here. First, the purposes of other automobile occupants are not collected in contemporary surveys, and their collection may prove to be cumbersome and difficult. Second, although such a stratagem may solve the problem of calibrating the automobile submodes correctly and would allow automobile occupancy to be estimated by mode-choice purpose for the calibration data, it does not solve the basic issue of calculating occupancy by purpose for multioccupant, multipurpose automobiles, nor does it solve the forecasting problems. Instead, it excludes them and replaces them with a loss of trips and information.

Therefore, alternative compromises to provide feasible solutions for practical transportation planning are proposed, which offer less overall error at the expense of varying levels of error by purpose. The compromises can be illustrated by considering two common situations in multioccupant, multipurpose automobile trips:

1. The driver is performing a serve-passenger trip (either home based or nonhome based) with a passenger(s) traveling to work or school; the driver will be classified as making either a home-based other or a non-home-based trip and the passenger(s) will be classified as making either a home-based work or home-based school trip; and
2. One occupant of the automobile is traveling to work or school, while another occupant is traveling to the same destination for a nonwork, nonschool purpose.

In both cases the use of occupancy by purpose will double count automobile trips, thereby obscuring the estimation of automobile occupancy by trip purpose. Three alternative compromises are defined. First, it could be assumed that all double counting occurs with at least one occupant traveling for work or school, so that estimated double counting is deducted from work and school purposes only. This solution will tend to understate the volume of automobile-vehicle trips for work and school and will most affect peak-hour assignments. Second, all double-counted automobile vehicles could be deducted from the home-based other and non-home-based trips. This is equivalent to assuming that every automobile user performing a serve-passenger trip has the same purpose as his passengers. If peak-hour assignments or policies concerning HOV lanes and carpooling are of primary concern, then this option, even though it overstates the number of vehicles affected, will be the best option.

Third, and arbitrarily, half of the double count for home-based work trips can be deducted from each of the home-based work trips and the two nonwork, nonschool purposes; and half of the double count for home-based school can be deducted from itself and the other half deducted from the two nonwork, nonschool trips. This is difficult to justify because the fraction of deduction is purely arbitrary. Yet it may also be interpretable as the least biased of the three compromise solutions.

CASE STUDY

The problems and solutions described in the preceding sections are demonstrated much more clearly with the case study, which illustrates all the problems previously mentioned. Furthermore, the home-based work (HBW) and home-based school (HBS) models were originally developed as multioccupant models, whereas the home-based other (HBO) and non-home-based (NHB) models were of the driver-passenger type, thereby demonstrating the pitfalls of this inconsistent treatment of the automobile mode. Two other items are of interest in the case study. First, evidence was uncovered that the reporting of automobile occupancy appears to be subject to a large reporting error, which serves to obscure the computation of corrections for double counting; and second, there was an initial incorrect assumption made about average occupancy for the 3-or-more-occupant automobiles, the effect of which turns out to be small compared with the effects of double counting.

The case study is for Honolulu, for which data were collected in the fall of 1981. The data were collected by means of a 24-hr travel diary in a procedure described in a paper by Ohstrom et al. elsewhere in this Record.

Reporting Automobile Occupancy and Purpose

An analysis of the survey data clearly indicates that the problematical mixed-purpose trips occur frequently, even though trip purposes of other automobile occupants were not requested. The results obtained from the survey data are given in Table 1. The last two categories show that there are a number of people who are engaged in serve-passenger trips, whereas the first two categories show an imbalance between car drivers and car passengers within the purposes. However, this latter issue of an imbalance is not conclusive evidence on its own. First, a question arises as to whether the small sample data produce a balance between automobile drivers and automobile passengers, which implies that for every two-occupant automobile driver there should be a two-occupant passenger; for every three-occupant driver there should be two three-occupant passengers; and so forth. This is far from what is found in the data, which indicate that there are far too few passengers or too many drivers at each occupancy level (Table 2).

TABLE 1 Drivers, Passengers, and Occupancy from Honolulu Survey Data

Mode	Occupancy	HBW	HBS	Serve Passenger: Home Based	Serve Passenger: Nonhome Based
Automobile driver	2	225	79	426	391
Automobile passenger	2	205	158	–	–
Automobile driver	⩾3	71	93	315	285
Automobile passenger	⩾3	74	374	–	–

Six reasons can be advanced for this:

1. The sample contains more drivers than passengers, thus representing a bias between passengers and drivers;
2. Many of the drivers misread the occupancy question and counted themselves as well (i.e., reporting one too many occupants);
3. The extra passengers are under 5 years old, who are correctly reported as occupants, but for whom there are no trip logs, thus producing no passenger reports;
4. Automobile passengers reported occupancy incorrectly;
5. There is a higher probability of forgetting to report an automobile-passenger trip than an automobile-driver trip; the 100 missing trip logs from the households that provided responses to the mail survey were from people making predominantly automobile-passenger trips; and
6. Automobile passengers misread the survey question and marked themselves down as automobile drivers in some cases.

TABLE 2 Drivers and Passengers by Reported Occupancy Level

	Occupants: 1	2	3	4	5	6-10	⩾11
Driver	6,001	2,422	867	359	124	78	4
Passenger	1	1,374	700	484	231	141	19
Total	6,002	3,796	1,567	843	355	219	23

Probably, part of the answer is to be found in each of these six reasons. It is unlikely that any one reason is solely responsible, or that any one has no effect. For example, that 6,001 drivers reported zero other occupants indicates that most drivers probably reported occupancy correctly. (If this question was consistently misread, there would be zero one-occupant automobiles.) That the question was misread sometimes is apparent because there is one automobile passenger who reported zero other occupants. Similarly, if all the automobile drivers were shifted to one lower occupancy, there would be serious imbalances in the opposite direction. Identical arguments can be made for automobile passengers.

The discrepancy is also not likely to be due entirely to children younger than 5 years old. If this were the case, there would be 3,288 trips by children younger than 5 years old as automobile passengers. Assuming that half of the surveyed households with two or more people in them have one child younger than 5 years old (which would appear to be an overestimate), then the survey households would have not more than 624 children younger than 5 years old. This would mean that these youngsters each make 5.27 trips per day compared with an average person-trip rate of 2.83 trips. Alternatively, every household with more than one person would have to have one child younger than 5 years old in the household to average the trip rate of all people older than 5 years old; this is equally unlikely.

Similar arguments apply to the 100 missing trip logs. These would have to have contained more than 32 automobile-passenger trips each to compensate for the missing automobile passengers. Assuming an average of 4 automobile-passenger trips per missing log would account for only 400 of the shortfall of automobile-passenger trips. Finally, although there is some evidence that respondents in the sample have a slightly higher income than the average, and that there were some intentional biases on household size, it appears unlikely that the sample could be biased to the extent that less than half of the automobile passengers that would be expected were found in the sample (3,854 sampled automobile-driver trips, where the number of passenger trips by occupancy would lead to the expectation of 1,965 trips). This would represent a large bias, and nothing else in the data supports such a supposition.

Given this, the sample should be adjusted so that it behaves consistently with the use of the model outputs. The models are used to estimate automobile use by occupancy, and every two-occupant automobile trip is assumed to generate 0.5 automobile-vehicle trips, while every three-or-more-occupant automobile trip generates 1/3.7 automobile-vehicle trips for HBW trips and 1/4.2 automobile-vehicle trips for HBS trips, as found empirically in these data.

Referring back to Table 2, there are 3,796 automobile trips with two occupants. These would be assumed to be split evenly between drivers and passengers, giving 1,898 of each. This generates a multiplier of 0.784 for two-occupant automobile drivers and 1.381 for two-occupant automobile passengers. By a similar process, 837 automobile drivers would have been estimated from the three-or-more-occupant categories out of 3,007 automobile trips, leaving 2,170 automobile passengers; but 1,432 drivers and 1,575 passengers were observed. Therefore, correction multipliers of 0.584 for automobile drivers and 1.378 for automobile passengers can be deduced. These figures yield an all-purposes average occupancy of 3.59 for the three-or-more-occupant automobiles.

The raw survey data indicate that there are 817 automobile drivers making serve-passenger trips with two occupants in the car. Factoring this, as indicated in the preceding paragraph, yields a total of 641 automobile-driver, two-occupant, serve-passenger trips. The data indicate that 15.06 percent of automobile passengers in two-occupant automobiles were making HBW trips, and 11.61 percent were making HBS trips. Assuming that the drivers making serve-passenger trips are distributed across all purposes in the same proportions as the automobile passengers, then 15.06 percent (97) HBW automobile passengers and 11.61 percent (74) HBS automobile passengers are being driven by serve-passenger drivers. In the HBW data there are 443 automobile trips with two occupants. By using the procedure applied to model forecasts, this would generate an estimate of 222 automobile-vehicle trips. But 97 of these automobile-vehicle trips are already counted in the MSO (HBO for modal split) and NHB purposes for automobile drivers. Therefore, only 125 automobile-vehicle trips from the 443 automobile-person trips should be counted to avoid double counting. This yields a factor of 1/3.54 instead of 1/2 for the two-occupant automobile-person trips to convert them to automobile-vehicle trips. This is a 43.6 percent reduction in the automobile-vehicle trips from those estimated without correction. Similarly, the school trips produced an observation of 241 automobile-person trips with two occupants, which would produce an estimate of 121 automobile-person trips. However, 74 of these are already counted in MSO and NHB trips. Therefore, the conversion factor from automobile-person trips to automobile-vehicle trips for two-occupant HBS automobile trips is (121 - 74)/241, or 1/5.13.

An identical procedure should be applied to the three-or-more-occupant automobile trips. The reader can readily confirm that this produces conversion factors to automobile-vehicle trips of 1/5.69 for HBW and 1/13.24 for HBS trips.

The next question is to determine the effect of this on the estimates of automobile-vehicle trips obtained for the 159 zones and 1985 data (Table 3). The original estimate of automobile-vehicle trips for these person trips was 421,112. Applying the new conversion factors yields an estimate of 393,338 automobile-vehicle trips. This shows a double counting of 27,774 automobile-vehicle trips, or 6.6 percent of the automobile-vehicle trips previously estimated for HBW trips. Results for the HBS trips are given in Table 4 and indicate a reduction of 22,394 automobile-vehicle trips, or 24.6 percent of the original estimate.

TABLE 3 Changes in Automobile-Vehicle Trip Estimation After Correction (HBW)

Automobile-Occupancy Category	Person	Original Vehicle	New Vehicle
Estimated one-occupant trips	356,837	356,837	356,837
Estimated two-occupant trips	104,312	52,156	29,466
Estimated three-or-more-occupant trips	40,027	12,129	7,035
Total	501,176	421,112	393,338

TABLE 4 Changes in Automobile-Vehicle Trip Estimation After Correction (HBS)

Automobile-Occupancy Category	Person	Original Vehicle	New Vehicle
Estimated one-occupant trips	58,571	58,571	58,571
Estimated two-occupant trips	29,470	14,735	5,745
Estimated three-or-more-occupant trips	58,918	17,854	4,450
Total	146,959	91,160	68,766

In total, there were 1,880,090 estimated automobile-vehicle trips for 1985, which these conversion factors would reduce to 1,829,901, a reduction of 2.67 percent of the original estimate. There were 2,414,755 automobile-person trips in the 1985 estimates, which yielded an average automobile occupancy of 1.28. The revised automobile-vehicle trips increases this to 1.32 persons per automobile.

The initial use of an average occupancy for three-or-more-occupant vehicles of 3.3, corrected subsequently to 3.7 for HBW and 4.2 for HBS trips, contributed about 10 percent to the change noted in these figures. Thus, although it is important to use a correct average occupancy for the highest occupancy grouping, the effects of an incorrect value are small compared with the problem of double counting.

It is reasonable to assume that the number of double-counted automobiles will be a function of the volume of HBW and HBS trips. Therefore, the correct procedure must always be to initially estimate the double count from these trips. However, there has to be some inconsistency in determining occupancy by purpose and in attributing automobile-vehicle trips to purposes because of the mixture of purposes represented in any multioccupant automobile. There are three alternatives that could be used with some justification from this analysis.

Alternative 1: Reduction of Work and School Trips

Automobile-vehicle trips are reduced solely in the HBW and HBS purposes. Therefore, the conversion factors defined earlier in this paper are used to compute vehicle trips from person trips. The conversion factors are given in the following table (the results are summarized in Table 5):

Purpose and Occupancy	Factor
HBW	
Two occupants	1/3.54
Three or more occupants	1/5.69
HBS	
Two occupants	1/5.13
Three or more occupants	1/13.24

Alternative 2: Reduction of Nonwork, Nonschool Trips

The additional automobile-vehicle trips are deducted from MSO and NHB instead of from HBW and HBS, after first calculating the double count from the HBW and HBS trips. This involves calculating the fraction of automobile-person trips for each of two occupants and three or more occupants that represent double-counted automobile-vehicle trips. If there were no double counting, then two-occupant vehicle trips would be obtained by using a conversion of 0.5 on automobile-person trips. The difference between this and the revised conversion factor of 1/3.54 for HBW is 0.2175. Thus there is a double count of 0.2175 times the 104,312 two-occupant automobile trips. In similar fashion, the factors that represent double counted automobile-vehicle trips for each occupancy of each purpose can be calculated, as noted in the following table:

Purpose and Occupancy	Factor
HBW	
Two occupants	0.2175
Three or more occupants	0.0945
HBS	
Two occupants	0.3051
Three or more occupants	0.1626

In the sample data, 64.68 percent of the automobile-driver, serve-passenger, multioccupant trips were home based and 35.32 percent were nonhome based. Therefore, after summing the total double-counted automobile-vehicle trips, 64.68 percent are deducted from MSO trips and 35.32 percent are deducted from NHB trips.

Applying this to the 1985 regional trip estimates, 45,042 automobile-vehicle trips are double counted. Deducting these from the MSO and NHB automobile-driver trips, by using the percentages given in the preceeding paragraph, reduces the number of automobile-driver trips (and therefore the number of automobile-vehicle trips) to 960,782 for MSO and to 361,982 for NHB. By using the corrected average occupancies for three or more occupants for HBW and HBS trips, new estimates of 419,814 vehicle trips for HBW and 87,342 for HBS are obtained.

Alternative 3: Reduction from All Trip Purposes

Although much less easy to justify, there is the proposition to deduct one-half of the double counts from each purpose. The double count for two-occupant HBW trips is 22,688 vehicle trips, of which 11,344 would then be deducted from the HBW trips and 11,344 from MSO and NHB trips together. Similarly, 1,891 vehicle trips would be deducted from HBW three-or-more-occupant automobile trips, 4,496 from HBS two-occupant trips, and 4,790 from HBS three-or-more-occupant trips. A total of 14,567 and 7,954 trips would be deducted from MSO and NHB trips, respectively, for a total of 22,521 trips.

It is instructive to see the effects of these alternatives against both the original estimates with no correction for double counting and the correction to a more correct average occupancy. These results are summarized in Table 5. It is also interesting to note the automobile occupancies by purpose that result from these various alternatives (Table 6). The results in Table 6 show some marked variations in automobile occupancy by purpose. Again, this serves to underline the problem of computing automobile occupancy by purpose.

TABLE 5 Comparison of Original Results and Alternative Solutions

Purpose and Occupancy	Uncorrected		Corrected Occupancy	Alternative		
	Person Trips	Vehicle Trips		1	2	3
HBW 1	356,837	356,837	356,837	356,837	356,837	356,837
HBW 2	104,312	52,156	52,156	29,467	52,156	40,812
HBW > 3	40,027	12,129	10,818	7,035	10,818	8,927
HBS 1	58,571	58,571	58,571	58,571	58,571	58,571
HBS 2	29,470	14,735	14,735	5,745	14,735	10,239
HBS > 3	58,918	17,854	14,028	4,450	14,028	9,238
MSO d	989,915	989,915	989,915	989,915	960,782	975,348
MSO p	314,454	314,454	314,454	314,454	314,454	314,454
NHB d	377,891	377,891	377,891	377,891	361,982	369,937
NHB p	84,360	84,360	84,360	84,360	84,360	84,360

TABLE 6 Vehicle Occupancy by Trip Purpose for Alternatives

Purpose	Uncorrected Model	Vehicle Occupancy: Three or More Occupancy Correction	Alternative		
			1	2	3
HBW	1.19	1,19	1.27	1.19	1.23
HBS	1.61	1.68	2.14	1.68	1.88
HBO	1.32	1.32	1.32	1.33	1.32
NHB	1.22	1.22	1.22	1.23	1.32

As a final note, the overall magnitude of the changes noted in this paper are of a similar order of magnitude to many of the other errors in the forecasting process. Nevertheless, it is worthwhile to seek a correction for at least four reasons. First, much of the existing error in forecasting models cannot currently be removed. Simply because the errors noted here appear no greater than those errors is no argument for ignoring correction and the possible improvement in accuracy to be obtained from improved methods to estimate automobile occupancy. Second, it is important to discern the inappropriateness of using automobile occupancy to assess the validity of data and models. Failure of data or models to reproduce observed automobile occupancy by purpose provides no information on validity. Third, when HOV lanes are of policy concern, the magnitude of the errors is large, proportionately. Depending on the method used, HOV lane volumes may range up to 100 percent greater with one method than with another. Fourth, the errors in automobile occupancy could be reduced by redefining some of the questions customarily asked in transportation surveys. In particular, attention should be given to determining whether or not a survey instrument design can be created that will both remove current potentials for misreporting or mismeasuring automobile occupancy, and permit data to be obtained on the purposes of all occupants in a multioccupant vehicle. With respect to the measurement problem, it is worth noting, anecdotally, that various designs of questions used by the authors that specify "including yourself" and various other terms designed to specify unambiguously how to count have all met with relatively similar rates of failure. Apparently, most people just do not bother to read the question properly and therefore are uninfluenced by any qualifiers on occupancy.

CONCLUSIONS

Two final comments are in order. First, one automobile-occupancy model specification should be applied across all trip purposes, with the occupancy model offering better information relative to current transportation planning issues. One model specification will simplify some of the problems in dealing with the multioccupant, multipurpose automobile. However, the analysis will still be necessary to adjust the derivation of automobile occupancy by trip purpose for use in calculating automobile-vehicle trips and for estimating the effects of HOV policies and similar issues.

Second, if policies relating to carpooling, HOV lanes, and similar concerns are to be examined, then alternative 2 should be used, which will provide a correct estimate of the number of three-or-more-occupant automobiles that are being used to work and to school, primarily in the peak period. Use of alternative 1 would result in ignoring a number of three-or-more-occupant automobiles because they are included in the MSO and NHB trips but are not explicit as to occupancy. When automobile occupancy by period or purpose is not critical, then alternative 1 (which is simpler) is probably the best procedure to use. Beyond this, the alternative procedures are a matter of the preference of the analyst. Of course, there is a danger that these various methods can be used to justify alternative strategies, and great care must be taken to select an alternative that is objectively justifiable and not subjectively convenient.

ACKNOWLEDGMENTS

The authors would like to thank the Oahu Metropolitan Planning Organization (OMPO) for permission to use the results described in the case study, and for their active support on the presentation of this paper. The authors particularly wish to thank Cathy Arthur and Gordon Lum of OMPO for their suggestions and help both in the original work and in the documentation of the results in this paper.

REFERENCES

1. M.L. Manheim. Fundamentals of Transportation in Systems Analysis, Volume 1: Basic Concepts. MIT Press, Cambridge, Mass., 1979, pp. 430-431, 456-457.
2. M.A. Kocus and M. Fuhrman. Assessment of State Emergency Energy Conservation Planning. In Transportation Research Record 870, TRB, National Research Council, Washington, D.C., 1982, pp. 60-66.
3. P.J. Valk. Commuter Demand for Ridesharing Services. In Transportation Research Record 823, TRB, National Research Council, Washington, D.C., 1981, pp. 17-21.
4. Oahu Model Update Study, Final Report, Volume II. Oahu Metropolitan Planning Organization, Honolulu, Hawaii, Dec. 1982.
5. B.D. Spear. Applications of New Travel Demand Forecasting Techniques to Transportation Planning. Case Study 2. FHWA, U.S. Department of Transportation, 1977, p. 47 ff.
6. Urban Transportation Planning Model Update--Phase III. Tech. Memorandum No. 1, Task H. Florida Department of Transportation, Tallahassee, Oct. 1982.
7. E.R. Ruiter and M. Ben-Akiva. Disaggregate Travel Demand Models for the San Francisco Bay Area. In Transportation Research Record 673, TRB, National Research Council, Washington, D.C., 1978, pp. 121-128.
8. P.R. Rassam, R.H. Ellis, and J.C. Bennett. The n-Dimensional Logit Model: Development and Application. In Highway Research Record 369, HRB, National Research Council, Washington, D.C., 1971, pp. 135-147.
9. Charles River Associates, Inc. A Disaggregated Behavioral Model of Urban Travel Demand. Final Report. FHWA, U.S. Department of Transportation, March 1972.

The opinions expressed in the paper are those of the authors and do not necessarily reflect the position or policies of the Oahu Metropolitan Planning Organization. The authors are solely responsible for any errors or omissions in the paper.

Successful Administration of a Mailed 24-Hour Travel Diary: A Case Study

ERIC G. OHSTROM, JOAN B. OHSTROM, and PETER R. STOPHER

ABSTRACT

Recent transportation survey research has shown that successful travel diaries can be constructed, and that these diaries can collect information on travel by individuals for a period of 24 hr or more. The successful diaries are comparatively expensive survey instruments and have been described primarily in terms of use in conjunction with a personal visit by an interviewer. The interviewer may collect some information at the time of the visit, but he plays an essential role in explaining the use of the diary. This interviewer visit has made the diary an expensive survey instrument. A case study of the administration of a travel diary survey conducted through a combination of telephone contact and mail-out, mail-back procedures is described. In the description of this case study it is shown that the diary can be administered successfully by this means, that the results obtained are of a high quality, and that a response rate significantly higher than that associated with most mail surveys can be obtained. A number of details of the administration methods used, which are believed to have contributed to the success of the instrument, are discussed. The procedure described produced a usable response rate of 58.5 percent of the mail sample of households, from which it was possible subsequently to calibrate new trip-generation and modal-split models. Some of the results obtained, including the higher trip rates for non-home-based trips, are described. It is suggested that refinements to the instrument and procedures could generate yet higher response rates.

Several papers have appeared recently extolling the virtues of a travel diary for use in collecting a 24-hr record of household members' travel (1-4). These travel diaries provide a means to have individuals record details about their travel and activities for a day in the future, rather than relying on recall. Brog et al. (1) and Stopher and Sheskin (2) claim that the data obtained are more complete than the data collected by the traditional recall surveys used for the past three decades in transportation planning activities. However, most transportation surveys that use the diary have made use of a face-to-face encounter between a survey person and one or more members of the household to administer the travel diaries (5).

Because of the need for careful design of the diary (i.e., the use of various devices such as color-keying, indented cuts, and special bindings), the diary is a comparatively expensive survey instrument. In versions that these authors have used in the United States, costs have varied between about $0.75 and $1.25 per diary. Given an average requirement of more than three diaries per household, the instrument alone can cost between $2.25 and $4.00 per household. In addition, many of the diaries will be returned spoiled or empty but unusable, or just not returned, thus increasing the cost per household for completed, usable diaries. It is a conservative estimate that the instrument cost alone for each completed household is approximately $10. If this cost is added to the cost of the labor-intensive activity of sending out survey personnel to deliver and explain the use of the diaries, and possibly also to retrieve completed diaries, the survey unit costs increase considerably. In a 1980 survey of this type in Michigan, Stopher and Sheskin (2) estimated the total per household cost (including data reduction) at approximately $125.

There has been a slow acceptance of the diary for urban area data collection. Some early efforts reported low response rates, which may have been a contributory factor to this slow acceptance. The cost of the diary procedure may also have much to do with this. However, the estimated travel-diary survey costs in excess of $100 must be set in the context of the cost of conventional home-interview surveys that cost anywhere from about $80 to more than $500 per household, depending on design, length of interview, response rates, and many other factors. Furthermore, more efficient sampling methods other than simple random sampling have been applied successfully, thereby increasing the efficiency of the survey personnel. Recent research (6,7) has indicated that large samples, on the order of 2 percent or more of regional households, are quite unnecessary for urban area updates; and that samples of considerably less than 5,000 households produce data of more than sufficient accuracy for virtually every transportation planning need. These characteristics have made the diary a practicable instrument, even at a cost of more than $100 per household. However, it is clear that, if the cost can be reduced, the procedure becomes more accessible to many urban areas and may offer a relatively low-cost method to update decades-old data or collect data needed for new types of models and forecasting procedures.

In this paper the use of the 24-hr travel diary is described. The diary used a combination of telephone and mail contacts that produced a high response rate, appears to have generated data that may be more complete than that obtained from more conventional methods, and that cost substantially less than $50 per completed household. The telephone contact provided an extremely effective means of random sampling, without the need to seek out and correct some form of household sampling frame.

As is usual in a survey effort of this nature, the procedures evolved as the survey proceeded. Rather than a chronology of developments of the technique, the procedure is described in the form in which it was administered. A detailed and extensive pilot survey was conducted but is not described herein. Without this pilot survey, many of the successful elements of the final design would not have been developed and implemented. Not all elements were tested in the pilot survey, but those that were not tested were introduced into the main survey to correct problems encountered in the pilot survey. In this respect, the pilot survey was not only indispensable to a good final survey, but succeeded in all of the primary areas that pilot surveys are designed to handle (8). One detraction from the case study is that the survey was not designed as a comparative exercise among alternative methods or instruments. Therefore, the success of the instrument and its administration must be based primarily on response rates, nonresponse biases, and measures of the quality of the data obtained.

DESIGN

Sample

The sample was designed to be selected in a two-phase process. The first phase was a simple random sample of telephone numbers generated by random-digit dialing. From the households contacted by telephone, the second-phase sample was selected on the basis of household size and automobile availability. Before the survey, certain combinations of these two variables were identified that should have encompassed more than 75 percent of households and more than 80 percent of daily regional trip making. The households contacted in the second-phase sample were asked to complete travel diaries.

Survey Instruments

The first-phase sample was given a 5-min telephone interview that established household size, automobile availability and ownership, number of workers in the household, number of adults, type of housing, and numbers of trips made by bus and car for each of work and nonwork purposes by the contacted respondent on the survey day. During the telephone interview the interviewer identified whether the household was eligible for the diary survey. (This was done by giving each interviewer a laminated selection grid that showed household size and automobile ownership. The interviewer first placed a penny on the column heading for the household size, and then moved the penny down the column to the appropriate value of the automobile availability. If the cell had an X in it, the household was not selected for the travel-diary survey; otherwise, it was.) If the household was eligible, the interviewer described the diary survey briefly and requested the address to which to send the diary materials. The contacted respondent was informed of the day to be used to complete the diaries.

The second-phase sample received a mail package that contained several items. First, there were the correct number of travel diaries (for all members of the household who were at least 5 years old), on the outside of each of which a sticker was attached indicating the day of the week on which the diary was to be filled out. The travel diary included not only a diary section as described by Stopher and Sheskin (2), but it also included a small booklet requesting details about the respondent (age, gender, relationship to other household members, education, driver's license status, and so forth) and details about one of the trips selected from the diary. These details included travel time components and cost for the trip selected, and equivalent data on up to two alternative travel modes for making that trip.

In addition to the diary, there was a one-page survey form asking for certain characteristics of the household. These details included the same vehicle availability and ownership questions used in the telephone survey, parking availability and cost at home, military or civilian status (because of the large number of military households on Oahu), household size, and income. The package also contained two signs indicating the travel-diary day, an envelope for collecting together and returning the survey forms (preaddressed and printed with a reply-paid postage license), and a cover letter from the director of the metropolitan planning organization (MPO) indicating the purpose of the survey, the importance of the household's response, and a telephone number to use for questions about the survey.

The entire package was mailed out in a large white envelope. Computer-generated address labels were fixed to the envelopes, using the contacted respondent's name, if given to the telephone interviewer. Postage stamps rather than metered or prepaid bulk mail were used for mailing, and these stamps included some attractive commemorative stamps (9).

Contact Procedure

Households were contacted initially by telephone, and those households that were eligible were then mailed a package of survey materials, as described in the preceding section. The travel-diary day was set as the same day of the week as the day of the telephone contact, but 1 week later. This was done in the belief that it would help respondents remember the day more easily, and because it would be less complicated for the telephone interviewers. The only exception to this procedure was for telephone interviews made on Saturday (no calling was done on Sundays). The interest was to obtain travel data from weekdays, so that Saturday interviews set the diary for the Wednesday or Thursday (10 to 11 days) following the telephone interview. (Use of the Wednesday or Thursday immediately following would not have provided sufficient time for the mail packages to arrive.)

A telephone follow-up was used with all households, based on a computer listing of the names, addresses, and telephone numbers for each travel-diary day. This contact was made on the evening before a household's travel-diary day by using specially trained interviewers. The purposes of this contact were to remind households of the agreed-on travel-diary day, to make sure that the survey package had been received and opened, and to answer any questions about the survey. In the few instances where a package had not been received, the address was verified and another package mailed with the request that the travel-diary day be the same weekday 1 week later. If the package had been received but not opened, the person called was asked to get the package and open it, and the interviewer explained what was in the package and how to use each item.

If a mail package had not been returned by 4 days after the travel-diary day, a reminder postcard was sent, urging completion on the same weekday of the week in which the reminder was received. Further follow-ups had been planned but were not executed because the response rate already achieved by these prior methods exceeded the clients' expectations and requirements. A limited follow-up and targeted remailing was undertaken and is described later in the paper. A "thank-you" letter and a copy of the State Highway Map were sent to all households that returned completed packages.

Logistics

The success of a multiple-contact survey of this type resides largely in an effective logistical design. The idea behind this is to make each respondent believe that his or her response to the survey is so important that the survey administration knows on precisely which day he or she is to complete travel diaries and knows whether the survey has been completed and returned.

The procedure for mailing out survey forms included a series of steps of checking, computerizing, packaging, and dispatching the forms. It is most easily described by considering a specific day's telephone interviews, that is, the first Wednesday. Telephone interviews were undertaken on Wednesday evening and were completed by about 8:30 p.m. On Thursday morning all telephone interviews, still in individual interviewer binders, were checked visually for completeness, correct designations of the household by cell of the trip-generation matrix, readability of the address, and correct identification of mail-out status. Specific errors were noted and the interviewer was informed of these and instructed on correcting problems before the start of the evening's interviewing. During checking, the interviews were tabulated by household size and vehicle availability categories to determine the distribution of surveys obtained and particularly to determine where terminations were occurring. This lead, for example, to discovery that the early days of the survey were experiencing an exceptionally high termination rate for one-person households. After sensitizing interviewers to this issue, the response of these households improved dramatically.

After checking, the eligible interviews were sorted by number of travel diaries to be sent out for the mail-out surveys, followed by ineligible households, and finally by terminations that were complete enough to keypunch. In this order, the interviews were then sequenced-numbered by using a numbering system beginning at 110001, where the first two digits designated the main interview survey. Each new day's interviewing began at the next hundred. Thus Wednesday, October 21 had interviews numbered 110001 through 110111. Thursday, October 22 then commenced at 110201. A log was maintained showing the beginning and ending number for each day and the assigned logging day and date for each.

The sequence-numbered forms were turned over to the keypuncher who completed a second visual check, looking specifically for problems likely to be encountered in the direct keypunching process. Usually this check was carried out in the late afternoon, after the interviewers reported for the evening's interviewing, so that any questions could be directed to the responsible interviewer. The complete answer set to the telephone interview was keypunched during the evening, checked for errors, and both a recontact listing and a set of address labels were generated.

For Wednesday evening's interviewing, the keypunching was completed on Thursday evening and address labels were available by Friday morning. The address labels included the sequence number of the household, the number of travel diaries to be mailed, and the diary day. On the morning that the address labels became available, the mail-out packages were assembled. This assembly included stamping the household number on each of the travel diaries, on the household-interview form, and on the return envelope. The package was made up for each household and mailed at a U.S. postal facility providing next-day delivery service. Thus Wednesday's interviews were processed and mail surveys were sent by Friday afternoon, with delivery probably occurring on Saturday and Monday. With the travel-diary day being the following Wednesday, most households would receive their survey packages about 3 days before the diary day. This procedure was followed throughout the survey period, except that interviews from Saturday were delayed by 1 day beyond this schedule.

The telephone recontacts were set up by using the computer listing produced when the telephone interview was keypunched, as noted earlier. The household sequence numbers were transferred to the recontact interview sheets for each evening's calls. The procedure was to work straight through the households in sequence order, making one attempt at each number. If the household was contacted successfully, the number was checked off on the computer listing, and the answer spaces were filled out on the recontact form. When one pass through the list was completed, the interviewers returned to the beginning of the list and reattempted each of the unsuccessful initial contacts. This procedure was repeated a third time during the evening, after which recontact was concluded. Requests for a later call back were accommodated if the call back was to be within the telephone-interviewing period, or only a short time beyond the end of it. In this way some 75 to 80 percent of all mail-out households were recontacted successfully on the evening of their travel-diary day.

The telephone recontact also represented a means of checking and verifying the computerized record of telephone numbers and addresses. Corrections were keypunched on the following evening and a dual set of labels produced from the corrected records, together with an extra mailing label for those cases where a remailing was to be done. The dual set of labels consisted of two consecutive labels for each household. The first had the word "card" printed at the top right and the second had the word "thanks" printed there. These were used to mail and control the subsequent follow-up.

As survey packages were received in the mail, each package was date-stamped, opened, and its contents examined. The travel diaries were opened to see if they had been filled out, and the number filled out was written on the outside of the return envelope in the space provided. The household survey form was checked to see if it was filled out, and the appropriate space was marked for this on the outside of the envelope. Returns were sorted into numerical order during this process, and the number of packages returned by day of original survey (indicated by the household number) was recorded. This provided a profile of the returns by time from the original interview, as discussed later in this paper.

For each survey day's responses, once the dual set of labels had been generated, a cross-check was made between returned packages and the labels. The labels showed both the household number and the number of travel diaries, while the return envelope now bore the number of returned, completed diaries. If missing diaries were detected by this check, this was marked on the envelope; and, in the event that not more than one diary was missing, a thank-you was sent to the household. If the survey package was processed before mailing of the reminder postcard, then the label marked "card" was crossed through and that marked "thanks" was used to send the thank-you package. If the package was too incomplete for a thank-you, both labels were crossed through.

On the day designated for postcard mailing, all the uncrossed "card" labels were used on reminder postcards. After the reminders were sent, "thanks" labels continued to be used to send out thank-you packages as complete returns were received, or were crossed through if an incomplete return was received. This procedure proved to be an effective way of keeping track of returns and reminders, and only a few errors (less than 10) were detected in which an incorrect thank-you or reminder was sent. (One household sent back an incorrectly sent thank-you package, with a note to say that they had not completed the survey forms and did not intend to, and therefore felt they should return the thank-you package.)

RESULTS OF THE SURVEY

Telephone Survey

A total of 2,883 telephone interviews were conducted, including 247 interviews of households that qualified for mailing but refused to give a mailing address. These are included in the 313 terminations in Table 1, not in the successful interviews. The rate of 65.5 percent of ineligible numbers called for interview is considered relatively low. Past telephone surveys have shown this rate to range between 75 and 85 percent. The lower rate in this survey is considered to be due to the sampling in proportion to numbers assigned by prefix (exchange) and to exclusion of the numbers outside the minimum and maximum currently assigned within each prefix.

TABLE 1 Disposition of Telephone Calls Made

Disposition	Telephone Numbers Called No.	Percent
Not in service	4,599	30.9
Business	863	5.8
Number changed to new listing[a]	380	2.6
No answer	2,773	18.7
Busy	1,060	7.1
Recorder[b]	64	0.4
Total ineligible	9,739	65.5
Terminated	313	2.1
Refused	1,364	9.2
Unsuccessful request for call back	562	3.8
Eligible nonresponses	2,239	15.1
Successful interviews	2,883	19.4
Total eligible	5,122	34.5

Note: Data are from Schimpeler-Corradino Associates.

[a]A telephone company recorded message indicating a new number assigned was considered equivalent to not-in-service status for numbers selected by the computer.

[b]A recorder was considered equivalent to no answer and tried again.

All interviewing was conducted in English, although there are many Oahu residents whose native language is not English. Translation problems and the expected difficulty of finding multilingual interviewers dictated a restriction to English. Of all telephone contacts, 191 households had language problems such that no telephone interview could be conducted. These are included in the terminated calls in Table 1. If answers could be obtained, but it was apparent that the household members would be unlikely to be able to deal with the written questionnaire, the interviewer was instructed to mark households otherwise eligible for a mailing as FLP (foreign language problem) and treat the household as ineligible. Only 5 such cases were encountered out of the 2,883 completed telephone interviews.

If the interviewer was unable to get a single question answered by the selected respondent, this was designated a refusal. The volume of refusals at 1,364, or 26.6 percent of eligible numbers, is considered high, but generally does not reflect on the skill of the interviewers. A call was considered to end in a termination if the interviewer succeeded in asking at least one question of the selected respondent and obtained an answer. The low rate of terminations, at 6.1 percent, is a reflection of the skill of the interviewer in obtaining responses once a respondent was contacted who could be persuaded to answer the first question. Furthermore, the number of respondents who terminated during the main questioning in this survey, as opposed to refusing to give an address for mailing, was only 66, or 1.3 percent. The unsuccessful requests for call back were those instances where contact was made with a household and the respondent requested a subsequent call back. Up to three attempts were made to recall the household, with each of these attempts being several days apart and with at least one on a weekday and one on a Saturday. Of these, 562 remained as failures to make a further contact by the end of the calling period.

Mail Survey

Of the 2,883 interviews conducted, 2,595 were with households eligible for a mail survey, 2,348 of which provided an address and were sent survey forms. A total of 1,485 mail forms were returned. The distribution of telephone and mail surveys by day of week is given in Table 2. The data show a fairly even distribution of survey effort by day of week, with only Thursday showing a significant drop below the other days, although this is compensated for in a higher eligibility rate and a higher response rate. Overall, about 90 percent of interviewed households qualified for the mail survey, and this varied from a low of 87.3 percent to a high of 92.6 percent. Of interviewed households, 81.4 percent were mailed surveys, and this varied by day of week from 78.0 to 83.7 percent. An average of 51.5 percent of all households contacted (57.2 percent of all eligible households, and 63.3 percent of all households mailed surveys) responded to the mail survey, with a variation from 48.5 to 55.2 percent by day of week.

The data in Tables 3-6 give the distributions of interviews by household size and vehicle availability. The zeroes in Tables 4-6 are in those cells where no mail surveys were designed to be sent out. Only 7 of the 12 cells of the matrix were designed

TABLE 2 Distribution of Telephone and Mail-Back Surveys by Day of Week Called

Day	Interviews	Eligible for Mail No.	Percent[a]	Sent Out No.	Percent[a]	Returned No.	Percent[a]
Monday	430	398	92.6	360	83.7	229	53.3
Tuesday	467	423	90.6	380	81.4	232	49.7
Wednesday	519	453	87.3	405	78.0	256	49.3
Thursday	382	350	91.6	312	81.6	207	54.2
Friday	524	472	90.1	435	83.0	289	55.2
Saturday	561	499	88.9	457	81.5	272	48.5
Total	2,883	2,595	90.5	2,340	81.5	1,485	51.5

Note: Data are from Schimpeler-Corradino Associates.

[a]Percentages are of interviews conducted.

TABLE 3 Distribution of Telephone Interviews Conducted

Vehicles per Household	Distribution by Persons per Household: 1 No.	Percent	2-3 No.	Percent	4 No.	Percent	≥5 No.	Percent	Total No.	Percent
0	65	2.25	98	3.40	14	0.49	16	0.55	193	6.69
1	188	6.52	588	20.40	164	5.69	126	4.37	1,066	36.98
≥2	24	0.83	653	22.65	462	16.02	485	16.82	1,624	56.33
Total	277	9.60	1,339	46.45	640	22.20	627	21.74	2,883	100.00

Note: Data are from Schimpeler-Corradino Associates.

TABLE 4 Distribution of Interviews Eligible for Mailing

Vehicles per Household	Distribution by Persons per Household: 1	2-3	4	≥5	Total
0	0	97	0	0	97
1	188	586	0	126	900
≥2	0	653	462	483	1,598
Total	188	1,336	462	609	2,595

Note: Data are from Schimpeler-Corradino Associates.

TABLE 5 Distribution of Interviews Mailed Out

Vehicles per Household	Distribution by Persons per Household: 1	2-3	4	≥5	Total
0	0	87	0	0	87
1	165	527	0	114	806
≥2	0	583	422	451	1,456
Total	165	1,197	422	565	2,349

Note: Data are from Schimpeler-Corradino Associates.

for mail out. From the data in Table 3 it can be seen that the omitted cells comprise 9.8 percent of the households interviewed by telephone. Primarily, the differences between Tables 4 and 5 are those households that refused to provide an address. In Table 6 the percentages of mail surveys returned in each cell are given. With the exception of the 2- and 3-person households with no vehicles, the rates are quite similar and show an even response over the matrix.

The high mail-back response to the survey is considered to have been achieved, at least in large measure, by the telephone recontact on the day before the travel-diary day for each household. In general, the reaction to recontact was positive. Many respondents indicated that they were ready to complete the forms and had no questions. An almost equal number either had not opened the package but did so under the prompting of the interviewer, or had opened it and had questions about the materials. A number of those contacted indicated initially that they did not plan to respond, but some of those appeared to be persuaded to do so by the interviewer. The remaining contacts generally indicated an assortment of problems, most of which occurred only once or twice in each evening and probably constituted not more than 1 to 2 percent of all mail outs, although a precise count was not maintained.

1. Some contacted households indicated they had not received the survey package, even though the address was verified as correct. No action was taken on those, because it was assumed that the Post Office had delayed delivery or the person contacted had overlooked the arrival of the package or was unaware of it.

2. Some contacted households indicated they had not received the survey and an error was found in the address. This error appeared to include the respondent having given an incorrect or incomplete address, the telephone interviewer making an error in transcribing the address, or a keypunch error in the address. These were corrected, and a new package was sent out.

3. In some instances the telephone number called was of someone completely different from the name and address recorded. Whenever possible, the name and address were then looked up in the telephone directory and the correct telephone number inserted. In many of these cases, however, the name and address were not listed. From a log kept that indicated the section of a page of computer-generated telephone numbers that an interviewer called each evening and from the interviewer number on the telephone-interview form, the telephone numbers called were searched. This search used a reverse directory to check each marked number for the name and address in question. Through this process, about half of these cases were recovered and correct telephone numbers appended. Some of these instances were recovered more simply, because a comparison between computer listing and original interview showed a simple keypunching error. Also, a few instances revealed a different name but the same address and subsequently were found to indicate a multifamily household. The remainder could not be traced and, for them, the telephone number on the computer record was removed.

The return profile for the mail-back survey is given in Table 7. Not unexpectedly, this profile shows that returns generally peaked two to three days after the designated diary day, suggesting that most respondents completed their travel diaries on the designated day. After the tenth day from the interview (thirteenth for Saturday, with its delayed diary days), the response declines quite rapidly, but there was a small increase around the fifteenth to sixteenth days following the postcard reminder and second diary day. There is, however, no way to

TABLE 6 Distribution of Interviews Returned

Vehicles per Household	Distribution by Persons per Household: 1 No.	Percent	2-3 No.	Percent	4 No.	Percent	≥5 No.	Percent	Total
0	0		41	47.1	0		0		41
1	115	69.7	340	64.5	0		65	57.0	520
≥2	0		386	66.2	266	63.0	272	60.3	924
Total	115		767		266		337		1,485

Note: Data are from Schimpeler-Corradino Associates.

TABLE 7 Return Profile for Mail-Back Surveys

Days from Interview	Return Profile (%) by Day of Week of Interview					
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
6	0.4	0.4	0	0.5	0.3	5.1
7[a]	0.4	1.3	1.4	1.0	1.7	1.1
8	15.7	10.3	0.4	11.6	1.7	8.4
9	31.9	15.9	15.2	11.1	19.4	0.4
10	7.9	25.0	17.0	24.6	14.6	1.8[b]
11	12.7	13.4	14.4	5.8	17.7	9.9[b]
12	9.6	3.0	5.1	11.6	17.0	11.0
13	1.3	6.5	6.1	10.6	2.4	18.3
14[c]	0.4	1.3	9.0	1.9	4.2	11.7
15	3.5	7.8	2.2	4.8	4.9	3.3
16	5.7	2.6	2.9	3.9	0	2.2
17	1.7	2.6	5.8	0.5	2.8	1.5[b]
18	2.2	2.6	0.4	1.9	2.1	4.8[b]
19	1.3	0	1.4	0.5	3.1	2.2
20	0	0.9	1.1	3.4	2.1	1.5
21[d]	0.4	0.4	0.7	0.5	0.7	5.5
22	1.3	0.4	0	0.5	0	0
23	0.4	0.4	0.4	0.5	0	1.5
24	0	0.9	1.8	0	0.7	1.8[b]
25	0.9	0.4	0	1.0	0.3	0.4[b]
Other	2.2	3.9	5.1	3.9	6.3	7.7

Note: Data are from Schimpeler-Corradino Associates.

[a]Diary day.
[b]Travel-diary days for Saturday interviews.
[c]Diary day 2.
[d]Diary day 3.

define how many responses were received as a result of the reminder postcard. It appears to be in the range of 8 to 12 percent of all returns. A small additional surge occurred after the third diary day, although the total volume of this was, as expected, small. Most of the remaining responses came from a targeted remailing of survey packages that occurred about 30 to 35 days after the original interviews. The remailing was a targeted remailing sent to households in certain zip codes and categories of household size and vehicle availability that were considered to have a response rate that was significantly below the general response rate. A total of 190 such mailings were sent out, of which 27 (14.2 percent) were returned.

With respect to this targeted remailing, it is interesting to speculate that, if the original plan to send a remailing to all nonresponding households had been executed, an extrapolation of this response might indicate the size of the final response that could have been achieved. A total of 863 remailings could have been made, given the nonresponding total, and a 14 percent response from this would have added a further 121 responses that might have been obtained, leading to an increase of 5.2 percent in the response rate for households receiving mail surveys. Such a reminder process should have achieved a final response rate of 67.5 percent. It is also reasonable to suppose that the targeted households for this remailing were inclined to be more nonresponsive than the average, so that it may also be speculated that this represents the low end of the potential response achievable.

Follow-up for Missing Data

Included in all of the response figures are all packages received by mail. Of these, 24 packages proved to be outright refusals, with the forms returned blank, which reduced the response total to 1,461 and the response rate by 1.6 percent. In addition, 37 of the 2,338 packages mailed were returned by the Post Office as undeliverable and no correct address was found from reverse directories, recontact telephone calls, or all other means available. These also are considered to constitute refusals, in that probably an intentional wrong address was provided. However, these 37 were not included in any of the reported returns. The refusals that were mailed back are evenly distributed over the household types defined by the trip-production matrix.

Subsequent analysis of the remaining returns revealed various elements of missing or conflicting data. It had been decided much earlier that a return would be considered complete if it was missing not more than one-third of the travel diaries that should be returned (i.e., no travel diaries missing for households sent one or two; one missing for households sent two through five; and two missing for those sent six through nine), and that critical questions on household size, vehicle availability, and household location had been answered on the mail-back forms. In those cases where the returned survey would be described as incomplete on this basis and, in addition, when any information was missing from the household survey form or any travel diaries were blank or missing, an attempt to complete the data by telephone was undertaken. A second category of responses requiring follow-up was identified: this was when critical data provided in the telephone interview differed from the data provided in the mail-back survey. Resolution of such conflicts was considered to warrant a telephone call. In many instances the conflicts were found to have arisen because of changes in the household between the original telephone interview and the travel-diary day, or because of an error in the information given to the telephone interviewer.

This follow-up procedure was reasonably successful in completing otherwise incomplete surveys and resolving conflicts, and it was relatively inexpensive at $2.00 per household. However, 90 responses were classified as too incomplete to be usable, reducing the final usable sample to 1,370 observations. The distribution of these complete surveys by the two primary categorization variables is given in Table 8.

USEFULNESS OF RESULTS

The data produced by this survey have been used subsequently to develop new models of trip generation and modal split for long-range regional transporta-

TABLE 8 Distribution of Usable Surveys by Household Size and Vehicle Availability

Vehicles per Household	Distribution by Persons per Household				
	1	2-3	4	≥5	Total
0	5	33	0	2	40
1	116	310	6	56	488
≥2	1	356	237	249	843
Total	122	699	243	307	1,371

Note: Data are from Schimpeler-Corradino Associates.

tion planning in Oahu. The data could be used, but have not been, for recalibrating the trip-length distributions for the gravity model. As a brief summary of the results obtained from the data, it can be noted that cross-classification models of trip production were produced for six purposes, and estimates of trip rates by households were produced that compared favorably with rates from other recent studies. For example, the weighted average person trip rate for Oahu was determined to be 3.08 motorized trips per day, compared with rates of 2.80 (1980) and 2.46 (1965) in southeast Michigan, 3.00 (1977) and 1.66 (1962) in Baltimore, and 1.57 (1977) in San Juan. Earlier studies in Oahu had also indicated a tendency for households on the island to show a higher trip-making rate than households on the mainland. It is also speculated that the travel-diary approach is more successful in obtaining a reasonably complete report of trip making.

Similarly, logit models of mode choice were calibrated for four purposes--home-based work, home-based school, home-based other, and nonhome based--with calibration data sets of 458, 329, 361, and 277 for the four purposes, respectively. Satisfactory models were obtained in each case, with coefficients that were within the expected ranges, t-scores that exceeded the 99 percent significance level, and satisfactory chi-square and rho-square statistics. For the selected models, the chi-square for home-based work was 355.3, with 9 degrees of freedom (df); for home-based school it was 134.1, with 8 df; for home-based other it was 34.0, with 6 df; and for non-home-based trips the chi-square was 113.8, also with 6 df. These all indicate reasonable fits to the data, and indicate that the data collected were clearly adequate for the job.

CONCLUSIONS

The case study reported in this paper demonstrates a procedure by which an intensive survey, based on travel diaries, was administered by telephone and mail and achieved a 50 percent saving in the survey cost per completed return, compared with the use of personal interviewers. The final result of this survey was the achievement of a mail-back response of 1,370 usable household returns, which represented a 58.5 percent response rate for the mail-back portion of the survey. Because the survey described here was conducted very much as a pioneering effort, it is considered that this response rate should be able to be improved further in subsequent refinements of the procedure.

The survey used some duplicate questioning so that it is also possible to deduce the nonresponse biases of the mail-back survey. This has not been explored in this paper, but it is an important element of the validity and value of a survey of this type.

The data produced have been used subsequently to develop new models of trip generation and modal split for long-range regional transportation planning on Oahu. The data could be used, but have not been, for recalibrating the trip-length distributions for the gravity model.

ACKNOWLEDGMENTS

The authors would like to thank the Oahu Metropolitan Planning Organization (OMPO) for permission to use the results described for the case study, and for their active support on the presentation of this paper. The authors particularly wish to thank Gordon Lum of OMPO for his suggestions and help both in the original work and in the documentation of the results of this paper.

REFERENCES

1. W. Brog, A.H. Meyburg, and M.J. Wermuth. Development of Survey Instruments Suitable for Determining Nonhome Activity Patterns. In Transportation Research Record 944, TRB, National Research Council, Washington, D.C., 1983, pp. 1-12.
2. P.R. Stopher and I.M. Sheskin. Toward Improved Collection of a 24-H Travel Record. In Transportation Research Record 891, TRB, National Research Council, Washington, D.C., 1982, pp. 10-17.
3. W. Brog, A.H. Meyburg, P.R. Stopher, and M.J. Wermuth. Collection of Household Travel and Activity Data: Development of an Instrument. Presented at 2nd International Conference on New Survey Methods in Transport, Hungerford Hill, Australia, Sept. 1983.
4. Sozialforschung Brog. KONTIV 75-77: Continuous Survey of Travel Behavior, Summary Volumes 1-3. Ministry of Transport, Federal Republic of Germany, Munich, 1977 (in German).
5. R. Parvataneni, P. Stopher, and C. Brown. Origin-Destination Travel Survey for Southeast Michigan. In Transportation Research Record 886, TRB, National Research Council, Washington, D.C., 1982, pp. 1-8.
6. M.E. Smith. Design of Small-Sample Home-Interview Travel Surveys. In Transportation Research Record 701, TRB, National Research Council, Washington, D.C., 1979, pp. 29-35.
7. P.R. Stopher. Small-Sample Home-Interview Travel Surveys: Application and Suggested Modifications. In Transportation Research Record 886, TRB, National Research Council, Washington, D.C., 1982, pp. 41-47.
8. I.M. Sheskin and P.R. Stopher. Pilot Testing of Alternative Administrative Procedures and Survey Instruments. In Transportation Research Record 886, TRB, National Research Council, Washington, D.C., 1982, pp. 8-22.
9. E.R. Babbie. Survey Research Methods. Wadsworth Publishing Company, Belmont, Calif., 1973.

The opinions expressed in this paper are those of the authors and do not necessarily reflect the position of the Oahu Metropolitan Planning Organization. The authors are solely responsible for any errors or omissions in the paper.

Analysis of Geographical and Temporal Variation in Vehicle Classification Count Statistics

DAVID L. GREENE, PATRICIA S. HU, and GLENN F. ROBERTS

ABSTRACT

The problem of estimating highway travel by vehicle type using available traffic vehicle classification count data is discussed. The data are analyzed by using techniques of discrete multivariate analysis. It is determined that vehicle type relative frequency distributions vary importantly across regions, highway systems, seasons, day of week, and time of day, but that interactions among these factors, which would complicate estimation of travel, are not of great importance. The only important two-way interactions involve highway system type; therefore it is possible to derive unbiased estimates of vehicle travel by vehicle type and highway system even from a nonrandom sample of classification count observations, provided that total travel by system is known. Some exploratory disaggregate vehicle travel estimates are presented.

The quantities of travel by type of vehicle and type of highway system are fundamental transportation data. Such information is important for analysis and forecasting of travel and energy use and for determining exposure rates in studies of highway safety. Vehicle survey data are useful for estimating travel by vehicle type, but not by highway system (1,2). In order to obtain travel estimates disaggregated in both dimensions, vehicle classification count statistics are needed. Classification count data consist of hourly counts of vehicles by type that are recorded at a particular location on the highway system network. Determining disaggregate travel by vehicle and highway type is thus a problem of inferring vehicle miles from vehicle counts.

If there were a sufficiently large, well-designed random sample of traffic counts, deriving unbiased estimates of vehicle travel would be, in principle, a rather simple exercise. Unfortunately, while there is a great volume of classification count data, none has been collected according to a statistically designed sampling plan. The problem is then one of removing, to the greatest extent possible, the bias inherent in the existing sample. To do this effectively, the variation in vehicle type distributions across time and space must be understood. If temporal and spatial dimensions affect the distribution of vehicle types independently, then sample bias can be corrected by a simple reweighting of the data.

This paper is divided into three parts. In the first part a probabilistic model of vehicle type relative frequencies, which helps to clarify the relationship between vehicle miles and vehicle counts by vehicle type, is presented. Second, the three major sources of vehicle classification data are described, and the results of an analysis of the structure of classification count data using loglinear models are presented. The implications for using available data to estimate disaggregate vehicle travel are discussed. Finally, several preliminary estimates of travel by 13 vehicle types and 10 highway system classes are presented and discussed. In the concluding section the interpretation of these estimates is discussed and important areas for further research are recommended.

STATISTICAL MODEL OF VEHICLE TYPE COUNTS AND TRAVEL

Traffic counts do not represent vehicle travel but rather represent density at a point on a road. Thus a set of assumptions must be specified by which vehicle travel estimates can be derived from vehicle count data. It is shown that if a functional class can be divided into homogeneous systems, then an unbiased estimate of vehicle type relative frequencies can be obtained as a weighted average of the estimated system relative frequencies. This result will be used in the section Exploratory Disaggregate Estimates of Vehicle Travel to estimate relative frequencies and travel by vehicle type for functional highway classes. The systems used will be regional functional classes classified by season, day of week, and time of day. The analysis of the variability of vehicle type relative frequencies across these systems in the next section will show that a particularly simple weighting scheme can be used that permits weights for temporal dimensions to be constant across systems. That the systems defined may in fact not be homogeneous is a persistent problem that can only be solved by improved random sampling strategies.

Assume that a functional highway class (see Table 1) is divided into segments that are sufficiently small and homogeneous that vehicle miles on the segment are equal to its length times a traffic count taken anywhere on the segment. The segment then forms the basic unit of analysis because there is nothing to be gained by subdividing it.

A collection of segments with identical (in practice, similar) traffic densities and vehicle type distributions are called a system in this paper. Clearly, a given functional highway class (e.g., urban Interstate) may be made up of several different systems. In fact, the same strip of road can be considered to belong to different systems, depending on the time of day or season of the year. In this sense a functional class has no underlying parameters of its own to be estimated, but rather is merely a sum of individual systems. Because the goal is to make inferences about vehicle miles of travel on functional highway classes, these will be derived from weighted averages of inferences about the systems that compose it. In particular, for the purposes of this study, the interest is in inferring the distribution of vehicle miles by vehicle type for each functional class.

Assume that an observer, standing at a roadside recording vehicle counts for a fixed time period such as an hour, is observing a random process. In particular, if N total counts are recorded during the period, assume that the probability of observing

TABLE 1 Variables and Categories

Variable	Number	Category
Quarter	4	1st
		2nd
		3rd
		4th
Region	4	Northwest
		South
		North Central
		West
Road type	10	Interstate, rural
		Other principal arterials, rural
		Minor arterials, rural
		Major collectors, rural
		Minor collectors, rural
		Interstate, urban
		Other freeways, urban
		Other principal arterials, urban
		Minor arterials, urban
		Collectors, urban
Day	2	Weekday
		Weekend
Time of day	5	5:00-9:00 a.m.
		9:00 a.m.-3:00 p.m.
		3:00-7:00 p.m.
		7:00-11:00 p.m.
		11:00 p.m.-5:00 a.m.
Vehicle type	13	Standard and compact cars
		Subcompact cars
		Motorcycles
		Buses
		Pickups, panels, and other two-axle, four-tire trucks
		Two-axle, six-tire single-unit trucks
		Two-or-more-axle single-unit trucks
		Three-axle combination trucks
		Four-axle tractor-semicombinations
		Other four-axle combinations
		Three-axle tractor, two-axle semicombinations
		Other five-axle combinations
		Six-or-more-axle combinations

C_1 vehicles of type 1, C_2 of type 2, up to C_m of type m is given by the multinomial distribution,

$$P(C_1, C_2, \ldots, C_m) = N! \prod_{k=1}^{m} p_k^{C_k} \Big/ \prod_{k=1}^{m} C_k! \quad (1)$$

The p_k's are the probabilities of observing a vehicle of type k in a sample of one, or alternatively, the relative frequencies of type k vehicles in the total population of vehicles traveling the given system. Also, it is required that

$$\sum_{k=1}^{m} C_k = N.$$

In general, the total number of counts recorded in an hour will itself be a random variable. Assume that the number of counts observed will follow a Poisson distribution. The Poisson is widely used both in traffic engineering and elsewhere to represent random arrivals (3):

$$P\lambda(N) = e^{-\lambda} \cdot (\lambda N/N!) \quad (2)$$

The Poisson distribution has expected value (mean) and variance both equal to λ. Compounding the Poisson and multinomial distributions in this way results in a distribution in which each of the vehicle type counts is distributed Poisson with parameter $\lambda_k = p_k\lambda$ (4),

$$P(C_1, C_2, \ldots C_m) = \prod_{k=1}^{m} \left\{ e^{-\lambda} p_k \left[(\lambda p_k) C_k / C_k! \right] \right\} \quad (3)$$

From this model some useful results concerning estimators of system traffic densities and vehicle frequencies can readily be derived.

Maximum likelihood estimators of λ and p_k can be obtained by taking derivatives of the log likelihood function,

$$\max_{\lambda, p_k, k} \log(P) = \sum_{k=1}^{m} \left[-\lambda p_k + C_k \log(\lambda p_k) - \log(C_k!) \right] \quad (4)$$

setting them equal to zero and solving for λ and p_k:

$$\partial \log(P)/\partial p_k = -\lambda + C_k (1/p_k)$$

$$\partial \log(P)/\partial \lambda = -\sum_{k=1}^{m} p_k + (1/\lambda) \sum_{k=1}^{m} C_k \quad (5)$$

Setting these equal to zero, and because

$$\sum_{k=1}^{m} p_k = 1,$$

then

$$\hat{p}_k = C_k/\lambda$$

$$\hat{\lambda} = \sum_{k=1}^{m} C_k \quad (6)$$

The unbiasedness of these estimators can be shown by taking expected values:

$$E(\hat{p}_k) = (1/\lambda) E(C_k) = (1/\lambda) \lambda p_k = p_k$$

$$E(\hat{\lambda}) = \sum_{k=1}^{m} E(C_k) = \sum_{k=1}^{m} \lambda p_k = \lambda \sum_{k=1}^{m} p_k = \lambda \quad (7)$$

In general, however, the actual λ will not be known in order to be able to estimate $\hat{p}_k$, and instead $\hat{\lambda}$ will have to be used. By using an approximation from Mood et al. (5) it can easily be shown that their quotient is unbiased at least up to a second-order Taylor series approximation,

$$E(p'_k) = E(C_k/\hat{\lambda}) \doteq (p_k\lambda/\lambda) - (1/\lambda 2)\, p_k\lambda + (p_k\lambda/\lambda 3)\, \lambda = p_k\lambda/\lambda = p_k \quad (8)$$

This result follows from showing that

$$\mathrm{Cov}\left(C_k, \sum_{j=1}^{m} C_j\right) = \mathrm{Var}(C_k) = p_k\lambda.$$

(The full proof is available on request from the authors.)

These estimators are appropriate for estimating the parameters of a system composed of essentially homogeneous road segments. If a random sample of segments is taken from the same system, then these estimators can be used to obtain maximum likelihood, unbiased estimates of the system parameters. A functional road class in a given region and time period will most likely be composed of several systems. In a sense, it has no underlying parameters of its own but, rather, is merely a summation of individual systems. In particular, it is clear that the relative vehicle mile frequencies (f_k) for vehicle types k = 1, ..., m are just the weighted averages of those of all systems in the class:

$$f_k = \left(1 \Big/ \sum_{i=1}^{S} T_i \right) \cdot \sum_{i=1}^{S} p_{ki} T_i = \sum_{i=1}^{S} r_i p_{ki} \quad (9)$$

Systems are indexed by i = 1, ..., S, T_i to represent total vehicle miles of travel on system i, and

r_i is the proportion of total functional class travel occurring on system i.

Unfortunately, the actual travel on a system is not generally known. However, it is known that on a segment j travel is

$$T_j = \lambda_i \ell_{ij} \tag{10}$$

where ℓ is the segment length and λ is the system average traffic count rate. For the system,

$$T_i = \lambda_i \sum_j \ell_{ij} \tag{11}$$

If $\Sigma \ell_{ij}$ is known, the maximum likelihood, unbiased estimator of λ can be used to estimate T_i (this estimator will also be unbiased). Then f_k can be straightforwardly estimated by substituting Equation 11 into Equation 9.

Suppose that N samples (a sample being, for example, a 1-hr vehicle count on a segment) are taken from different systems, where n_i is the number of samples from system i. To obtain an unbiased estimate of the true weighted average for the functional class, it follows from Equation 9 that parameter estimates from each sample must be weighted in proportion to total vehicle miles from each system. This is readily done by using counts and system mileage as demonstrated in Equation 11. The important result here is that to obtain an unbiased estimate of the vehicle type distribution, it is only necessary to have traffic counts and system lengths.

STRUCTURE OF VEHICLE CLASSIFICATION COUNT DATA

Four data bases were supplied by FHWA. One data base contains estimates of total vehicle miles of travel (VMT) by state and highway class (corresponding to FHWA, table VM-2). The remaining three data bases contain vehicle type count records from (a) the Highway Performance Monitoring System (HPMS) case study (6), (b) various truck weight study (TWS) counts, and (c) various traffic counts conducted by states for their own purposes.

The TWS and HPMS data are both large data bases of equivalent size. The HPMS contains 27,070 usable hourly records and the TWS contains 32,650 such records. The distribution of these records by functional class, however, is extremely different. The HPMS cases are divided about equally: 13,246 rural and 13,824 urban. The TWS, on the other hand, is heavily biased toward rural roads, with 27,158 rural cases and only 5,492 urban ones. Geographically, the TWS used for this study is more comprehensive, with data from 22 states, with at least 1 in each of the 9 census regions. Because it is a case study, the HPMS includes data from only four states and one planning region: Arkansas, Iowa, Minnesota, Washington, and the Delaware Valley. In terms of traffic counts, the two data bases are roughly equal in size, with each having just more than 10 million counts.

The chief problem with vehicle count data is that it has not generally been gathered in accordance with statistical sampling procedures designed to produce comprehensive coverage for the entire United States. Instead, counts have been taken for different purposes and at different times under varying conditions. In short, what has been produced is a nonrandom sample. Most techniques of statistical inference are designed to be applied to a random sample. The challenge in working with a nonrandom sample lies in discovering ways to eliminate the bias inherent in the sample (e.g., weekdays may be oversampled relative to weekends, or daytime hours oversampled relative to nighttime hours). One aspect of the sample bias that cannot be corrected within the scope of this project is the choice of traffic count observation locations on the road network. In terms of theory offered in this paper, this is to say that the authors may not be able to work with homogeneous systems. From the viewpoint of this analysis, the choice of traffic count locations must be assumed to be representative of or a random sample within a particular functional class and region.

Each data base was reorganized into a table of total count frequencies classified by quarter, day, time, region, functional class, and vehicle type. The categories of each variable used are given in Table 1. Thus the cell labeled spring, weekday, 9:00 a.m.-3:00 p.m., region 4, rural Interstate, motorcycles, would contain the sum of all motorcycle counts from all observations having those attributes in the data base in question. Certain vehicle categories were combined so that no variable had more than 10 categories, a requirement of the statistical software that was used. The result is a six-dimensional table with a total of 16,000 cells, many of which are empty for any given data base. In terms of the theory, each cell is considered to be a homogeneous system.

The technique of discrete multivariate analysis using log-linear models is used to analyze tables of frequency data cross-classified by categorical variables. Consider a three-way table of traffic counts by vehicle type (V), functional highway class (C), and region (R). The lower case letters i, j, k are used to index the levels (or categories) of the variables V, C, R; and I, J, K are the number of levels in each category. Let f_{ijk} be the observed frequency (count) in cell i, j, k of the table (matrix). Log-linear modeling assumes that the logarithm of the expected cell count $[E(f_{ijk}) = F_{ijk}]$ is a linear function of certain parameters associated with individual effects of each variable and interactions of variables. If the variable symbols are used as superscripts and the variable indices are used as subscripts to indicate the level of each variable, the model can be written as

$$\ell_n F_{ijk} = \theta + \lambda_i^V + \lambda_j^C + \lambda_k^R + \lambda_{ij}^{VC} + \lambda_{ik}^{VR} + \lambda_{jk}^{CR} + \lambda_{ijk}^{VCR} \tag{12}$$

The λ's are usually called effects and the superscript identifies to which variable or interaction of variables the effect pertains. In Equation 12, λ^V, λ^C, λ^R are the main effects of variables V, C, R, in which λ^{VC}, λ^{VR}, λ^{CR} are their two-way interaction and λ^{VCR} is their three-way interaction. Clearly, the table of frequency counts contains only IJK cells, whereas Equation 12 specifies (1 + I + J + K + IJ + IK + JK + IJK) parameters. To eliminate this parameter redundancy, the following constraints are imposed:

$$\begin{aligned}
&\sum_i \lambda_i^V = 0, \sum_j \lambda_j^C = 0, \sum_k \lambda_k^R = 0 \\
&\sum_i \lambda_{ij}^{VC} = \sum_j \lambda_{ij}^{VC} = \sum_i \lambda_{ik}^{VR} = \ldots = \sum_k \lambda_{jk}^{CR} = 0 \\
&\sum_i \lambda_{ijk}^{VCR} = \sum_j \lambda_{ijk}^{VCR} = \sum_k \lambda_{ijk}^{VCR} = 0
\end{aligned} \tag{13}$$

With the constraints of Equation 13, the model (Equation 12) has exactly as many parameters as there are cells in the table. If all parameters were estimated, the model would fit the table exactly. The model (Equation 12) is termed the saturated model because it includes all possible effects. In general, all effects are not statistically significant, and thus the identification of a log-linear model consists of determining which effects are needed, and which λ terms are superfluous.

Generally, only hierarchical models are considered. In a hierarchical model, a higher-order interaction effect is included only if all lower-order effects involving the variables in the higher-order effects are also included. Thus if λ^{VR} is included, λ^V and λ^R must also be included. When only hierarchical models are considered, each model can be described as a minimal set of higher-order effects. For example, specifying the hierarchical model (VC, CR) is equivalent to the model (θ, V, C, R, VC, CR). In this fitted model, the marginal sums associated with V, C, R, VC, CR, and the table total will exactly equal those of the original table. Thus in the hierarchical model including the parameter VR is equivalent to exactly fitting the IxK marginal table formed by summing over j (the levels of the variable C).

Log-linear models are useful for understanding the relationships between variables in a table and for estimating a table of expected frequency counts using a fitted model. What needs to be known are the important relationships among functional highway class (C), season (Q), day of week (D), region (R), and time of day (T), and also the distribution of traffic counts by vehicle type (V). The procedure consists of estimating a new table that fits a subset of the six-way table margins and measuring the degree to which it fits the original table. The degree of fit is measured by means of the likelihood ratio χ^2 statistic,

$$\chi^2 = 2 \sum_{IJK} f_{ijk} \ln(f_{ijk}/F_{ijk}) \tag{14}$$

which is asymptotically distributed as χ^2 with degrees of freedom equal to the number of cells minus the number of parameters to be estimated.

In order to test the significance of a particular parameter (e.g., λ^{AB}) in the model (Equation 12), the difference in χ^2 is computed between the hierarchical model that includes this term,

$$\ln F_{ijk} = \theta + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} \tag{15}$$

and the model that includes all the same terms except λ^{AB},

$$\ln F_{ijk} = \theta + \lambda_i^A + \lambda_j^B + \lambda_k^C \tag{16}$$

The difference in the two models' χ^2 is also distributed χ^2 with degrees of freedom equal to the difference in degrees of freedom of the two models [here (I-1)(J-1)]. By testing Equation 15 versus Equation 16, it is actually a test of whether A and B influence cell counts independently or whether they interact in determining cell counts.

Log-linear analysis allows simultaneous interaction of all variables. In some cases it is reasonable to consider one variable a dependent variable that is affected by the other variables but does not influence them. In the present case vehicle type should be considered the dependent variable (e.g., vehicle type does not influence the number of counts on weekends versus weekdays, rather the reverse). When one variable is considered the dependent variable and all others are independent variables, the joint marginal of the independent variables must always be fitted. In the six-way traffic count table the CDTQR margin must always be fitted. Given this, the interest is in testing hypotheses about only those terms involving V.

In general, the technique of log-linear model analysis is applied to a random sample of data. When the data have not been collected by means of a simple random sample, it is necessary to fit additional marginals to control for the fact that the sample size (marginal sums) in some combinations of categories has been determined exogenously. In the case of the traffic count data, the only factor that is in fact random is the number of counts by vehicle type for a given observation. Everything else has been determined by the peculiarities of the traffic count sample frame. This requires that the CDTQR margin be fitted exactly. Fortunately, this is the same requirement imposed when V is considered the dependent variable.

The analysis of the traffic count data proceeds by adding terms to the null model (CDTQR) to form successively more complex models involving V. A stepwise procedure of the BMDP4F statistical software package was used. Each of the three traffic count data bases (HPMS, TWS, and state data) were analyzed separately. Because of the extremely large sample sizes of these data bases (on the order of 5 to 10 million counts), every conceivable effect is significant at commonly used significance levels (e.g., 0.05, 0.01). The reason for this is that, in a very large sample, even the most trivial differences can be detected with great accuracy. To determine which parameters are important and which are trivial, some other measure is needed. Goodman (7) has suggested a quasi-R^2 (coefficient of multiple determination) based on the percentage reduction in χ^2 brought about by introducing an additional parameter. In the present case the interest is in percentage reductions in χ^2 over the null model (CDTQR, V) brought about by adding interaction terms involving V.

Because only hierarchical models are considered, a shorthand notation is used in which only the highest-order terms are mentioned. For example, the following two are equivalent:

CDTQR, VCT, VCR

and

C, D, T, Q, R, CD, CT, CQ, CR, DT, DQ, DR, TQ, TR, QR, CDT, CDQ, CDR, CTQ, CTR, CQR, DTQ, DTR, DQR, TQR, CDTQ, CTQR, CDQR, DTQR, CDTR, CDTQR, VCT, VC, VT, VCR, VR, V.

A limitation of the BMDP software (8) is that no more than 10 categories can be defined for a single variable. It was therefore necessary to combine three vehicle type categories. Single-unit truck counts, except pickups and so forth, were combined into one class, as were all four-axle combinations and five-axle combinations.

The stepwise procedure begins with the basic (null) model CDTQR, V and adds terms. Results for the HPMS data are given in Table 2. The individual effect of each variable on the vehicle type distribution is captured by the two-way interactions with V. Region and road class appear to be the most important influences. Day and time effects are only about one-third as potent and the quarter effect is almost negligible. When all the two-way effects are included in the model, the percentage of χ^2 accounted for increases to 83. Interestingly, this is almost exactly equal to the sum of χ^2 reductions the individual effects (84), an indication that interactions of higher order may not be important.

Examination of χ^2 reduction due to three factor interactions indicates that only the class-region interaction reduces χ^2 by more than 1 percent.

TABLE 2 Stepwise Analysis of HPMS Traffic Count Data

Model	Degrees of Freedom	Likelihood-ratio χ^2	Quasi-R^2
CDTQR, V	9,171	1,417,160	0.0
Individual Two-Way Interactions			
CDTQR, VC	9,054	1,018,124	0.28
CDTQR, VD	9,162	1,288,042	0.09
CDTQR, VT	9,135	1,281,126	0.10
CDTQR, VQ	9,144	1,370,291	0.03
CDTQR, VR	9,153	931,127	0.34
CDTQR, VC, VD, VT, VQ, VR	8,964	245,229	0.83
Individual Three-Way Interactions			
CDTQR, VT, VQ, VR, VCD	8,865	239,917	0.83
CDTQR, VD, VQ, VR, VCT	8,598	227,791	0.84
CDTQR, VD, VT, VR, VCQ	8,676	231,348	0.84
CDTQR, VD, VT, VQ, VCR	8,838	182,694	0.87
CDTQR, VC, VQ, VR, VDT	8,928	229,811	0.84
CDTQR, VC, VT, VR, VDQ	8,937	242,053	0.83
CDTQR, VC, VT, VQ, VDR	8,946	229,074	0.84
CDTQR, VC, VD, VR, VTQ	8,856	239,964	0.83
CDTQR, VC, VD, VQ, VTR	8,892	237,322	0.83
CDTQR, VC, VD, VT, VQR	8,910	237,410	0.83

TABLE 3 Stepwise Analysis of TWS Traffic Count Data

Model	Degrees of Freedom	Likelihood-ratio χ^2	Quasi-R^2
CDTQR, V	3,330	870,986	0.0
Individual Two-Way Interactions			
CDTQR, VC	3,224	601,158	0.31
CDTQR, VD	3,321	868,682	0.003
CDTQR, VT	3,294	690,654	0.21
CDTQR, VQ	3,303	825,040	0.05
CDTQR, VR	3,303	710,485	0.18
CDTQR, VC, VD, VT, VQ, VR	3,125	227,259	0.74
Individual Three-Way Interactions			
CDTQR, VT, VQ, VR, VCD	3,070	224,809	0.74
CDTQR, VD, VQ, VR, VCT	2,802	208,984	0.76
CDTQR, VD, VT, VR, VCQ	2,968	203,739	0.77
CDTQR, VD, VT, VQ, VCR	2,915	151,177	0.83
CDTQR, VC, VQ, VR, VDT	3,089	226,411	0.74
CDTQR, VC, VT, VR, VDQ	3,106	225,176	0.74
CDTQR, VC, VT, VQ, VDR	3,107	223,694	0.74
CDTQR, VC, VD, VR, VTQ	3,017	224,779	0.74
CDTQR, VC, VD, VQ, VTR	3,017	219,974	0.75
CDTQR, VC, VD, VT, VQR	3,035	140,305	0.84
CDTQR, VD, VT, VQR, VCR	2,825	74,329	0.92

These results suggest that spatial variation in traffic distributions dominates temporal variation. Simple two-way region and road class interactions with vehicle type reduce χ^2 by 34 and 28 percent, respectively. The effect of season only reduces χ^2 by 3 percent. This suggests that for the HPMS data base, at least, little would be lost by ignoring the seasonal variation in vehicle type relative frequencies (not total counts, because these have been accounted for by the CDTQR terms). The results also suggest that each factor, class, day, time, region, and road class can be considered approximately independent of the others in its effect on vehicle type relative frequency.

Because HPMS includes only five states covering four regions, the importance of region might be expected to be greater than in the other data sets where each regional effect is the average of several possibly different states Second, the HPMS has by far the most complete coverage across all other variables. This is simply a result of the fact that the HPMS is a systematic data-gathering program. The sampling system ensured good coverage by day of week, season, time of day, and road class. The other data sets are not systematic and generally have large gaps (e.g., weekends at night are sparsely sampled). In brief, it should be expected that the variable region in the HPMS data base is in fact representing particular states. On the other hand, variables such as time or day in the other data sets could possibly be assumed for particular states that reported data for odd times while others did not.

Log-linear model analyses of the TWS counts are summarized in Table 3. The general pattern is similar to that of the HPMS. Functional highway system class is the most important single factor. Region and time of day are considerably less important, and day of week and quarter are almost negligible. The VC term alone accounts for a 31 percent reduction in χ^2. All two-way interactions account for 74 percent as compared with 83 percent in the HPMS. Interactions are somewhat more important. The road class-region interaction and quarter-region interaction appear to be most important. Including both reduces the lack of fit by 92 percent.

The state data base analysis results show strong similarity to that of the HPMS (Table 4). Road class and region appear to be the most influential

TABLE 4 Stepwise Analysis of State Traffic Count Data

Model	Degrees of Freedom	Likelihood-ratio χ^2	Quasi-R^2
CDTQR, V	3,942	1,251,697	0.0
Individual Two-Way Interactions			
CDTQR, VC	3,799	751,830	0.40
CDTQR, VD	3,933	1,226,458	0.02
CDTQR, VT	3,906	1,138,397	0.09
CDTQR, VQ	3,915	1,186,096	0.05
CDTQR, VR	3,856	844,404	0.33
CDTQR, VC, VD, VT, VQ, VR	3,641	258,141	0.79
Individual Three-Way Interactions			
CDTQR, VT, VQ, VR, VCD	3,595	252,306	0.80
CDTQR, VD, VQ, VR, VCT	3,278	163,581	0.87
CDTQR, VD, VT, VR, VCQ	3,395	220,831	0.82
CDTQR, VD, VT, VQ, VCR	3,406	222,791	0.82
CDTQR, VC, VQ, VR, VDT	3,605	255,681	0.80
CDTQR, VC, VT, VR, VDQ	3,615	256,210	0.80
CDTQR, VC, VT, VQ, VDR	3,624	253,473	0.80
CDTQR, VC, VD, VR, VTQ	3,533	247,105	0.80
CDTQR, VC, VD, VQ, VTR	3,531	233,086	0.81
CDTQR, VCR, VTR, VD, VQ	3,310	205,838	0.84

factors. Time of day is considerably less important, and day of week and quarter are again almost negligible. Two-way road class and region interactions with vehicle type reduce χ^2 by 40 and 33 percent, respectively. Including all two-way interactions accounts for 79 percent of reduction in χ^2. The road class-time interaction with vehicle type reduces χ^2 by 87 percent. This suggests that road class and time of day depend on each other in their effect on vehicle type relative frequencies.

EXPLORATORY DISAGGREGATE ESTIMATES OF VEHICLE TRAVEL

The results of the log-linear analysis imply a simple basic structure to traffic count data. The distribution of vehicle traffic among vehicle types does vary across time and space. But the only important interaction effects are two-way effects that include highway class. In terms of the theory, this means that the effect of time of day does not vary across systems within a highway class. Therefore, to develop estimates of travel by type of vehicle disaggregated by highway class, the remaining dimensions can be weighted independently. This result was used to produce some experimental estimates of disaggregate regional travel. These estimates are experimental because biases not controlled in this analysis (e.g., network location) can, and probably do, still influence the results. It should also be noted that the authors did not have complete coverage of states in this sample.

The estimation process consists of two components: (a) weighting and collapsing categories and combining data bases, and (b) using the final processed data to estimate travel. To appreciate the role of preprocessing and weighting in the estimation process, it is useful to begin with a description of the second and final step. Assuming that there is either a random sample or that the vehicle classification count data are weighted to correct for sample bias, the estimation of disaggregate vehicle travel is relatively straightforward. Let c represent traffic counts that are indexed by $i = 1, 2, \ldots, I$ for vehicle types; $j = 1, 2, \ldots, J$ for regions; and $k = 1, 2, \ldots, K$ for highway functional class. All other dimensions (e.g., time of day, day of week, season) have been eliminated in the weighting and dimension collapsing process. The absence of a subscript will be used to signify that counts have been summed over that dimension. For example,

$$\sum_i c_{ijk} = c_{jk},$$

which is the total count for all vehicles in region j on functional class k. From the c_{ijk} and c_{jk}, the relative frequencies are computed for each vehicle type, region, and functional class, which is represented by f_{ijk},

$$f_{ijk} = c_{ijk}/c_{jk}.$$

Recall that if there is a random sample or if the bias in the sample has been eliminated through weighting, then the vehicle miles by each vehicle type i should be proportional to f_{ijk} for all $i = 1, 2, \ldots, I$. This of course applies only to the appropriate region and road system. Given this fact, and the fact that

$$\sum_i f_{ijk} = \sum_i (c_{ijk}/c_{jk}) = c_{jk}/c_{jk} = 1,$$

the f_{ijk} can be used to distribute total VMT, on a given functional class in a particular region, among the various types of vehicles. Let T_{jk} denote travel in region j on functional class k; then

$$T_{ijk} = f_{ijk}T_{jk}$$

is the estimate of disaggregate vehicle travel. If summed across vehicle types, the analyst will get back the total vehicle travel in region j, functional class k, with which he began:

$$\sum_i T_{ijk} = T_{jk} \sum_i f_{ijk} = T_{jk} \cdot 1.$$

The key assumption made is that once the three-dimensional array of traffic counts c_{ijk} is arrived at, any bias in the data has already been removed. In general, this will not be true unless there is a reasonably well-designed sample to begin with. Bias in the sample may arise from three principal sources, only one of which can be corrected:

1. Location bias, which results from collecting counts on an atypical location on the road network;
2. Time-space bias, which results from a nonrandom allocation of observations over time, across functional classes, across regions, and even across states within a region; and
3. Missing data for any category, especially states or highway classes.

Only the second kind of bias can be mitigated. This can be done by the weighting of categories in advance.

The weighting process is best illustrated by example. Suppose that the dimensions and categories given in Table 1 are used. Statistical analyses of the three major vehicle classification count data bases indicated that vehicle type frequency distributions vary across all these dimensions and categories. For example, for any given day of week, season, region, and functional class the distribution of traffic by vehicle type will be different at different times of the day. Therefore if there are twice as many daytime as nighttime observations, the final estimate of total travel by vehicle type will be biased toward the daytime pattern. It was also noted that the vehicle type frequency distribution varies jointly by functional class and time of day and by functional class and region. Because the final estimates will be by functional class and region, this does not complicate matters. To get unbiased estimates, the weights of observations by time, day, and season need only to be corrected independently.

Suppose that half of the observations (records, not counts) were taken on weekends and half on weekdays. This represents sample bias because a uniform distribution over time would give 2/7 on weekends and 5/7 on weekdays. To correct this bias a weight of 2 for weekends and 5 for weekdays can be specified. Because it is known in advance what the distribution of samples over time in an unbiased sample should look like, it is simple to weight categories of temporal dimensions.

Unfortunately, by weighting the sample observations, there is a trade-off of a reduction in bias for a loss in efficiency. To see this, imagine that there were 10^5 weekday observations in the data but only 10 weekend observations. By using a 5:2 weighting, the bias is reduced in theory but the variance (decrease in reliability) of the estimate is greatly increased. The reason is that while there is a great deal of information about weekday travel, next to nothing is known about weekend travel, and yet the data are used as if the analyst had 0.4×10^5 weekend observations. In practice, caution should be exercised when weighting observations when the input data are extremely maldistributed. In such cases it may be better not to try to correct for sample bias at all.

In the same way that categories of a dimension can be weighted and summed, data from different data sets can also be assigned weights and combined. The weights may reflect the analyst's confidence in a particular data set or simply the actual number of observations in each. This allows several data sets to be processed (categories weighted and dimensions collapsed) individually, combined at any desired point, and then further processed as a combined set.

Three sets of disaggregate vehicle travel estimates by region and functional class were produced based on 1980 VMT by state and functional class. [Note that these data are from the FHWA, U.S. Department of Transportation (1982). Used were tables of "Vehicle miles of travel classified by state and functional class highway category" for 1980 and 1981, table VM-1, "Annual Vehicle Miles of Travel and Related Data--1981," and three traffic count data tapes supplied by Paul Svercl of Highway Planning, Highway Statistics Branch.] The first two sets are based on the traffic count data from the TWS and HPMS data bases. The state data base was not used because inconsistencies in its method of vehicle classification could not be resolved. For each state, traffic counts are given a weight proportional to total state VMT. In general, this changed the results little in comparison with counts not weighted by state. Finally, a combined set of estimates was produced based on the state VMT-weighted data from both data sets. Observations for a given region from the HPMS and TWS data were given equal weight, even though the TWS always represented more states.

The estimates based on weighted traffic counts are given in Tables 5-7. Vehicle categories have been combined to reduce the size of the tables and also because four vehicle types--large cars; small cars; two-axle, four-tire trucks; and 3S-2 semitrailers (18 wheelers)--account for virtually all the vehicle travel. Some vehicle categories never achieve as much as 1 percent of total travel in any region.

Some general patterns of vehicle travel hold up across regions and data bases. For example, combinations or semitrailers are always most prevalent on rural Interstates and are less common the lower the order of the road system. Also, the distribution of total regional travel among vehicle types varies importantly, but not drastically, across regions and data bases. For example, the West and South always show the most single-unit trucks, mostly two-axle, four-tire (pickup) trucks. Finally, it appears from these data that combination trucks may account for a greater percentage of total vehicle miles than previously thought, possibly by as much as a factor of 2 (it should be noted that the tables do not include local roads, which account for 14 percent of the 1980 VMT). This holds for both the HPMS and TWS data bases. In 1980, the FHWA estimated that 3.7 percent of total U.S. highway miles were by combination trucks. The exploratory estimates from this research are considerably higher.

The travel estimates represented in these tables represent direct empirical estimates based on the available data. Because of problems with these data,

TABLE 5 Estimates of VMT by Highway Category and Vehicle Type (1981), HPMS Data Only

	Vehicle Type							
					Trucks			
	Cars and Motorcycles		Buses		Single Unit		Combination	
Road Type	VMT (10^9)	Percent	VMT (10^9)	Percent	VMT (10^9)	Percent	VMT (10^9)	Percent
Rural Interstate	70.1	59.3	0.4	0.3	24.5	20.7	23.2	19.6
Rural arterial	133.1	59.7	0.9	0.4	70.0	31.4	18.9	8.5
Rural other	79.4	56.1	1.1	0.7	51.8	36.6	9.4	6.6
Urban Interstate	109.4	68.2	0.4	0.2	33.9	21.1	16.7	10.4
Urban other	385.6	74.4	2.0	0.4	117.7	22.7	12.9	2.5
Total	777.6	67.0	4.8	0.4	297.9	25.7	81.1	6.9

TABLE 6 Estimates of VMT by Highway Category and Vehicle Type (1981), TWS Data Only

	Vehicle Type							
					Trucks			
	Cars and Motorcycles		Buses		Singe Unit		Combination	
Road Type	VMT (10^9)	Percent	VMT (10^9)	Percent	VMT (10^9)	Percent	VMT (10^9)	Percent
Rural Interstate	77.7	57.6	0.4	0.3	26.7	19.8	30.1	22.3
Rural arterial	177.9	67.8	0.5	0.2	60.6	23.1	23.3	8.9
Rural other	110.7	70.1	0.6	0.4	34.8	22.0	11.9	7.5
Urban Interstate	117.8	73.4	0.3	0.2	26.4	16.5	15.9	9.9
Urban other	355.7	76.1	0.9	0.2	96.6	20.7	14.0	3.0
Total	839.8	71.0	2.7	0.2	245.1	20.7	95.2	8.1

TABLE 7 Estimates of VMT by Region and Vehicle Type (1981), HPMS and TWS Data Combined

	Vehicle Type							
					Trucks			
	Cars and Motorcycles		Buses		Single Unit		Combination	
Region	VMT (10^9)	Percent	VMT (10^9)	Percent	VMT (10^9)	Percent	VMT (10^9)	Percent
Northeast	132.9	79.2	0.9	0.5	23.7	14.1	10.3	6.1
South	286.6	63.7	1.4	0.3	118.0	26.2	43.8	9.7
North	242.3	72.5	1.0	0.3	66.7	20.0	24.1	7.2
West	147.0	66.6	0.6	0.3	63.2	28.6	10.0	4.5
Total	808.8	67.0	3.9	0.3	271.6	23.2	88.2	7.5

it is not possible to quantify the accuracy of these estimates with any precision. Three of the four regions are missing data for one road type. The Northeast and South are missing minor rural collector data, and the West is missing data for other urban expressways. In addition, not all states are represented, and there are good reasons to believe that routes high in truck traffic were oversampled.

CONCLUSIONS

Vehicle classification count data are the sole source of information on vehicle travel by type of vehicle, highway system class, and geographical area. Although a great deal of classification count data has been collected, it has not been collected according to statistically unbiased sampling procedures, and this presents serious problems for estimation of vehicle travel. Discrete multivariate analysis of the classification count data has revealed a simple structure to the variation in vehicle type distributions across time and space. Vehicle type relative frequencies vary by region, highway system, day of week, time of day, and season. There are also important interactions between highway class and region, and highway class and time. Vehicle type relative frequencies vary most across the geographical dimensions (regions and highway systems), although temporal variations are also important. The combination of all main effects and two two-way interaction effects accounts for about 90 percent of the variation (as measured by reduction in χ^2) in vehicle type relative frequencies in three different vehicle classification count data bases.

This result implies that sample bias in classification count data along these five dimensions can be corrected relatively easily if vehicle travel by highway class and region, as well as by vehicle type, is being estimated. This can be done by appropriately weighting observations according to the time-space distribution of the road network. Region and functional highway system were the only geographic dimensions used in this analysis. Because it is not necessary to aggregate over these dimensions, there is no need to develop weights for them. Weights could easily have been computed, however, based on highway system mileage by region.

An important geographic factor not controlled in this analysis is the particular location of the traffic count on the given highway class. In principle, to obtain unbiased estimates of vehicle type relative frequencies, locations for observing classification counts should be randomly distributed on the highway system. This is the most important unknown factor in estimating vehicle travel from available classification count data. Another important issue deserving further attention is the fact that although weighting factors can remove sample bias, they also tend to increase the variance of estimators, especially when the sample is extremely maldistributed.

Experimental estimates of disaggregate vehicle travel by 4 census regions and 10 FHWA highway system classes were derived by using data from the HPMS and TWS data bases. The estimates suggest a much higher level of combination truck travel than official FHWA estimates. Because of the way the data were collected, there is reason to believe that the classification count estimates may be biased by a selection of locations on the highway network with above-average levels of truck traffic. This question deserves further attention.

The ability to estimate highway travel by vehicle type is limited by (a) a lack of comparable data for all states, (b) the gross spatial and temporal biases of existing traffic classification count samples, and (c) the unknown bias due to choice of observation location on the network. Some of the problems caused by a and b can be ameliorated and to some extent quantified by further analysis. The problem of locational bias and the final resolution of other data problems can ultimately be solved only by the use of statistically valid sampling techniques.

ACKNOWLEDGMENTS

The authors wish to thank David McElhaney, Frank Jarema, and Paul Svercl for their support and guidance from the inception of this research to its conclusion. The authors also thank Stephanie Floyd for her customary diligent yet expeditious preparation of this manuscript.

This research was sponsored by the FHWA under an Interagency Agreement under Martin Marietta Energy Systems, Inc., with the U.S. Department of Energy.

REFERENCES

1. U.S. Department of Transportation. 1977 Census of Transportation Truck Inventory and Use Survey. Report TC77-T-52. Bureau of the Census, U.S. Department of Commerce, May 1980.
2. Consumption Patterns of Household Vehicles. Report DOE/EIA-0328, Feb. 1983; Report DOE/EIA-0319, April 1982. Energy Information Administration, U.S. Department of Energy, 1983.
3. D.L. Gerlough and M.J. Huber. Statistics with Applications to Highway Traffic Analyses. Eno Foundation for Transportation, Inc., Westport, Conn., 1978.
4. W. Feller. An Introduction to Probability Theory and Its Applications, Volume 1. Wiley, New York, 1968.
5. A.M. Mood, F.A. Graybill, and D.C. Boes. Introduction to the Theory of Statistics. McGraw-Hill, New York, 1974.
6. Highway Performance Monitoring System, Vehicle Classification Case Study. Office of Highway Planning, FHWA, U.S. Department of Transportation, Aug. 1982.
7. L.A. Goodman. Analyzing Qualitative/Categorical Data. Abt Books, Cambridge, Mass., 1978.
8. M.B. Brown. Frequency Tables--P4F. In BMDP Statistical Software (W.J. Dixon, ed.), University of California Press, Berkeley, 1981.

Trip Chains and Activity Sequences: Test of Temporal Stability

LIDIA P. KOSTYNIUK and RYUICHI KITAMURA

ABSTRACT

A study of the temporal stability of urban travel patterns is reported. Daily trip records of individuals from southeast Michigan, obtained from origin-destination survey data sets of 1965 and 1980, are compared and analyzed for temporal changes. In addition to the traditional indicators of travel patterns such as trip rates, trip durations, and travel time budgets, the distribution of trips within trip chains, the sequencing and linkages of activities, and the time-of-day dependency of out-of-home activity participation are considered in the analysis. A series of hypotheses about the temporal stability of these indicators are tested by using log-linear models of contingency table analysis. The results indicate that, generally, these aspects of travel patterns are not stable over time. However, temporal stability is identified in the linkages and sequencing of activities and in the time-of-day dependencies of the decision to return home.

The patterns of person movement in an urban area are the result of the travel behavior of a large number of individuals. Travel choices of an individual arise from a fundamental set of activity choices that represent the individual's needs and desires. These choices are at the same time subject to a set of spatial and temporal constraints attributable to the individual's obligations and commitments, technologies and authorities available to him, and interpersonal linkages (1-5). The choices also reflect the interdependent nature of his activity participation decisions. An individual's current decision is influenced by previous as well as future decisions (6).

From this viewpoint, it is logical to characterize the environment of an individual in which his activity and travel choices are made in terms of activity-related variables in addition to the traditional transportation network and land use variables (e.g., the amount of time allocated to a set of activities required for the maintenance of a household varies, depending on the technologies available to it). The available technologies may also induce substitutions between out-of-home and in-home activities. These factors undoubtedly affect the type of activities the individual pursues, the amount of time allocated for the activities, the locations where they are pursued, and hence the activity and travel pattern over time and space.

The process of forecasting future travel patterns is based on the assumption that there is stability in the relationships that quantify trip making. Models of these relationships are developed from cross-sectional travel data and credibly reproduce the travel patterns from which they were derived. However, the models are generally not based on theory about the motivation of trip making and are not causal in the formal sense. They are descriptive and may be confounded by the environment in which they were developed.

In the 30 years since the introduction of travel forecasting models, many changes have occurred in the activity and travel environments of urban residents of the United States. The freeway and highway networks have been expanded and most urban areas have become decentralized. The number of people licensed to drive has increased continuously, as has the number of vehicles owned by households. Many labor-saving appliances and home entertainment devices have been introduced into households. Sociodemographic changes include the decreasing size of the households, the steady increase of single parent households, and the increasing participation of women in the labor force.

In light of these changes it is reasonable to expect that changes in travel behavior have also occurred during this time period. The viewpoint of "organizationalism" (7) may be taken, and it may be argued that it is not the environment that influences an individual's behavior, but it is the individual who chooses and modifies his environment. Nevertheless, it cannot be denied that the impacts of the changes have, in many cases, expanded tremendously the range of choices available to the individual, and also have eliminated some of the choices that were once available. In any event, the changes that have been observed in the past 30 years point to the need for reviewing the interrelationship between the travel environment and behavior. The first question that must be addressed is: Are there properties of trip making that have not changed over time, and if there are, are these the ones that are reflected in the forecasting procedure?

The results of an ongoing investigation into the stability of activity and travel patterns are summarized herein. Emphases are placed here on those aspects of urban travel behavior that closely represent the daily activity and travel pattern, but to which relatively little attention has been paid in the past. Specifically, examined in this study are the stability in the way a given number of sojourns are combined into trip chains, stability in the linkages of activity types in the daily pattern, and stability in the time-of-day dependencies of activity choices. The objective of the study is to infer, on the basis of statistical observations made on survey data, whether there is some regularity in the way a given set of activities is pursued and in the way trips are organized over the 1-day period.

APPROACH

Most studies of temporal stability in travel patterns (8-18) focused on the stability of travel forecasting models, especially trip generation and trip distribution models (8-10,12). A small number addressed the stability of limited aspects of travel patterns such as daily person trip rates, trip lengths, and travel budgets (13-16). Several studies claim stability of trip generation models at an aggregate level (8,11), whereas most studies cannot conclude that this kind of stability exists. Some (15) argue that it is not the separate components of trip making such as the number of trips or length of trips that remain stable over time, but rather a time budget for travel. They show that at least at the aggregate level travel time per traveler does appear to be stable. Others (16) found temporal differences in trip rates and trip lengths in two cities in western New York State, and claim that, although there is some empirical support for a travel budget at the household level, there was no evidence of such a budget at the individual level. There is evidence for temporal stability in the work trip (12), especially in the work trip of males (14).

In their previous efforts (19,20), the authors have examined temporal stabilities in various indicators of individuals' daily travel patterns, particularly the validity of the assumption that travel behavior of population subgroups remains stable over time. The results indicated that the subgroups defined in terms of the traditional variables--car ownership and household size--do not exhibit stability in their behavior, and also that the life cycle combined with car ownership yields a set of subgroups with relatively stable behavior. The latter result is perhaps because the life-cycle variable is most strongly correlated with the patterns and constraints of the daily activity and travel (3,21-23). Overall, it appears that trip making by only some specific subgroups of the population or for very limited trip purposes has been shown to have temporal stability. The state of the knowledge on the temporal stability of travel behavior and patterns can be best characterized as inconclusive [further discussions on previous studies can be found elsewhere (20)].

The conflicting findings as to the stability of the traditional indicators of travel patterns, such as the number of trips, trip length, and daily travel time budget, suggest that assuming such stability is at best groundless. The use of models based on such assumptions in stability analysis may limit the scope of the investigation rather than aid it. Accordingly, an approach with minimal assumptions as to the stability or variability of travel behavior is selected for empirical examination of this study.

The statistical tool chosen for the temporal stability analysis of this study is the log-linear model of multidimensional classification analysis (24). The log-linear model uses observations organized into a multiway frequency table according to a set of categorical variables. The analysis does not assume any relation about the effect of each variable, and if a nonlinear relationship exists, the model will depict it as such. Furthermore, the model is capable of representing interaction effects of arbitrary order, and it is extremely effective in travel behavior analysis where many variables in the survey data are discrete (e.g., number of trips) or categorical (e.g., sex and occupation).

Stability in behavior is tested by fitting a log-linear model that represents a given behavioral hypothesis. Application of the model to hypothesis testing is described elsewhere (19). The flexibility in specifying the log-linear model with higher-order interaction effects allows comprehensive investigation of the nature of the stability in travel behavior. One focus of the examination of this study is on the relative magnitude of the variation in travel patterns over time compared with the cross-sectional variations due to mode usage and work participation. Another interesting aspect to be examined is the stability in activity scheduling over the 1-day period. It is the intention of this study to conduct an extensive explorative analysis of the behavior through statistical examination of alternative behavioral hypothesis and to infer the stability that may exist in travel behavior.

SAMPLE

The results of two origin-destination surveys in southeast Michigan are used in this study. The first survey, which was a conventional large-scale home-interview survey, was conducted in 1965, and the second was conducted in 1980. The latter used a 1-day trip diary to collect trip records; its sample size is much smaller than that of the 1965 survey. A detailed comparison of these two surveys was made by the authors (20). This comparison indicates that it is reasonable to assume that the trip records obtained in both data sets are comparable in accuracy.

The same set of screening criteria was applied to both the 1965 and 1980 data files. This process eliminated from the sample of this study those individuals whose trip records were incomplete or inconsistent, whose paths on the survey day did not originate and terminate at home, who made trips outside the study area, and who were less than 18 years old. The sample of this study is thus strictly controlled and is very different than those used in other stability analyses. This would cause a problem if predicting areawide demand was the study objective. The use of controlled samples simplifies the design of data tabulations and makes the interpretation of their results more straightforward. The screening process resulted in in a sample of 218,284 trip records of 53,928 individuals from the 1965 file, and 8,248 trip records of 2,351 individuals from the 1980 file.

Individuals who did not make a trip on the survey day are not analyzed in this study because the original 1965 file does not contain records of such individuals. Accordingly, all averages are taken per tripmaker rather than per person. The tripmaker in this study is defined as an adult individual who, according to the trip records, made at least two trips on the survey day. This will not affect the analysis here if the probability that an individual will take a trip on a given day has not changed between 1965 and 1980. This appears to be a reasonable assumption for employed individuals' weekday travel patterns. If any difference exists in the probability, it is believed that the difference resulted in the underestimation in this study of the changes between the two time points. This possible bias in the result due to the limitation in the data sets must be kept in mind in interpreting the results of this study.

The individuals are classified into two groups according to the presence or absence of work trips in their activity schedules; those who made work trips on the survey day will be referred to as workers and the others will be called nonworkers. The individuals also are classified by a set of conditions collectively termed "mode usage." The "car users" include those individuals who held a driver's license, who had at least one car available to the household, and who made all trips by car (either as driver or passenger) or on foot. Those who did not satisfy all these conditions are referred to as "other" individuals. The type of out-of-home activity is defined in this study in terms of trip purpose categories.

TEMPORAL CHANGES IN OVERALL TRAVEL PATTERN INDICATORS

In this section the changes in traditional travel pattern indicators, such as trip rates and trip durations that are found between the 1965 and 1980 samples, are summarized. The data in Table 1 give the changes in the number of trips, number of sojourns, number of trip chains, and average number of sojourns per chain for four sample subgroups defined by mode usage and work participation. Overall, the number of trips made by a tripmaker decreased by 15.1 percent in 1980. Similar changes can be found for the number of sojourns and chains. It is notable that the declines are in general larger in the car-user subgroups, whereas the other mode user subgroups exhibit smaller changes. The other workers show small (and statistically not significant) increases in the indicators of mobility examined in this table. The data indicate that no stabilities exist in these basic indicators, and also that the temporal changes did not take place uniformly across the four subgroups.

The data in Table 2 give the mean number of trips made for respective out-of-home activity types. Quite notable are the decreases in 1980 of shopping, social-recreation, and serve-passenger activities that can be found irrespective of work participation or mode usage (the only exception is the slightly increased rate of serving passengers by the other mode group). The decline in shopping may be attributable to the decreased state of economy that the area was undergoing in 1980. The decrease in social-recreational activity may be an indication of the substitution of in-home activities for out-of-home activities in 1980 as a result of television sets and other home-entertainment appliances. The decreased rate of serve-passenger trips is perhaps due to the increased fraction of individuals with driver's licenses and also to increased car ownership. Other notable changes include increased school activity across the subgroups, and the increase in eating meals out of home, especially by nonworkers. Although the data show a slight decrease in work trips, the average number of work trips per employed person remained unchanged at 1.24.

The rapid decentralization that took place in the Detroit metropolitan area after 1965 is reflected in the general increase in the mean trip time (Table 3). This increase, however, is by no means uniform across the sample subgroups or trip types. For example, the data indicate that some types of trips show decreases in average duration. The home-to-work trips show a slight decrease, regardless of mode usage, and the non-home-based trips show relatively small differences between the two data sets, except for the nonworkers with other mode usage, who show a 43.2 percent increase. The temporal differences in the mean trip time vary widely from an increase of 35.8 percent to a decrease by 0.5 percent, depending on work participation and mode usage.

The variation across the sample subgroups is quite notable and is in fact statistically more significant than the variations over time. The mean total travel time per tripmaker (mean time budget) again varies widely across subgroups and over time from 60.48 to 87.07 min. The mean time budget is more stable between 1965 and 1980 than the number of trips, sojourns, or trip chains among the car user subgroups (Table 1). This, however, cannot be concluded for the other nonworker subgroup. The assumption of stability in trip rates or travel time budgets is not supported by these comparisons.

TABLE 1 Average Number of Trips, Sojourns, Chains, and Sojourns per Chain per Tripmaker: 1965 Versus 1980

Mode Usage	Work Participation		Subgroup Mean 1965*	1980	%Change
Car Users	Nonworkers	No. of Trips	4.51	3.70	-18.0
		No. of Sojourns	2.85	2.28	-19.9
		No. of Chains	1.66	1.41	-14.8
		Sojourns/Chain	1.72	1.62	-6.1
		Sample Size	16,121	620	
	Workers	No. of Trips	4.20	3.54	-15.7
		No. of Sojourns	2.67	2.19	-17.9
		No. of Chains	1.54	1.35	-11.9
		Sojourns/Chain	1.74	1.62	-6.8
		Sample Size	27,578	1,099	
Others	Nonworkers	No. of Trips	2.96	2.84	-4.2
		No. of Sojourns	1.77	1.63	-8.0
		No. of Chains	1.19	1.21	1.6
		Sojourns/Chain	1.49	1.35	-9.5
		Sample Size	5,299	318	
	Workers	No. of Trips	2.85	2.95	3.7
		No. of Sojourns	1.63	1.70	4.2
		No. of Chains	1.21	1.25	3.1
		Sojourns/Chain	1.35	1.36	1.0
		Sample Size	4,390	184	
Total	Total	No. of Trips	4.05	3.44	-15.1
		No. of Sojourns	2.54	2.10	-17.5
		No. of Chains	1.51	1.34	-11.1
		Sojourns/Chain	1.68	1.56	-7.1
		Sample Size	53,928	2,221	

*The 1965 file does not contain records of walk trips. In order to make the comparison more direct, the walk trips in the 1980 file are excluded from this tabulation.

TABLE 2 Average Number of Sojourns by Activity Type, Work Participation, and Mode Usage: 1965 Versus 1980

By Work Participation

	Workers 1965	Workers 1980	Nonworkers 1965	Nonworkers 1980
Work	1.34	1.24	--	--
School	0.02	0.07	0.06	0.23
Eat Meal	0.12	0.13	0.08	0.13
Personal Business	0.23	0.23	0.54	0.50
Shopping	0.27	0.19	0.86	0.64
Social-Recreation	0.24	0.14	0.59	0.35
Serve Passengers	0.28	0.13	0.45	0.20
Total	2.51	2.12	2.58	2.06
Sample Size	32,508	1,283	21,420	938

By Mode Usage

	Car Users 1965	Car Users 1980	Others 1965	Others 1980
Work	0.88	0.81	0.52	0.40
School	0.04	0.11	0.04	0.24
Eat Meal	0.12	0.15	0.05	0.06
Personal Business	0.36	0.34	0.31	0.35
Shopping	0.54	0.40	0.39	0.32
Social-Recreation	0.39	0.23	0.35	0.22
Serve Passengers	0.42	0.19	0.05	0.07
Total	2.74	2.22	1.71	1.66
Sample Size	43,699	1,719	10,229	502

Total

	1965	1980
Work	0.81	0.72
School	0.04	0.14
Eat Meal	0.11	0.13
Personal Business	0.35	0.34
Shopping	0.51	0.38
Social-Recreation	0.38	0.23
Serve Passengers	0.35	0.16
Total	2.54	2.10
Sample Size	53,928	2,221

Note: Excludes sojourns made by walk trips. The averages are per tripmaker.

TABLE 3 Average Trip Duration and Total Travel Time per Tripmaker

Mode Usage	Work Participation	Trip Type	Average Duration (min) 1965	1980	%Change
Car Users	Nonworkers	To Home	14.3	17.1	19.4
		Home to Other	14.2	17.0	19.7
		Non-Home Based	14.4	15.7	9.3
		Weighted Avg.	14.3	16.7	17.0
		Total Travel Time	64.4	61.6	-4.3
	Workers	To Home	21.4	22.8	6.4
		Home to Work	24.6	24.3	-1.2
		Home to Other	13.4	16.1	20.3
		Non-Home Based	17.1	18.5	8.1
		Weighted Avg.	19.8	21.4	8.3
		Total Travel Time	83.2	75.9	-8.8
Others	Nonworkers	To Home	21.2	28.7	35.6
		Home to Other	20.6	27.2	32.4
		Non-Home Based	18.7	26.8	43.2
		Weighted Avg.	20.4	27.8	35.8
		Total Travel Time	60.5	78.0	29.0
	Workers	To Home	31.9	31.9	0.0
		Home to Work	33.8	32.4	-4.3
		Home to Other	15.5	23.5	51.1
		Non-Home Based	24.1	22.3	-7.5
		Weighted Avg.	29.8	29.6	-0.5
		Total Travel Time	84.7	87.1	2.8
Total	Total	Weighted Avg.	18.6	21.3	14.5
		Total Travel Time	75.5	73.1	-3.1

Note: Excludes walk trips.

CHANGES IN TRIP-CHAINING BEHAVIOR

The data in Table 1 indicated that, overall, the average number of sojourns per trip chain has decreased in 1980 together with the number of sojourns and the number of trip chains. The aggregate tabulation result, however, is misleading, as the detailed examination in this section of the tendency in chaining trips shows. The distribution of individuals by the number of sojourns and chains made on the survey day is given in Table 4. The overall changes documented in Table 1 are also presented here as the differences between 1965 and 1980 in the marginal distributions of the number of sojourns and chains for both nonworkers and workers.

Further inspection of the data in Table 4 indicates that, in spite of the decreased average number of sojourns per trip chain, the individuals with a large number of sojourns pursued them in fewer trip chains in 1980. For example, 30 percent of the nonworkers who made six or more sojourns combined them into four or more trip chains in 1965. This percentage decreased to 5 percent in 1980, whereas the percentage of nonworkers who combined six or more sojourns into one or two trip chains increased to 72.5 from 43.4 percent. A similar tendency can also be found among workers. Given that four or more sojourns are pursued, the individuals in the 1980 sample consolidated them into fewer trip chains. Presumably, after the two energy crises, people are more concerned with energy and trip costs, and they plan and schedule daily out-of-home activities more conscientiously in 1980 than in 1965.

Statistical examination of the nature of the apparent instability in trip chaining is carried out by applying the log-linear model while considering mode usage and work participation as contributing factors. The results are summarized in Table 5. The hypothesis testing of this study examines the significance of interaction terms of the log-linear model and infers the magnitudes of the effects of the year and the other factors on the stability. The first model of Table 5 does not include any interaction terms and represents the null hypothesis that all factors are independent. Therefore, model 1 assumes that the distributions of the number of sojourns (S), chains (C), mode usage (M), and work participation (W) do not vary with year (Y); that is, across the two surveys. The large chi-square value indicates that the model does not fit the observation and the hypothesis is rejected.

Model 2 represents the null hypothesis that the distribution of work participation and mode usage has changed between 1965 and 1980 [i.e., the interaction effect involving the three factors (WMY) is significant], but the distributions of the number of chains and sojourns and their combinations have not

TABLE 4 Distribution of Individuals by Number of Sojourns and Number of Chains: 1965 Versus 1980

Work Trip	Year	No. of Sojourns	No. of Chains 1	2	3	≥4	Total	
Nonworkers	1965	1	100.0				100.0	[39.8]
		2	57.2	42.8			100.0	[22.8]
		3	43.9	39.6	16.5		100.0	[14.5]
		4	28.2	42.3	22.2	7.4	100.0	[8.7]
		5	21.4	39.9	25.1	13.5	100.0	[5.4]
		≥6	12.7	30.6	26.7	30.0	100.0	[8.9]
		Total	63.9	24.0	8.0	4.0	100.0	[100.0]
	1980	1	100.0				100.0	[48.2]
		2	51.4	48.6			100.0	[23.5]
		3	45.1	31.0	23.9		100.0	[13.6]
		4	37.0	34.2	20.5	8.2	100.0	[7.0]
		5	40.0	20.0	35.0	5.0	100.0	[3.8]
		≥6	22.5	50.0	22.5	5.0	100.0	[3.8]
		Total	71.4	20.7	6.9	1.0	100.0	[100.0]
Workers	1965	1	100.0				100.0	[41.6]
		2	33.4	66.6			100.0	[23.0]
		3	41.8	41.8	16.4		100.0	[14.3]
		4	27.2	53.8	15.0	3.9	100.0	[8.1]
		5	27.1	47.3	19.2	6.4	100.0	[5.1]
		≥6	27.4	41.7	20.1	10.9	100.0	[7.9]
		Total	61.0	31.3	6.1	1.5	100.0	[100.0]
	1980	1	100.0				100.0	[48.7]
		2	39.4	60.6			100.0	[20.6]
		3	48.0	34.1	17.9		100.0	[13.7]
		4	34.3	48.0	14.7	2.9	100.0	[7.8]
		5	44.2	42.3	9.6	3.8	100.0	[4.0]
		≥6	48.6	38.6	10.0	2.9	100.0	[5.4]
		Total	70.4	24.6	4.5	0.5	100.0	[100.0]

[ ]: Percentage of the row total to the grand total

TABLE 5 Testing the Variations in Chain-Sojourn Combinations by Year, Work Participation, and Mode Usage

Model	Hypothesis Tested	χ^2	df	α
1. C,S,W,M,Y	All factors are independent	7944.6	132	.0000
2. CS,WMY	WORK and MODE depend on YEAR, but CHAIN and SOJOURN remained unchanged	5186.0	119	.0000
3. CS,CWMY,SWMY	Distributions of CHAIN and SOJOURN changed, but not their combinations (CS)	509.8	63	.0000
4. CSW,CWMY,SWMY	CS combination varies by WORK, but not by MODE or YEAR	93.8	54	.0006
5. CSM,CWMY,SWMY	CS combination varies by MODE, but not by WORK or YEAR	485.5	54	.0000
6. CSY,CWMY,SWMY	CS combination varies by YEAR, but not by WORK or MODE	478.8	54	.0000
7. CSWM,CWMY,SWMY	CS combination varies by WORK and MODE, but not by YEAR	67.5	36	.0011
8. CSMY,CWMY,SWMY	CS combination varies by MODE and YEAR, but not by WORK	441.8	36	.0000
9. CSWY,CWMY,SWMY	CS combination varies by WORK and YEAR, but not by MODE	50.7	36	.0526

CHAIN (C) = No. of chains (1, 2, 3, ≥4).
SOJOURN (S) = No. of sojourns (1, 2, 3, 4, 5, ≥6).
YEAR (Y) = Year (1965, 1980).
MODE (M) = Mode usage (car only, others).
WORK (W) = Work participation (worker, nonworker).

changed. The significant chi-square value in the table is not surprising in light of the differences in the average number of chains, sojourns, and sojourns per chain found in Table 1.

Model 3 assumes that the expected cell frequencies of the multiway table vary depending on mode usage, work participation, as well as year, but these three factors affect only the marginal distributions of the number of chains and sojourns; the interaction effects that influence the expected frequencies of respective chain-sojourn combinations, given the marginal distributions, are not affected by the three factors. In other words, model 3 eliminates the hypothesis of the stability in the marginal distributions of the number of chains and sojourns (C an WMY interact, so CWMY is significant; similarly, SWMY is assumed to be significant), but assumes that their combinations are unaffected by work participation, mode usage, or year. This hypothesis is again rejected.

Models 4-6 assume that chain-sojourn combination patterns differ depending on, respectively, work participation, mode usage, and year. CSM, CSW, and CSY represent the three-way interaction terms. The small chi-square value of model 4, which assumes that chain-sojourn combinations depend on work participation, is quite notable. The result indicates that models 1-3 showed poor fits because the effect of work participation on chain-sojourn combinations was not represented in them. Inclusion of the year effect into the model (model 6 with CSY), on the other hand, does not show any remarkable improvement. The effect of the added work participation can be evaluated by taking the difference in the chi-square values between model 3 (with CS) and model 4 (with CSW). The result is a chi-square value of 509.77 - 93.80 = 415.97, with 9 degrees of freedom (df), a highly significant result. The difference between model 3 and model 6, on the other hand, indicates that the year effect is only marginally significant (509.77 - 478.82 = 30.95, with df = 9). Thus it can be concluded that the year difference does affect the chain-sojourn combination patterns, but the effect is much less substantial compared with that of work participation.

It is evident from the tabulation that the marginal distributions of the number of sojourns and the number of chains significantly differ between 1965 and 1980. The changes have occurred interactively with work participation and mode usage, as the importance of effects CWMY and SWMY in Table 5 indicates. The statistical examination of this section further indicates that, given the differences in their marginal distributions, combinations of the number of sojourns and chains depend not so much on the year as on work participation. The last model (model 9), which includes an interaction term of chain-sojourn combination, work participation, and year (CSWY), captures the differences between the two time points discussed at the beginning of this section and is not significantly different from the observation. However, the majority of the variation in the observation is already explained by model 4, and the addition of the year effect in model 9 provides a significant but marginal improvement to the goodness of fit. The patterns of complex daily travel involving multiple sojourn chains remained relatively stable between the two time points. In fact, variations within the year across the sample subgroups are more substantial than the variations across the surveys.

STABILITY IN ACTIVITY SEQUENCING AND LINKAGES

It seems logical to assume that there exist some patterns in the way a set of activities to be pursued on a given day are organized into an activity and travel schedule. Previous studies showed that activities with less flexibilities tend to be pursued earlier in the day (25) and also earlier in a trip chain before more flexible activities (26). Figure 1 shows the temporal stability of the tendency in sequencing activities in a trip chain. Similar tendencies were observed in data sets from several urban areas (26). The trip to serve passengers, which is often subject to tight interpersonal constraints (3,27), tends to be pursued before other activities by both workers and nonworkers. Work and school activities that follow serve-passenger trips are in general accompanied with rigidly fixed schedules. More discretionary and flexible activities such as social-recreation tend to be pursued last in the chain. While Figure 1 shows some differences among the significant sequencing relations in 1965 and those in 1980, they are caused by the small size of the 1980 sample. There exists no evidence that the way individuals schedule their daily activities have changed between 1965 and 1980.

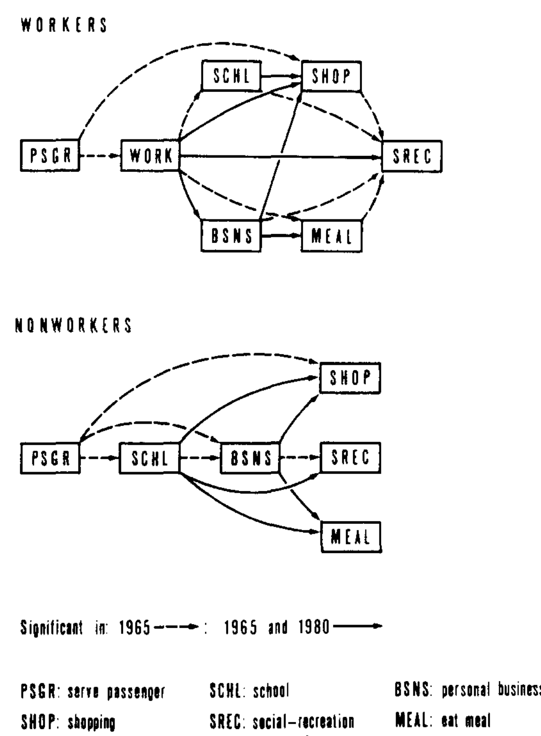


FIGURE 1 Hierarchy in activity sequencing in trip chains.

The patterns of out-of-home activity linkages also remained unchanged between 1965 and 1980. The data in Table 6 indicate this by presenting salient flows in activity transition matrices. The salient flow is defined in this study as the cell of a transition matrix whose observed frequency is significantly larger than the expected frequency. The chi-square value corresponding to $\alpha = 0.05$ (df = 1) is used as the criterion of significance. A transition matrix is developed by organizing into the matrix form the observed frequencies of transitions from one activity type to another. Only directly linked activities are analyzed in this table.

Most of the diagonal cells are significant, which indicates that the same types of activities tend to be pursued successively. The work-to-eating-meal and

TABLE 6 Salient Flows in Activity Transitions by Year

	Origin Activity	Destination Activity 1	2	3	4	5	6	7	Tmp. Home	Prm. Home
Nonworkers	2 School	-	▣	□					▣	□
	3 Eat Meal	-	□				□	□		
	4 Personal Business	-		□	▣	▣				
	5 Shopping	-				▣				□
	6 Social-Recreation	-		□			▣			▣
	7 Serve Passengers	-	□					▣	▣	
Workers	1 Work			▣					▣	▣
	2 School								▣	
	3 Eat Meal	▣					□			
	4 Personal Business	□			▣	▣			□	
	5 Shopping					▣	□		□	□
	6 Social-Recreation						▣			▣
	7 Serve Passengers	▣	▣			○		□	□	

□:Salient in the 1965 sample.
○:Salient in the 1980 sample.
▣:Salient in both 1965 and 1980 samples.

Tmp. Home: Temporary return to home
Prm. Home: Permanent return to home

eating-meal-to-work transitions salient in both 1965 and 1980 represent typical activity linkages in workers' daily activity schedules and can be found in analyses of other data sets (28). It is quite notable that the salient flows of the 1980 matrix form a subset of the 1965 flows. This is again because of the much smaller size of the 1980 sample. The only exception to this is the serve-passengers-to-shopping transition by the workers salient in 1980. This is perhaps a result of the increased participation of women in the labor force, who tend to pursue these activities more frequently than the male counterpart (21,29).

The quantitative similarity in the activity transition between 1965 and 1980 can be seen by the data in Table 7, which give the transition matrices by

TABLE 7 Transition Matrices of 1965 and 1980 by Work Participation

NONWORKERS

Year	Origin Activity	Destination Activity 1	2	3	4	5	6	7	8	Prm. Home	Total
1965	2 School		.028	.037	.088	.145	.092	.052	.153	.405	1.00
	3 Eat Meal		.012	.011	.074	.134	.185	.086	.098	.400	1.00
	4 Personal Business		.004	.021	.145	.182	.085	.045	.208	.310	1.00
	5 Shopping		.001	.012	.051	.190	.072	.039	.210	.429	1.00
	6 Social-Recreation		.002	.025	.042	.102	.151	.056	.132	.490	1.00
	7 Serve Passengers		.009	.014	.063	.106	.076	.158	.322	.252	1.00
	8 Home		.013	.032	.181	.273	.247	.254	-	-	1.00
	Total		.006	.020	.090	.170	.122	.098	.173	.320	1.00
1980	2 School		.036	.022	.032	.058	.047	.018	.252	.536	1.00
	3 Eat Meal		.008	.016	.062	.117	.094	.055	.187	.461	1.00
	4 Personal Business		.012	.039	.120	.153	.054	.025	.159	.437	1.00
	5 Shopping		.007	.032	.065	.169	.054	.017	.172	.484	1.00
	6 Social-Recreation		.005	.040	.069	.079	.098	.045	.117	.548	1.00
	7 Serve Passengers		.026	.015	.077	.077	.092	.092	.265	.357	1.00
	8 Home		.110	.072	.159	.267	.233	.159	-	-	1.00
	Total		.028	.037	.088	.145	.092	.052	.152	.405	1.00

WORKERS

Year	Origin Activity	1	2	3	4	5	6	7	8	Prm. Home	Total
1965	1 Work	.151	.003	.056	.050	.037	.029	.041	.212	.422	1.00
	2 School	.135	.010	.029	.035	.042	.059	.045	.253	.403	1.00
	3 Eat Meal	.559	.003	.004	.040	.039	.079	.045	.046	.185	1.00
	4 Personal Business	.178	.002	.028	.104	.090	.051	.037	.214	.295	1.00
	5 Shopping	.042	.001	.019	.037	.126	.054	.025	.223	.472	1.00
	6 Social Recreation	.054	.003	.034	.028	.046	.126	.052	.096	.562	1.00
	7 Serve Passenger	.252	.004	.020	.033	.042	.044	.151	.203	.250	1.00
	8 Home	.192	.015	.037	.146	.245	.230	.137	-	-	1.00
	Total	.168	.005	.040	.064	.084	.077	.066	.162	.334	1.00
1980	1 Work	.093	.008	.069	.062	.041	.020	.030	.182	.495	1.00
	2 School	.186	.010	.029	.000	.049	.039	.020	.294	.373	1.00
	3 Eat Meal	.497	.016	.011	.021	.048	.048	.021	.079	.259	1.00
	4 Personal Business	.168	.007	.040	.141	.111	.037	.024	.145	.327	1.00
	5 Shopping	.110	.004	.034	.046	.129	.046	.015	.114	.502	1.00
	6 Social-Recreation	.049	.005	.022	.038	.030	.082	.044	.077	.621	1.00
	7 Serve Passengers	.264	.031	.012	.049	.055	.031	.055	.166	.337	1.00
	8 Home	.227	.080	.080	.184	.171	.190	.067	-	-	1.00
	Total	.152	.019	.055	.079	.075	.053	.035	.139	.394	1.00

Note: Prm. Home = permanent return to home.

year and by work participation (nonworkers and workers cannot be analyzed together here because, by definition, the nonworkers' transition probabilities to work are zero). The intensity of activity linkages is now represented by transition probabilities. The larger probabilities of permanently returning home and the smaller probabilities of temporarily returning home in 1980 show the decreased number of sojourns and chains in the latter survey. Decreased probabilities of transitions involving social-recreation travel in 1980 are also notable.

The stability of the transition matrices is again examined by applying the log-linear model. The results are summarized in Table 8. The four factors considered in the analysis--activity at origin (O), activity at destination (D), mode usage (M), and year (Y)--are by no means independent for both nonworkers and workers (model 1). Introduction of the interaction term, which involves the activity at origin and activity at destination (OD) in models 2 and 3, represents the hypothesis that the distributions of activities at origin and destination are unaffected by the other factors and also that there exists a unique linkage intensity between each pair of activity types, and its intensity does not depend on the mode usage or year. Model 4 relaxes this hypothesis by allowing the marginal distribution of activities at origin and destination to vary, depending on the other two factors. These hypotheses are rejected.

Model 5, where the activity linkages are assumed to vary by mode usage but not by year, shows a relatively good fit, but it is significantly different from the observation for the workers [χ^2 = 198.6, df = 108, the probability (α) that the model represents the population from which the observations were obtained is less than 0.00005]. The result indicates a better fit of the model to the nonworkers' activity transitions (χ^2 = 130.0, df = 80, α = 0.0004). This result is rather counterintuitive because workers presumably have less flexibility as to their daily activity scheduling because of the rigid work schedule; therefore, changes of lesser magnitude in their daily activity patterns would be expected. It may be the case that workers are in general more mobile and their travel patterns tend to vary with the environment more than do the patterns of nonworkers.

Model 6 assumes that activity linkages are dependent on the year but not on the mode usage. This model exhibits poorer fits than model 5, especially for workers. Again, the variations in the transition matrix over time are less in their magnitudes than the cross-sectional variations across sample subgroups. The results of models 5 and 6 for workers together suggest that the changes in the activity linkage patterns cannot be explained by either mode usage or year, but their combined effects are important. It implies that the behavior of the two mode usage subgroups of workers changed differently between 1965 and 1980. This result is consistent with the earlier finding (19) that travel behavior of the respective car ownership subgroups has changed between the two time points.

STABILITY IN TIME-OF-DAY DEPENDENCY OF ACTIVITY CHOICE

The time-of-day dependencies of activity patterns arise from physiological requirements (e.g., sleeping), institutional elements (e.g., business and work hours), and perhaps purely habitual choices of the individuals (e.g., 3 o'clock tea). The institutional elements are broadly defined here to include factors that affect the supply of opportunities for activities (e.g., television show). The time-of-day dependencies of activities are naturally affected by these factors. The time-of-day dependencies of out-of-home activities are affected by additional factors, including the substitution of in-home and out-of-home activities (3,30).

Figure 2 shows the activity patterns of nonworkers and workers over the 1-day period by presenting the distribution of activity occurrence times by activity types for 1965 and 1980. Only those activity types with a sufficient number of observations are presented in the figure. The distributions clearly show the tendency mentioned earlier, in that activities of a more obligatory nature with less flexibility are pursued first earlier in the day. This can be seen most clearly in the workers' distributions. Naturally, work activity typically commences in the morning, and the figure shows a sharp peak around 8:00 a.m. The frequency of personal business increases in the afternoon after 3:00 p.m., with peaks between 3:00 and 6:00 p.m. Compared with personal business, shopping activity is less frequently engaged in during the morning

TABLE 8 Testing the Variation in Activity Linkages by Year and Mode Usage

		Nonworkers			Workers		
Model	Hypothesis Tested	χ^2	df	α	χ^2	df	α
1. O, D, M, Y	All factors are independent.	13656.9	200	.0000	22268.9	262	.0000
2. OD, M, Y	Activity linkages (OD), MODE, and YEAR are independent.	6653.5	160	.0000	5036.3	208	.0000
3. OD, MY	MODE varies over YEAR, but not the distributions of ORIGIN, DESTINATION, and OD linkages.	6286.8	159	.0000	4975.5	207	.0000
4. OD, OMY, DMY	Distributions of ORIGIN and DESTINATION activities vary depending on MODE and YEAR, not OD linkages.	260.9	120	.0000	1089.7	162	.0000
5. ODM, OMY, DMY	OD linkages vary by MODE, but not by YEAR.	130.0	80	.0004	198.6	108	.0000
6. ODY, OMY, DMY	OD linkages vary by YEAR, but not by MODE.	180.0	80	.0000	947.0	108	.0000

ORIGIN (O) = Origin activity categories; see Table 6.
DESTINATION (D) = Destination activity categories.
YEAR (Y) = Year (1965, 1980).
MODE (M) = Mode usage (car only, others).

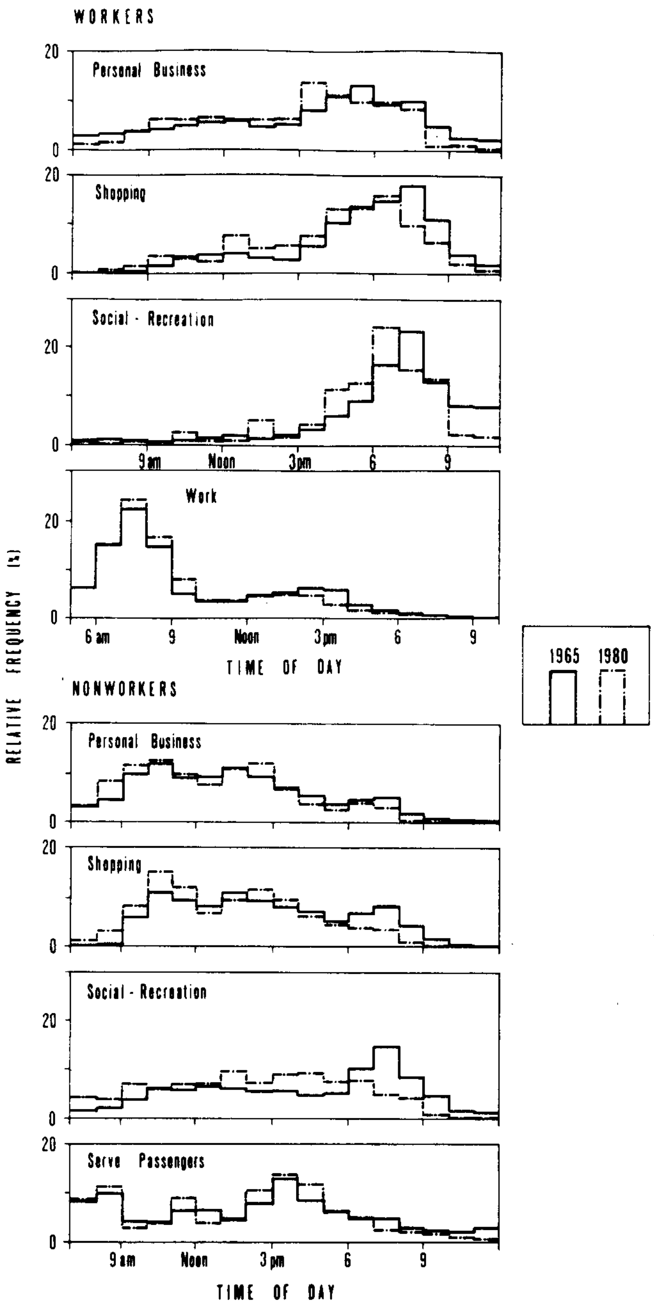


FIGURE 2 Distribution of activity occurrence times.

period, and the peaks after 6:00 p.m. are sharper. Social-recreation is the last activity to be pursued, with a small probability of engagement until 3:00 p.m. and sharp concentrations between 6:00 and 9:00 p.m. Although less pronounced, similar tendencies can be found from the distributions of nonworkers. The rather irregular patterns of serve-passenger activity reflect the typical time periods when chauffeuring of workers and children takes place.

There are certain differences between the 1965 and 1980 distributions. In general, the figure indicates that out-of-home activity engagement during the evening period has declined in 1980. This is most noticeable for social-recreation trips: The sharp peak of the nonworkers' distribution in 1965 has completely disappeared in 1980, the workers' 1980 peak is 1 hr earlier than in 1965, and the workers' engagement in this activity after 9:00 p.m. has declined markedly in 1980. It is not possible to identify the reasons for this change from the statistical analyses of the survey results. However, it appears logical to conjecture that the change is at least in part caused by the substitution of out-of-home activities by in-home activities as a result of widespread ownership of the television set and other entertainment devices in 1980.

The stability in the time-of-day dependencies in activity engagement is statistically examined by using the log-linear model. Only non-home-based choices, where the origin location is outside the home base, are analyzed here, and activities are represented by simplified activity type classification. Three categories are used for the analysis of nonworkers' patterns: out-of-home activity, returning home temporarily, and returning home permanently. The last category implies that the out-of-home activity schedule of the day is completed. The analysis of the workers used four categories consisting of the three activities just mentioned and work activity. The results are summarized in Table 9.

Models 8 and 9 show good fits to the observation. Model 8 assumes that activity choice by time of day (CT combination) depends on the mode usage but not on the year, whereas the distributions of activity choices (C) and their occurrences over the time of day (T) vary by mode and year. The good fit implies that the activity choice given the time of day when it is made has not changed between 1965 and 1980. The marginal distributions of the choices and their occurrence times, however, did vary and resulted in the overall differences in the out-of-home activity engagement, as seen in Figure 2. It may well be the case that the time-of-day dependencies of activity choice have not changed if all activities, including the in-home activities, are taken into account. The apparent changes in the out-of-home activity and travel patterns discussed in this paper may be attributable to this substitution effect.

CONCLUSIONS

Travel patterns are not stable over time. Changes can be found in many aspects of the travel behavior reported in the survey results. A few aspects of the behavior that were found in this study to possess temporal stability include the patterns in linking and sequencing activities. In both the 1965 and the 1980 data sets, individuals were found to pursue less-flexible activities before flexible activities and also to pursue activities of the same type successively. Time-of-day dependencies of activity choice also showed certain similarities, which indicates that obligatory and less-flexible activities tend to be pursued earlier in the day. These qualitative similarities, however, do not imply that the relationships that quantify these tendencies remained unchanged. For example, the transition probabilities among activity types are not stable over time even after the differences in the distribution of activity types between the two time points are taken into consideration. The way a given number of sojourns are organized into trip chains also showed clear differences, which suggests that the individuals in the 1980 sample organized a larger number of sojourns into a fewer number of trip chains when they pursued many sojourns.

Some log-linear models involving the time factor were concluded to fit the observations well and indicated that the probability of certain behavior, given a particular condition, is stable over time. For example, the probability of returning home or pursuing additional out-of-home activities given the time of day was found to be stable over time. One

TABLE 9 Testing the Variations in Non-Home-Based Activity Choice by Time of Day and Mode Usage

Model	Hypothesis Tested	Nonworkers χ^2	Nonworkers df	Nonworkers α	Workers χ^2	Workers df	Workers α
1. C, T, M, Y	All factors are independent.	14010.0	183	.0000	48835.2	250	.0000
2. CT, M, Y	CHOICE-TIME (CT) combinations, MODE, and YEAR are independent.	2754.3	151	.0000	3314.0	202	.0000
3. CT, MY	MODE varies over YEAR, but the distributions of CHOICE and TIME and CT combination do not depend on MODE or YEAR.	2489.2	150	.0000	3270.2	201	.0000
4. CT, CMY, TMY	Distributions of CHOICE and YEAR vary by MODE and YEAR, but not CT combinations.	203.9	96	.0000	486.7	144	.0000
5. CTM, MY	Distributions of CHOICE and YEAR and CT combinations vary by MODE, but not by YEAR.	439.6	100	.0000	645.7	134	.0000
6. CTM, CMY	Distribution of CHOICE varies by MODE and YEAR, in addition to the variations in Model 5.	371.0	96	.0000	577.4	128	.0000
7. CTM, TMY	Distribution of TIME varies by MODE and YEAR, in addition to the variations in Model 5.	189.8	68	.0000	317.7	102	.0000
8. CTM, CMY, TMY	CT combinations vary by MODE, but not by YEAR.	63.9	64	.4789	94.4	96	.5266
9. CTM, CMY, TY	CT combinations vary by MODE, but not by YEAR; Distribution of TIME does not depend on YEAR.	104.3	80	.0355	128.3	112	.1387
10. CTM, TMY, CY	CT combinations vary by MODE, but not by YEAR; Distribution of CHOICE does not depend on YEAR.	84.2	66	.0129	133.0	99	.0129
11. CTY, CMY, TMY	CT combinations vary by YEAR, but not by MODE.	164.3	64	.0000	431.6	96	.0000

CHOICE (C) = Non-home-based activity choice (out-of-home activity, return home temporarily, return home permanently) for nonworkers; (work, out-of-home activity, return home temporarily, return home permanently) for workers.
TIME (T) = Time of day (-8 am, 8-9 am, ..., 10-11 pm, 11 pm-).
YEAR (Y) = Year (1965, 1980).
MODE (M) = Mode Usage (car only, others).

conjecture that can be developed from the study results is that the way an individual develops his daily activity schedule is stable, but the outcome of the process (i.e., the out-of-home activity and travel pattern) varies, depending on the travel environment as the input to the scheduling process.

The sharp decline of the out-of-home social-recreational activities in the evening period observed in the 1980 sample also suggests that an unstable travel pattern may result from a stable activity pattern because of the substitution between in-home and out-of-home activities. The qualitative similarities in activity scheduling and sequencing also suggest that what the individuals do may not have changed very much over time, but how and where they do it--the concern of transportation planners--have. More extensive analysis that involves not only activity and travel patterns, the individual's attributes, and network and land use variables, but also more comprehensive representation of the travel environment, is a future task.

REFERENCES

1. T. Hargerstrand. What About People in Regional Science? Regional Science Association Papers, Vol. 24, 1970, pp. 7-21.
2. D. Damm. Theory and Empirical Results: A Comparison of Recent Activity Based Approach. In Recent Advances in Travel Demand Analysis (S. Carpenter and P.M. Jones, eds.), Gower, Aldershot, England, 1983, pp. 3-33.
3. M.I. Clarke, M.C. Dix, P.M. Jones, and I.G. Heggie. Some Recent Developments in Activity-Travel Analysis and Modeling. In Transportation Research Record 794, TRB, National Research Council, Washington, D.C., 1981, pp. 1-8.
4. R. Kitamura, L.P. Kostyniuk, and M.J. Uyeno. Basic Properties of Urban Time-Space Paths: Empirical Tests. In Transportation Research Record 794, TRB, National Research Council, Washington, D.C., 1981, pp. 8-19.
5. W.W. Recker, M.G. McNally, and G.S. Root. Chaining Behavior in Urban Tripmaking. Interim Report. Institute of Transportation Studies, University of California, Irvine, 1983.
6. R. Kitamura. Incorporating Trip Chaining into Analysis of Destination Choice. Transportation Research, Vol. 18B, No. 1, 1984.
7. R. Herz. The Influence of Environmental Factors on Daily Behavior. Environment and Planning A, Vol. 14, 1982, pp. 1175-1193.
8. D. Hill and N. Dodd. Studies of Trends of Travel Between 1954 and 1964 in a Large Metropolitan Area. In Highway Research Record 141, HRB, National Research Council, Washington, D.C., 1966, pp. 1-23.
9. E.J. Kannel and K.W. Heathington. Temporal Stability of Trip Generation Relationships. In Highway Research Record 472, HRB, National

Research Council, Washington, D.C., 1973, pp. 17-27.

10. H. Ashford and F. Holloway. Time Stability of Zonal Trip Production Models. Transportation Engineering Journal, Vol. 98, 1972, pp. 799-806.
11. W.J. Ladd, R.H. Larson, and K. Boka. A Trip Generation Model for the Detroit Area. Transportation and Land Use Study. Southeast Michigan Council of Governments, Detroit, 1969.
12. R.L. Smith, Jr., and D.E. Cleveland. Time-Stability Analysis of Trip Generation and Predistribution Modal-Choice Models. In Transportation Research Record 569, TRB, National Research Council, Washington, D.C., 1976, pp. 76-86.
13. S.J. Bellomo, R.B. Dial, and A.M. Voorhees. Factors, Trends, and Guidelines Related to Trip Length. NCHRP Report 89. HRB, National Research Council, Washington, D.C., 1970, 59 pp.
14. C. Doubleday. Some Studies of the Temporal Stability of Trip Generation Models. Transportation Research, Vol. 11, 1977, pp. 255-263.
15. Y. Zahavi and J.M. Ryan. Stability of Travel Components Over Time. In Transportation Research Record 750, TRB, National Research Council, Washington, D.C., 1980, pp. 19-26.
16. G.S. Cohen and M.A. Kocis. Components of Change in Urban Travel. Prelim. Res. Report 159. New York State Department of Transportation, Albany, 1979, 46 pp.
17. J.D. Downes, S. Johnsen and D. Morrell. Household and Person Trip Generation Models. TRRL Supplementary Report 401. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1978.
18. F.W. Memmott III. Home-Interview Survey and Data Collection Procedures. In Highway Research Record 41, HRB, National Research Council, Washington, D.C., 1963, pp. 7-12.
19. L.P. Kostyniuk and R. Kitamura. Temporal Stability of Urban Travel Patterns. Presented at World Conference on Transport Research, Hamburg, Germany, 1983.
20. L.P. Kostyniuk and R. Kitamura. Inferring Temporal Changes in Travel Behavior from Origin-Destination Survey Data. Presented at 2nd International Conference on New Survey Methods in Transport, Pikolbin, New South Wales, Australia, Sept. 1983.
21. L.P. Kostyniuk and R. Kitamura. Life Cycle and Household Time-Space Paths: Empirical Investigation. In Transportation Research Record 879, TRB, National Research Council, Washington, D.C., 1982, pp. 28-37.
22. D.K. Boyle and J.E. Chicoine. Application of the Life Cycle Concept to Trip Generation Procedures. Planning Division, New York State Department of Transportation, Albany, 1982, 123 pp.
23. I. Salomon and M. Ben-Akiva. Life-Style Segmentation in Travel Demand Analysis. In Transportation Research Record 879, TRB, National Research Council, Washington, D.C., 1982, pp. 37-45.
24. M.B. Brown. Screening Effects in Multidimensional Contingency Tables. Applied Statistics, Vol. 25, 1976, pp. 37-46.
25. R. Maw. Analysing Demand for Leisure Facilities. Built Environment, Vol. 1, 1972, pp. 114-118.
26. R. Kitamura. Sequential, History-Dependent Approach to Trip-Chaining Behavior. In Transportation Research Record 944, TRB, National Research Council, Washington, D.C., 1983, pp. 13-22.
27. R. Kitamura. "Serve-Passenger" Trips as a Determinant of Travel Patterns. In Recent Advances in Travel Demand Analysis (S. Carpenter and P.M. Jones, eds.), Gower, Aldershot, England, 1983, pp. 137-162.
28. F.E. Horton and W.E. Wagner. A Markovian Analysis of Urban Travel Behavior: Pattern Response by Socioeconomic-Occupational Groups. In Highway Research Record 283, HRB, National Research Council, Washington, D.C., 1969, pp. 19-29.
29. S. Hanson and P. Hanson. The Impact of Women's Employment on Household Travel Patterns: A Swedish Example. In Women's Travel Issues: Research Needs and Priorities (S. Rosenbloom, ed.), Office of University Research, U.S. Department of Transportation, 1978, pp. 127-169.
30. J. Jacobson. Models of Non-Work Activity Duration. Ph.D. dissertation. Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, 1978.

Comparative Analysis of the Transferability of Disaggregate Automobile-Ownership and Mode-Choice Models

ERIC I. PAS and FRANK S. KOPPELMAN

ABSTRACT

In this paper the study of model transferability is extended to disaggregate models of automobile-ownership level. Models of automobile ownership and mode to work are estimated and transferred among sectors of a metropolitan region. The transfer effectiveness of these models is evaluated by using previously developed disaggregate and aggregate measures of model transfer effectiveness. The automobile-ownership models are found to have a high degree of transfer effectiveness in this context, higher than the transfer effectiveness of mode-choice models in the same context. It is concluded that previous findings about the effectiveness of model transfer, based on studies of mode-choice models, can be extended to automobile-ownership models.

The application of travel demand models estimated on observed data for prediction of conditional future behavior in the same or other context is commonly undertaken as part of the transportation systems analysis process (1). The application of a model in a context other than that in which it was originally estimated is described as model transfer. Model transfer is likely to be effective in predicting behavior in the application context if the transferred model will contain useful information about the behavioral phenomenon of interest in the application context. Models that contain such useful information are described as transferable. Model transferability is necessarily conditional on similarity of the underlying behavioral process in the estimation and application contexts and the adequacy of the model to represent that behavior (2). A number of studies of transferability of disaggregate travel choice models have been undertaken in recent years. Most of these studies consider mode choice (2-5), whereas some examine frequency choice (6,7).

The goal of this study is to extend the analysis of the transferability of travel choice models to the related choice of automobile ownership. The transferability of automobile-ownership choice models is analyzed and the transferability of these models is compared to that of mode-choice models. These analyses were undertaken in the context of an artificial transfer situation created by dividing the Washington, D.C., region into three geographically distinct sectors. These sectors are distinctly different in terms of the demographic characteristics of their populations, such as household size, household income, and automobile ownership, and with respect to travel time and cost to the central business district (CBD) by both car and bus transit (2).

Automobile-ownership and mode-choice models are estimated for each sector, and the transfer effectiveness of each model to the other two sectors is examined. This analysis was undertaken within a single urbanized area to reduce the confounding effect of differences in variable definition, measurement of level-of-service variables, and sampling procedures between metropolitan areas. Previous studies of the transferability of disaggregate mode-choice models suggest that the results of intra-area transfer studies are indicative of inter-area transfer effectiveness.

MODEL STRUCTURE AND ESTIMATION

Models of Travel and Related Choices

Travel behavior is commonly analyzed in the four steps embodied in the traditional aggregate urban transportation model system: trip generation, trip distribution, modal split, and network assignment (1,8,9). The comparable choices for disaggregate analysis are trip frequency (whether or not to make a trip), destination, mode, and path choice. An important issue in travel analysis revolves around the structure of these choices and the models that represent them.

Charles River Associates (10,11) derived a sequential formulation of the choice process and applied it to estimation of choices of shopping trip frequency, mode, destination, and time of day. Ben-Akiva (12) argued that certain of these choices are behaviorally joint and that they should be represented by a joint or simultaneous choice model. He also demonstrated that sequential model estimations may be quite different from those obtained by estimation of the corresponding simultaneous model. However, the differences in parameter estimates reported were not statistically significant at any reasonable level, and the goodness-of-fit measures for the simultaneous and sequential models were essentially the same. Ben-Akiva and Lerman (13) extended the individual choice structure to form a hierarchical model of travel and travel-related choices. In this hierarchy mobility choices, including residential location, automobile ownership level, and breadwinner mode choice to work, are assumed to be made jointly. Decisions on trip frequency, destination, and mode for nonwork trips are assumed to be made jointly but conditional on the higher-level mobility choices.

The discussion of choice model structure is based on behavioral conjecture about the sequence of the (unobserved) decision process employed by the tripmaker. More recently, McFadden (14) suggested an alternative theoretical basis for mathematically structuring multidimensional choice models. Specifically, he formally derived the nested logit model that takes account of similarity among alternatives with respect to excluded variables. In this structure, the mathematical form of the choice model represents an interdependence among a subset of alternatives due to the sharing of common unobserved attributes rather than a sequential dependence among choices. This theoretical approach leads to a similar mathematical form of the choice model as that obtained based on choice sequence.

Choice of Automobile Ownership and Mode to Work

These concepts were applied to the choice of automobile-ownership level and breadwinner mode to work. In this paper the component models of a sequential choice model, with mode choice conditional on automobile ownership, are examined. Excluded as conceptually unreasonable were mutual independence of these choices and the sequential model with automobile ownership conditional on mode choice to work. In a previous paper (15) the authors estimate and evaluate the joint choice model of automobile ownership and mode to work, and compare transferability of the joint and sequential model structures. The utility of a joint automobile ownership/mode to work alternative is defined by

$$U_{A,M} = V_{A,M} + \epsilon_{A,M} \quad (1)$$

where

$U_{A,M}$ = utility of automobile ownership A and mode M,
$V_{A,M}$ = systematic portion of that utility, and
$\epsilon_{A,M}$ = unobserved stochastic portion of that utility.

A sequential model of the choice of automobile ownership and mode to work can be developed by assuming that the stochastic component of utility in Equation 1 can be additively separated. The nested logit model is obtained under the assumption that

$$\epsilon_{A,M} = \epsilon_A + \epsilon_{AM} \quad (2)$$

where ϵ_{AM} is that portion of the stochastic utility that jointly varies over automobile ownership and mode and is Gumbel distributed with parameter λ^{-1}, and ϵ_A is that portion of the stochastic utility that varies only over automobile ownership and is distributed such that the sum $\epsilon_M + \epsilon_{AM}$ is Gumbel distributed with parameter 1.

In this case the conditional mode and marginal automobile-ownership choice models are of the form

$$P(M/A) = \exp[(V_M + V_{AM})/\lambda] / \sum_{M'/A} \exp[(V_{M'} + V_{AM'})/\lambda] \quad (3)$$

and

$$P(A) = \exp(V_A + \lambda\tilde{\Gamma}_A)/\sum_{A'} \exp(V_{A'} + \lambda\tilde{\Gamma}_{A'}) \quad (4)$$

where

$P(M/A)$ = probability of choosing mode M conditional on automobile ownership A,
$P(A)$ = marginal probability of choosing automobile ownership A,
V_A = that portion of observed utility that is strictly related to automobile-ownership level,
V_M = that portion of observed utility that is strictly related to mode,
V_{AM} = remaining portion of observed utility that is determined jointy by automobile ownership and mode,
λ = measure of dissimilarity between pairs of mode alternatives conditional on automobile ownership, and
$\tilde{\Gamma}_A$ = expected value of choosing the best mode given automobile ownership A.

The mathematical definition of $\tilde{\Gamma}_A$ is given by

$$\tilde{\Gamma}_A = \ln \sum_{M'/A} \exp[(V_{M'} + V_{AM'})/\lambda] \quad (5)$$

The estimation procedures for the sequential model structure are well developed and are documented in the literature (12,14,16). The basic procedure is to

1. Estimate the conditional portion of the model described in Equation 3 (note that λ cannot be estimated, but ratios of β/λ can be estimated, where β is a parameter in the utility function),
2. Compute the expected value of the set of conditional alternatives by using Equation 5, and
3. Estimate the marginal choice model as represented in Equation 4.

The estimation process is based on maximum likelihood procedures in steps 1 and 3.

RESEARCH DESIGN

Data and Model Specification

The data used were collected by the Washington Council of Governments in 1968 as part of a general effort to develop models of travel demand and transport system operations. A portion of these data was used, which describes breadwinners who made a work trip from their residence to work place in the CBD. (Note that breadwinners are defined as the household member working in the highest job category.) The data set includes a total of 2,654 persons and includes characteristics of the individual and household; level-of-service data for the work trip by drive alone, shared ride, and transit; and the mode chosen.

Previous studies of disaggregate choice models employed data from the Washington, D.C., data set. In particular, Lerman and Ben-Akiva (17) used these data to estimate joint choice models of automobile ownership (zero, one, two cars) and mode to work (car, transit). The specifications used in the present research are based on this previous work. The specification of a joint choice model is selected initially and compatible specifications are developed for the conditional and marginal choice models.

The choices of interest in this study are automobile-ownership level and breadwinner mode to work. The alternatives for automobile ownership are defined as zero, one, or two or more cars. The alternatives for mode to work include drive alone, shared ride, and transit. Two assumptions are made about the availability of particular alternatives. First, it is assumed that a household with no licensed drivers cannot choose to own an automobile. Second, if the work tripmaker does not have a driver's license, he is assumed not to be able to choose the drive-alone alternative.

There are no other assumed restrictions on alternative availability. The data set includes only individuals living in areas served by transit. Thus the transit alternative is available to everyone. Shared ride is assumed to be available to everyone. It is not assumed that the level of household income places any restriction on the maximum number of cars owned or available to the household.

Next, the utility function for each alternative is formulated. It is expected that the joint choice of automobile ownership and mode choice to work will be influenced by the level-of-service characteristics of the work trip by ride alone, shared ride, and transit; the differential travel capabilities of the household with different levels of automobile ownership; and the socioeconomic characteristics of the individual and household.

The general specification adopted by Lerman and Ben-Akiva (17) was followed, but modified to account for differences in alternatives (three mode-choice alternatives were included in this research) and limitations in the data available to the authors. First, transportation level-of-service variables were included. These are in-vehicle and out-of-vehicle travel time and out-of-pocket travel cost. Second, housing attributes are represented in terms of whether the residence is a single-family house. This characteristic is selected to take account of the availability of parking space, and this variable is associated with the two-or-more-automobile ownership alternative. Third, three socioeconomic variables were included. Household income is used to modify the importance of out-of-pocket travel costs. Number of licensed drivers is used to modify the utility of different levels of automobile ownership (the utility of owning increased numbers of vehicles increases with the number of drivers in the household). An indication that an individual is a government worker is used to represent the effect of work place incentives on the value of the shared-ride mode. Finally, the average effect of excluded variables is represented by constants for different automobile-ownership levels and different mode choices.

These specifications exclude two variables used by Lerman and Ben-Akiva (17): automobile-ownership costs and accessibility to nonwork locations for households with and without automobiles. The Washington data set does not include information on automobile-ownership costs. It was preferred to exclude this variable rather than include a fixed average annual cost per vehicle that is invariant across households. The accessibility measure used by Lerman and Ben-Akiva (17) represents the value of increased automobile ownership in improving household access to the opportunities other than work in the spatial environment. Although this is a useful variable, the data necessary to formulate it were not available to the authors.

A description of each variable included in the specifications of the automobile-ownership and mode-choice models is presented in Table 1. The generalized price variable (Equation 5) is included to capture the effect of modal utilities on automobile-ownership choice.

TABLE 1 Specification of Conditional Mode and Marginal Automobile-Ownership Choice Models

Explanatory Variable	Description of Variable	Conditional Mode-Choice Model	Marginal Automobile-Ownership Choice Model
DUMA and DUMSR	Dummy variables, specific to drive-alone and shared-ride alternatives	X	
DUM1CAR and DUM2CAR	Dummy variables, specific to the one- and two-car alternatives		X
CDA and CSR	Number of cars, drive-alone and shared-ride interaction variables	X	
GWSR	Dummy variable that indicates if the breadwinner is a government worker; specific to the shared-ride alternatives	X	
STRDUM	Dummy variable that indicates whether the household resides in a single-family structure; specific to the one- and two-car alternatives		X
IDLIC	The inverse of the number of driver's licenses in the household for the one-car alternatives; twice the inverse of the number of driver's licenses for the two-car alternatives		X
TTT	Round trip total travel time (min)	X	
OVTTD	Round trip out-of-vehicle travel time (min) divided by one-way distance (miles)	X	
OPTCINC	Round trip out-of-pocket travel cost (cents) divided by annual household income ($000s)	X	
GENPRICE	Generalized price of mode of travel for a given level of automobile ownership		X

Note: An X indicates that the explanatory variable is included in the particular model.

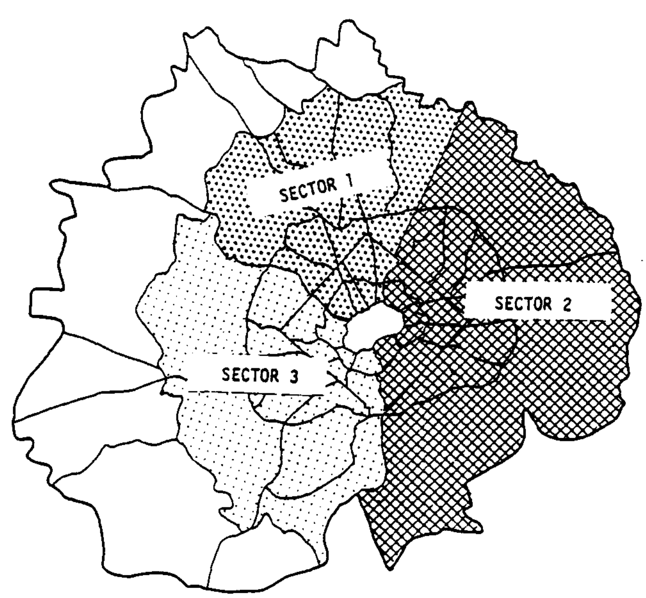


FIGURE 1 Estimation sectors in Washington region.

Analysis of Model Transferability

An artificial environment was created for transferability analysis by dividing the Washington area into three geographically distinct sectors, as shown in Figure 1. That is, the opportunity to examine transferability was created in a situation where there are no differences in variable definitions, data-collection methods, and characteristics of the metropolitan area environment. These advantages are important in developing an understanding of transferability. It is recognized that the issue of intraregional transferability is less of a concern than that of interregional transferability. However, earlier studies indicate that intraregional transfer results are indicative of interregional transfer effectiveness (18,19). The marginal automobile-ownership and conditional mode-choice models were estimated for each of these three sectors, and the transferability of each model to the other two sectors was examined.

The transferability of the different models was evaluated in terms of the ability of the transferred model to describe the observed behavior in the application context. This is accomplished by examining the accuracy of disaggregate and aggregate predictions using the transferred model in the application context in absolute terms and relative to the predictive accuracy of the corresponding locally estimated model. The specific measures to be used and their properties are developed in earlier work (2). A summary description of these measures is presented here. The disaggregate transferability measures (Table 2) are based on the likelihood that the data observed in the application environment were generated by the choice process described by the transferred model. The transfer likelihood ratio index is analogous to the conventional likelihood ratio index or rho-square measure (20). It compares the log likelihood of the transferred model to the log likelihood of a base (equally likely or market-shares) model. The transfer index compares the prediction effectiveness of the transferred model over the base model relative to the prediction effectiveness of a locally estimated model.

The aggregate measures of transferability (Table 3) evaluate the ability of the model to replicate observed choice frequencies in prediction for aggregate groups, using the explicit enumeration aggregation procedure (21). This is done by measuring the difference between the observed and predicted number of individuals selecting each alternative in each aggregate group. Specifically, the root-mean-square-error (RMSE) measure is used to represent the expected relative or proportional error in a typical aggregate prediction (22), and the relative aggregate transfer error is the ratio of transfer and local RMSE.

The disaggregate and aggregate transfer test statistics developed by Koppelman and Wilmot (2) are not reported here because these statistics were found to be less useful in the analysis of transferability than the index measures previously discussed. The transfer test statistics are reported in Koppelman and Pas (15).

EMPIRICAL RESULTS

Estimation Results

Models of mode choice conditional on automobile ownership and of marginal automobile-ownership choice are estimated for each of the three sectors by using the specifications previously described. The estimation results are given in Tables 4 and 5, respectively. These models are all significant at high levels relative to both the equally likely and market-share base models and account for a reasonable proportion of the behavioral variation in the data. Note that the marginal automobile-choice models have substantially higher likelihood ratio index (rho-square) values than the mode-choice models, despite the limited specification of the automobile-ownership model.

All the parameters in the conditional mode-choice models are highly significant ($p < 0.01$), except those associated with out-of-pocket travel cost and out-of-vehicle travel time. All the parameters in the marginal automobile-ownership choice models are statistically significant ($p < 0.01$), except the parameter of the inclusive price of travel mode in the automobile-ownership model for sector 2. Thus, from a statistical perspective, the models are extremely satisfactory. Furthermore, all parameter estimates that are statistically different from zero have acceptable signs. The parameters for the generalized price of mode of travel in the automobile-ownership models are expected to be between zero and one. Although the parameters obtained in two sectors are greater than one, they are not significantly different from one.

TABLE 2 Disaggregate Indices of Transferability

Measure	Definition	Description
Transfer likelihood ratio index, $\rho_i^2(\beta_j)$	$\rho_i^2(\beta_j) = 1 - [LL_i(\beta_j)/LL_i(\text{BASE})]$ where $LL_i(\beta_j)$ is the log likelihood that the behavior observed in context i was generated by the model estimated in context j (with parameters β_j)	This index is similar in form to the commonly used rho-square measure proposed by McFadden *(20)*; the index is bounded by one; the base model may be an equal-shares or market-shares model
Transfer index, $TI_i(\beta_j)$	$TI_i(\beta_j) = [LL_i(\beta_j) - LL_i(\text{BASE})]/[LL_i(\beta_i) - LL_i(\text{BASE})]$	This index measures the predictive accuracy of the transferred model relative to a locally developed model; the index has an upper limit of unity; the base model may be an equal-shares or market-shares model; the transfer index is related to the transfer likelihood ratio index by $TI_i(\beta_j) = \rho_i^2(\beta_j)/\rho_i^2(\beta_i)$

TABLE 3 Aggregate Indices of Transferability

Measure	Definition	Description
Root-mean-square error (RMSE)	$RMSE_i(\beta_j) = \left(\sum_{m,z} \hat{N}_{mz} REM_{mz}^2 / \sum_{m,z} \hat{N}_{mz}\right)^{1/2}$ where REM_{mz} is the relative error measure in prediction alternative m in group z; i.e., $REM_{mz} = (\hat{N}_{mz} - N_{mz})/\hat{N}_{mz}$ where $\hat{N}_{mz}$ is the number of persons in group z predicted to choose alternative m, and N_{mz} is the number of per sons in group z observed to choose alternative m.	This index measures the average relative error in prediction weighted by the size of the prediction element
Relative aggregate transfer error (RATE)	$RATE_i(\beta_j) = RMSE_i(\hat{\beta}_j)/RMSE_i(\hat{\beta}_i)$	This index measures the aggregate error of the transferred model relative to the local model

TABLE 4 Estimation Results: Conditional Mode-Choice Model

Variable	Estimated Parameter Values (t statistics) Sector 1		Sector 2		Sector 3	
DUMDA	-2.71	(7.31)	-1.79	(4.81)	-3.19	(7.26)
DUMSR	-2.35	(10.91)	-1.87	(9.63)	-2.36	(7.78)
CDA	1.67	(8.35)	1.57	(7.35)	2.08	(8.45)
CSR	1.20	(7.72)	1.33	(9.23)	1.43	(6.75)
GWSR	.77	(5.01)	.48	(3.33)	.60	(3.77)
TTT	- .038	(6.06)	- .018	(3.53)	- .021	(3.84)
OVTTD	.78	(.13)	- .052	(.88)	- .096	(1.23)
OPTCINC	.19	(1.44)	.0018	(.17)	.014	(.84)
Number of Cases	944		961[a]		746	
Number of Observations	2648		2582		2165	
Log Likelihood						
At Zero	-962.5		-933.7		-790.0	
At Market Share	-904.4		-899.7		-771.6	
At Convergence	-778.0		-812.6		-690.5	
Likelihood Ratio Statistic						
Zero Base	368.9		242.3		198.9	
Market Share Base	252.8		174.2		162.2	
Likelihood Ratio Index (ρ^2)						
Zero Base	.192		.130		.126	
Market Share Base	.140		.097		.105	

[a]There were three cases in the data set in which the household reported having zero drivers and also reported having one car available. Because these cases selected a non-feasible alternative, they were omitted from the analysis.

TABLE 5 Estimation Results: Marginal Automobile-Ownership Model

Variable	Estimated Parameter Values (t statistics) Sector 1		Sector 2		Sector 3	
DUM1CAR	4.46	(8.83)	4.79	(9.01)	6.24	(7.83)
DUM2CAR	4.50	(4.72)	5.59	(5.44)	5.47	(3.31)
STRDUM	1.00	(4.96)	.92	(5.12)	1.19	(5.63)
IDLIC	-4.60	(10.73)	-4.23	(11.57)	-5.64	(8.95)
GENPRICE	1.32	(3.92) (0.95)[a]	.40	(1.16) (1.74)[a]	1.79	(3.02) (1.33)[a]
Number of Cases	855		832		718	
Number of Observations	2565		2496		2154	
Log Likelihood						
At Zero	-939.3		-914.0		-788.8	
At Market Share	-781.1		-776.6		-577.4	
At Convergence	-596.6		-622.6		-426.6	
Likelihood Ratio Statistic						
Zero Base	685.4		582.9		724.3	
Market Shares Base	369.0		308.0		301.6	
Likelihood Ratio Index (ρ^2)						
Zero Base	.365		.319		.459	
Market Share Base	.236		.198		.261	

[a]T-statistics for the generalized price variable are formulated against the null hypotheses of $\beta = 0$ and $\beta = 1.0$.

Transferability Analysis

The transferability of the conditional mode and marginal automobile-ownership choice models is examined through use of the measures previously outlined. The transferability of the estimated models is evaluated in terms of parameter transferability, disaggregate prediction accuracy, and aggregate prediction accuracy. Examination of the hypothesis that the estimated model parameters describe the population behavior in the application context (15) rejects the transferability of the alternative specific constants in both the automobile-ownership and mode-choice models. Thus in this paper partial, rather than full, model transfer is considered. That is, the transferability analysis results that follow are based on models in which the alternative specific constants are adjusted to match the aggregate choice shares in the application context.

Disaggregate Transferability Prediction Indices

The ability of the conditional mode and marginal automobile-ownership choice models estimated in each sector to predict the disaggregate behavior observed in each of the other sectors is examined by use of the transfer likelihood ratio index and the transferability index evaluated against a market-share reference. These results are given in Tables 6 and 7 for each sector pair and with pooled values across all transfers (19). The transfer rho-square measures highlight two interesting facets of this analysis. First, it is observed that the rho-square values are highest for transfers into contexts that have high rho-square values for locally estimated models. For example, the automobile-ownership model provides the best fit to the sector 3 data, and the transfer rho-square measures are higher for transfers into sector 3 than into sectors 1 and 2. Second, it is observed that the transfer rho-square measures for the marginal automobile-choice model are consistently higher than those for the conditional mode-choice model, despite the apparently limited specification used for automobile ownership.

The transfer indices reported for the different models across sector pairs are generally quite high (greater than 0.86 in every case). The transfer indices for the marginal automobile-choice model are generally higher (four of six cases) than for the conditional mode-choice model. The pattern of transfer indices among sector pairs (which directional pairs have higher or lower transferability) varies between the two models. However, it appears that high transferability index values are obtained for transfer into sectors with a high local rho-square for the corresponding model. That is, it appears that model transferability measured by the transfer index is best in contexts in which behavior can be most effectively described by the particular model specification.

Overall, the disaggregate transferability prediction indices indicate that both conditional mode-choice and marginal automobile-ownership models are highly transferable between sector pairs. These results also indicate that transferability is generally higher for transfer into sectors that have high local rho-square values.

Aggregate Transfer Prediction Indices

RMSE is used to summarize the aggregate prediction error in both local and transfer prediction, and the relative values of RMSE are used to describe the degree to which transferred models increase aggre-

TABLE 6 Disaggregate Transferability Prediction Indices: Conditional Mode-Choice Model

		PREDICTING ON		
		Sector 1	Sector 2	Sector 3
ESTIMATED ON	Sector 1	.140 (1.00)	.083 (0.86)	.097 (0.92)
	Sector 2	.130 (0.93)	.097 (1.00)	.100 (0.95)
	Sector 3	.133 (0.95)	.092 (0.95)	.105 (1.00)

Composite Measures[a]

Transfer Likelihood Ratio Index = .106

Transfer Index = .93

Note: The base for computation of the transfer likelihood ratio index and the transfer index measures reported here is the market-shares model.

[a]Composite measures are weighted averages of the corresponding measures across multiple transfers *(19)*.

TABLE 7 Disaggregate Transferability Prediction Indices: Marginal Automobile-Ownership Choice Model

		PREDICTING ON		
		Sector 1	Sector 2	Sector 3
ESTIMATED ON	Sector 1	.236 (1.00)	.186 (0.94)	.258 (0.99)
	Sector 2	.228 (0.97)	.198 (1.00)	.246 (.094)
	Sector 3	.230 (0.98)	.168 (0.86)	.261 (1.00)

Composite Measures[a]

Transfer Likelihood Ratio Index = .216

Transfer Index = .94

Note: The base for computation of the transfer likelihood ratio index and the transfer index measures reported here is the market-shares model.

[a]Composite measures are weighted averages of the corresponding measures across multiple transfers *(19)*.

gate prediction error over that produced by the locally estimated model. The aggregate prediction groups employed in this study are the traffic analyses districts identified in the study area. Sectors 1 and 3 contain 16 districts and sector 2 contains 19 districts.

RMSE and the relative aggregate transfer error for the mode to work and automobile-ownership choice models are given in Tables 8 and 9. The RMSEs average 22 and 24 percent for the conditional mode-choice and marginal automobile-ownership choice models, respectively.

It is interesting to observe that the best (lowest) measures of RMSE for local prediction occur in those sectors for which the locally estimated model had the best (highest) rho-square values in Tables 4 and 5. These results suggest a reasonable level of consistency between these different measures.

TABLE 8 Aggregate Transferability Prediction Indices: Conditional Mode-Choice Model

		PREDICTING ON		
		Sector 1	Sector 2	Sector 3
ESTIMATED ON	Sector 1	.186 (1.00)	.241 (1.08)	.219 (1.01)
	Sector 2	.202 (1.09)	.222 (1.00)	.227 (1.04)
	Sector 3	.197 (1.06)	.224 (1.01)	.219 (1.00)

Composite Transfer Measures[a]

Transfer Root Mean Square Error = .219

Relative Aggregate Transfer Error = 1.05

[a]Composite measures are weighted averages of the corresponding measures across multiple transfers *(19)*.

TABLE 9 Aggregate Transferability Prediction Indices: Marginal Automobile-Ownership Choice Model

		PREDICTING ON		
		Sector 1	Sector 2	Sector 3
ESTIMATED ON	Sector 1	.245 (1.00)	.248 (1.01)	.171 (1.04)
	Sector 2	.281 (1.15)	.246 (1.00)	.205 (1.24)
	Sector 3	.238 (0.97)	.250 (1.02)	.165 (1.00)

Composite Transfer Measures[a]

Transfer Root Mean Square Error = .237

Relative Aggregate Transfer Error = 1.00

[a]Composite measures are weighted averages of the corresponding measures across multiple transfers *(19)*.

The relative aggregate transfer errors are low for all model transfers. They are less than 1.1, except for two transfers of the marginal automobile-ownership model. Further, the pooled values for this measure (1.05, 1.00) indicate a small increase in aggregate prediction error attributable to model transfer.

These results suggest that the use of disaggregate models for aggregate prediction is quite satisfactory. More important, for the purpose of this study, the increased error in aggregate prediction associated with use of transferred models is relatively small.

Overall, both the absolute and relative aggregate prediction measures indicate that transferred disaggregate choice models are effective in predicting aggregate choice shares.

DISCUSSION AND CONCLUSIONS

The mode and automobile-ownership choice models estimated in each sector are statistically significant and account for a reasonable proportion of the variation in the observed choices. An interesting feature of the estimation results is that the automobile-ownership models have substantially better likelihood ratio index (rho-square) values than the mode-choice models, despite the somewhat limited specification of the automobile-ownership model. Specifically, the rho-square values for the automobile-ownership models are generally twice as large as for the mode-choice models. This observation raises the question of whether the better fit of the automobile-ownership model has any impact on the relative transferability of the automobile-ownership and mode-choice models. This question is addressed in the following paragraphs, where the discussion centers on the transferability of models in which the alternative specific constants are adjusted to match the aggregate choice shares in the application environment.

The disaggregate transferability results are evaluated in absolute terms by the transfer likelihood ratio index and in relative terms by the transfer index. The transfer likelihood ratio index values for both the automobile-ownership and mode-choice models are in the same magnitude range as for the corresponding locally estimated models. That is, (a) the transferability for both sets of models is good and (b) the transferred automobile-ownership models are roughly twice as effective as the mode-choice models. On the other hand, the transfer index results indicate that, relative to locally estimated models, the mode-choice and automobile-ownership choice models are equally transferable. The result that improved fit of a model in the estimation environment appears to lead to improved transferability in absolute but not relative terms parallels the results reported by Koppelman and Wilmot (23) in connection with the impact of improved specification on model transferability.

The disaggregate transferability analyses also indicate that transferability is generally higher for transfer into sectors that have high local rho-square values. For example, the automobile-ownership model fits the observed data in sector 3 better than in the other two sectors. The transfer rho-square values reported in Table 7 indicate that the automobile-ownership model is more transferable into sector 3 than into sectors 1 and 2.

These results all indicate that model transfers are most effective when the transferred model is one that would be highly satisfactory if it were estimated in the application environment. Unfortunately, the only way to obtain this information is to estimate the corresponding model in the application environment, which eliminates the need for model transfer. However, the comparative results of the transferability of mode-choice and automobile-ownership models indicate that if there is evidence to suggest that models of particular choice behaviors are generally satisfactory, it is reasonable to infer that such models could be transferred effectively.

The aggregate transfer prediction analyses show little discrimination between the transferability of mode-choice and automobile-ownership models. These results do indicate, however, that the increased error in aggregate prediction associated with the use of transferred models is small (less than 10 percent in 10 of 12 transfers reported). Thus transferred disaggregate mode and automobile-ownership choice models appear to be able to predict aggregate shares satisfactorily, both in absolute terms and relative to locally estimated models.

The transferability analyses reported in this paper provide no clear indication of which sector pairs provide better estimation transfer contexts for transfer of disaggregate choice models in Washington, D.C. This result is not surprising, given that model transfer appears to depend on the fit of a locally estimated model in the application context, and the fact that the mode-choice model provides the best estimation goodness-of-fit in sector 1, whereas the automobile-ownership model provides the best estimation goodness-of-fit in sector 3.

The study reported in this paper leads to two basic conclusions. First, it is concluded that the findings of earlier research concerning the transferability of disaggregate mode-choice models can be extended to automobile-ownership choice models. Both automobile-ownership and mode-choice models exhibit a high degree of transferability at the disaggregate and aggregate levels in the intraurban transfer situations examined in this study.

The second basic conclusion reached in this study is that model transfer is more effective in those choice situations where behavior can be explained better by the mathematical model used to describe choice behavior. That is, if a given choice behavior can, in general, be well represented by a model, transfer of that model will generally be satisfactory. Although this conclusion is consistent with prior expectations, it is valuable that such expectations be confirmed empirically. Further, this study indicates that automobile-ownership level choice is predicted well by a relatively simple disaggregate choice model specification.

ACKNOWLEDGMENTS

The study reported in this paper is a portion of a research project investigating "Conditions for the Effective Transferability of Discrete Choice Models." The project is being undertaken at the Transportation Center of Northwestern University under contract with the U.S. Department of Transportation. James Ryan of the Urban Mass Transportation Administration, who is contract monitor, has provided important advice, encouragement, and support to the research team. Important contributions have been made to the research project by Chester G. Wilmot, Geok-Koon Kuah, Geoffrey Rose, and Albert Lunde. Jo Ann Scott of the Civil and Environmental Engineering Department at Duke University prepared the manuscript.

REFERENCES

1. M.L. Manheim. Fundamentals of Transportation Systems Analysis. MIT Press, Cambridge, Mass., 1979.
2. F.S. Koppelman and C.G. Wilmot. Transferability Analysis of Disaggregate Choice Models. In Transportation Research Record 895, TRB, National Research Council, Washington, D.C., 1982, pp. 18-24.
3. T.J. Atherton and M.E. Ben-Akiva. Transferability and Updating of Disaggregate Travel Demand Models. In Transportation Research Record 610, TRB, National Research Council, Washington, D.C., 1976, pp. 12-18.
4. A.P. Talvitie and D. Kirshner. Specification, Transferability, and the Effect of Data Outliers in Modeling the Choice of Mode in Urban Travel. Transportation, Vol. 7, No. 3, Sept. 1978.
5. R. Galbraith and D. Hensher. Intra-Metropolitan Transferability of Mode Choice Models. Journal of Transport Economics and Policy, Vol. 16, No. 1, Jan. 1982.
6. L.C. Caldwell III and M.J. Demetsky. Transferability of Trip Generation Models. In Transportation Research Record 751, TRB, National Re-

search Council,, Washington, D.C., 1980, pp. 56-62.
7. C. Doubleday. Some Studies of the Temporal Stability of Person Trip Generation Models. Transportation Research, Vol. 11, No. 4, Aug. 1977.
8. P.R. Stopher and A.H. Meyburg. Urban Transportation Modeling and Planning. Lexington Books, Lexington, Mass., 1975.
9. A. Kanafani. Transportation Demand Analysis. McGraw-Hill, New York, 1983.
10. Charles River Associates. A Disaggregated Behavioral Model of Urban Travel Demand. Final Report. FHWA, U.S. Department of Transportation, 1973.
11. T.A. Domencich and D. McFadden. Urban Travel Demand: A Behavioral Analysis. North-Holland, Amsterdam, 1975.
12. M.E. Ben-Akiva. Structure of Passenger Travel Demand Models. In Transportation Research Record 526, TRB, National Research Council, Washington, D.C., 1974, pp. 26-42.
13. M.E. Ben-Akiva and S.R. Lerman. Some Estimation Results of a Simultaneous Model of Auto Ownership and Mode Choice to Work. Transportation, Vol. 3, No. 4, 1974.
14. D. McFadden. Modeling the Choice of Residential Location. In Transportation Research Record 673, TRB, National Research Council, Washington, D.C., 1978, pp. 72-77.
15. F.S. Koppelman and E.I. Pas. Multidimensional Choice Model Transferability. Office of University Research, U.S. Department of Transportation, 1983.
16. K.L. Sobel. Travel Demand Forecasting with the Nested Multinomial Logit Model. In Transportation Research Record 775, TRB, National Research Council, Washington, D.C., 1980, pp. 48-55.
17. S.R. Lerman and M.E. Ben-Akiva. Disaggregate Behavioral Model of Automobile Ownership. In Transportation Research Record 569, TRB, National Research Council, Washington, D.C., 1976, pp. 34-55.
18. C.G. Wilmot. The Impact of Market Segmentation on the Transferability of Discrete Choice Models. Ph.D. dissertation. Northwestern University, Evanston, Ill., Aug. 1983.
19. F.S. Koppelman and G. Rose. Geographic Transfer of Travel Choice Models: Evaluation and Procedures. Presented at International Symposium on New Directions in Urban Modeling, University of Waterloo, Waterloo, Ontario, Canada, July 1983.
20. D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers in Econometrics (P. Zarembka, ed.), Academic Press, New York, 1973, Chapter 4.
21. F.S. Koppelman. Guidelines for Aggregate Travel Prediction Using Disaggregate Choice Models. In Transportation Research Record 610, TRB, National Research Council, Washington, D.C., 1976, pp. 19-24.
22. F.S. Koppelman. Methodology for Analyzing Errors in Prediction with Disaggregate Choice Models. In Transportation Research Record 592, TRB, National Research Council, Washington, D.C., 1976, pp. 17-23.
23. F.S. Koppelman and C.G. Wilmot. Effect of Model Specification Improvement on Transferability. Presented at 63rd Annual Meeting, Transportation Research Board, Washington, D.C., Jan. 1984.

Travel Regularities and Their Interpretations: A Discussion Paper

JANUSZ SUPERNAK

ABSTRACT

The regularities in travel behavior analyses are examined in this paper. Reasons are investigated for different interpretations of travel regularities caused by (1) differences in basic assumptions, model specification, and selection of analysis unit; (b) differences in selection and evaluation of empirical material; and (c) differences in data used. Criteria for evaluation of meaningfulness and applicability of travel regularities are proposed. Travel-time budget analyses and studies of travel behavior of homogeneous groups of persons are compared as alternative approaches to investigate differences in travel regularities and diversity of their interpretations.

Detecting regularities and establishing relationships in any analyzed phenomenon, process, or behavior is always an important and interesting part of any research effort. Discovering regularities is normally a first sign of understanding the analyzed problem. Often these regularities have useful applications. In human travel behavior, regularities confirmed by several studies from different metropolitan areas can constitute a basis for geographically transferable models and can be used in travel demand forecasts and policy analyses.

In travel behavior analysis, as in other fields of research, there are several ways to analyze different regularities and many ways to interpret them. This diversity in travel regularities and their interpretations does not necessarily mean that any one version must be wrong. However, in many cases where the research conclusions are divergent, it is natural to ask why. What methodological differences are responsible for different regularities? Are these regularities meaningful, consistent, or useful in practical applications? Although the final judgment about quality of different regularities and respective research approaches will never be fully objective, evaluation using these criteria would be in order.

There are several reasons potentially responsible for different interpretations of regularities in travel behavior analyses: (a) conceptual differences (different approaches, analysis units, unit stratifications, model specifications, and so forth); (b) differences in selection, presentation, evaluation, and interpretation of empirical findings; and (c) overall quality of the data used (i.e., its completeness, adequacy, accuracy, compatibility).

An attempt is made to address some of the issues in this paper. The content, form, and scope of this paper were prompted by the comments of Zahavi (1) printed with the author's article "Travel-Time Budget: A Critique" published in Transportation Research Record 879 (2). At that time there was no opportunity to respond in an author's closing statement. However, because Zahavi raised many important issues, both on the subject of travel-time budgets and on the broader questions of interpreting differences among researchers' analyses of data, many of his comments and other works (1,3) are used as the basis for this paper. Over several years, Zahavi has contributed many innovative ideas in the study of travel budgets, but possibly just as significant are the important methodological issues that have been generated by his research. Today it may be neither important nor appropriate to debate the validity or nonvalidity of the concept of travel-time budgets. Although this article is not intended as a response to the late Zahavi's comments on travel-time budgets, it does refer to them in an effort to illustrate the issues and questions he raised in the field of travel behavior analysis.

ANALYSIS OF HUMAN TRAVEL BEHAVIOR: CONCEPTUAL DIFFERENCES

Background

The concept of stability of a travel-time budget is well-known to a majority of researchers in the field and does not need to be introduced in detail in this paper. A good summary of Zahavi's work is given elsewhere: "Analysis of Zahavi's work on the subject revealed the evolution of the concept, from an overall average daily travel time for vehicles, to average values per traveler systematically influenced by socioeconomic factors, to a final relationship with the average speed of the transportation system" (4). This last version is given in a report by Zahavi: "The mean daily TT-budget per travel is an inverse function of speed, decreasing as speed increases, to an asymptote of about 1.1 hours per day" (3,p.IV). Full details of the concept are given in several reports (3,5,6).

The author's critique of travel-time concepts in general (not specifically Zahavi's work) presented in TRR 879 (2) (a) questioned the meaningfulness and applicability of travel-time concepts; (b) postulated some methodological improvements in travel-time budget studies; (c) found trip rates of homogeneous groups of persons more stable than the respective travel budgets; and (d) revealed regularities in several travel characteristics of these groups, with a potential for geographic transferability of outside-the-home activity budgets of homogeneous groups of persons. Details are given in that paper (2).

Regularities: Behavior of an Individual or a Group?

Any regularity in travel behavior refers to either the entire population (of persons or travelers) or to some clearly specified subgroups of the population. In any disaggregate approach where, for example, individual i is used as the analysis unit, the implied assumptions should be that the results can be generalized over a larger group of persons represented by individual i. Also, whichever period is used as an analyzed time duration (e.g., a day), the behavior expressed by such characteristics as trip rates, travel-time budget, and so forth, should not be expected to be identical each and every day. For example, a travel-time budget of 60 min means that the average daily travel time of an average representative of an analyzed group (G) is 1 hr.

Averaging travel characteristics in order to generalize the travel behavior of the population under study encounters some problems represented by the following questions: Because the human population is heterogeneous, should the average behavior of the entire population or rather its more homogeneous subpopulations be described? What is the geographic and temporal stability of travel characteristics if it is known that the population structure is subject to significant changes in both space and time? How will changes in the population structure influence the validity of certain transportation policies that may apply differently to different population subgroups?

Heterogeneity of the Population and Importance of its Proper Stratification

It is interesting to note that, for any heterogeneous population under study (human population is certainly just this with respect to outside-the-home activities and travel patterns), meaningful regularities can be found only after meaningful, crucial variations are found. For example, more is known about dogs than about mammals as a whole, and more about bulldogs than about dogs as a whole. The differences among biological species were the reasons for stratifying them into more homogeneous groups whose average physical outlooks, behaviors, and so forth, could already be found to be quite regular. Any analysis based on an average traveler (i.e., a person who just happened to travel during the survey day by motorized modes) fails to recognize crucial variations within heterogeneous groups; thus by averaging over unidentifiable units, the analysis fails to discover really meaningful regularities.

Acknowledging an existence of a high heterogeneity of the population (persons or travelers) with respect to its travel behavior (employed husbands versus housewives, or groups between the ages of 20 to 30 versus 70 to 80) implies certain methodological consequences. If travel behavior is predominantly differentiated by age and employment status, these variables should be the primary candidates for consideration in any analysis of travel patterns. They should also be a basis for meaningful stratification of a heterogeneous population into more homogeneous subpopulations.

Why should the analyst care about a proper stratification of the population in any study of travel behavior? The reasons are as follows:

1. The analyst wants to identify groups of distinctly different travel behaviors that could be caused by differences in objective needs for traveling, options available, and travel constraints;
2. The analyst wants to assure a proper representation of each group while designing, for example, cross-sectional or longitudinal surveys of travel behavior;
3. The analyst would like to capture dynamic changes in representations of each group, and their consequences on the population treated as a whole; and
4. The analyst is interested in identifying differences among groups in reaction to relevant outside changes, both natural (e.g., the changing energy situation) and imposed (policies).

What criteria should be followed for a proper stratification of the population due to an analyzed issue? Theoretically, the desired criterion could be formulated as a formal minimization of the within-group variance for each group. In analyses of trip rates it could be formally done by stratifying groups due to their number of daily trips: those with zero, one, two, and so on. Thus the within-group variance would be zero, and the total variation would be explained by the between-group variance. However, this grouping would be quite useless because the groups could not be identified.

Therefore, a much more complicated formula is required: stratification into homogeneous groups has to result from some kind of multivariate analysis. This should reduce the within-group variance to the extent possible, and result in a relatively small number of homogeneous groups that are relevant to the analyzed issue, easy to identify, and whose populations will be relatively easy to predict.

An excellent guide for creating homogeneous groups can be found in two reports (7,8). It has to be noted that any arbitrary stratifications, one-dimensional or multidimensional, may appear ineffective or even totally irrelevant, and that "segmentation along an irrelevant dimension will result in inaccurate prediction results" (8).

The Significance of Homogeneity of Groups of Persons

The importance of homogeneity of groups of similar travel behavior can be demonstrated by at least two points.

1. Homogeneous groups should have smaller variability than the population as a whole; this can reduce the desired size of the travel survey if a stratified sampling scheme is chosen.
2. Homogeneity of the groups should reduce the divergence of results between different survey techniques (e.g., more units and shorter observation time versus fewer units and longer observation time for travel behavior). This may be crucial in justifying, for example, a 1-day transportation survey procedure from which judgments about an average daily behavior are made.

Homogeneity of the groups does not denote even distributions. Quite often coefficients of variance of observations of either trip rates or travel budgets will still remain relatively high. There are some reasons for this:

1. Including zeros to represent nontravelers and low numbers for nonvehicular trips increases the tail part of data and results in higher variances;
2. A short observation period (normally 1 day) is responsible for a large number of zero observations if travel is regular but sporadic [a simple numerical example in a previous paper (2) shows that coefficient of variance can drop dramatically if the observation period increases]; and
3. The coefficient of variance is not always an absolute measure of variability in data.

Transferability of Travel Characteristics

Is transferability the primary criterion for the evaluation of meaningfulness and applicability of regularities? Some authors appear to suggest that the answer to this question is yes. Zahavi writes that "the primary test for different approaches is whether or not the model is transferable in both space between cities and in time in one city" (1).

Note that the transferability criterion is a demanding one and clearly it is quite risky. If models are expected to be fully transferable in both space and time, then any single empirical test that proves against transferability could jeopardize the final conclusions, even if all previous tests supported the notion of transferability.

The problem of transferability appears to be more complex than the preceding quote from Zahavi (1) might suggest. First, it is clear that some travel characteristics should not be expected to be spatially transferable. For example, average daily travel times to and from work vary widely among cities because of their different distributions of residential areas and work places, and differences in sizes, shapes, types of industry, transportation infrastructures, and so forth. Thus it would be unreasonable to expect the obligatory part of travel-time budget (T^{obl}) to be transferable. The overall travel-time budget (T) could be transferable, but this would impose a regulatory role on the discretionary part of the travel-time budget (T^{disc}) because $T = T^{obl} + T^{disc}$, a notion that was questioned in a previous paper (2). Second, it is not clear whether the spatial transferability is a prerequisite for the temporal transferability: some authors disagree with this notion. On the other hand, the existence of geographic transferability in some characteristics may not imply meaningfulness of this single regularity. This issue will be discussed in more detail later in the paper.

Therefore, the following criteria for evaluating the meaningfulness of regularities can be proposed:

1. The subsets of the population to which regularities are applicable should be clearly specified,
2. Regularities should be adaptable for another urban environment [an absolute transferability (e.g., trip rate N_i = const) may be possible but is not strictly required],
3. Regularity should provide a logical and consistent explanation (at least signs of relationships should always be the same),
4. Regularity should properly illustrate major trends observed in analyzed phenomenon or behavior, and
5. Regularity (or set of regularities) should be easily applicable.

Problem of Partial Regularities

The concept of stability of the travel-time budget per traveler (TT/TR) is examined. In order to reliably estimate the amount of traveling in the system [total travel time (T) or distance (D) in this concept], it is not enough to confirm transferability of the daily travel time per traveler (TT/TR) or the relationship TT/TR as a function of speed V,

$$TT/TR = b + (a/V) \quad (a, b = \text{constants}) \tag{1}$$

In order to obtain the estimation of T, at least two more relationships have to be transferable: (a) percentage of traveling households (β^{HH}) as a function of household characteristics, and (b) average number of travelers per household (TR/HH) as a function of household characteristics. Thus

$$T = (L/HS) \cdot \beta^{HH} \cdot (TR/HH) \cdot (TT/TR) \tag{2}$$

where L is the population size and HS is the average household size.

Thus the concept of stability of the travel-time budget requires simultaneous transferability of regularities in all three characteristics: TT/TR, β^{HH}, and TR/HH. [In the entire UMOT interaction process (3), stability of the daily household expenditure on travel (M) as a share (C) of household income (I) has to be assumed, i.e., M = C(I/HH).] The temporal stability of the travel-time budget per traveler (TT/TR), even if fully confirmed, will be useless if at least one of the other relationships previously mentioned appears nontransferable. It is worth noting that these relationships are virtually ignored in the travel-time budget literature, even though they deserve the same attention as does the TT/TR relationship.

Analysis Unit Controversy: Person Versus Household

The analysis units used in the stability of the activity budget concept versus the travel-time budget concept are examined. In the first case, the unit is an average representative of homogeneous group i, whereas in the second case it is a motorized traveler that is representative of an average traveling household H. Averaging over unidentified household members has one important disadvantage: it ignores the high heterogeneity of the family.

The household versus person (or traveler) controversy was commented on in some works (2,9,10). Here, only the main points are presented to explain why an individual level of data aggregation was chosen for the analysis made in a previous paper (2).

1. An individual is the only true travel decision maker; travel choices of an average household member (or traveler) have virtually no interpretation.

2. A reasonably small number of homogeneous groups (categories) can be created only at the individual level. Applying the unit "an average representative of a homogeneous group of households" is virtually impossible or at least impractical; it would require hundreds of different types of households, and yet a vast majority of these units will have to remain highly heterogeneous.

3. References to a person's household environment can be introduced at this level if needed. The need can be disclosed by peforming a multivariate analysis of significance of the variables. In some cases household-oriented variables can be individualized [e.g., car availability (11)]. Household references can take the form of a hybrid approach (12), but the bottom line is that effective stratifications of the population need not follow a person's family affiliations (2,7,13). A family (in a transportation sense) is one of the most heterogeneous sets of three, four, or five persons one can think of (9).

4. The effect of household size is not unobservable at the person level of data aggregation. Moreover, the individual approach addresses another important issue: it identifies the person who constitutes the additional family member. The daily travel time per person drops sharply with family size not because of any magic power of the household size variable, but because family members number 1 and number 5 are, as a rule, very different people (e.g., employed father versus his preschool child). In a person approach family members will belong to different homogeneous groups and possess different travel characteristics. If multivariate analysis reveals that the household-size variable is needed at the person level, it can be introduced into the model (e.g., by distinguishing housewives from families with children from housewives without children). Finally, the interactions and trade-offs among family members are difficult to describe at any level, even at the family level. On the other hand, some effect of these trade-offs can be observed at the person level (i.e., employed husbands spend less time on family shopping than their nonemployed wives).

Controversy: Person Versus Traveler

The discussion about the analysis unit in travel budget studies is well documented in the literature (1,2,4,14). The majority of researchers base their calculations on all persons, independently of whether they traveled or not during the survey day.

If the concept of traveler is applied, then an arbitrary 1-day observation period will become the reference point. Theoretically, however, any time period can be chosen to define the traveler. Travel surveys today are not necessarily based on 1-day data: the observation period can be 1 month, 1 week, 2 days, 1 day, or peak period. For each of these periods both the definition of traveler and the percentage of nontravelers will be different. Over longer periods of time virtually everyone becomes a traveler.

There are several consequences of this choice of the analysis unit.

1. The concept of traveler has no clear reference to the frequency of traveling; it treats someone traveling every day in the same way as someone traveling once a week (if he happened to travel during the survey day).

2. The consequence of 1 is that regularities per traveler may contradict those of per person, with a potential for confusion and misinterpretation of resulting relationships. This point can be illustrated by a (simplified) numerical example. Three groups of American television watchers are investigated: group A consists of people who regularly watch daily news and practically nothing else. Group B watches only "60 Minutes," a popular weekly news magazine program. Group C watches only main sport events such as the Super Bowl in football and final play-offs in basketball. These results are summarized in Table 1. Which group watches more television: A, B, or C? Group C watches the most on a daily basis if they watch (the importance of "if" is crucial). The order is reversed if how much time the representatives of

TABLE 1 Example of Television Watching Time Budgets for Groups A, B, and C

	Television Watching Time Budget (hr)		
	Group A	Group B	Group C
Yearly per person	183.0	52.0	12.0
Daily per watcher[a]	0.5	1.0	3.0

[a]Watcher (similar to the definition of traveler) is the person who watches television during a given day.

these groups devote to watching television over longer periods of time (e.g., a year) is analyzed.

3. Ignoring nonmotorized travel appears to be a more serious problem than indicated by Zahavi (1): "It should be noted at this stage that walking, as a mode, was found to be a small proportion of travel in Baltimore; walking comprised only 3-12 percent of the total travel time of the above travelers belonging to high- and low-income households, respectively. As for distance, the proportions were only 1-5 percent, respectively." First, these proportions are much higher in other cities and, especially, in city centers. Second, the concept of motorized modes also excludes cycling, an important way of traveling in several countries in Europe and Asia. For example, Brog and Erl (15) report that the importance of bicycle as a mode of travel in West Germany is growing. Finally, the decision to exclude nonmotorized modes is difficult to accept on conceptual grounds. The time spent on traveling by nonmotorized modes has to be taken, as in the case of travel by motorized modes, from the total budget of disposable time. If the proportion of nonmotorized modes is indeed only marginal, why not include these trips into travel-budget considerations? Gunn (14) warns that "it is dangerous to assume that trends and relationships based on travel by mechanized modes alone can be given any general behavioral interpretation."

By using regularities that are valid "per homogeneous group of persons," the analyst can (a) easily generalize this regularity over any longer period of time, (b) capture temporal trade-offs many persons make for their activities (e.g., to do more traveling during one day in order to have more time left for within-home activities the next day), and (c) substitute a series of partial regularities to illustrate travel behavior (e.g., TT/TR, β^{HH}, and TR/HH) by a single regularity per person.

DIFFERENCES IN INTERPRETATIONS OF EMPIRICAL FINDINGS

Examples of Regularities

To illustrate some points mentioned previously, some examples of regularities discussed by Zahavi (1) and Supernak (2) are examined in more detail. Figures 1-3 and Table 2 (2) represent a sample of regularities relating to behavior of homogeneous groups of persons, whereas Figures 4-9 and Table 3 represent a sample of regularities referring to the travel-time budget concept.

The interpretation of results shown in Figures 1-9 and Tables 2 and 3 will be made according to the proposed criteria for evaluation of the meaningfulness of regularities (see section on transferability).

Regularities: Application for Specific Population Groups

Regularities presented in Figures 1-3 (2) apply to clearly specified segments (categories) of the population. All regularities appear to be category specific; differences between categories in all characteristics analyzed are significant, thus supporting the relevance of the stratification of the population into eight groups due to age, employment status, and automobile availability. Regularities presented in Figure 4 (1) also apply to a specific population subgroup: average traveler representative of traveling household H.

In either approach a major problem is the ability to predict the representation of either (a) population of a given person category or (b) the population of travelers. In the first case, prediction of person categories is based on projection of age and employment, as well as on the forecast of the automobile availability made separately for employed and nonemployed persons. The desired level of automobile availability was found to depend primarily on population density (11).

The ability to predict population of travelers depends on the consistency of the relationships shown in Figures 5 and 6.

Adaptability of Regularities into Another Urban Environment

The adaptability of regularities into another urban environment was tested for trip rates within Baltimore and for traveler trip rates between Baltimore

TABLE 2 Basic Travel Characteristics of Person Categories 1-8 in Baltimore *(2)*

No.	Category Description	α_i (%)	$\gamma_i^{nontrav}$ (%)	β_i^{walk} (%)	N_i Mean	N_i SD	T_i (min) Mean	T_i (min) SD	t_i (min) Mean	t_i (min) SD
1	Person <18 years old	18.1	14.8	48.8	2.98	2.10	51.6	36.9	17.3	10.9
2	Employed, 18-65 years old, car never available	9.1	9.9	26.7	2.50	1.72	62.7	42.6	25.1	17.8
3	Employed, 18-65 years old, car sometimes available	13.5	6.3	8.5	3.17	1.91	63.8	38.8	20.1	13.5
4	Employed, 18-65 years old, car always available	18.5	4.3	4.8	3.48	2.00	69.8	38.1	20.0	12.8
5	Nonemployed, 18-65 years old, car never available	17.4	50.6	51.2	1.33	1.73	22.8	35.9	16.7	16.5
6	Nonemployed, 18-65 years old, car sometimes available	6.8	25.2	16.5	2.55	2.22	40.6	38.9	15.9	10.8
7	Nonemployed, 18-65 years old, car always available	6.4	18.1	4.7	2.99	2.36	44.1	37.2	14.8	10.4
8	Persons >65 years old	10.3	35.2	27.2	1.48	1.65	22.8	34.8	15.4	16.3
Entire population		100.0	20.5	22.4	2.59	2.10	48.3	41.8	18.7	14.3

Note: α_i = percentage in the sample, $\gamma_i^{nontrav}$ = percentage of nontravelers (nontraveler = person making no trip during the survey day), β_i^{walk} = percentage of walking trips, N_i = daily trip rate, T_i = time spent on traveling during the day, and t_i = average trip duration.

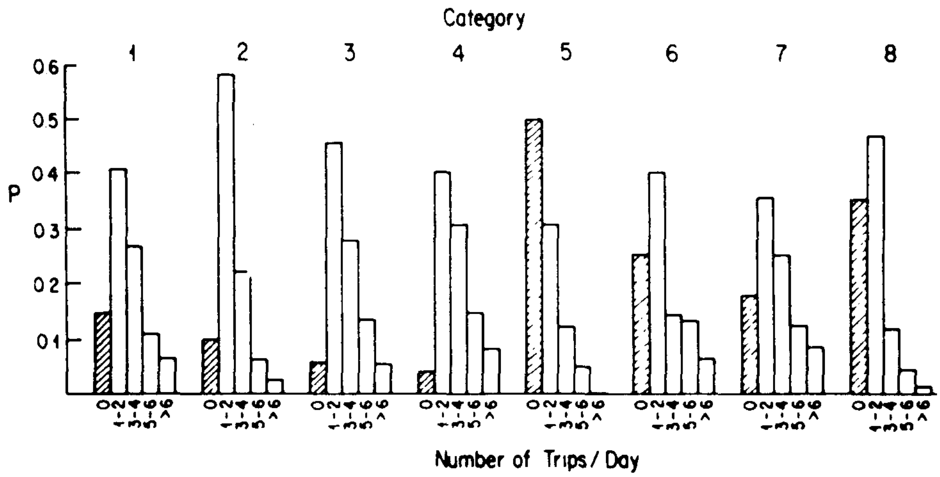


FIGURE 1 Trip frequency distributions for person categories 1-8 in Baltimore *(2)*.

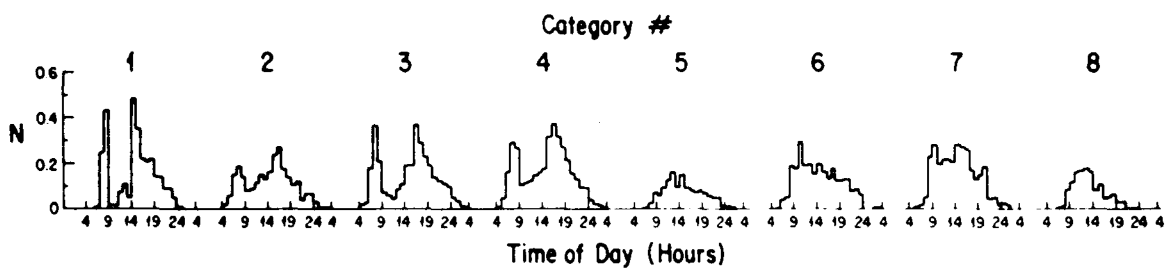


FIGURE 2 Hourly trip histograms for person categories 1-8 in Baltimore *(2)*.

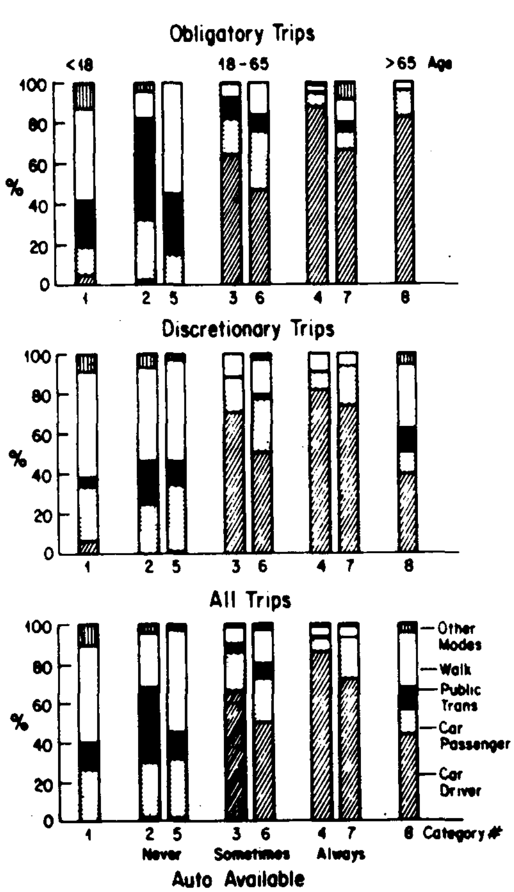


FIGURE 3 Basic modal splits for person categories 1-8 in Baltimore *(2)*.

and Minneapolis with reasonable success. Trip rates appeared to be geographically more stable than respective travel-time budgets for person categories. The claim of any universal geographic transferability of mobility characteristics of homogeneous groups of persons (N_i = const) can not be made yet, and it is not likely to occur. The condition of adaptability of regularity will be satisfied if the relationship $N_i = f$ (city characteristics) appears consistent and transferable. More compatible data sets are needed to perform the necessary tests.

As for the concept of travel-time budget, Figures 4-6 (1) provide satisfactory evidence of stability and transferability of all three relationships crucial to the success of the concept (see Equation 3 presented later): (a) regularity of travel-time budgets per motorized traveler distribution (Figure 4); (b) regularity of relationship explaining the percentage of traveling household (Figure 5); and (c) regularity of relationship explaining the percentage of travelers per household (Figure 6). Four cities,

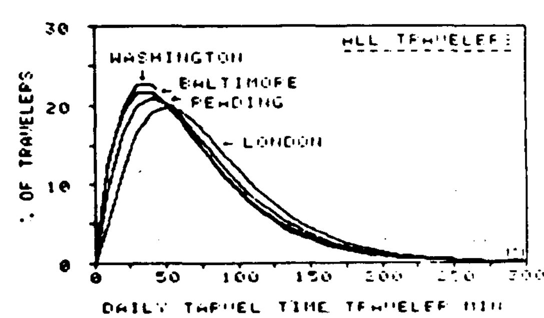


FIGURE 4 Travel time per traveler distribution: all travelers in four cities *(1)*.

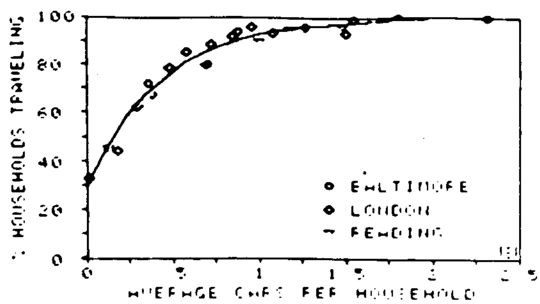


FIGURE 5 Percentage of households traveling versus cars per household *(1)*.

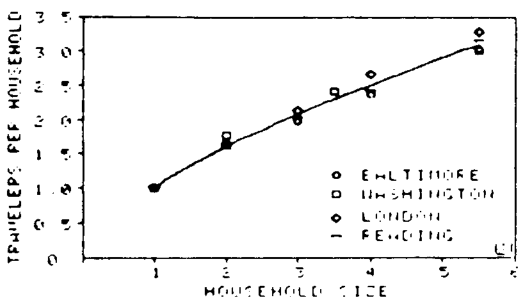


FIGURE 6 Travelers per household versus household size *(1)*.

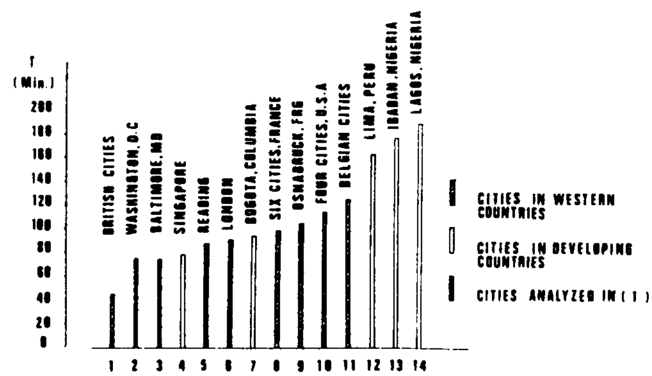


FIGURE 7 Mean daily travel times for selected locations.

Baltimore and Washington, D.C., in the United States and London and Reading in the United Kingdom, were selected for these transferability tests. Travel-time frequency distributions appear to be "transferable among the four cities when accounting for travel speed" (1) (see Figure 4).

The ability to generalize findings shown in Figure 4 over a larger number of cities from several countries around the world can be tested by analyzing data provided in several papers (2,3,16). If daily travel-time means were to be ranked in increasing order, the picture would look like Figure 7. Four cities analyzed by Zahavi (1) happened to be "neighbors" in this large spectrum of different results. Conclusions about transferability of the travel-time frequency will have to be questioned if some other cities were selected for this comparison (e.g., No. 1, 3, 11, and 14).

The relationship presented in Figure 5 also creates some problems when generalized over populations in other cities. In many cities around the world there are more carless households than those with cars. It is not likely that in these countries only one-third of carless households will travel by motorized modes during an average 24-hr period, considering that a vast majority of carless households has at least one employed member who has to work every working day.

The relationship shown in Figure 6 explains the number of travelers per traveling household by household size. Figure 1 shows that there are significant differences in percentages of nontravelers among different household members. Therefore, the number of employed family members and the number of students should be seen as a primary explanatory variable to estimate the number of travelers per household. In Table 3 a sample of results of the relationship Trav/HH = a + b_1 (HH size) + b_2 (cars/HH) is presented for American and West German cities. The results appear nontransferable and inconsistent.

Consistency of Regularities

The regularities in travel behavior should be logical (i.e., signs of relationships should be as expected and consistent). For example, Figures 1-3 show that if the percentage of employed persons increases, there is more travel in general, by car, and during rush hours, as expected. More of these regularities are presented elsewhere (10,11). One of them is the increasing role of the automobile in areas of low population density.

Another example verifies the postulated inverse relationship between daily travel time per traveler and speed. The best relationships for distance per traveler versus door-to-door speed for Munich, West Germany, were found by Zahavi (3,p.138) as follows:

$$\text{Dist/Traveler} = -7.184 + 1.738\ (\text{Speed}),\ \text{for carless households} \quad (3)$$

$$\text{Dist/Traveler} = -0.739 + 1.173\ (\text{Speed}),\ \text{for car owning households} \quad (4)$$

For north and south corridors of Washington these relationships are, respectively, as follows (5,p.35):

$$\text{Dist/Traveler} = 1.841 - 1.002\ (\text{Speed}) \quad (5)$$

$$\text{Dist/Traveler} = -1.639 + 1.277\ (\text{Speed}) \quad (6)$$

TABLE 3 Travelers per Household by Household Size and Car Ownership

City	Year	a	b_1	b_2
Washington, DC	1955	0.917	0.192 (3.23)	0.471 (3.79)
Washington, DC	1968	0.643	0.231 (4.50)	0.503 (7.62)
Twin Cities	1958	0.024	0.325 (5.28)	0.870 (5.74)
Nurenburg	1975	0.205	0.547 (14.06)	0.275 (3.67)

The relationships cited do not appear consistent. The values of coefficient estimates vary dramatically, and even the signs of the relationships are divergent. For example, Equations 3, 4, and 6 indicate that if speed increases, daily travel time per traveler increases, too (which is contrary to the postulated form of the relationship given by Equation 1) whereas Equation 5 would support the opposite conclusion.

Regularities: Ability to Capture Major Trends

Any regularity should properly illustrate major trends in the analyzed issue. This is a condition for a satisfactory forecasting ability of any model that is based on this regularity. For example, in several countries (including the United States), two trends have had a profound effect on the situation on the highways: (a) increasing female participation in the labor force and (b) process of surbanization.

Figures 1-3 show the results of the first trend. There is more travel in general, by car, and during rush hours as a result of an increase in female employment. Also, it can be shown (11) that the trend of the population moving into the suburbs consistently causes an increasing need for higher automobile availability and, consequently, an increase in automobile use. The person category approach also appears convenient to illustrate major demographic trends such as the increasing percentage of older people in the population.

Figure 6, on the other hand, can illustrate the effect of shrinking household size on the number of travelers in a household, but cannot capture the effect of increasing female participation in the labor force.

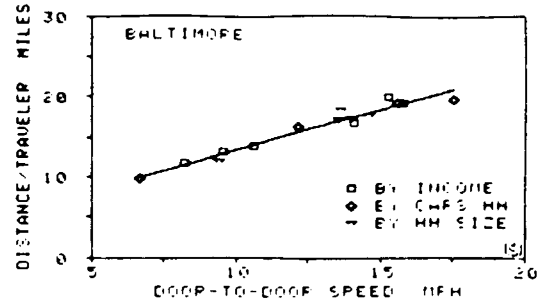


FIGURE 8 Distance per traveler versus speed *(1)*.

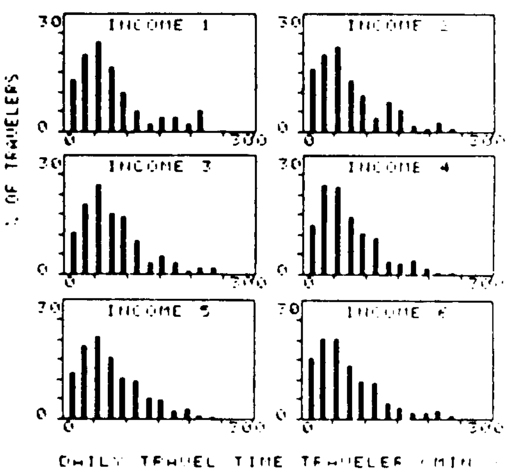


FIGURE 9 Travel time per traveler distributions by household income *(1)*.

Applicability of Regularities

Both approaches discussed here--the one based on homogeneous categories of persons and the one based on average traveler--are easily applicable and require a limited amount of basic data. Both approaches directly refer to several modeling stages such as automobile ownership and availability, trip and travel generation, and modal split.

The advantage of the person category approach is its consistency in using the same analysis unit through all modeling stages (11,12).

Regularities: Alternative Interpretations

It is not uncommon that different researchers can make different interpretations of the same regularity. For example, Figure 8 (1) can be interpreted to mean that "travelers at higher speed spend less daily time for more travel distance" (1). Alternative interpretations could be that (a) longer travel distances, even in the aggregate, are normally traversed with higher door-to-door speeds (by using expressways more often or by increasing the fast in-vehicle time part of travel by public transportation), and, more importantly, that (b) travelers and their characteristics may be seen as irrelevant here because the relationship illustrates the operation of the transportation system rather than traveler behavior.

Figure 9 (1) can be treated as an illustration of consistency or regularity. An alternative interpretation could be that stratification by income appears irrelevant. It can be argued that six distributions for six income groups are in fact equivalent to a single distribution for the entire population. Thus Figure 9 can be treated as an example of an irrelevant stratification. Similarly, stratification by income, car ownership, and household size appears to be irrelevant for the distance per traveler relationship (Figure 8). If an irrelevant variable is left in the forecast model, it can lead to a wrong prediction because the true explanatory variables are more likely to be outside the model.

Consequences of Differences in Interpretations of Travel Regularities

Analysis of regularities is often associated with a testing of some more general concepts and theories. Specific interpretations of these regularities influence these concepts and may lead to conclusions that are different than those of other researchers and that are sometimes counterintuitive. Often the validity of a given interpretation can be tested by applying some boundary conditions. Sometimes a common sense, overall understanding of the field and experience can be quite useful evaluation tools, as well.

A quote from Zahavi et al. (5,pp.78-79) is a good example: "Exercises carried out with the UMOT travel process produced some results which appeared to be counterintuitive at first sight. For example, the scenario which provides a free transit system resulted in an increase in travel distance by both transit and private modes." This counterintuitive finding was recently criticized by Downes and Emmerson (17). It could be interesting to analyze to what extent did the methodological issues discussed in this paper contribute to this result.

DATA INFLUENCE ON REGULARITIES AND THEIR INTERPRETATIONS

The discussion about different interpretations of travel regularities has yet another dimension. If the overall quality of data is bad, the entire verification of the empirical findings becomes virtually impossible or meaningless. The analyst would not know what was responsible for the lack of regularities: an irregular original, a poor model, or just poor data.

Data Quality: What Requirements?

There is an obvious interdependence among the design of the data-collection process, the gathering of data, the data analysis, and the presentation of results. All can contribute to the overall quality of the data and to the validity of the interpretation.

There are several elements useful for evaluation of data quality. The data sets should be accurate, complete, representative, flexible for different uses, and compatible. Data quality issues have been covered by several recent publications (18), and it will not be discussed here. Rather, the data compatibility issue, which is crucial for the validity of transferability tests, will be discussed.

Data Compatibility: A Fundamental Requirement

In order to be compatible, data sets have to be consistent in the following elements: subject subsystem records, object subsystem records, and travel process records.

Subject subsystem refers to an individual as a potential traveler and his relevant characteristics. The most common problems with data records about travelers are (a) completeness of the record (all persons, not only travelers, and all relevant personal characteristics), (b) flexibility of the record (avoiding prestratification according to age groups, for example), and (c) subjective versus objective perception, biases, errors, and so forth.

The object subsystem should cover all land use characteristics and transportation infrastructure records. Uniform network coding, compatible ways to introduce parameters of a given transportation system, and uniform records of land use patterns (residential densities) are samples of data problems associated with the object subsystem.

Trip records have to be given special attention. All modes, including walking, biking, and so forth, should be recorded. Clear definition of the trip, distinction between intracity and intercity travel, definition of the shortest trips, and so on should be made compatible. Work-day travel and weekend travel should be separated. Uniform, or at least compatible, definitions of trip purposes should be made. These problems are only some examples of potential discrepancies.

Consequences of Data Adjustments

The problem of data noncompatibility in travel demand analyses is both serious and common. A comparative analysis of trip patterns in Baltimore and the Twin Cities (2) is a typical example of difficulties with data compatibility. Data sets from these cities differed significantly because of both the records of traveler characteristics (e.g., different age brackets) and trip records (e.g., different definitions of the shortest trips). Also, data records had to be checked for errors (e.g., whether a trip-chaining pattern was logical). Therefore, careful and systematic data adjustment had to be made to assure compatibility of both sets. Only after this process was finished could the results from Baltimore and the Twin Cities be compared at all.

One of the consequences of data adjustment is that the results based on the processed data should vary from the results based on raw data. The need for data adjustment was the reason why, for example, results of the Twin Cities travel-time budgets presented by Supernak (2) varied from some previous results cited by Zahavi (1).

FINAL REFLECTION

Final recommendations are not offered in this paper because it is intended as a discussion paper. Examples of alternative approaches, different results, and diversified interpretations of travel regularities have been presented. Also, insight into the reasons why these differences do happen was provided. Differences in interpretations of results do not necessarily prove anyone wrong; instead they illustrate a healthy diversity of research approaches, assumptions, and conclusions. Different views are often helpful for better understanding the analyzed field. It is hoped that this paper will stimulate some more thoughts and discussion. It is often through this process that progress in any field is made.

ACKNOWLEDGMENTS

The author would like to acknowledge the contribution of Y. Zahavi to the subject of travel analysis and particularly travel-time budgets, and to thank the Transportation Research Board for publishing his article "Travel-Time Budget: A Critique" and permitting him to copy various charts and figures from that article and Zahavi's discussion. The author appreciates the help of Preston Luitweiler, David Schoendorfer, and Marilyn Macklin.

REFERENCES

1. Y. Zahavi. Discussion on "Travel-Time Budget: A Critique". In Transportation Research Record 879, TRB, National Research Council, Washington, D.C., 1982, pp. 25-27.
2. J. Supernak. Travel-Time Budget: A Critique. In Transportation Research Record 879, TRB, National Research Council, Washington, D.C., 1982, pp. 15-28.
3. Y. Zahavi. The UMOT Project. Report DOT-RSPA-DPB-20-79-3. U.S. Department of Transportation and Ministry of Transport, Federal Republic of Germany, 1979.
4. O. Strambi, A.H. Meyburg, and M.A. Turnquist. Travel Time Budgets: A Reassessment of the Concept of Its Empirical Evidence. Proc., World Conference on Transport Research, SNV Studiengesellschaft Nahverker mbH, Hamburg, Federal Republic of Germany, 1983.
5. Y. Zahavi, M.J. Beckmann, and T.F. Golob. The "UMOT"/Urban Interactions. Report DOT-RSPA-DPD-10-7. U.S. Department of Transportation, 1981.
6. Y. Zahavi. Travel Transferability Between Four Cities. Traffic Engineering and Control, Vol. 23, No. 4, 1982.
7. E.I. Pas. Analytically Derived Classification of Daily Travel-Activity Behavior: Description, Evaluation, and Interpretation. In Transporta-

tion Research Record 879, TRB, National Research Council, Washington, D.C., 1982, pp. 9-15.
8. I. Solomon and M. Ben-Akiva. Life-Style Segmentation in Travel-Demand Analysis. In Transportation Research Record 879, TRB, National Research Council, Washington, D.C., 1982, pp. 37-45.
9. J. Supernak. Transportation Modeling: Lessons from the Past and Tasks for the Future. Transportation, Vol. 12, 1983, pp. 79-90.
10. J. Supernak, A. Talvitie, and A. DeJohn. Person-Category Trip-Generation Model. In Transportation Research Record 944, TRB, National Research Council, Washington, D.C., 1983, pp. 74-83.
11. J. Supernak. Car Availability Versus Car Ownership Modeling: Theoretical Discussion and Empirical Findings. Proc., World Conference on Transport Research, SNV Studiengesellschaft Nahverkehr mbH, Hamburg, Federal Republic of Germany, 1983.
12. D. Zumkeller. Are Persons or Households the Basic Unit of Travel Demand Simulation?--The Concept of a Hybrid Model. Presented at Planning and Transport Research and Computation Company, Ltd., Summer Annual Meeting, Brighton, England, 1983.
13. R. Herz. The Influence of Environmental Factors on Daily Behavior. Environment and Planning A, Vol. 14, 1982.
14. H.F. Gunn. Travel Budgets--A Review of Experience and Modeling Implications. Transportation Research, Vol. 15A, 1981.
15. W. Brog and E. Erl. The Development of Bicycle Transportation in Selected Areas of the Federal Republic of Germany Between 1975/76 and 1980/81. Presented at 10th Planning and Transport Research and Computation Company, Ltd., Summer Annual Meeting, University of Warwick, Conventry, England, 1982.
16. G.A. Banjo and P.J. Brown. Patterns of Time-Budget Expenditure in Nigeria. In Transportation Research Record 879, TRB, National Research Council, Washington, D.C., 1982, pp. 1-9.
17. J.D. Downes and P. Emmerson. Urban Transport Modelling with Fixed Travel Budgets (An Evaluation of the UMOT Process). TRRL Report SR 799. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1983.
18. W. Brog, A.H. Meyburg, and M.J. Wermuth. Development of Survey Instruments for Determining Nonhome Activity Patterns. In Transportation Research Record 944, TRB, National Research Council, Washington, D.C., 1983, pp. 1-12.

The views expressed in this paper are solely those of the author.

Changes in Regional Travel Characteristics in the San Francisco Bay Area: 1960-1981

HANNA P. H. KOLLO and CHARLES L. PURVIS

ABSTRACT

The results of updating a travel survey in the nine-county San Francisco Bay Area are reported. The trip-making characteristics from the 1965 home-interview survey are compared with those from the 1981 telephone survey. The comparison is complemented with work trip modal shares from 1960, 1970, and 1980 census journey-to-work data. The observed changes in travel habits are traced to changes in demographic and economic characteristics in the region. Household trip rates are summarized by trip purpose, mode of travel, household size, automobile ownership, income, and housing structure type. The significance of the changes in trip rates is assessed intuitively and verified by simple statistical tests. The comparative analysis indicates that the total household trip rates are stable over long periods of time. However, there are significant shifts in the frequency of trip making by trip purpose: Households make fewer home-based shopping and personal business trips and more non-home-based trips now relative to 1965. Although some trip rates by socieconomic stratifications are significantly different in the two surveys, the overall effect on aggregate regional rates are tempered by shifts in the distribution of households by socioeconomic stratifications. Regional transit shares for work trips were found to be on the decline between 1960 and 1970, and were constant between 1970 and 1980. For those urban counties where significant transit service improvement took place between 1970 and 1980, transit work trip shares increased significantly. Public transportation appears to be absorbing more of the nonwork trip market now relative to 1965.

The purpose of this paper is to report an update of a travel survey and to investigate changes in trip characteristics since 1965 in the San Francisco Bay Area. This update was done in 1981 by using a relatively small-sample telephone survey of about 7,100 households. The earlier home-interview survey was conducted in 1965 and consisted of about 30,000 households. The survey results are corroborated by and complemented with 1960, 1970, and 1980 census journey-to-work data. The changes in travel characteristics are traced to changes in demographic, economic, and car-ownership variables.

Updating travel data for use in transportation planning has been a subject of much concern in the decades of the 1970s and the 1980s. In this era of fiscal constraints, planners and researchers have questioned the undertaking of large-scale home-interview surveys similar to those of the 1950s and the 1960s. At the same time, an equally important concern has been the use of old travel data in travel demand model development, travel forecasting, and in the day-to-day activities of metropolitan planning organizations (MPOs).

The concept of small-sample surveys grew not only out of financial necessity, but it also had popularity because of advances in the development of travel demand models. A new breed of models was in the research stages and in limited application in the early 1970s (1-3). These disaggregate behavioral models require a small sample of households, tripmakers, and trip observations for their estimation. In the San Francisco Bay Area it was found that their application in the traditional urban travel forecasting process requires aggregate validation (4,5). Furthermore, their transferability from one urban area to another hinges on a recent base year disaggregate and aggregate adaptation, where model coefficients are reestimated or adjusted to replicate known or estimated trip patterns (6).

The introduction of the journey-to-work questions in the 1970 and the 1980 Census of Population and Housing provided a valuable complement to the regional travel data bases in metropolitan areas. However, a gap still remained with regard to the need for updating nonwork travel data. It was with this realization that the San Francisco Bay Area Metropolitan Transportation Commission (MTC) embarked on its 1981 small-sample survey (7) to complement the 1980 Urban Transportation Planning Package (UTPP) data for work trips and to update the 1965 survey.

The 1965 survey was expanded by MTC in 1976 by using updated estimates of socioeconomic variables. The expansion was to total households by housing structure type and 290 zones. The sample included about 20,500 households and their weekday trips.

The 1981 household travel survey was a telephone survey of 7,091 households selected disproportionately throughout the region. About one-half of the surveyed households were residents of San Francisco County, at a sampling rate of 1.2 percent. The other eight counties had a sampling rate of 0.22 percent. Beyond this sample control total, households were selected by using telephone directory-based random digit dialing in such a way that unlisted households could be selected. The weekday component of the sample was 6,209 households. This weekday sample was weighted to the 1980 census count of households by three household-size groups and 45 districts of residence. Trip expansion combined household weighting with minor adjustment factors for missing trip data (8).

Any changes that are discerned from a comparative analysis of this type are bound to be colored by inherent biases in the data. These biases arise because of incompatible definitions, unrepresentative samples, different survey instruments and data-collection methods, data preparation approaches, and otherwise imprecise data base estimates. A special effort was made in the present analysis to prepare and report data that are as compatible as possible. For the 1965 data, the files were reprocessed by using the same trip purpose and mode aggregations as those used in the 1981 survey. The research described herein proceeded as though the data base is solid and representative. However, this may not be the case, and the readers are forewarned about such issues.

A number of points should be kept in mind as the comparisons are made and generalizations are drawn. First, the 1981 survey had a carefully selected small sample, with a follow-up for nonresponse. In contrast, the 1965 sample was much larger but had about 45 percent nonresponse or incomplete interviews, without any follow-up. Second, the 1981 survey preparation was more carefully conducted than the 1965 survey. Sample expansion used more behavioral stratifications. The 1981 survey had a better census sample frame to expand to, relative to 1965. Third, the census journey-to-work data are based on reported travel for the most frequent work trip location and mode for the week before April 1 of the census year. Survey trips are the actual weekday trips made by the respondents.

The regional travel patterns are, to a large extent, dependent on demographic and economic characteristics. Therefore, any investigation of changes in travel has to take into consideration the changes over time in such variables as household size, household income, employed persons per household, and car ownership. Reported here are regional data summarized from Bureau of the Census tapes and reports, estimates of the Association of Bay Area Governments (ABAG), and from household travel surveys conducted in the region. These are used for interpreting changes in trip characteristics. The summary data place the changes in trip making into a demographic and economic context and shed some light on the possible biases regarding representation of these variables in the surveys.

The San Francisco Bay Area consists of nine counties surrounding the Bay. About 5 million people in some two million households live in this vast region of 4.5 million acres. About 2.5 million jobs provide employment opportunities for its residents (9).

A summary of aggregate regional growth from 1960 to 1980 is given in Table 1. Between 1960 and 1970 the growth was 27 percent in total population, 32 percent in the number of households, 31 percent in employed residents, 53 percent in total school enrollments (ages 3 to 34), 149 percent in college enrollments, and 27 percent in kindergarten and elementary school (grades 1 through 8) enrollments. The decade of the 1970s recorded a growth of 12 percent in total population, 27 percent in the number of households, 36 percent in employed residents, 6 percent in total school enrollments, 81 percent in college enrollments, and -18 percent in kindergarten and elementary school enrollments. The decline over time in household size is evident from the data in Table 1. This is accompanied by an increase in the number of employed persons per household, income per household, drivers per household, and cars per household. These are important variables that influence regional travel in the aggregate and by market segment.

CHANGES IN REGIONAL HOUSEHOLD TRIP RATES, 1965-1981

A comparative analysis is undertaken here for trip rates by trip purpose, mode of travel, and household stratifications commonly used in travel analyses.

TABLE 1 Regional Demographic and Economic Characteristics, 1960-1980

Variable	1960 Census	1965 (ABAG)	1965 BATSC Survey[a]	1970 Census	1980 Census	1980-1981 Survey
Total population (000s)	3,639	4,216		4,628	5,180	
Population in households (000s)	3,515	4,106	4,331	4,501	5,059	5,051
Households (000s)	1,174	1,387	1,387	1,553	1,971	1,970
Employed residents (000s)	1,433	1,664	1,697	1,882	2,555	2,639
School enrollment (000s)						
Total	904			1,380	1,464	
Kindergarten and grades 1-8	616			782	642	
High school	195			326	333	
College	93			232	419	
Mean household income ($)		9,353	9,592	11,251	24,350	26,517
Household size	2.99	2.96	3.12	2.90	2.57	2.56
Employed persons per household	1.22	1.20	1.22	1.21	1.30	1.34
Drivers per household			1.67			1.75
Automobiles per household	1.12		1.40	1.33	1.68	1.70
Automobile ownership (%)						
Households with no car	20		14	16	12	11
Households with one car	53		44	44	36	35
Households with two cars	24		34	33	33	36
Households with three or more cars	3		8	7	19	18

[a]BATSC = Bay Area Transportation Study Commission.

The definitions of trip purposes and modes follow the traditional terms. Home-based work trips are those to and from work and work-related business. Home-based shop is a catchall category that includes shopping, personal business, and other trip purposes. Person mode is the summation of vehicle driver, vehicle passenger, and transit passenger. Other mode includes motorcycle, moped, and bicycle trips.

It should be noted that data used in this paper are taken from an array of census reports, MTC reports, and special tabulations. An MTC report (10) contains most of the 1981 survey data cited. The 1965 data are in special tabulations recently completed by MTC staff.

Before discussing the specifics of the comparison, the importance of household trip rates as prime determinants of total travel in transportation planning is stressed. Any changes in the rates from past surveys are of prime concern to transportation analysts. Such changes are not only important for updating trip-generation models, but are also used in microanalyses in subarea and facility planning.

Household Trip Rates by Purpose and Mode

The comparison between 1965 and 1981 trip rates is given in Table 2 by trip purpose and mode. Overall, total trips per household decreased by about 1 percent. This small change suggests that the effects of energy shortages in 1973 and 1979 on trip making have stabilized. By trip purpose, the change in trip rates ranges between -17 percent for home-based shop and +23 percent for non-home-based trips. Work trip rates increase by 2 percent, social-recreation trip rates increase by 7 percent, and school trips decrease by 13 percent. The increase in work trip rates is insignificant. The decrease in school trips is traced to drops in school enrollments for kindergarten and elementary school grades 1 through 8. A comparison of the 1970 and 1980 census data on

TABLE 2 Weekday Regional Trips per Household by Purpose and Mode, 1965 Versus 1981

Mode	Home Based Work	Home Based Shop	Home Based Social-Recreation	Home Based School	Nonhome Based	Total
In-vehicle person						
1965	1.518	2.307	0.915	0.295	1.499	6.535
1981	1.558	1.964	1.011	0.387	1.894	6.814
Difference (%)	3	-15	10	31	26	4
Transit						
1965	0.220	0.085	0.035	0.086	0.060	0.486
1981	0.206	0.085	0.044	0.126	0.097	0.558
Difference (%)	-6	0	26	47	62	15
School bus						
1965				0.146		0.146
1981				0.089		0.089
Difference (%)				-39		-39
Walk						
1965	0.090	0.286	0.177	0.514	0.281	1.348
1981	0.076	0.188	0.143	0.285	0.303	0.995
Difference (%)	-16	-34	-19	-45	8	-26
Other						
1965	0.031	0.053	0.057	0.057	0.065	0.263
1981	0.050	0.037	0.063	0.065	0.042	0.257
Difference (%)	61	-30	11	14	-35	-2
Total						
1965	1.858	2.732	1.184	1.097	1.906	8.777
1981	1.89	2.274	1.262	0.952	2.335	8.713
Difference (%)	2	-17	7	-13	23	-1

school enrollments for these grades shows that the decrease in enrollments was about 18 percent. This is a symptom of the demographic changes in the household composition of the past decade. (See Table 1 for changes in household size and school enrollments.)

A significant change appears to have occurred in travel behavior between 1965 and 1981. Household members have switched their travel habits from home-based shopping trips to non-home-based trips. This change is interpreted intuitively as a response to increases in travel costs and gasoline shortages of the past decade. It appears that households have switched from their frequent home to shop and personal business activities to combining their chores into multileg tours, thus increasing the number of non-home-based trips.

By mode, the range of variation in trip rates is between -39 percent for school bus passengers and +15 percent for transit passengers and vehicle drivers. Person trips increase 5 percent and walk trips decrease by 26 percent. The decrease in walk trips is universal over all trip purposes: although non-home-based walk trips increase by 8 percent, its share of total non-home-based trips drops from 15 to 13 percent. The largest drop in walk trip rates is for school trips. This is caused by the drop in enrollments for kindergarten and elementary schools, as previously noted. It is reasonable to assume that walk to school is largely a market for students in kindergarten through 8th grade, and therefore a drop in such enrollments will cause a drop in walk to school. The changes in the walk mode for other trip purposes appears to be symptomatic of more multileg tours where the walk mode cannot compete with other modes for such a diversified market of trip purposes. The substantial change in the school bus passenger mode is due to the passage of Proposition 13 in California in 1978. This change in real property taxation yielded major reductions in local government revenues, including school bus programs. The slack was taken by higher patronage for automobile and public transportation.

A number of studies (11-14) have addressed the stability of trip frequency, trip-generation models, and travel time characteristics. A few of these studies (11,12) present trip rate data comparable with those reported here. Furthermore, their comparison is for a much earlier time span before the 1973 and 1979 gasoline shortages. Seven U.S. cities studied by an ITE committee (11) show an average increase from 6.5 to 7.7 person trips per household. This 18 percent increase over an average period of 12.4 years is in contrast to the results in this study of 5 percent (7.021 to 7.372) over a period of 16 years. On the basis of this comparison it appears that the energy shortages of the 1970s have moderated the increases in trip rates.

Household Trip Rates by Household Size

In travel forecasting, trips are sometimes generated by household size. Alternatively, some travel demand models incorporate average household size as an explanatory variable in linear-regression models. The average effect of household size on trip making is assessed here by analyzing trips per household by household size (Table 3). Trips per person can also be computed, but the percentage change will be the same.

For total trips, all household-size groups experience an increase in trips per household. However, the average household trip rate remains unchanged. This is because of a major shift in the regional distribution of households by household size, as shown in Table 3. There is now a much larger proportion of households in the one-person group, and much less in the five-or-more-person group. This is supported not only by the two surveys but by the 1970 and 1980 censuses as well.

For work trips, the small household-size groups experience little change in trips per household. As household size increases, the change in trips per household increases. This is because larger households have a higher number of employed persons now as compared with 1965.

For shopping trips, all households experience a drop in trips except for the one-person group. For the balance of the trip purposes, all household-size groups increase their trip making. However, the net effect on school trips is a reduction in the regional trip rate. This is also due to changes in

TABLE 3 Weekday Regional Trips per Household by Household Size, 1965 Versus 1981

Household Size	Percentage of Households	Home Based: Work	Home Based: Shop	Home Based: Social-Recreation	Home Based: School	Nonhome Based	Total
1 person							
1965	15	0.883	0.903	0.523	0.060	0.966	3.335
1981	26	0.889	0.966	0.622	0.086	1.390	3.953
Difference (%)		1	7	19	43	44	19
2 persons							
1965	30	1.734	1.874	0.823	0.164	1.574	6.169
1981	33	1.767	1.868	1.075	0.267	2.103	7.079
Difference (%)		2	0	31	63	34	15
3 persons							
1965	18	2.137	2.618	1.104	0.762	1.821	8.443
1981	16	2.262	2.539	1.310	0.937	2.598	9.646
Difference (%)		6	-3	19	23	43	14
4 persons							
1965	17	2.193	3.666	1.479	1.613	2.358	11.309
1981	15	2.646	3.612	1.896	2.087	3.293	13.533
Difference (%)		21	-1	28	29	40	20
≥5 persons							
1965	20	2.273	4.796	2.083	3.220	2.849	15.222
1981	10	3.183	4.585	2.506	3.729	3.715	17.717
Difference (%)		40	-4	20	16	30	16
All households							
1965	100	1.858	2.732	1.184	1.098	1.906	8.778
1981	100	1.890	2.274	1.262	0.952	2.335	8.713
Difference (%)		2	-17	7	-13	23	-1

the regional distribution of households by household size.

Household Trip Rates by Automobile Ownership

Automobile ownership is an important household characteristic that determines mobility and trip making. A comparison of trips per household by automobile-ownership group is given in Table 4. As can be seen, household trip rates increase as automobile ownership increases. This is due, in part, to the high correlation between automobiles owned and household size. The changes are minimal for total trips, except for the one- and two-automobile households.

For work trips there is a decrease in trips per household for the zero- and one-car owners. This is balanced by an increase for the four-or-more-automobile group. These shifts can be interpreted as symptoms of the high unemployment in 1981 relative to 1965 for the low automobile-ownership group.

For shopping trips the reduction in the rates is in contrast to the increase in non-home-based rates, as noted earlier. This holds true for most automobile-ownership groups.

For social-recreation trips there are modest decreases in the trip rates for households who own cars, in contrast to the increase for those who do not own cars. The increase can be inferred from the increase in the number of persons in old or retired households who have more leisure time. This group also increased its transit share for social-recreation trips from 20 percent in 1965 to 27 percent in 1981.

School trip rates drop for the medium automobile-ownership groups and rise for the high automobile-ownership group. The reduction is due to a drop in walk and school bus passengers more than the automobile modes. The increase in the high automobile-ownership trip rate is due to increased college enrollments, which is related more to the automobile mode than other modes.

Household Trip Rates by Housing Structure Type

Housing structure type has been used in many travel demand analyses as a stratification for trip generation. It is a surrogate variable for household size, income, and automobile ownership. With the changes occurring in household preferences, prompted by the high costs of housing, there are some questions regarding the use of this variable in place of more behavioral variables that it purports to represent. The increase in apartment conversion to condominiums and the introduction of townhouse developments have provided opportunities for a change in household composition of those families who choose to or are forced to occupy multifamily structures. Condominiums and townhouses are used nowadays by wealthy households and households of medium size. Their trip-making characteristics may not coincide with apartment dwellers. Therefore, an investigation of their trip characteristics is in order.

The data in Table 5 give a comparison of household trip rates by housing structure type and the changes that have occurred. For condominiums and townhouses, the trip rates given are from the 1981 survey only because they were not reported in the 1965 survey. As can be seen, condominium and townhouse dwellers have higher total trip rates than apartment dwellers, lower rates than single-family dwellers, and rates close to duplex dwellers.

The change between 1965 and 1981 for single-family structure type is small for total trips. Home-based work trip rates increase by 8 percent because of increases in employed persons per household. Home-based shopping trip rates decrease and non-home-based trip rates increase. Social-recreation trips increase and school trips decrease. All these changes are manifestations of the phenomena observed earlier. Apartment dwellers decrease their trip-making rates for work and total trips more than any other housing structure type. The drop in work trips is attributed to higher unemployment in 1981 relative to 1965. The drop in social-recreation trips is a sign of the hard economic times the region is experiencing. The increase in school trips is small and is attributed to larger households (with children) shifting to apartment housing.

Household Trip Rates by Income

Household income continues to be a significant vari-

TABLE 4 Weekday Regional Trips per Household by Automobile Ownership, 1965 Versus 1981

Automobile Ownership	Percentage of Households	Home Based: Work	Home Based: Shop	Home Based: Social-Recreation	Home Based: School	Nonhome Based	Total
No car							
1965	14	0.940	1.139	0.562	0.493	0.778	3.912
1981	11	0.724	1.159	0.629	0.541	0.942	3.996
Difference (%)		-23	2	12	10	21	2
1 car							
1965	44	1.669	2.430	1.022	0.891	1.606	7.618
1981	35	1.298	1.738	0.974	0.577	1.713	6.301
Difference (%)		-22	-28	-5	-35	7	-17
2 cars							
1965	34	2.233	3.452	1.469	1.488	2.471	11.113
1981	36	2.211	2.661	1.412	1.062	2.754	10.101
Difference (%)		-1	-23	-4	-29	11	-9
3 cars							
1965	6	2.782	3.992	1.905	1.616	2.979	13.274
1981	12	2.789	3.246	1.827	1.600	3.314	12.776
Difference (%)		0	-19	-4	-1	11	-4
≥4 cars							
1965	2	3.214	4.291	2.002	1.647	3.509	14.663
1981	6	3.684	3.054	2.008	1.880	3.954	14.580
Difference (%)		15	-29	0	14	13	-1
All households							
1965	100	1.858	2.731	1.184	1.097	1.906	8.777
1981	100	1.890	2.274	1.262	0.952	2.335	8.713
Difference (%)		2	-17	7	-13	23	-1

TABLE 5 Weekday Trips per Household by Housing Structure Type, 1965 Versus 1981

Housing Structure Type	Percentage of Households	Home Based				Nonhome Based	Total
		Work	Shop	Social-Recreation	School		
Single family							
1965	66	1.978	3.257	1.353	1.380	2.164	10.131
1981	64	2.134	2.727	1.467	1.230	2.639	10.196
Difference (%)		8	-16	8	-11	22	1
Condominium or townhouse							
1965	0	NA	NA	NA	NA	NA	NA
1981	5	1.924	1.807	1.148	0.454	2.339	7.672
Difference (%)		NA	NA	NA	NA	NA	NA
Duplex							
1965	8	1.713	2.068	0.930	0.778	1.377	6.866
1981	6	1.665	1.902	1.131	0.626	2.038	7.362
Difference (%)		-3	-8	22	-20	48	7
Apartment							
1965	26	1.632	1.667	0.861	0.509	1.459	6.128
1981	25	1.428	1.411	0.809	0.521	1.667	5.835
Difference (%)		-12	-15	-6	2	14	-5
All							
1965	100	1.859	2.731	1.184	1.097	1.906	8.778
1981	100	1.890	2.274	1.262	0.952	2.335	8.713
Difference (%)		2	-17	7	-13	23	-1

able in determining trip-making characteristics. Comparative trip rates between 1965 and 1981 are given in Table 6 by income group. The low, medium, and high groups are defined by selecting households from the two surveys to form approximately equal proportions based on the income distribution of households in the two surveys.

The data indicate that total household trip rates have dropped by about 4 percent for the medium-income group. For the high-income group total trip rates have increased by 3 percent. Work trips are down by 6 percent for the low-income group and are stable for the medium-income group. Work trips for the high-income group have increased by 11 percent, an indication of an increase in employed persons per household. Shopping trip rates are down significantly for all groups, except for those households that refused to report their income. Social-recreation trip rates have not changed for the medium-income group, but have increased 7 and 10 percent for low and high income, respectively. School trip rates decreased across the board and non-home-based trip rates increased significantly.

Statistical Tests of Significance for Changes in Trip Rates

Differences between 1965 and 1981 trip rates per household were assessed in the previous sections by inspecting the percentage changes by trip purpose and mode for the two surveys. Intuitive judgments and interpretations were made by analyzing the changes in demographic and economic variables over the same period of time. In contrast, the statistical measures associated with the trip rates are summarized in Table 7 for selected trip purposes and modes.

Sample means, standard deviations, and standard error of the means are calculated in Table 7. These sample descriptors are estimates of the true population statistics. The standard error of the mean is the standard deviation of the sampling distribution of the mean trip rates. Confidence intervals around the means were established at the 0.05 level for a two-tailed test. Standard errors of the difference between means were estimated manually. A t-statistic for the difference between sample means was con-

TABLE 6 Weekday Regional Trips per Household by Income, 1965 Versus 1981

Income Group	Home Based				Nonhome Based	Total
	Work	Shop	Social-Recreation	School		
Low						
1965	1.067	2.003	0.895	0.713	1.248	5.925
1981	1.004	1.829	0.958	0.677	1.524	5.992
Difference (%)	-6	-9	7	-5	22	1
Medium						
1965	1.971	3.030	1.292	1.174	1.951	9.418
1981	2.018	2.333	1.291	0.984	2.432	9.058
Difference (%)	2	-23	0	-16	25	-4
High						
1965	2.490	3.419	1.521	1.396	2.764	11.590
1981	2.772	2.795	1.668	1.228	3.422	11.886
Difference (%)	11	-18	10	-12	24	3
Income not reported						
1965	1.633	1.913	0.753	0.960	1.145	6.403
1981	1.594	2.022	1.030	0.869	1.692	7.207
Difference (%)	-2	6	37	-9	48	13
All households						
1965	1.858	2.731	1.184	1.097	1.906	8.777
1981	1.890	2.274	1.262	0.952	2.335	8.713
Difference (%)	2	-17	7	-13	23	-1

TABLE 7 Statistical Analysis of Variation in Average Trip Rates, 1965 Versus 1981, Regional Weekday Trips by Purpose and Mode

		1965 Survey			1981 Survey					
Market Stratifier	Market Segment	Mean	Standard Deviation	Standard Error of Mean	Mean	Standard Deviation	Standard Error of Mean	Standard Error of Difference Between Means	t-Score	Significant Difference
Trips per Household										
Trip purpose	HBW	1.858	1.774	0.012	1.890	1.876	0.024	0.026	1.23	no
	HBSH	2.731	3.716	0.026	2.274	2.778	0.035	0.051	8.96	yes
	HBSR	1.184	2.445	0.017	1.262	2.034	0.026	0.034	2.28	yes
	HBSK	1.097	2.280	0.016	0.952	1.883	0.024	0.032	4.56	yes
	NHB	1.906	3.481	0.024	2.335	3.225	0.041	0.050	8.65	yes
Trip mode	Vehicle driver	4.534	4.549	0.032	5.231	4.765	0.060	0.067	10.46	yes
	In-vehicle	6.535	6.981	0.049	6.814	6.237	0.079	0.099	2.83	yes
	Transit	0.486	1.071	0.007	0.558	2.028	0.026	0.020	3.67	yes
	Person	7.021	6.942	0.048	7.372	6.223	0.079	0.098	3.57	yes
Total		8.777	8.129	0.057	8.713	7.091	0.090	0.114	0.56	no
Trips per Person										
Trip purpose	HBW	0.595	0.709	0.005	0.737	0.806	0.010	0.011	3.77	yes
	HBSH	0.875	1.068	0.007	0.887	1.097	0.014	0.016	0.58	no
	HBSR	0.379	0.731	0.005	0.492	0.858	0.011	0.011	8.06	yes
	HBSK	0.351	0.480	0.003	0.371	0.481	0.006	0.007	16.82	yes
	NHB	0.610	1.212	0.008	0.911	1.506	0.019	0.019	11.80	yes
Trip mode	Vehicle driver	1.452	1.681	0.012	2.041	2.074	0.026	0.026	22.84	yes
	In-vehicle	2.092	2.074	0.014	2.658	2.369	0.030	0.031	18.20	yes
	Transit	0.156	0.461	0.003	0.218	0.858	0.011	0.008	7.40	yes
	Person	2.248	2.037	0.014	2.876	2.273	0.029	0.030	20.70	yes
Total		2.810	2.172	0.015	3.399	2.531	0.032	0.033	13.98	yes

Note: HBW = home-based work, HBSH = home-based shop, HBSR = home-based social-recreation, HBSK = home-based school, and NHB = nonhome based.

structed by using standard statistical formulas (15). This assumes random independent samples that have a normal sampling distribution of the mean trip rates. The judgment about the significance of the differences between 1965 and 1981 trip rates is based on the computed and tabled t-statistics. When the computed t-statistic is greater than 1.960 (table t-statistic at 0.05 level), the null hypothesis that the two means are equal is rejected. Therefore, the significant difference is labeled yes. If the computed t-statistic is less than 1.960, the null hypothesis that the two means are equal is not rejected. Therefore, the significant difference is labeled no.

The summary statistics of Table 7 suggest that total trips and home-based work trips per household from the two surveys are not significantly different. The other trip purposes are. Trip rates per person are significantly different for all trip purposes, except home-based shopping. This shows the effect of changes in household size on household trip rates. By mode, the trip rate per household and per person are significantly different for the drivers, the in-vehicle person, the transit, and the person mode.

The comparison between the statistical test performed here and the percentage changes reported earlier indicates that changes in total trip rates (by purpose or mode) of 5 percent and over can be considered significant. Changes of less than 5 percent are insignificant.

Note that assessing the significance of the differences, statistically or intuitively, should be taken for what it is. The size of the sample by cell, the magnitude of the trip rate, and the proportion of trips by a market segment should also be considered as judgments are made about the change and in the use of rates for forecasting.

CHANGES IN AGGREGATE TRIP CHARACTERISTICS, 1960-1980

Aggregate data are areawide estimates derived from expanded survey, expanded census, or 100 percent counts. Aggregate trip characteristics discussed in this section represent average regional weekday travel. Their value lies in understanding the overall composition of the travel market or in data factoring. The data are referred to interchangeably as 1980/1981 travel. This is because the 1981 survey is expanded to 1980 households and, therefore, it represents 1980 travel. The assumption is that the household trip-making characteristics did not change between 1980 and 1981.

Distribution of Trips by Trip Purpose and Mode

The data in Table 8 give the trip purpose shares by mode for the regional trips in the two surveys. Between 1965 and 1980 work trips hold their share of the market, social-recreation trips remain relatively stable, school trips drop their share, and non-home-based trips increase by the same amount that shopping trips decrease (5 percent). This shows similar signs of change as those observed earlier in the trip rate analysis.

The trip purpose shares by mode fluctuate more than the total. The direction of shift between shopping and non-home-based trips is consistent across all modes. Another important change shown in Table 8 relates to public transportation. Of the total transit trips, work trip purpose share drops from 45 to 37 percent.

The regional modal shares for work trips from 1960 to 1981 are given in Table 9. Two estimates are shown for the 1965 and 1981 surveys. Home-based work (HBW) is the traditional definition. Home-based-work census-comparable (HWC) is an estimate that takes

TABLE 8 Regional Trip Purpose Shares (%) by Mode of Travel, 1965 Versus 1981

Mode of Travel	Home Based Work	Shop	Social-Recreation	School	Nonhome Based	Total
In-vehicle person						
1965	23.2	35.3	14.0	4.5	22.9	100
1981	22.9	28.8	14.8	5.7	27.8	100
Transit						
1965	45.3	17.6	7.2	17.6	12.3	100
1981	36.9	15.2	8.0	22.6	17.3	100
Total						
1965	21.2	31.1	13.5	12.5	21.7	100
1981	21.7	26.1	14.5	10.9	26.8	100

TABLE 9 Regional Modal Shares (%) for Work Trips, 1960 to 1980

Mode of Travel to Work	1960 Census	1965 Survey HWC	1965 Survey HBW	1970 Census	1980 Census	1981 Survey HWC	1981 Survey HBW
Vehicle driver	NA	68.4	69.2	70.9	71.3	72.0	73.0
Vehicle passenger	NA	12.6	12.4	8.9	9.4	9.6	9.4
In-vehicle person	73.4	81.0	81.6	79.8	80.7	81.6	82.4
Transit	16.2	12.7	11.9	11.6	11.6	11.7	10.9
Walk	8.2	5.0	4.8	5.9	4.5	4.2	4.0
Other	2.3	1.3	1.6	2.8	3.2	2.5	2.6

Note: The modal shares for total travelers for all columns equal 100 percent.

into consideration the modal components used in the 1980 census. The census shares are from data in published reports (16-18). The data in Table 9 indicate that in-vehicle person share to work increased by 6 percent from 1960 to 1970 and increased another 1 percent by 1980. Transit shares were on the decline between 1960 and 1970, and remained stable between 1970 and 1980. Walk trip shares continue to decline since 1960. This decrease is a sign of continued suburbanization in the region, where residences are increasingly farther from jobs for the walk mode to hold its own.

In Table 10 the modal percentage shares by trip purpose from the two surveys are given. In-vehicle person trip share increases moderately for all trip purposes except school, which increases sharply. Transit work trip share shows a decline between 1965 and 1981. Nonwork transit trip shares show an increase by a moderate amount, except for school trips, where the share doubles. The moderate increase in nonwork transit shares are understated because the decline between 1960 and 1970 of work trip transit share (shown in Table 9) probably applies to nonwork trips as well. This means that between 1970 and 1980 nonwork trip transit shares increased more than indicated by the data in Table 10. Walk and other mode trip shares declined for all trip purposes.

Comparison of county transit shares (not reported here) indicates that, for total trips, all nine counties increased their transit share between 1965 and 1980. For work trips, the urban counties that had improvements in bus and rail service increased their transit share significantly. Taking all these statistics together (Tables 8-10), it is reasonable to assume that transit is now absorbing more of the nonwork trip market.

Car Occupancy by Trip Purpose

Car occupancy is an important variable for assessing trends and for converting automobile-person trips to vehicle trips. The use of such an average is predicated by the absence of reliable car-occupancy models.

TABLE 10 Regional Modal Shares (%) by Trip Purpose, 1965 Versus 1980/1981

Mode of Travel	Home Based Work	Shop	Social-Recreation	School	Nonhome Based	Total
In-vehicle person						
1965	81.6	84.4	77.3	26.9	78.7	74.5
1981	82.4	86.4	80.1	40.6	81.1	78.2
Transit						
1965	11.9	3.1	3.0	7.8	3.2	5.5
1981	10.9	3.7	3.5	13.3	4.1	6.4
School bus						
1965	-	-	-	13.3	-	1.7
1981	-	-	-	9.3	-	1.0
Walk and other						
1965	6.4	12.4	19.8	52.0	18.1	18.4
1981	6.7	9.9	16.3	36.8	14.8	14.4
Total						
1965	100.0	100.0	100.0	100.0	100.0	100.0
1981	100.0	100.0	100.0	100.0	100.0	100.0

The data in Table 11 give comparative regional occupancies by trip purpose from 1965 to 1981. They are computed from aggregate data for vehicle driver and vehicle passenger modes. Between 1970 and 1980 work trip car occupancies remain constant. Comparison of 1965 and 1970 data suggests that work trip occupancies were on the decline during the 1960s. It is unfortunate that there are no data from the 1960 census to verify this apparent declining trend.

For nonwork trips the data show a decline in vehicle occupancies of 4 to 20 percent. The 20 percent decline in school trip occupancy may be due to a decline in school enrollments for grades 1 through 8. The students in these grades are a potential market for carpooling (children driven) to school. Another factor here is the increase in college enrollments, a potential low car-occupancy group for school trips.

TABLE 11 Comparative Regional Weekday Car Occupancies by Trip Purpose

Trip Purpose	1965	1970	1980	1981
Home-based work	1.18	1.13	1.13	1.13
Home-based shop	1.44			1.24
Home-based social-recreation	1.81			1.73
Home-based school	2.78			2.23
Nonhome based	1.45			1.25
Total	1.44			1.30

The decline in shopping and non-home-based car occupancies may be due to the combined effect of a decrease in household size and the making of fewer home-based trips in favor of more non-home-based trips. Obviously, because there were fewer household members during the 1970s relative to the 1960s, car occupancy for home-based shopping trips was lower. As more trips are combined into multileg tours, there is less of a chance for carrying passengers to the diversified activities conducted in non-home-based locations.

It should be pointed out that aggregate regional data do not necessarily reflect specific corridor or local highway car occupancies. Whereas the average occupancies may be stable or declining, major-corridor occupancies are on the increase for peak commuting periods in the San Francisco Bay Area.

Reported Trip Duration by Trip Purpose and Mode

In both the 1965 and the 1981 surveys, respondents were asked to record the times at the beginning and at the end of their trips. The resulting door-to-door one-way trip times showed minor changes for total trip purposes or total modes. By purpose, work trips are longest and shopping trips are shortest. By mode, transit trips are longer in 1981 than in 1965 by about 5 to 8 min for work, shop, school, and non-home-based trips. Social-recreation transit trip lengths are longer by about 14 min. This is an indication that residents of the region are using available transit to farther destinations relative to 1965.

The trip length frequency distributions by purpose and mode were also compared and found to be quite similar. The distributions were not smooth, but had kinks at 5-min intervals for all trip purposes. This is a well-known phenomena, where respondents tend to report the times to the nearest 5 min. Because of this, smooth network travel times are used in most travel demand analyses instead of survey-reported travel times.

The 1980 census data indicate that the average regional home-to-work trip length is 24 min (18). The 1981 home-based work trip length was found to be 27 min, 13 percent higher than the census data. Because of the differences between sample sizes and definitions, the 1981 estimate may not be unreasonable.

SUMMARY AND CONCLUSIONS

Updating large-scale, old home-interview travel surveys with a small sample is worthwhile. It provides up-to-date information, comparative trip characteristics for investigating changes over time, and valuable data sets for disaggregate model development.

Household trip rates were found to be constant for total weekday trips. However, a shift has occurred between trip purposes: households made fewer home-based shopping and personal business trips and more non-home-based trips in 1981 relative to 1965. This is an indication that the frequent home-based trips are being combined into multileg tours, thus increasing the number of non-home-based trips.

Household trip rates by socioeconomic stratifications have undergone some change. However, changes for the average regional household are much less due to shifts in the distribution of households by these stratifications.

The work trip transit share for the region from the 1981 survey was the same as that reported in the 1980 census journey-to-work data. This share declined between 1960 and 1970. Between 1970 and 1980 the regional transit share was constant, but increased in those counties where transit service improvements were introduced.

Between 1965 and 1981 transit shares for nonwork trips increased for every county. The statistics suggest that public transportation is now absorbing more of the nonwork travel market relative to 1965.

Average regional car occupancies for work trips declined during the 1960s and remained stable in the 1970s. For nonwork trips, average occupancies declined between 1965 and 1981 because of changes in household size and combining of trips into multileg tours.

Regional trip length frequency distributions reported by the respondents in the two surveys were found to be grouped into 5-min intervals. The changes in regional trip lengths between 1965 and 1981 were negligible.

ACKNOWLEDGMENTS

The work reported in this paper was financed in part through grants from UMTA, U.S. Department of Transportation, under the Urban Mass Transportation Act of 1964, as amended; and through grants from FHWA and the California Department of Transportation. The authors are indebted to Patrick Hackett of MTC for file preparation, computer programming, and data processing. Joanna Grant assisted in the computations. Shelley Roberts and Brenda Germany typed the manuscript.

REFERENCES

1. Charles River Associates, Inc. A Disaggregate Behavioral Model of Urban Travel Demand. FHWA, U.S. Department of Transportation, 1973.
2. M.E. Ben-Akiva. Structure of Passenger Travel Demand Models. Ph.D. thesis. Department of

Civil Engineering, Massachusetts Institute of Technology, Cambridge, 1973.
3. D.L. McFadden. The Measurement of Urban Travel Demand. Journal of Public Economics, Vol. 3, 1974, pp. 303-328.
4. Cambridge Systematics. Travel Model Development Project: Phase 2 Final Report. Metropolitan Transportation Commission, Berkeley, Calif., 1980, Vol. 1-3.
5. Travel Model Validation Summary Report: Santa Clara Valley Corridor Evaluation. Metropolitan Transportation Commission, Berkeley, Calif., 1978.
6. T.J. Atherton and M.E. Ben-Akiva. Transferability and Updating of Disaggregate Travel Demand Models. In Transportation Research Record 610, TRB, National Research Council, Washington, D.C., 1976, pp. 12-18.
7. M.M. Reynolds, S.M. Flynn, and D.B. Reinke. 1981 San Francisco Bay Area Travel Survey. In Transportation Research Record 877, TRB, National Research Council, Washington, D.C., 1982, pp. 51-58.
8. Sample Weighting and Trip Expansion: 1981 MTC Travel Survey. Working Paper 4A. Metropolitan Transportation Commission, Berkeley, Calif., 1983.
9. Projections 83: Forecasts for the San Francisco Bay Area, 1985-2000: Population, Employment, and Housing. Association of Bay Area Governments, Berkeley, Calif., 1983.
10. 1980 Regional Travel Characteristics: 1981 MTC Travel Survey. Working Paper 8. Metropolitan Transportation Commission, Berkeley, Calif., 1983.
11. ITE Technical Council Committee 6F-12. Consistency of Origin-Destination (O-D) Characteristics Through Time. ITE Journal, Vol. 49, No. 10, 1979, pp. 32-39.
12. G.S. Cohen and M.A. Kocis. Components of Change in Urban Travel. In Transportation Research Record 775, TRB, National Research Council, Washington, D.C., 1980, pp. 42-47.
13. E.J. Kannel and K.W. Heathington. Temporal Stability of Trip Generation Relations. In Highway Research Record 472, HRB, National Research Council, Washington, D.C., 1973, pp. 17-27.
14. Y. Zahavi and J.M. Ryan. Stability of Travel Components Over Time. In Transportation Research Record 750, TRB, National Research Council, Washington, D.C., 1980, pp. 19-26.
15. H.M. Blalock, Jr. Social Statistics, 2nd ed. McGraw-Hill, New York, 1979.
16. U.S. Census of Population, 1960--Volume I: Characteristics of the Population, Part 6, California. Bureau of the Census, U.S. Department of Commerce, 1963, Table 82.
17. U.S. Census of Population and Housing, 1970. Census Tracts PHC (1)-189, 190, 192, 223. Bureau of the Census, U.S. Department of Commerce, 1972, Table P-2.
18. Travel and Mobility Characteristics: 1980 Census. Working Paper 1. Metropolitan Transportation Commission, Berkeley, Calif., 1983.

The opinions, findings, and conclusions expressed herein are those of the authors and do not necessarily reflect those of MTC or the funding agencies.

An Update on Household-Reported Trip-Generation Rates

JOANNA M. BRUNSO and DAVID T. HARTGEN

ABSTRACT

In this study person trip rates determined in a statewide telephone survey in New York State during January 1983 are reviewed. The results of the study indicated that the average adult in New York makes 2.8 one-way trips per day, and lives 4.5 miles from work and 2.2 miles from shopping. There were no significant differences in average trip rates between upstate and downstate New York, or between urban, suburban, small town, or rural areas of the state. Trip rates vary with income, employment status, sex, number of household vehicles, and presence of children. It was concluded that trip-generation rates are largely transferable between geographic areas, if demographic differences are accounted for, and that transportation planners can have confidence in applying person trip rates from this and other surveys.

In order to plan intelligently for transportation, transportation planners must anticipate changes in travel. Many cities are losing population, as are some established suburbs; others are growing rapidly (1). Such changes can be more easily studied if existing trip information from one region can be transferred to another region. But this presumes stable trip-generation rates over time--an untested assumption. In the 12-year period between 1962 and 1974, trip-generation rates for home-based New York State households were found to be largely stable (2), both in the aggregate and for demographic groups. Changes in travel observed in that period were accounted for almost entirely by changes in the number of households in each group and changes in the total magnitude of households.

The results of a telephone survey conducted in New York State in 1983 are described in this paper. The purpose of the survey was to learn how residents of the state were conserving energy, but current information on trip rates for various activities was also obtained. The differences in trip rates by different demographic and geographic groups are reviewed for weekends and weekdays by trip purpose and mode. Although the question of stability of trip rates over time was not thoroughly investigated, the relative stability of trip rates over place was established, thereby substantiating previous studies that conclude that the primary determinants of trip rates are demographic, not geographic.

BACKGROUND

In the 1960s most urban areas (more than 50,000 population) conducted home-interview travel surveys in which data on trip-generation rates were collected (3); many of these surveys were updated in the 1970s. Trip-generation rates were generally expressed by zonal or cell (aggregated) data, either with cross-classified or regression models (4) and usually at the household level. As early as 1977, Dobson and McGarvey (5) demonstrated the empirical equivalence of regression and cross-classification models of home-based travel. Recent work by Stopher and McDonald (6) extends the analysis of variance (ANOVA) approach. A major compilation (7) compared home-based and non-home-based rates across different cities, within cells, by income level. Recent comparisons of rates from different areas (8,9) showed general stability over space and time within cells defined by income, automobile ownership, or family size. More recent studies (10) have identified discrepancies between home-generated travel and non-home-generated travel, which have not been resolved. Recent research has focused on the life cycle of the individual or household (11-15). The life cycle uses age and employment status of the household head and spouse and the number of children. McDonald and Stopher (16), however, found little empirical justification in the use of such variables. Person-level analyses (9,17) have recently been proposed. In analyzing the trip data, both points of view will be considered.

METHODOLOGY

A random sample telephone survey of 1,503 New York State residents 18 years of age or older was conducted between January 9 and February 2, 1983. Only one adult per household was interviewed; thus the trip rates presented here are person trip rates. The sample was stratified so that men and women in each county were sampled in proportion to 1980 population. The survey slightly underrepresented New York City, zero vehicle households, and low-income households (Table 1). However, it must be remembered that the survey excluded all those younger than 18 years of age as well as those who were institutionalized or without household telephones.

TABLE 1 Comparison of Survey Response with 1980 Census

	Survey		1980 Census
	No.	Percent	(%)
Region			
New York City	600	40.0	43.9
Long Island	223	14.8	14.8
Westchester/Rockland	95	6.3	6.4
Upstate	585	38.9	34.8
Sex			
Male	642	42.7	47.5
Female	861	57.3	52.5
Vehicles per household			
0	274	18.2	38.1
1	609	40.5	33.2
2	418	27.8	21.3
>3	195	13.0	7.4
NA		0.5	
Household size			
1	346	23.0	25.9
2	417	27.7	29.1
3-4	508	33.8	31.3
>5	225	14.9	13.7
Missing	7	0.5	
Income			
<$10,000	309	20.6	30.4
$10,000-$15,000	212	14.1	14.8
$15,000-$25,000	353	23.5	25.4
>$25,000	496	33.0	29.4
Missing	133	8.8	

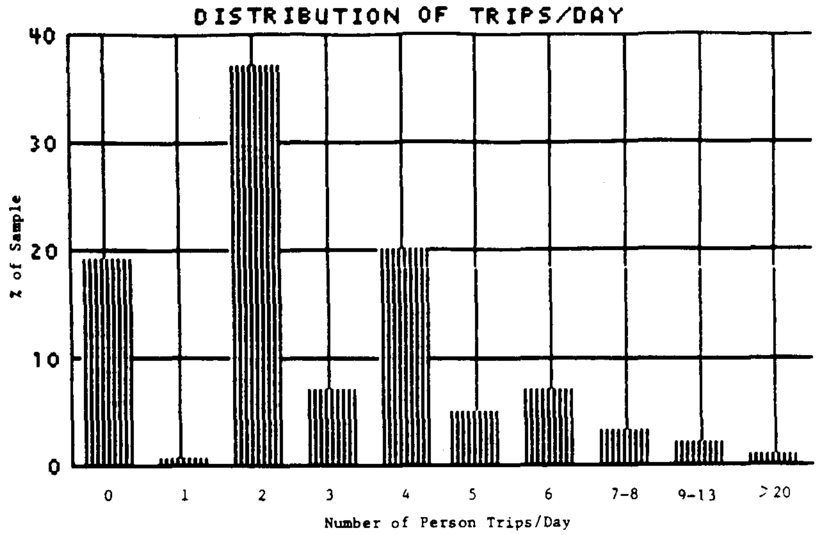


FIGURE 1 Frequency distribution of trips per day.

FACT FINDERS
N.Y.S. Transportation Survey

NEXT, WE WOULD LIKE TO ASK YOU ABOUT THE TRAVELING YOU DID YESTERDAY. WE'RE INTERESTED IN ANY TRIPS YOU TOOK YESTERDAY NO MATTER HOW SHORT OR LONG, OR WHETHER YOU WALKED OR RODE. THINK ABOUT YOUR ENTIRE DAY YESTERDAY AND ALL THE SEPARATE TRIPS YOU MIGHT HAVE TAKEN.

2A) DID YOU LEAVE YOUR HOME AT ALL YESTERDAY ?
1=Yes (continue trip questions)
2=No (go to next section)
16

2B) HOW MANY SEPARATE TRIPS DID YOU TAKE YESTERDAY ? record #
(trips refers to <u>one-way</u> movements)
17 18

2C) FOR WHAT PURPOSES, DID YOU TAKE YOUR (1st, 2nd, 3rd, etc.) TRIP YESTERDAY?

<u>PURPOSE OF TRIP</u>
0=No Yesterday Trip
1=Work/Work Related
2=Shopping/Household Business
3=Social
4=Recreational
5=Civic/Religious
6=School
7=Drop Off, Pick Up-Other(s)
8=Other
9=Missing

2D) ON THIS (1st, 2nd, 3rd, etc.) TRIP, HOW DID YOU TRAVEL ?

<u>MODE OF TRIP</u>
0=No Yesterday Trip
1=Walk
2=Bicycle, Motorcycle, Moped
3=Taxi
4=Local or Commuter Bus
5=Train
6=Subway
7=Car, Van, Light Truck
8=Other
9=Missing

(If travel is by [#7] Car, Van, Light Truck, Ask Question 2E.)
2E) DID YOU TRAVELALONE, WITH FAMILY MEMBERS, OR WITH FRIENDS ?
0=No Yesterday Trip
1=Alone
2=Family Members
3=Neighbors, Friends
9=Missing

TRIP CHART

	Purpose	Mode 1	Mode 2	Mode 3	Status					
						19	20	21	22	23
TRIP 1:	___	___	___	___	___	24	25	26	27	28
TRIP 2:	___	___	___	___	___	29	30	31	32	33
TRIP 3:	___	___	___	___	___	34	35	36	37	38
TRIP 4:	___	___	___	___	___	39	40	41	42	43
TRIP 5:	___	___	___	___	___					

FIGURE 2 Travel questions asked on New York State travel and energy survey.

The respondents were asked to recall "trips made yesterday." The survey takers were careful to point out that each trip was to be considered a one-way leg of a journey made outside the home, even if it was by walking. Although this type of survey may appear to be constrained by the recall of the respondent and brevity of the survey telephone space, most home-interview surveys were also recalls.

The number of person trips ranged from 0 to 65 (Figure 1). A total of 10 persons making 20 or more trips per day were all found to be engaged in work duties: the person making 65 trips was a package deliveryman. One person making 13 trips was a homemaker making a series of personal business and shopping trips. Because no respondent reported more than 13 and less than 20 trips, it was decided to treat those 10 persons making more than 13 trips per day (0.6 percent) as outliers; this leaves a sample of 1,493. Questions concerning trip rates are shown in Figure 2. The respondent was first asked to list the total number of one-way trips made the previous day (question 2B). For the first five of these, purpose, mode, and occupancy were also recorded. The total number of separate trips is used to determine the person trip rates.

RESULTS

Differences by Area

For this survey, New York State was divided into downstate and upstate areas (Figure 3). The respondents were also asked to describe the type of area in which they lived as big city, suburban, small town, or rural. The analysis indicates that adults in New York State average about 2.8 one-way trips per day (Table 2); there is no significant difference in the mean number of person trips per day between upstate and downstate New York or between the various regions of the state. This number is slightly higher in the suburbs, where respondents report an average of 3.1 trips upstate and 2.9 trips downstate, but these numbers are not statistically significant. This is also true for small towns (2.7) and rural areas (2.6), both upstate and downstate, but again neither difference is significant. This finding is particularly important for New York State as a whole because it indicates that average person trip rates are largely similar throughout the state.

TABLE 2 Average Trips per Day by State Regions

	City	Sub-urban	Small Town	Rural	Total
Upstate					
Avg trips	2.7	3.1	2.6	2.6	2.7
Sample size	168	158	139	115	580
Avg distance to work (miles)	2.6	3.8	3.9	5.5	3.8
Avg distance to shop (miles)	1.5	3.6	3.2	5.8	3.0
Downstate					
Avg trips	2.8	2.9	2.7	2.6	2.8
Sample size	434	367	88	24	913
Avg distance to work (miles)	4.1	6.1	5.5	5.6	5.0
Avg distance to shop (miles)	1.1	1.9	2.1	3.8	1.6
Statewide					
Avg trips	2.8	2.9	2.7	2.6	2.8
Sample size	602	525	227	139	1,493
Avg distance to work (miles)	3.7	5.4	4.6	5.6	4.5
Avg distance to shop (miles)	1.2	2.1	2.8	5.5	2.2

Weekday and Weekend

The average weekday trip rate of 3.0 trips per adult

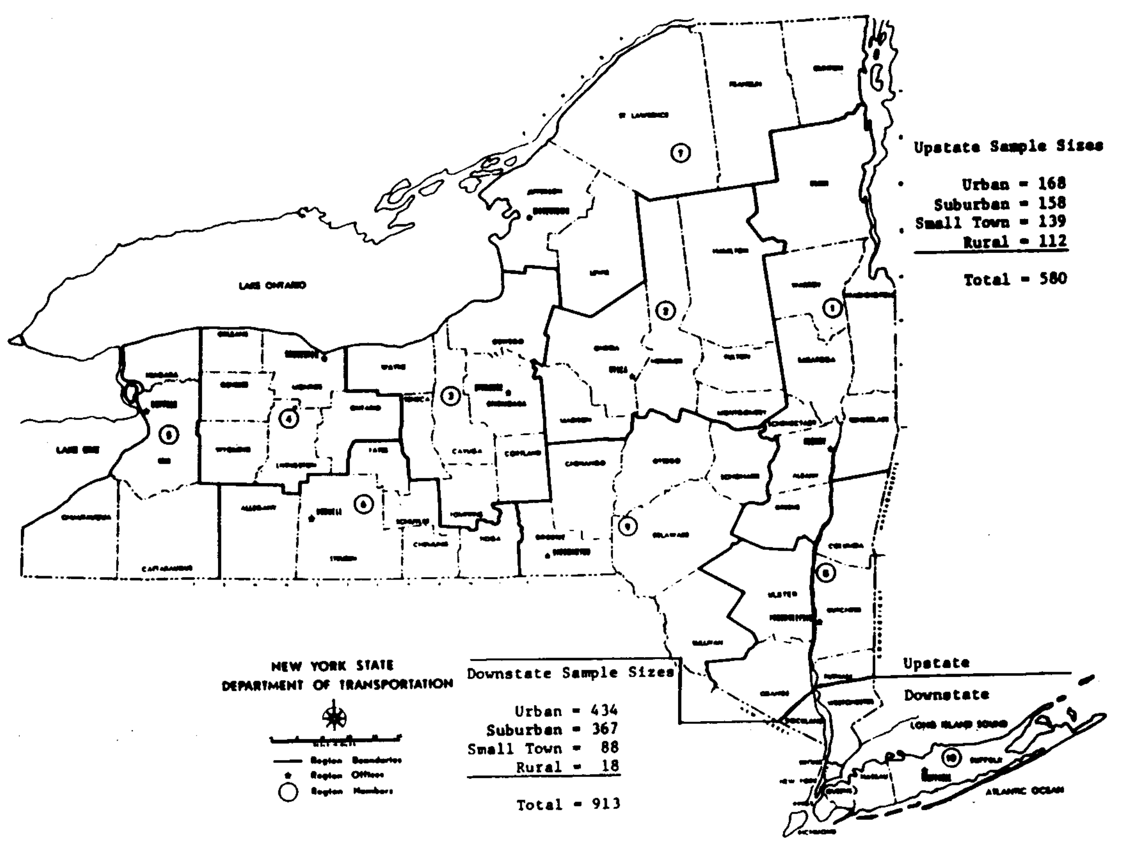


FIGURE 3 New York State map with survey sample sizes.

corresponds precisely with the 1973 Buffalo home-interview survey and the 1974 Rochester survey of 3.0 and 3.1 trips, respectively (Table 3). Both of these surveys were conducted on weekdays only.

There are significant differences in average weekday trips versus average weekend traffic (Table 4). This is particularly true for suburban areas. But 22 days of the survey occurred in January after holiday shopping and during a time not conducive to recreational travel; this could have lowered weekend trip rates. The greatest range of average trip rates is shown in small towns. Overall, Saturday, Sunday, and Monday have the lowest trip rates per person, whereas Thursday and Friday trip rates are highest (Table 5).

Demographic Effects

The largest average trip rate is by persons employed part-time (3.3 trips per day) and those in the highest income (3.4 trips per day). When these two factors are combined, the average trip rate is 3.9 trips. Homemakers and retired persons tend to travel the least--an average of 2.4 and 1.7 person trips, respectively (Table 6).

Men make more trips per day than women in all categories of income and employment (Table 7). On average, males make 3.1 trips per day and females make 2.6 trips per day. Both male and female respondents whose total household income is greater than $25,000 have similar trip rates: 3.4 for men and 3.3 for women. Men employed full-time and women employed part-time make the most trips--3.4 and 3.2, respectively. Retired women (1.4), male students (1.3), and low-income women (1.9) make the fewest trips.

The greater the number of vehicles there are per household, the higher the average person trip rate (Table 8). The greatest average number of person trips (3.8) is made in households where there are three or more vehicles and two drivers per household. The fewest number of trips is made by households in which there are no vehicles and no licensed drivers.

The number of children younger than 18 years old within one household has an increasing impact on person trip rates of the respondents. Persons in households with four or more children make 3.8 trips per day, those with one to three children make approximately 3.0 trips per day, and those without children make an average of 2.6 trips per day (Table 9). Households with four adults and no children younger than 18 make as many trips as households

TABLE 3 Average Weekday Trips per Person 18 Years of Age and Older

Survey	Sample Size	Total Trips	Avg Trips per Person
1974, Rochester	4,861	15,138	3.1
1973, Buffalo	4,197	12,592	3.0
1983, statewide	1,068	3,204	3.0

TABLE 4 Average Trips per Day on Weekdays and Weekends by State Region

	City[a]	Suburbs[a]	Small Town	Rural	All[a]
Weekday					
Avg trips	2.9	3.1	2.8	2.7	3.0
Sample size	438	370	157	103	1,068
SD	2.11	3.15	2.78	2.07	2.27
Weekend					
Avg trips	2.5	2.3	2.4	2.4	2.4
Sample size	164	155	70	36	425
SD	2.05	2.35	2.36	1.76	2.06
z	2.1	15.6			4.6

[a]Significant difference for data in column.

TABLE 5 Average Number of Trips per Day of Week

	Monday		Tuesday		Wednesday		Thursday		Friday		Saturday		Sunday	
	Trips	Sample Size	Trips	Sample Size	Trips	Sample Size	Trips	Sample Size	Trips	Sample Size	Trips	Sample Size	Trips	Sample Size
Upstate	2.2	86	2.9	83	3.0	83	3.3	75	3.2	87	2.9	72	2.0	94
Downstate	2.9	130	2.9	130	2.9	126	3.0	140	3.2	128	2.6	138	2.2	121
City	2.7	91	3.0	76	2.8	91	3.1	92	3.1	88	2.9	82	2.1	82
Suburban	2.9	73	2.9	80	3.3	65	3.3	77	3.4	75	2.6	76	2.1	79
Small town	1.8	29	2.9	33	2.8	38	3.4	28	3.0	29	2.4	33	2.3	37
Rural	2.4	23	2.7	24	3.1	15	2.1	18	3.3	23	3.1	19	1.5	17
Statewide	2.6	216	2.9	213	2.9	209	3.1	215	3.2	215	2.7	210	2.1	215

TABLE 6 Effect of Income and Employment Status on Average Trips per Person

Employment Status	Avg Trips per Person by Total Household Income: <$10,000	$10,000-$15,000	$15,000-$20,000	>$25,000	All
Employed full-time	2.9	2.8	3.3	3.4	3.2
Employed part-time	2.8	2.7	3.3	3.9	3.3
Unemployed	2.0	2.0[a]	3.1[a]	1.8[a]	2.2
Homemaker	1.8	0.7[a]	2.0	3.5	2.4
Retired	1.3	1.6[a]	2.5	3.2[a]	1.7
Student	2.7[a]	3.5[a]	2.4[a]	2.6[a]	2.9
All	2.0	2.4	3.0	3.4	

[a]Sample size <30.

TABLE 7 Effect of Gender on Trip Rates

	Avg Trip Rate by Sex of Respondent: Male	Female
Income		
<$10,000	2.2	1.9
$10,000-$15,000	2.8	2.2
$15,000-$25,000	3.4	2.7
>$25,000	3.4	3.3
Employment status		
Employed full-time	3.4	3.0
Employed part-time	3.3	3.2
Unemployed	2.6	1.8
Homemaker		2.4
Retired	2.1	1.4
Student	1.3	3.0
Total	3.1	2.6

Note: Sample sizes are 632 male and 861 female.

TABLE 8 Effect of Vehicles per Household on Trip Rates

	Avg Trips by Vehicles Owned				
	0	1	2	≥3	All
Licensed drivers					
0	1.8	1.7[a]	-	-	1.8
1	2.6	2.6	2.8	2.7	2.6
2	3.0	2.9	3.2	3.8	3.1
3	2.3[a]	3.0	2.8	3.2	3.0
Household size					
1	2.2	2.3	2.8	-	2.4
2	2.3	2.5	3.2	2.9	2.7
3	1.7	2.8	2.8	3.1	2.7
4	2.7	2.9	3.5	3.4	3.3
≥5	2.3	3.2	2.8	3.7	3.1
Statewide	2.2	2.7	3.1	3.3	2.8

[a]Sample size <10.

TABLE 9 Effect of Household Size on Trip Rates

Children per Household	Avg Trips per Respondent by Adults per Household					Total	
	1	2	3	4	≥5	Trips (avg)	Sample Size
0	2.4	2.7	2.5	3.7	2.7	2.6	909
1	2.6	3.0	2.8	3.1[a]	3.5[a]	3.0	250
2	2.4	3.3	3.1	2.4[a]		3.1	198
3	2.7[a]	3.0	2.9[a]			3.0	77
≥4	3.75[a]	3.8[a]	3.5[a]			3.8	48
Total sample size	419	699	214	96	54		

[a]Sample size <30.

with four or more children. Some of these household members may in fact be dependents older than 18, but because complete household data were not available, it was difficult to determine the composition of the household. However, the data in Table 9 tend to confirm the findings of Boyle and Chicoine (18) on the influence of children on trip rates.

To determine what factors are most influential in effecting trips per day, the procedure known as Automatic Interaction Detector (AID) was used. AID is a statistical procedure that partitions data sequentially, ascending to the most important classifications. Several of these analyses were done with various subsets of the independent variables. None of the coefficients of correlation exceeded 0.137 (Figure 4), and it was not possible to demonstrate any effect of the life-cycle state on the number of person trips. This supports the findings of McDonald and Stopher (16) regarding the strength of life-cycle variables. However, it is possible that there was not enough household-level data within the survey for household interactions and life cycles to show their influence on the trip rates.

Figure 4 shows the influence of income, employment status, and distance from work as the primary determinants of the number of trips made. An income of less than $15,000 first divides the data set; the highest average trip rate (3.7) is attributed to the person who belongs to a household with one more licensed driver than vehicles, who lives less than 5 miles from work, and who is employed either full-time or part-time. Household sizes of four or more also influence the respondent to make more trips ($\bar{y}$ = 3.6). Also of interest in this analysis are factors

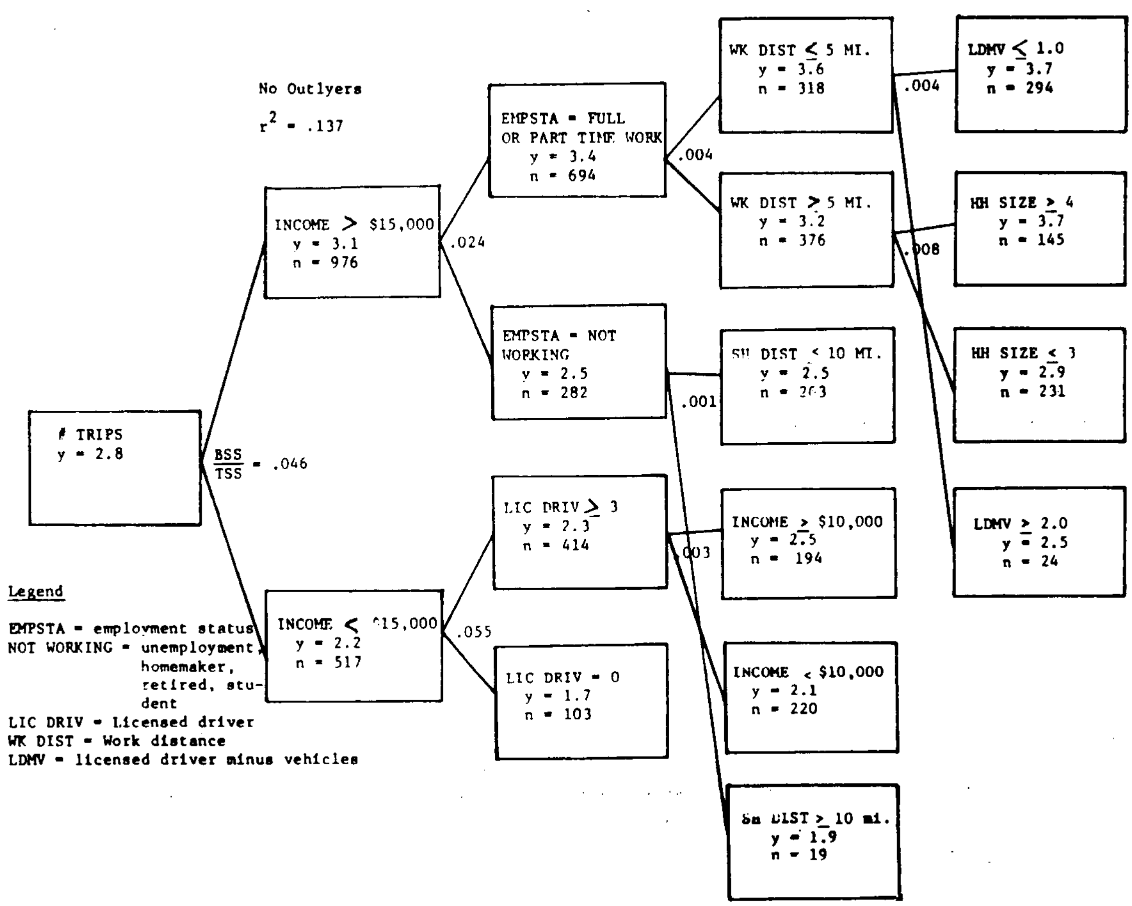


FIGURE 4 AID diagram of influence of demographic variables on trips per person.

that do not appear in the data. Codes for New York State regions and the upstate and downstate split were available to determine the split, but neither factor was able to influence the split in any of the analyses. The split among city, suburban, small town, or rural occurs only for those persons living 5 or more miles from work with three or more cars at their disposal, usually an indication of substantial income. Thus this analysis would appear to confirm that the determinants of travel are similar across the state, and that those variables that do influence travel are largely demographic.

Analysis of Trip Tables

Because the purpose of the survey was to develop trip rates useful for energy use calculations of specific types of activities, trip purposes were classified as destination purposes such as work, shop, social, recreation, civic, and so forth, rather than the more familiar terms used in modeling such as home-based work, work-based shop, and so on. Thus the rates developed are more easily compared with the National Personal Transportation Study (NPTS) analysis (19) rather than other trip-generation analyses, such as those by Stopher and McDonald (6). Only 4.4 percent of the sample made more than six trips [this is little more than that found by Stopher and Sheskin (9) in their investigation of 24-hr travel records]. If it is assumed that the sixth trip for those making six trips was to return home, then the trip rates to specific destinations may be considered fairly representative. Only small differences are apparent between upstate and downstate, but there are greater differences between days of the week. Most social and recreation activity occurs on weekends, most shopping and household business occurs on Friday and Saturday, and most travel occurs during the week (Table 10).

The mode of travel for eight specific destinations is given in Table 11. In this table "return home" is not allocated to the specific purposes, as is done in the NPTS study (19). This survey, taken during winter weather in January 1983, is reasonably close to the 1980 census figures for usual mode to work collected during April 1980.

Ridesharing data are given in Table 12. As discussed in other research (20), there is a problem with defining ridesharing because many people do not regard traveling with family or friends to be ridesharing. By avoiding the term carpooling or ridesharing, and instead asking whether the trip was made with family, friends, or neighbors, the degree of ridesharing is easier to determine. Sharing rides is the common mode for social, recreational, and religious trips (i.e., rides are most often shared with family). The greatest percentage of ridesharing with neighbors or friends occurs for social reasons, but this is still less than family ridesharing. Nevertheless, it appears that ridesharing is the norm for nonwork travel.

TABLE 10 Purposes of Trips by Day of Week

	Percentage of Trips to Specific Destinations by Day of Week							
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Total
Work or work related								
Upstate	35	43	41	42	37	17[a]	16[a]	33
Downstate	42	44	39	52	40	8[a]	15[a]	36
Statewide	39	44	40	48	39	15	16	35
Shop or household business								
Upstate	35	29	34	25	40	37	34	34
Downstate	37	28	31	27	40	46	32	34
Statewide	37	29	32	27	40	41	33	34
Social-recreation								
Upstate	21	9[a]	18	17[a]	16[a]	36	36	22
Downstate	13[a]	17	21	14[a]	16	40	38	22
Statewide	16	14	20	15	16	37	37	22
Other								
Upstate	9[a]	19[a]	8[a]	15[a]	7[a]	9[a]	13[a]	11
Downstate	8[a]	10[a]	8[a]	7[a]	4[a]	6[a]	15[a]	8
Statewide	8	13	8	10	5	7	14	9

Note: Sample used is only for those people making ≤ 6 trips per day. The percentages for all categories (upstate, downstate, and statewide) equal 100 percent.
[a]Sample size <30.

TABLE 11 Mode of Travel by Purpose

	Percentage of Trips by Mode						
Purpose	Walk	Bike	Taxi	Public Transit	Car Only	Total[a]	Multimodal[b]
Work	6.6	0.1	1.1	22.9	67.8	98.5	6.1
Shopping and personal business	21.0	0.1	0.5	8.3	70.0	99.9	0.9
Social	13.5	-	3.0	11.1	72.8		1.6
Recreation	29.1	-	2.6	4.2	62.4	99.9	1.7
Civic and religious	12.9	-	-	7.4	79.6	99.9	-
School	15.9	-	-	30.5	53.7	100.0	3.7
Drop off	5.3	-	-	3.2	91.5	100.0	-
Return home	**13.9**	**0.01**	**1.3**	**16.3**	**67.8**	**99.3**	**2.5**
1980 census for New York State, means of travel to work	8.6	← 1.0	→	27.1	63.3		

[a]Columns may not add to 100 due to rounding or other category.
[b]Excludes walking.

TABLE 12 Ridesharing

	Trips by Automobile or Light Truck (%)			
Destination Purpose	Drive Alone (1)	Drive with Family Members (2)	Drive with Neighbors, Friends (3)	Rideshare (2, > 3)
Work	74.9	10.5	14.6	25.1
Shopping and household business	56.0	35.7	8.3	44.0
Social	33.7	37.0	29.3	66.3
Recreational	39.7	42.5	17.8	60.3
Civic and religious	25.6	67.4	7.0	74.4
School	51.1	40.0	8.9	48.9
Drop-off, pick-up, and other	15.7	62.7	21.7	84.4
Home	56.9	29.3	13.7	43.0

TABLE 13 One-Way Trip Rates by Automobiles per Household

	Trip Rates by Automobiles Owned				
Destination Purpose	0	1	2	3	All
Work or work related					
One-way trips	0.41[a]	0.53[b]	0.64	0.67	0.56
SD	0.82	0.58	0.81	1.04	
Shop or personal business, one-way trips	0.48	0.59	0.66	0.66	0.60
Shop or household business					
One-way trips	0.46	0.53	0.57	0.55	0.54
SD	0.77	0.79	0.82	0.84	
Serve passengers					
One-way trips	0.02	0.06	0.9	0.9	0.6
SD	0.02	0.06	0.9	0.8	
Social-recreation, one-way trips	0.31	0.32	0.35	0.37	0.33
Social					
One-way trips	0.24	0.23	0.27	0.30	0.25
SD	0.57	0.53	0.56	0.60	
Recreation					
One-way trips	0.07	0.09	0.08	0.07	0.08
SD	0.26	0.33	0.32	0.25	
Civic, education, and religious, one-way trips	0.09	0.07	0.11	0.13	0.09
Civic and religious					
One-way trips	0.03	0.03	0.05	0.04	0.04
SD	0.16	0.17	0.23	0.20	
Education					
One-way trips	0.06	0.04	0.06	0.09	0.06
SD	0.26	0.22	0.26	0.32	
Total	1.28[c]	1.51[d]	1.75[e]	1.82	1.5
Sample size	274	608	416	188	
Percent of sample	18.4	40.7	27.9	12.6	

[a]Significant difference at 0.95 confidence level between 0.41 and 0.53, z = 2.28.
[b]Significant difference at 0.95 confidence level between 0.53 and 0.64, z = 2.1.
[c]Significant difference at 0.95 confidence level between 1.28 and 1.51, z = 24.3.
[d]Significant difference at 0.95 confidence level between 1.51 and 1.75, z = 3.1.
[e]Significant difference at 0.95 confidence level between 1.75 and 1.82, z = 7.0.

Trip rates vary significantly with the increase in automobiles owned by the household. The data in Table 13 indicate that, for all purposes, the number of person trips by all modes increases as the number of household vehicles, automobiles, and light trucks increases. The difference between the trip rates are significant only for work trips and total trips.

(Data on rates by income, age, sex, and automobiles owned versus purpose are available from the authors.)

SUMMARY AND CONCLUSIONS

The person trip rates collected during a statewide telephone survey in New York State during January 1983 have been analyzed. Findings of interest or significance from this study are as follows.

1. The average person in New York State makes 2.8 trips per day, and lives 4.5 miles from work and 2.2 miles from shopping.
2. There is no significant difference between person trip rates upstate and downstate or between persons residing in areas designated as urban, suburban, small town, and rural. Thus trip rates as such can be applied statewide.
3. There are differences in trip rates between weekday and weekend travel as well as between specific days of the week. Wednesday, Thursday, and Friday are the heaviest days of travel; Saturday and Sunday are the lowest. Most social and recreation trips occur on Saturday and Sunday.
4. Factors that influence the number of trips made per day per adult are income, sex, employment status, number of household vehicles, and presence of children younger than 18 years old in the household. In general, women make fewer trips than men, but this difference tends to disappear as household income increases. However, the life-cycle influence on trip rates could not be confirmed for person trip rates.
5. Two-thirds of all trips are made by automobile. The percentage of trips made by automobile is greatest for nonwork trip purposes. Work travel, however, has the highest rate of solo-occupant travel. Ridesharing (family or friends) is the usual mode for social, recreation, civic, educational, and religious destinations; approximately 44 percent of shopping trips involve ridesharing. However, the majority of nonwork ridesharing involves travel with family.
6. Nonwork trip purposes represent approximately 65 percent of all trip destinations made by New York State consumers. These trips are divided approximately into 34 percent for shopping and household business, 22 percent for social or recreational purposes, and 9 percent for all other purposes. Work represents only 35 percent of all travel.

This analysis of trip rates collected from a statewide telephone survey has shown that while variables such as income, employment status, household size, and presence of children do affect individual trip rates, there is no evidence that geographic location within the state affects trip rates. Results from this and other travel surveys therefore appear transferable to any study area within the state. This hypothesis was investigated as early as 1967 by the Bureau of Public Roads (21). Remarkably, the relative importance of various demographic parameters in accounting for variance in travel (i.e., income and work status) was generally confirmed. It was shown, however, that automobile ownership and household size also influence travel considerably. These findings are consistent with many transportation studies that segregate trip-generation data into one or more of these key parameters.

These findings increase the confidence that transportation analysts may have in using trip-generation rates developed from other cities or earlier studies. Although transferability of trip-generation rates is a subject of considerable concern, the findings here suggest that transferability may be more possible than previously thought. In addition, the findings suggest that transferability across space may be equally as likely as transferability over time. Obviously, adjustments should be made for the number of households or persons in different demographic cells, but application of existing trip-generation rates within these cells through estimated future households or persons is nonetheless a reasonably valid procedure. Should the analyst be concerned with the possibility of errors introduced by such an assumption, a sensitivity analysis varying the trip-generation rates or the forecasts of households or persons per cell would determine the magnitude of likely error.

Further analysis of the nature of these trip-generation rates should be undertaken. For instance, it is possible that the net small differences between upstate and downstate New York trip-generation rates are the combined effect of significant differences in income (which would tend to increase trip generation downstate compared with upstate) and density and automobile ownership (which would tend to have the reverse effect). A more carefully structured tabular analysis would identify whether either or both of these hypotheses are working in the data that have been presented.

ACKNOWLEDGMENTS

This study was supported by a grant from UMTA as part of a larger study of energy use reductions in nonwork travel. The authors wish to thank Charron Many and Tedi Toca for the preparation of the manuscript.

REFERENCES

1. A. Rapoli. New York State: Profile of the 1980s. Transportation Analysis Report 8. Planning Division, New York State Department of Transportation, Albany, Dec. 1981.
2. G.S. Cohen and M. Kocis. Components of Change in Urban Travel. Prelim. Res. Report 159. Planning Division, New York State Department of Transportation, Albany, July 1979.
3. Urban Origin-Destination Surveys. FHWA, U.S. Department of Transportation, 1973.
4. Trip Generation Analysis. FHWA, U.S. Department of Transportation, Aug. 1975.
5. R. Dobson and W.E. McGarvey. A Comparative Analysis of Category and Regression Trip Generation Models. Transportation, 1977.
6. P.R. Stopher and K.G. McDonald. Trip Generation by Cross-Classification: An Alternative Methodology. In Transportation Research Record 944, TRB, National Research Council, Washington, D.C., 1983, pp. 84-91.
7. A.B. Sosslau, A.B. Hassan, M.M. Carter, and G.V. Wickstrom. Quick-Response Urban Travel Estimation Techniques and Transferable Parameters: User's Guide. NCHRP Report 187. TRB, National Research Council, Washington, D.C., 1978, 229 pp.
8. D. Klinger, R. Kusmyak, and S. Liss. Household Travel. Report 9, 1977 National Personal Transportation Study. FHWA, U.S. Department of Transportation, July 1982.
9. P.R. Stopher and I.R. Sheskin. Toward Improved Collection of 24-H Travel Records. In Transportation Research Record 891, TRB, National Research Council, Washington, D.C., 1982, pp. 10-17.
10. F.A. Reid. Critique of ITE Trip Generation Rates and an Alternative Basis for Estimating New Area Traffic. In Transportation Research Record 874, TRB, National Research Council, Washington, D.C., 1982, pp. 1-5.
11. P.M. Allaman, T.J. Tardiff, and F.C. Dunbar. New Approaches to Understanding Travel Behavior. NCHRP Report 250. TRB, National Research Council, Washington, D.C., Sept. 1982, 142 pp.
12. A.L. Rhodes. Household Travel Patterns and Expenditures, Links with Societal Structure. Report DOT-RSPA-DMA-50/83-8. Research and Special Programs Administration, U.S. Department of Transportation, Jan. 1983.
13. C. Zimmerman. Life Cycle Travel by the American Family. Report 7, 1977 National Personal Transportation Study. FHWA, U.S. Department of Transportation, July 1981.
14. A. Denk and D.K. Boyle. Life Cycle Characteristics of High Travel Households. Transportation Analysis Report 15. Planning Division, New York State Department of Transportation, Albany, Aug. 1982.
15. J.E. Chicoine. The Life Cycle Concept: A Practical Application of Transportation Planning. Transportation Analysis Report 22. Planning Division, New York State Department of Transportation, Albany, Aug. 1982.
16. K.G. McDonald and P.R. Stopher. Some Contrary Indications for the Use of Household Structure in Trip-Generation Analysis. In Transportation Research Record 944, TRB, National Research Council, Washington, D.C., 1983, pp. 92-100.
17. J. Supernak, A. Talvitie, and A. DeJohn. Person-Category Trip-Generation Model. In Transportation Research Record 944, TRB, National Research Council, Washington, D.C., 1983, pp. 74-83.
18. D.K. Boyle and J. Chicoine. Application of the Life-Cycle Concept to Trip Generation Procedures. Transportation Analysis Report 24. Planning Division, New York State Department of Transportation, Albany, Dec. 1982.
19. M.E. Roskin. Purposes of Vehicle Trips and Travel. Report 3, 1977 National Personal Transportation Study. FHWA, U.S. Department of Transportation, Dec. 1980.
20. J.M. Brunso and D.T. Hartgen. Carpool Coordinator Demonstration Study: Overview and Analysis of "Before" Data. Prelim. Res. Report 150. Planning Division, New York State Department of Transportation, Albany, March 1979.
21. Guidelines for Trip Generation Analysis. Bureau of Public Roads, June 1967.

The views expressed in this paper are those of the authors, including any errors of fact or omission.

Logit Mode-Choice Models for Nonwork Trips

PETER R. STOPHER, ERIC G. OHSTROM, KENNETH D. KALTENBACH, and DONALD L. CLOUSE

ABSTRACT

Most research on logit models of mode choice has concentrated on the work trip, a fact frequently commented on by critics for some years. With the increasingly widespread adoption of the logit model as the basic mode-choice model of practical transportation planning, more logit models for nonwork purposes are being installed in travel forecasting procedures. In this paper the form that most of these models take and the assumptions on which they are based are examined. It is shown that the majority of these are not calibrated, but are updated from the work models. The inappropriateness of this is demonstrated through selected case studies, and the types of models that can be built are described. It is shown that calibration of nonwork models is feasible and presents no new problems over the work mode-choice models, and that the relative weights of cost and time components in work models are different from those found for fully calibrated nonwork models. The data requirements and calibration needs are also discussed.

Throughout most of the development of disaggregate models of mode choice, research concentrated almost exclusively on developing models of choices for the work trip. This was justified on a number of grounds, including the importance of the work trip in planning and policy decisions, and the convenience and appropriateness of the work trip for research. In this respect, it was often pointed out that collecting data on work trips presents a relatively simple and inexpensive data-collection activity; and that, because of the habitual nature of the trip, there is a greater chance that the work trip represents a rational choice of mode and that knowledge may exist about the alternatives. It is not the purpose of this paper to deliberate over these reasons or to produce evidence as to whether or not there exist foundations for them. Suffice it to say that there are published research results that cast some doubt on each of these basic assumptions and reasons, but that these still appear to have been insufficient to generate any significant change in the direction of research.

Of course the authors do not claim that there has been no research on nonwork models. There are several published papers about models for shopping trips (1-4), and a few instances of other nonwork models as well (5,6). However, the total number of such publications is insignificant in comparison with those on work trips. Furthermore, the logit model for the work trip has remained relatively simple, certainly in the perception of practicing transportation planners, whereas much of the research on nonwork models has generated more complex model forms and has tied the mode-choice models to other models in the stream, such as destination choice (trip distribution) or route choice. Given the added complexity stemming from this, the fact that most practical travel forecasters are reasonably content with existing aggregate trip-distribution models, and that aggregate versions of these more complex models are largely unknown, the few nonwork models that have been developed have largely failed, so far, to penetrate practice.

In this paper the pros and cons of substituting aggregate or disaggregate mode-choice models in the standard travel-forecasting process, as opposed to making radical changes in the modeling process and its structure, are not discussed. Rather, it is accepted that the majority of planning regions in the United States use the conventional four-step modeling process for travel forecasting, as exemplified by the Urban Transportation Planning System (UTPS) program of the U.S. Department of Transportation, and they have simply chosen to replace or update the modal-split models in this process. Also, it should be noted that the authors use the term "modal split" to refer to models that are conceptually and structurally aggregate, while using the term "mode choice" for models and procedures that are either disaggregate entirely or are based on use of disaggregate data for their development.

PRACTICAL IMPLEMENTATION

For more than two decades of modern regional transportation planning, no agreement could be reached on the form and structure of the modal-split model. It was frequently stated that, although only two types of trip-distribution models (gravity and intervening opportunity) were to be found in use, there were as many different modal-split models as there were urban areas that had completed a long-range transportation planning activity. Documentation of modal-split models tended to demonstrate the range of different types and structures of models (7,8). In the past few years this situation has changed quite dramatically. Almost every urbanized area that has updated or improved their model stream, and every area that has considered seriously the potential building of a line-haul transit service, has introduced a set of logit models of modal split. Such models are currently in use in Los Angeles and San Francisco, California; Washington, D.C.; Miami, Florida; Honolulu, Hawaii; Detroit, Michigan; Minneapolis-St. Paul, Minnesota; New Orleans, Louisiana; and San Juan, Puerto Rico, to name a few.

As noted previously, there has been considerable research on the mode-choice logit model for work trips, but relatively little for any other trip purposes. In applying logit methods to the standard travel forecasting stream, models are required to cover all purposes. In practical transportation planning, the emerging standard appears to be to use about six trip purposes for trip generation and trip distribution, but to aggregate these purposes to three or four for mode choice. In most of the cities previously mentioned, there are three models for the purposes of home-based work (HBW), home-based other (HBO), and non-home-based (NHB) trips. In one or two instances, an additional model exists for home-based school (HBS) trips, but these are more usually left as part of the HBO trips or excluded altogether, and dealt with in some other estimation procedure that includes an allocation of trips by school bus.

Clearly, then, every locality that has introduced logit models of modal split has had the need to build not only the well-researched, reasonably well-understood work trip model for HBW trips, but has also had to develop models for at least two other purposes, HBO and NHB, neither of which has been researched nor understood to any great extent. Knowledge of how to build a model for shopping trips also has not helped the definition of models for these much more aggregate purposes. For some reason, not widely reported, transportation planners and planning agencies appear to have decided that the lack of research on these models also indicates that they would not be possible to calibrate in the normal sense.

Against this situation, two primary methods have been used to build models for HBO and NHB purposes, neither of which represents true calibration (i.e., free fitting of all model parameters to current or recent data). The first method that has been used--quasi-updating--is to define the HBO and NHB models in terms of the relative coefficients found for the work trip model and to seek to determine an overall multiplying factor for the utility from the work model. This assumes that the relative weights of components of travel time, travel cost, and any user characteristics in the models are the same for all trip purposes. There is no research or other literature to support this position, but it is widely held. In some instances the models so developed are even further removed from calibration, because the work model may in some cases have been built with predetermined relationships between some of the variables. Illustrations of this are discussed later in the paper.

The second method of building the needed additional models--factoring--is to build factor models that use the zonal market shares from the work model and apply this, usually through some factoring procedure, to NHB trips. In many respects this differs from quasi-updating only in that the factor is derived by a different procedure.

One may question to what extent this treatment of nonwork trips is of any real importance. It is clear that most conventional bus systems derive most of their ridership from the peak periods, carrying primarily work and school trips. Even systems that include some form of rapid transit are still likely to carry significantly more trips in the peak period and to derive a large portion of their patronage from the work trip. Nevertheless, these statistics do not indicate that the nonwork, nonpeak trips can be dismissed and can be treated substantially less accurately than the work trips. In most large urban areas work trips represent about 20 to 25 percent of total daily trips. Home-based nonwork trips generally constitute a further 50 to 55 percent of trips, whereas NHB trips make up the balance (20 to 30 percent) of regional person trips. In a typical medium or large urban area in the United States, the transit share of the market ranges from 2 to 15 percent of all trips, and about 50 percent of this transit share comes from the work trip.

As examples of these figures, 1980 statistics for the Los Angeles region show that work trips constitute about 18 percent of daily person trips, home-based nonwork trips are about 52 percent, and NHB trips are 30 percent. The bus system carries about 3 percent of these trips, with 45 percent of transit trips being HBW trips. Overall, transit carries 7.5 percent of HBW trips, 2.4 percent of HBO trips, and 1.3 percent of NHB trips.

In Honolulu, it is estimated that 16 percent of regional trips are HBW trips, 48 percent are HBO trips (including HBS trips), with 36 percent being NHB trips. Transit carries about 14.9 percent of the HBW trips, 7.9 percent of HBO trips, and 5.4 percent of NHB trips. Because of the high use of the public bus system for HBS trips, which are included in the HBO total, Honolulu buses derive only 30 percent of their resident (not including the substantial tourist ridership in Honolulu) patronage from the work trip. If school trips are added to this, most of which also occur in the peak periods, the percentage of patronage for HBW and HBS trips becomes 53 percent. The Honolulu bus system carries 8.2 percent of the resident person trips plus an additional 29,000 tourist trips on an average weekday.

Finally, in Miami the regional split of trips among purposes is 26 percent for HBW trips, 60 percent HBO trips, and 14 percent NHB trips. The regional transit share is 4.2 percent, consisting of 7.8 percent of HBW trips, 1.7 percent of HBO trips, and 8.1 percent of NHB trips (the latter being high because of the relatively high proportion of NHB trips for Miami Beach and the high transit share of all trips in Miami Beach) (9).

ILLUSTRATIVE EXAMPLES

It is useful to see the form of the models that are produced by the alternative methods of building HBO and NHB mode-choice models. Several examples have been selected from reported models that are in current use in several different locations.

Minneapolis-St. Paul

This is one of the earliest models to have been developed and applied for regional travel forecasting (10). The coefficients for these models are given in Table 1. The ratio of out-of-vehicle time coefficients to in-vehicle time coefficients in the HBO model is exactly 2.5, and the ratio of the cost and in-vehicle time coefficients is 1.5. Neither of these ratios appears as such in the work model, although both represent values that have been stated frequently to represent the conventional wisdom of the relative values of these in logit models. Overall, these ratios appear to have been established and only the absolute values of the coefficients and the values of the modal constants were fitted to transit share data. In the NHB model the ratio of 2.5 between out-of-vehicle time and in-vehicle time is maintained, generating coefficients of -0.025 for out-of-vehicle time components and -0.01 for in-vehicle time. The cost coefficient is -0.0039, which appears as almost the same ratio as the ratio of HBW in-vehicle time to cost. Although this is not a pure example of the types described earlier, these models appear to be generally of the form of the ones that define the HBO and NHB models from the HBW models, calibrating only an overall multiplier to fit ob-observed transit shares.

Miami

The Miami model was built in 1976 and revised in 1978 (9). It was built under difficult circumstances in that no calibration data were available for constructing it. Therefore, it was built from existing trip tables, estimated modal splits, and information from other logit models, principally those for Wash-

TABLE 1 Cost and Time Coefficients of Models for Minneapolis-St. Paul

Purpose	Wait Time	Walk Time	Out-of-Vehicle Time	In-Vehicle Time	Parking Cost	Running Cost	Total Cost
HBW	−0.044	−0.030	–[a]	−0.031	–	–	−0.014
HB-nonwork	−0.020	−0.020	–[a]	−0.008	–	–	−0.012
NHB	−0.025	−0.025	–[a]	−0.0100	–	–	−0.0039

[a]Several alternative coefficients are used for out-of-vehicle time for automobile, depending on occupancy.

ington, D.C. The coefficients for these models are given in Table 2. In every case the ratio between the excess-time coefficient and the in-vehicle time coefficient is 2.5, and the ratio between the in-vehicle time coefficient and the cost coefficient is -0.3333. The cost coefficient is, in this case, the coefficient for a variable of cost divided by income.

This model is an excellent example of the first type of construction, in which the ratios among the coefficients are prespecified, and fitting of the model is concerned only with an overall factor for the model coefficients and any mode-specific constants.

New Orleans

This model was built in 1981 and incorporates some additional sophistications not apparent in the previous two models. (Note that the data for this model are from unpublished reports by Barton-Aschman Associates, Inc.) These sophistications include using different coefficients for walk time and wait time and introducing yet a further coefficient for automobile time when used as access to transit. The coefficients for these models are given in Table 3. In the HBW model the ratios between each of walk time and wait time and in-vehicle time are approximately 2.3 and 5.3; whereas the ratio between cost and in-vehicle time is 0.53. Notwithstanding these values, the model reverts to a 2.5 ratio for both walk and wait times to in-vehicle times for both the HBO and NHB models. The cost coefficients demonstrate almost exactly the same relationship to in-vehicle time as the Minneapolis models, which suggests that this model may have been used as the basis for the cost coefficient, with additional modifications being made to the cost coefficient to replicate observed transit shares more accurately.

Los Angeles I

The first Los Angeles model to be described is the one built for the Los Angeles Rapid Transit System in 1976. The time and cost coefficients for the HBW model are as follows (11): out-of-vehicle travel time/distance = 24.37, in-vehicle travel time = -0.01465, cost/income = -0.1860, and the factor = 2.332. This model, which was never adopted for regional forecasts by the local agencies, consisted of a logit work mode-choice model and a factoring procedure for nonwork trips. The factoring procedure is based on the observation that approximately 43 percent of transit trips are work trips. After estimating the HBW trips, the trip interchange totals of transit trips generated by the work model are multiplied by 2.332, which represents the inverse of the proportion of transit trips that are work trips. This is an excellent example of the second method of developing nonwork mode-choice models.

Los Angeles II

The second Los Angeles model was built in 1982. The coefficients are given in Table 4. (Note that these data are from unpublished reports for the Southern California Association of Governments by Cambridge Systematics, Inc., 1982). This model represents an exception to the previous ones, insofar as the HBO model is concerned. This model was calibrated to data, and no use was made of relationships between coefficients in the work model for devising this model. The ratio of the coefficients of excess time and in-vehicle time is 5.6 for the HBW model and 3.1 for the HBO model. In these models cost is divided by income, thus making comparison with some of the other models more difficult. However, the ratio of the cost coefficient to in-vehicle travel time is 2.01 for the HBW model and 3.17 for the HBO model.

TABLE 2 Cost and Time Coefficients of Models for Miami

Purpose	Wait Time	Walk Time	Out-of-Vehicle Time	In-Vehicle Time	Parking Cost	Running Cost	Total Cost
HBW	–	–	−0.0515	−0.0206	–	–	−0.0618
HB-nonwork	–	–	−0.0415	−0.0166	–	–	−0.0498
NHB	–	–	−0.0193	−0.0077	–	–	−0.0231

Note: The cost and time coefficients are for transit, nonbeach traffic only. Models exist for each of transit and highway for both beach and nonbeach zones. Each model contains different coefficients, but the ratios among coefficients are the same.

TABLE 3 Cost and Time Coefficients of Models for New Orleans

Purpose	Wait Time	Walk Time	Out-of-Vehicle Time	In-Vehicle Time	Parking Cost	Running Cost	Total Cost
HBW	−0.0332	−0.0769	–[a]	−0.0145	–	–	−0.0078
HB-nonwork	−0.0165	−0.0165	–[a]	−0.0066	–	–	−0.0116
NHB	−0.0328	−0.0328	−0.3048	−0.0131	–	–	−0.0047

[a]Out-of-vehicle time is for automobile only, and several coefficients exist for the occupancy levels for HBO and NHB.

TABLE 4 Cost and Time Coefficients of Models for Los Angeles (1982)

Purpose	Wait Time	Walk Time	Out-of-Vehicle Time	In-Vehicle Time	Parking Cost	Running Cost	Total Cost
HBW	-0.157	-0.0329	-0.0557	-0.0111	–	–	-0.019
HB-nonwork	–	–	-0.0746	-0.0256	–	–	-0.0293
NHB	–	–	–	–	–	–	–

Note: The NHB transit share is factored from the work modal split.

Again, these values serve primarily to demonstrate that the HBO model was calibrated freely and that the assumed values from the earlier models do not appear to be replicated by these calibrated values. This is discussed at more length later in the paper.

In this model set the NHB transit trips are estimated by multiplying the HBO share of transit trips (expressed as a fraction) by a fractional constant to determine the transit share of NHB trips. NHB trips are subdivided into other-to-work and other-to-other trips. For the former, the fractional multiplier of the HBO modal split is 0.2608, and for the latter it is 0.3431. In the event that a trip interchange has no HBO trips, the NHB transit market shares are set at 0.0182 for other-to-work trips and at 0.0156 for other-to-other trips. These values are approximately the regional modal splits for these two purposes.

The NHB model is an example of the factor model, whereas the HBO model represents one of the still-few instances of the free calibration of a model for nonwork trips.

More examples could be drawn from those that are in current use, but those documented in the preceding paragraphs provide adequate illustrations of the types of models that are in current use and that are based on the noted methods of calibration.

FULL CALIBRATION

The alternative to the foregoing procedures is to calibrate the home-based nonwork and NHB models directly from available data. As noted earlier in the paper, there appear to be certain myths surrounding full calibration of these models that have led to the preponderance of the model-fitting procedures described in the previous section of the paper. In this section two case studies are described that should expose the myths. The first of these case studies deals with what is likely to be the most common case for practical transportation planning, in which the region does not have household data that have been collected recently with calibration of logit mode-choice models in mind. Rather, the data are likely to be of the form required for updating earlier types of forecasting models. In the second case study data were collected expressly to allow calibration of logit models of mode choice for all purposes. This is closer to the ideal situation, but is likely to occur far less often than the first case.

Case Study 1

This case study is for San Juan, Puerto Rico (12). New modal-split models were to be constructed for use in a conventional UTPS-based forecasting procedure, but the modal-split models were to be aggregate logit models. The work plan for this activity did not include either time or money to permit collection of data for constructing new models. However, a data set existed that had been collected in 1977 for updating a fully conventional set of home-interview data. The data set consisted of 1,178 households, from which standard trip data for 24-hr, household demographics, and locational data had been obtained. The trip data consisted primarily of the mode of travel, the origin and destination, the time of day, and the purpose of the trip. Information existed on whether or not the household had automobiles available and how many automobiles were available. The number of licensed drivers was not included in the data.

A calibration data set was developed for mode choice by subdividing the reported trips into the purposes of HBW, HBO, NHB, and HBS. Data were compiled for each trip from the path characteristics of the highway and transit networks to represent the travel characteristics for each trip. For HBW and the HBS trips, the travel characteristics were developed from the peak networks, whereas the characteristics for HBO and NHB trips were drawn from the midday or 24-hr networks. Paths were defined for three primary mode alternatives: automobile, bus, and publico (jitney). It was assumed that access to bus was by walk only, whereas publico could be accessed by either walk or walk and bus. No distinction was obtainable in the travel characteristics for automobile based on the occupancy, except to divide the cost among the occupants. The trip characteristics obtained from the path files and zonal characteristics were walking time, waiting time, in-vehicle time, parking cost, and running cost (running cost is total out-of-pocket costs, not including parking).

The calibration was achieved by using ULOGIT in the UTPS program package. This model required that trips be deleted from the calibration file if any of the alternatives had no path and therefore no trip characteristics. From the 1,178 households, the calibration data sets consisted of 864 HBW trips, 579 HBS trips, 798 HBO trips, and 346 NHB trips. The lack of captivity data prevented removal of captives from the calibration data. The coefficients of the models are given in Table 5.

First, it may be observed that the models for all four purposes produced sensible results in terms of the signs and magnitudes of the coefficients. Hence concerns that models for nonwork trip purposes cannot be calibrated from conventional data appear to be unfounded. Second, note that the relative values of the coefficients differ from those described in the quasi-updated models. In the HBW model walking time and waiting time each have about the same coefficient, and it is more than 3 times the value of the in-vehicle time coefficient. The cost coefficient is about 0.31 of the in-vehicle time coefficient. For HBO trips, the coefficients of walking and waiting time are again similar, but are 12 times the value of the in-vehicle time. The cost coefficient is equal to the in-vehicle time coefficient in this case.

The HBS model is substantially different. In this case the in-vehicle time coefficient was so insignificant and small that the variable was not used in the final model. The walking time coefficient was

TABLE 5 Cost and Time Coefficients of Models for San Juan

Purpose	Wait Time	Walk Time	Out-of-Vehicle Time	In-Vehicle Time	Parking Cost	Running Cost	Total Cost
HBW	−0.049	−0.040	–	−0.013	–	–	−0.004
HBS	−0.053	−0.025	–	–	−0.014	−0.003	–
HBO	−0.060	−0.061	–	−0.005	–	–	−0.005
NHB	−0.119	−0.026	–	−0.010	−0.016	−0.002	–

more than twice the size of the waiting time coefficient, and is 4 times the size of the in-vehicle time coefficient for the work model. Parking cost has a coefficient that is nearly 5 times the size of running cost. The latter coefficient is about 0.2 of the work model in-vehicle travel time coefficient, and is about 0.12 of the waiting time coefficient of this HBS model. Finally, the NHB model shows a further set of different relationships. In this case walking time is weighted 4.5 times more heavily than waiting time and almost 12 times as heavily as in-vehicle time. The cost variable is again divided into the two components of parking and running cost, with the former having a coefficient that is 6 times the value of the latter, and 1.6 times the in-vehicle time coefficient. The ratio of the cost coefficient to the in-vehicle time coefficient is 0.2.

Generally, there is little support from this model for the ratios assumed in many of the noncalibrated models. The work mode-choice model exhibits coefficient relationships that are well within the range of those that have been reported in a variety of other localities. The lack of importance of in-vehicle travel time for school trips is reasonably acceptable, suggesting that, given the necessity to go to school and the relative lack of choice in school location, in-vehicle travel time is of little consequence in choosing among available travel modes. In all models both walking and waiting times are weighted much more heavily than in-vehicle travel time, although walking is considered far more onerous for HBO and NHB trips than for the other purposes.

Case Study 2

The second case study is from Honolulu, Hawaii (13). In this study data were collected expressly for calibration of a set of logit mode-choice models, although it was decided that network (aggregate) data should be used for the calibration data set. Data were collected by means of a travel diary from 1,370 households (see paper by Ohstrom et al. elsewhere in this Record), and the calibration data set was developed by geocoding the origins and destinations of the trips and again extracting the travel characteristics from the path files. The models were structured around the alternatives of automobile (with three occupancy levels), local bus, and express bus. Express bus could be accessed by walk or local bus, while local bus had walk access alone. Express bus was available for only HBW and HBS trips, and both of these purposes again used the peak transit network characteristics, with congested highway speeds, whereas midday transit network characteristics and free-flow highway conditions were used for the HBO and NHB models. As with the San Juan model, no distinction in the characteristics of multioccupant automobile trips could be obtained beyond the division of cost among the occupants. Again, the characteristics used were walking time, waiting time, in-vehicle time, parking cost, and running cost. Sociodemographic variables were also tested, but the only one found to affect the models significantly was the ratio of available vehicles to licensed drivers (minimum value of 0.0 and maximum value of 1.0). This variable was not retained in the final models because of concerns about the ability of local agencies to forecast it. Retention of the calibration values, in place of forecasts, would leave the variable as little more than a constant term.

Calibration was achieved by using the QUAIL program developed at the University of California at Berkeley (14), which permits calibration data to contain a variety of subsets of alternative modes. Therefore, the only discarded data were for any trips where only one mode had a path between a pair of zones or where the trip was totally within the zone. From the 1,370 households, the calibration data sets consisted of 458 HBW trips, 329 HBS trips, 361 HBO trips, and 277 NHB trips. In this case the data included information on captivity, and captives were excluded from the calibration data. In addition, a number of data points were lost because the network characteristics created outliers that would bias the calibration results. An outlier was defined as arising when the chosen mode had travel characteristics (times and costs) that were all inferior to those of any of the nonchosen modes and the sum of the time components was more than 20 min in excess of the worst alternative not chosen. Alternatively, if the total travel time for the chosen mode was more than 3 times the travel time of the next alternative, it was also considered an outlier. The results of the calibration are given in Table 6.

The conclusions to be drawn from these models are similar to those from the San Juan models in their essential points for this paper. Again, the results clearly show that logit models can be calibrated satisfactorily for all of the purposes. Likewise, the relative values of coefficients differ substantially from the assumed values, and show significant differences from purpose to purpose. In the HBW model walking time is weighted by 3.3 times in-vehicle time, whereas waiting time has a coefficient of 5.7 times that of in-vehicle time. Costs are split, with running cost having a coefficient that is 0.17 of in-vehicle time and parking cost a coefficient that is 0.72 of in-vehicle time. In the HBS model walking time is valued at 6.6 times in-vehicle time and waiting time is valued at 4.5 times, whereas cost is 0.47 times the in-vehicle time. In this case in-vehicle travel time did not appear in either the HBO or NHB models. This may signify a problem with the midday and uncongested networks, but it also may be a realistic reflection of behavior. For the HBO model, walking time is considered about 2.5 times as onerous as waiting time, and about 3.5 times as onerous as in-vehicle time for the work trip. Parking cost is 3.5 times as important as running cost, whereas the latter has a coefficient somewhat smaller than for the HBW model.

Finally, the NHB model shows walking time to be more than 3 times as onerous as waiting time and has cost coefficients for both parking and running costs that are almost identical to the HBO values. The

TABLE 6 Cost and Time Coefficients of Models for Honolulu

Purpose	Wait Time	Walk Time	Out-of-Vehicle Time	In-Vehicle Time	Parking Cost	Running Cost	Total Cost
HBS	–0.099	–0.068	–	–0.015	–	–	–0.007
HBO	–0.101	–0.041	–	–	–0.007	–0.002	–
NHB	–0.126	–0.040	–	–	–0.006	–0.003	–

results of the NHB and HBO models are similar to those of San Juan, and suggest a radically different weighting of coefficients to any of the noncalibrated models discussed. (It should be noted that the Honolulu models have been recalibrated subsequently, with minor changes in certain inputs, and some changes have occurred in final coefficient values.)

CONCLUSIONS

Three conclusions are in order from the cases discussed in this paper. First, planning agencies and their consultants should not conclude that the lack of reported research on nonwork models is in any way indicative of potential problems in fitting the models. Although not discussed here, it is appropriate to observe that the statistics of goodness-of-fit for the NHB models are generally inferior to those of the HBW models, which is consistent with experience in fitting trip-generation models for NHB trips. Nevertheless, the values of these statistics are adequate to indicate a useful model. This is further borne out by obtaining coefficients that are reasonable and that also show consistency between two localities described herein. The statistics for HBO models were found to be comparable with the HBW models. In all cases coefficients were found to have t-scores well in excess of 2.0 for included variables, and chi-square values were, as usual, far larger than any table values for the appropriate degrees of freedom. For details, however, the reader is referred to the original reports.

Second, although it is clearly desirable that data be collected that are designed for the purpose of calibrating logit models, it is possible to obtain adequate fits from data that may have been collected several years previously and that were not collected specifically for logit modeling. A cautionary note is appropriate to the effect that use of network-derived characteristics requires great care in path building, a topic that is too extensive to deal with in this paper.

Finally, the transferability of HBW logit-model coefficients that have been assumed in building models for other purposes is not borne out by true calibration. The relationships tend to be significantly different from those in the work models, and may exhibit variation from locality to locality. Similar local differences are also to be found among HBW models when unconstrained calibration is performed. Furthermore, not all of the travel characteristics found to be significant in work mode-choice models are significant in nonwork models. Therefore, factoring from work models, or defining a multiplier for a predefined combination of times and costs for nonwork models, is not an appropriate procedure to use.

ACKNOWLEDGMENTS

The authors would like to thank the Oahu Metropolitan Planning Organization (OMPO) for permission to use the results described for the Honolulu case study and for their active support on the presentation of this paper. The authors particularly would like to thank Cathy Arthur and Gordon Lum of OMPO for their suggestions and help both in the original work and in the documentation of the results in this paper. The authors would also like to thank the Department of Transportation and Public Works of the Commonwealth of Puerto Rico for permission to use the results described for the San Juan case study.

REFERENCES

1. T.E. Parody. Technique for Determining Travel Choices for a Model of Nonwork Travel. In Transportation Research Record 673, TRB, National Research Council, Washington, D.C., 1978, pp. 47-53.
2. T.A. Domencich and D. McFadden. Urban Travel Demand: A Behavioral Analysis. North Holland/American Elsevier, New York, 1975, Chapter 7.
3. F.S. Koppelman and J.R. Hauser. Destination Choice Behavior for Nongrocery Shopping Trips. In Transportation Research Record 673, TRB, National Research Council, Washington, D.C., 1978, pp. 157-165.
4. T.J. Adler and M.E. Ben-Akiva. Joint Choice Model for Frequency, Destination, and Travel Mode for Shopping Trips. In Transportation Research Record 569, TRB, National Research Council, Washington, D.C., 1976, pp. 136-150.
5. J. Horowitz. Disaggregate Demand Model for Nonwork Travel. In Transportation Research Record 673, TRB, National Research Council, Washington, D.C., 1978, pp. 65-71.
6. E.R. Ruiter and M.E. Ben-Akiva. Disaggregate Travel Demand Models for the San Francisco Bay Area. In Transportation Research Record 673, TRB, National Research Council, Washington, D.C., 1978, pp. 121-137.
7. M.J. Fertal et al. Modal Split: Documentation of Nine Methods for Estimating Transit Usage. Bureau of Public Roads, U.S. Department of Commerce, 1966.
8. P.R. Stopher and A.H. Meyburg. Urban Transportation Modeling and Planning. Lexington Books, Lexington, Mass., 1975, Chapter 9.
9. Kaiser Transit Group. Prime Modal Split Model Recalibration Report. Metropolitan Dade County, Miami, Fla., Dec. 1978.
10. R.H. Pratt Associates, Inc., and DTM, Inc. Development and Calibration of Mode Choice Models for the Twin Cities Area. Metropolitan Council of the Twin Cities, Minneapolis, Minn., Aug. 1976.
11. Cambridge Systematics, Inc. The Development of a Disaggregate Behavioral Work Mode Choice Model. California Department of Transportation, Sacramento, and Southern California Association of Governments, Los Angeles, Jan. 1977.

12. Schimpeler-Corradino Associates. San Juan Transportation Planning Models. Final Report. Department of Transportation and Public Works, San Juan, P.R., Sept. 1982.
13. Schimpeler-Corradino Associates. Oahu Model Update Study. Final Report, Volume II. Oahu Metropolitan Planning Organization, Honolulu, Hawaii, Dec. 1982.
14. J. Berkman; D. Brownstone and Associates. QUAIL 4.0 Programmer's Manual. Department of Economics, University of California, Berkeley, March 1979.

The opinions expressed in the paper are those of the authors, and do not necessarily reflect the position or policies of the Oahu Metropolitan Planning Organization or the Commonwealth of Puerto Rico. The authors are solely responsible for any errors or omissions in the paper.

Sequential Model of Interdependent Activity and Destination Choices

RYUICHI KITAMURA and MOHAMMAD KERMANSHAH

ABSTRACT

A sequential model of daily travel patterns that consists of activity and destination choice submodels is developed in this study. The model development takes into account the interdependencies among the choices and the constraints imposed on the movement in time and space. The empirical analysis indicates that non-home-based destination choice is critically dependent on the residence location of the individual and that activity choice is influenced only marginally by the accessibility of the origin location. As a practical and immediate modification of non-home-based destination choice models, it is proposed in this study that destination-to-home travel time be included as a factor that enables a more realistic depiction of spatial travel patterns.

In previous efforts (1,2) the authors have examined the properties of activity choice that are directly related to generation of trips and their temporal distribution over a 1-day period. The results have revealed the characteristics of time-of-day dependencies of activity choice and revealed patterns in sequencing activities in trip chains. Analysis of the dependence of activity choice on its own history indicated that activity history may be represented in a simple manner for use in travel behavior analysis. This study draws on the previous efforts and expands it by introducing the spatial dimension into its scope.

The ultimate objective of this continuing effort is to develop a practical model system that makes possible a more realistic depiction of complex daily travel behavior. The effort and the resulting models can be characterized by the following two aspects. The first is its explicit recognition and incorporation into the model structure of the fact that trips made by an individual are linked to each other. This leads to the emphasis in this study of the interdependencies among choices that underlie the entire daily travel and activity pattern. In other words, this study does not isolate a trip or a travel choice from the rest to be analyzed independently. Second, the effort acknowledges that the movement of an individual is constrained in time and space because of various factors, including the social commitments, obligations, limited transportation capabilities, and physiological needs of the individual (3-8). The constraints are most typically associated with activities that allow little scheduling flexibilities such as work, chauffeuring children to school, or having lunch during a lunch break. This study therefore emphasizes, among others, time-of-day dependencies of activities and trips.

A system of models is developed in this study. It consists of home-based and non-home-based destination choice models that incorporate the effects of trip continuity together with those of time of day. The activity choice models of this study are expanded to include, in addition to the variables used in the previous study (2), spatial factors such as the travel time between the home base and the origin activity location and the accessibility from that location.

The objective of this study is, first, to identify the extent to which destination choice is influenced by factors other than the traditional variables (i.e., the origin-destination travel time and the attributes of alternative destination locations). More specifically, the study is an endeavor to show that the location of an individual's home and the locations of alternative destinations relative to the home location critically influence non-home-based destination choice. The second objective is to identify the effects that spatial factors have on activity choice, either independently or jointly with other factors, including time of day, activity history, and socioeconomic characteristics of the individual. Note that the effects of the latter group of variables have been studied earlier (2), and accessibility indices as spatial factors have often been used in previous travel behavior analyses (9,10). The intention of this study is to achieve a more comprehensive treatment of these factors in analyzing daily travel patterns. Their intricate interactive effects are examined through statistical hypothesis testing that involves specification and estimation of alternative destination and activity choice models. Based on the results of the study, a practical modification that can be made to destination choice models for improved depiction of spatial travel patterns is proposed.

BACKGROUND

Formulations of destination choice models are typically based on the assumption that the trip is made from the home base and that only one destination location will be visited after the individual leaves home. Non-home-based choice, where the origin of the trip is not the home base, is analyzed while isolating the trip from the rest as an independent unit of analysis. Accordingly, the behavior of linking trips into a multiple-sojourn chain is not appropriately taken into consideration in the conventional analyses. This simplification is implicit in the behavioral or statistical derivations of commonly used trip distribution models such as the gravity model (11-13). The simplification also makes possible formulation of spatial choice models while using as explanatory variables only the attributes of respective destination alternatives and the spatial separation between the origin and destination. The models thus developed appear to capture the observed tendencies in spatial travel patterns with their simple model structure and with a relatively small set of explanatory variables. Nevertheless, this simplification may impose serious limitations when attempting to expand the scope of the analysis to include multiple-sojourn trip chains. Further discussions of the limitations and problems arising from the assumption can be found in Hanson (14,15). [This study focuses on trip linkages and constraints in its effort of extending the framework of destination choice analysis. Possible alternative developments are discussed elsewhere (16-18), with emphases on additional factors and behavioral aspects.]

An alternative approach is to acknowledge that choices underlying daily travel and activity patterns are interdependent (19). This can be done by analyzing travel choices as a simultaneous decision that is concerned with the entire daily activity and travel pattern (20-22), or by analyzing the series of choices sequentially (23,24). In the latter case, interdependencies can be accounted for by specifying the choices as dependent on the past history of activities (1,2), by viewing them as dependent on possible future behavior (25), or possibly on both. The interdependencies are reflected in the models of this study through activity choices that are assumed to be history dependent, and destination choices that are specified as, to an extent, dependent on the future.

By viewing the destination choices in an individual's daily travel pattern as interrelated choices and recognizing the fact that his travel pattern develops around the home base, it is hypothesized that the residence location of the individual is of critical importance in explaining the non-home-based destination choice. Note that the residence location has not been included in previous analyses of destination choice. However, the very fact that the individual sooner or later returns home in the future suggests that the choice is influenced by the location of the home.

For example, consider the choice of a shopping opportunity by a worker on the way back to home from the work place. This destination choice for the non-home-based shopping trip is influenced by the location of the home because it is dependent on the intended future behavior, in this case, returning home. Accordingly, the choice cannot be explained by the conventional factors alone, but its explanation requires that additional factors be introduced into the analysis. The distance between the alternative destination and home appears to be a promising candidate variable that may well explain this type of future dependency.

The importance of the residence location as a factor in non-home-based destination choice models can be seen in the following discussion, which emphasizes the constrained nature of urban travel choice. Individuals are typically subjected to certain constraints as to the locations where they can be at various time periods of the day. In other words, the range of locations where the individuals can exist is confined within a limited region in the time-space coordinates, which is often called a prism (3). This constraint will affect the choice of both activities and their locations.

Suppose that an individual located outside the home wishes to visit another location for an out-of-home activity, but he must return home by time T. The time available for the out-of-home activity and travel is T - t, where t is the present time. Let i be the location where the individual is currently located, and j be the potential destination. Then the following relation must be satisfied for location j to be accessible:

$$d_{ij} + d_{jh} < T - t \qquad (1)$$

where d_{ij} is the travel time between locations i and j, and d_{jh} is the travel time between j and the home base. The inequality indicates that the destination-to-home travel time (d_{jh}) is an important element in destination choice under the prism constraint.

Additional evidence for the importance of the residence location is given by the following empirical observation of the series of destination choices in a trip chain. By applying the log-linear model of contingency table analysis to a large-scale origin-destination survey data set, Kermanshah (26) found that there exists a predominant pattern into which a set of destination locations to be visited are frequently arranged in a trip chain: The individuals tend to visit farther locations first, and subsequent destinations tend to be closer to home or cluster in the vicinity of the preceding locations. The finding implies that the home location is again of critical importance in adequately capturing the pattern of sequencing the locations visited in a trip chain.

MODEL FORMULATION

The activity and destination choice models of this study are formulated by using a two-stage approach, where activity choice and destination choice are separately modeled; choice of destinations given the out-of-home activity type is first modeled, and then activity choice models are developed. Accordingly, the destination choice models include as alternatives only nonhome destination opportunities. The structural framework of the model system of this study is described in detail elsewhere (24). It is worthy to note that a similar activity-location model system has been developed by van der Hoorn (27) with emphasis on determining trip generation based on temporal tendencies in activity engagement and also on differentiating in-home and out-of-home activities.

The non-home-based destination choice model of this study is formulated as

$$\left.\begin{aligned} P_a(i,j,t) &= \exp[V_a(i,j,t)]/\Sigma_k \exp[V_a(i,k,t)] \\ V_a(i,j,t) &= V(d_{ij}, d_{jh}, A_j, t, y) \end{aligned}\right\} \text{for } j = 1 \ldots, J \qquad (2)$$

where

J = number of destination alternatives,
a = type of the activity for which the choice is made,
$P_a(i,j,t)$ = probability that destination j will be chosen by individual i at time t to pursue an activity of type a,
$V_a(i,j,t)$ = measure of attractiveness of destination j when visited from i at time t to pursue an activity of type a,
A_j = vector of attributes of destination j,
t = time of day,
y = activity history,
d_{ij} = travel time between origin i and destination j, and
d_{jh} = travel time between home h and location j.

The multinomial logit model, which has been used in many previous analyses of spatial choice (28-30), is used here as the model structure. The representative utility or attractiveness measure of destination j $[V_a(i,j,t)]$ is time-of-day dependent and is formulated with the distance measure (d_{jh}), the travel time between the home base and destination j. This is in addition to the conventional origin-destination travel time (d_{ij}). Other factors considered in the model development are activity history, time of day, attributes of destination locations, and socioeconomic attributes of the individual. The variables used are summarized in Table 1. Not all of the variables in the table appear in the final models selected in this study.

Noting that the individual's time budget for activity and travel becomes tighter as the day proceeds, it is expected that the valuation of travel time varies depending on the time of day; presumably the individual is less willing to take a long trip at the end of the day than in the beginning of the day. Such a time-dependent nature in destination choice can be represented in the model by introducing an interaction term that involves time-of-day and travel time variables. Similar terms can be used to represent a possible history dependency in destination choice.

The emphasis placed in this study on temporal dependencies of activity and travel requires that time of day be explicitly incorporated into the framework of the model. This leads to the formulation of the model where the attraction measure of a destination is defined as a function of the time of day as well as its attributes, such as retail employment. This is based on the belief that activity

TABLE 1 Variables Considered in Model Development

Variable Group	Abbreviation	Definition
Destination attributes (A_j)		
Population	POP	ln[(zonal population)/1,000]
Retail employment	REMP	ln[(zonal retail employment)/1,000]
Nonretail employment	NREMP	ln[(zonal nonretail employment/1,000]
Travel time (d)		
Origin-destination travel time	d_{ij}	Time (min) obtained from off-peak network skim trees
Home-origin travel time	d_{ih}	
Home-destination travel time	d_{jh}	
0-1 dummy for $d_{ih} - d_{jh}$	d_s	1 if $d_{ih} - d_{jh} > 0$; 0, otherwise.
Accessibility index (I_a)		
Accessibility of zone i for activity type a at time t	$I_a(i,t)$	$\ln \Sigma_j \exp[V_a(i,j,t)]$
Time of day (t)		
Function of time of day		t, t^2, $\exp(t)$, $\exp(-t)$, $\ln(t)$; t is in hours
Store hours (0-1)	$D_s(t)$	1 if t is between 9:00 a.m. and 9:00 p.m.; 0, otherwise
Business hours (0-1)	$D_b(t)$	1 if t is between 8:00 a.m. and 5:00 p.m.; 0, otherwise
Activity history (y)		
Activity engagement in previous chains in		Binary variable: 1 if activities of the indicated type were pursued in the trip chains previously made
Personal business	PBNS01H	
Social recreation	SREC01H	
Shopping	SHOP01H	
Serving passengers	SRVP01H	
Activity engagement in the current chain in		Binary variable: 1 if an activity of the indicated type has been pursued in the current trip chain
Personal business	PBNS01C	
Social recreation	SREC01C	
Shopping	SHOP01C	
Serving passengers	SRVP01C	
Current activity		Binary variable: 1 if the current activity is of the indicated type
Personal business	PBNS	
Social recreation	SREC	
Shopping	SHOP	
Serving passengers	SRVP	
Out-of-home time	OHTIME	Cumulative amount of time spent so far outside home for both trips and activities
No. of chains	CHAINS	Cumulative number of home-based trip chains made so far
Socioeconomic attributes (e)		
School-age children	SCHLAG	Binary variable: 1 if the age of youngest child in the household is between 5 and 12; 0, otherwise
Household role	ROLE	Binary variable: 1 if an individual is female and not employed; 0, otherwise
No. of children	CHLDRN	Number of household members who are 17 years old or younger and not married
Household income	INCOME	Median value of the household's annual gross income category ($)
No. of cars	CARS	Number of cars available to the household

and destination choices are made on the basis of the availability of functions that accommodate and facilitate the pursuit of intended activities, but not the physical existence of the facilities themselves (31). For example, a department store after it has closed in the evening should not be counted as a destination opportunity. In order to represent such temporal variations in the availability of opportunities, variables were developed that represent typical business and store hours. Note that the inclusion of the time-of-day-dependent attraction measures in the model offers a mechanism for evaluating the changes in activity and travel patterns that correspond to changes in the availability over the 1-day period.

The alternatives of the non-home-based activity choice model include four activity types (personal business, shopping, social recreation, and serving passengers) and two returning-home options (i.e., returning home temporarily, and returning home permanently). The last alternative implies that the out-of-home activity schedule of the day will be terminated. This study hypothesizes that choice of activity type depends on the distribution of opportunities around the origin location. For example, if the individual who has just completed an out-of-home activity is located in an area with intense commercial development, the individual may be more likely to pursue additional shopping activities. This effect is represented by the following accessibility index defined for location i (9,10):

$$I_a(i,t) = \ln \left\{ \Sigma_j \exp \left[V_a(i,j,t)\right] \right\} \tag{3}$$

where the $V_a(i,j,t)$'s are obtained from the non-home-based destination choice models. This index represents the expected maximum utility; that is, the expected utility of that destination that is most attractive to the individual who intends to pursue activity of type a and is located at i at time t. Inclusion of the $I_a(i,t)$'s for all activity types would indicate the relative attractiveness of the respected types of activities. Note that the accessibility measure is a function of the travel time to opportunities from i, and may be viewed as a proxy variable for travel cost for activity engagement from that location. Also note that the measure is time-of-day dependent, and that the activity choice model takes on the form of the nested logit model. Another spatial factor considered in the non-home-based activity choice model is the distance of the origin location from the home base.

The home-based destination choice model has the same logit form. The model development effort considers the traditional factors (d_{ij} and A_j) and also the variables representing the past history of activity and travel as well as time of day. The home-based activity choice model is similar to the one developed in the earlier effort (2). The types of variables included in the four types of activity and destination choice models are given in Table 2.

DATA SET

The statistical analysis of this study uses a subsample of the 1977 Baltimore travel demand data set. The subsample is almost identical to the one used in the previous effort of activity choice model formulation (2), and includes adult individuals whose daily trip records are complete and consistent, and whose households had access to a car. Only those individuals who did not make work trips on the survey day are analyzed in this study. The activity choice analysis excludes weekend trip records because of the obvious differences in time use patterns between weekdays and weekends. The sample screening criteria, which are similar to the ones used in previous studies (1,2,6,8,26,32), are used here with the intention of controlling the sample so that the travel environment within which the individuals' activity and travel patterns develop will be relatively homogeneous. Such a controlled sample and the resulting internal homogeneity are believed to aid in the effort of interpreting the results and inferring causal relationships by simplifying these tasks. The current sample is slightly smaller than the one used in the previous study (2) because a new set of screening criteria, which are concerned with the consistency of spatial information, is introduced in this study. Because only aggregate measures of the attributes of destination alternatives are available in the data file, the analysis uses 70 planning districts as the alternatives of destination choice.

The resulting sample used in the development of activity choice models includes 343 home-based choices and 550 non-home-based choices in 343 trip chains made by 209 individuals. Unfortunately, the sample size is not large enough for estimating destination choice models by activity types, and weekend observations had to be included in order to facilitate the estimation process. The sample used for the development of the destination choice models of this study includes 647 home-based choices and 354 non-home-based choices with nonhome destinations.

TABLE 2 Variables Examined in Development of Activity and Destination Choice Models

	Destination Choice Model		Activity Choice Model	
Variable Group	Home Based	Nonhome Based	Home Based	Nonhome Based
Destination attributes (A_j)	X	X		
Travel Time				
d_{ij}	X	X		
d_{ih}				X
d_{jh}		X		
Accessibility index				
$[I_a(i,t)]$			X	X
Time of day (t)	X	X	X	X
Activity history (y)	X	X	X	X
Socioeconomic attributes (e)	X	X	X	X

Note: X indicates that the variable group is examined in the model development.

ESTIMATION RESULTS

The key question in the empirical analysis is whether the traditional destination attraction measures and origin-destination travel time adequately explain destination choice behavior, or whether additional factors, such as the distance between an alternative destination and the home base, should be introduced into the model. Another interesting aspect to be examined is the interplay of temporal and spatial factors. The temporal variables may influence destination choice, and the temporal and spatial factors may jointly or independently affect activity choice.

Non-Home-Based Destination Choice Models

The model coefficients are estimated by using, as the choice set, 12 randomly selected destination alternatives and the destination that was actually

TABLE 3 Non-Home-Based Destination Choice Models

	Activity Type					
	Personal Business[a]		Social-Recreation		Shopping	
Variable	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
d_{ij}			-0.0824	-3.87		
$\ln(t)d_{ij}$	-0.0592	-6.02			-0.0640	-6.29
d_{jh}	-0.1391	-5.89			-0.1792	-6.46
$\ln(t)d_{jh}$			-0.0617	-7.38		
POP	0.3363	1.55	0.5410	2.85		
$(REMP)D_s(t)$	0.3557	2.37			0.6871	5.18

Note: Variables are defined in Table 1.
[a]Includes serving passengers.

TABLE 4 Summary Statistics for Table 3

	Activity Type		
	Personal Business[a]	Social-Recreation	Shopping
$L(0)$	-307.79	-266.76	-333.44
$L(\beta)$	-151.84	-153.58	-134.79
Sample size	120	104	130
$\rho^2 = 1 - L(\beta)/L(0)$	0.507	0.424	0.596
χ^2	311.90	226.36	397.30
df	4	3	3

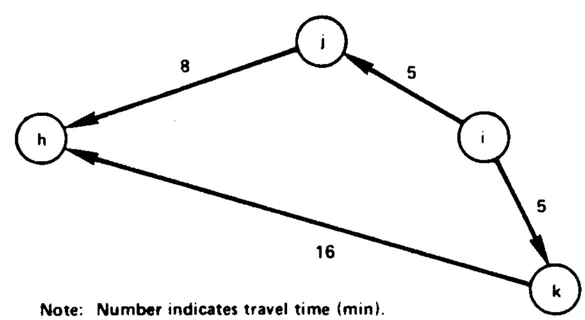
Note: $L(\beta)$ = log-likelihood with the model coefficients; $L(0)$ = log-likelihood without any coefficients; and the chi-square values presented are defined as $-2[L(0) - L(\beta)]$.
[a]Includes serving passengers.

chosen. Because of the insufficient sample size, two activity types--personal business and serving passengers--had to be grouped together in this non-home-based destination choice modeling.

The final models selected (Tables 3 and 4), after examination of a large number of alternative model formulations, are rather simple and involve only three groups of variables: time of day, travel time, and attraction measures of the destination. Models with interaction terms consisting of travel time measures and history variables or socioeconomic attributes were estimated to evaluate the effects of the latter variables on destination choice, especially on the trip length. Effects of the socioeconomic attributes and activity history, however, were not evident from the model specification effort of this study.

The estimation results confirm the hypothesized importance of the travel time between the destination and home. Inspection of the t-statistics indicates that this variable is at least as significant as the traditional origin-destination travel time. Its significance is especially notable for the social-recreation activity. The same conclusion can be obtained from the data in Table 5. The table presents another set of destination choice models that were estimated without the time-of-day effects in order to make the comparison of the relative effects of d_{ij} and d_{jh} easier. It can be seen that d_{jh} has a coefficient value and t-statistic close to those of d_{ij} in the models for personal business and shopping. In the model for social-recreation, both its coefficient and t-statistic are twice as much as those of d_{ij}.

The estimated effect of this variable is illustrated here by using the example discussed earlier. Suppose that an individual at a nonhome location (i) is making a destination choice for shopping. There are two opportunities, j and k, with identical attributes (i.e., $A_j = A_k$) and the same distance away from i ($d_{ij} = d_{ik}$). Opportunity k, however, is twice as far from the home base as opportunity j (d_{kh} = 16 min, and d_{jh} = 8 min). This is shown in Figure 1. The conventional destination choice model would predict the identical choice probability for the two opportunities. The estimated shopping des-



Note: Number indicates travel time (min).

FIGURE 1 Effect of residence location on non-home-based destination choice.

TABLE 5 Alternative Non-Home-Based Destination Choice Models Without Time-of-Day Effects

	Activity Type					
	Personal Business		Social-Recreation		Shopping	
Variable	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
d_{ij}	-0.1532	-6.01	-0.0814	-3.83	-0.1674	-6.27
d_{jh}	-0.1367	-5.79	-0.1684	-7.38	-0.1803	-6.49
POP	0.3423	1.59	0.5493	2.90		
$(REMP)D_s(t)$	0.3548	2.37			0.6888	5.26

Note: Variables are defined in Table 1.

tination choice model of Table 3, on the other hand, yields the predicted choice odds of

$$P_a(i,j,t)/P_a(i,k,t) = \exp[-0.1792(8 - 16)] = 4.2,$$

namely, the opportunity closer to home is more than 4 times likely to be chosen than the other.

The way the individual chooses his destinations in a series of trips cannot be characterized as the movement of a frog jumping between lily pads, and the location of the destination relative to the home base is an important concern to the individual. This conjecture, now supported by the empirical result, has not been incorporated into the standard destination choice or trip distribution analysis. It is proposed in this study that the destination-to-home travel time be considered in formulating non-home-based destination choice models, such that the individuals' movements can be characterized appropriately as human behavior, not as the random movement of a frog.

Another new feature of the models developed here is the inclusion of time-of-day variables. This is based on the belief that the time of day influences not only activity choice (2,31) but also the choice of the location to pursue the activity. Only few studies (33) have examined the temporal dependencies of destination choice behavior. The present estimation results indicate that, as the day proceeds and the time constraint becomes tighter, the negative effect of origin-destination travel time increases for personal business (including serving passengers) and shopping. In other words, the individuals tend to make shorter non-home-based trips for these two activity types toward the end of the day. For the social-recreation activity, the time variable is combined with the destination-to-home travel time, implying a somewhat different effect of time that social-recreational activity locations tend to cluster around the home base in the later part of the day.

Non-Home-Based Activity Choice Model

As activity choice models have been developed in the previous study in an aspatial context (2), the present effort concentrates on the introduction of spatial elements into the model and examination of their effects on activity choice. The discussion on the estimated coefficients of those variables that are included in the previous model development effort is not repeated in this paper. The interested reader is referred to the work by Kitamura and Kermanshah (2). The spatial variables considered in modeling the non-home-based choice are accessibility indices [$I_a(i,t)$'s] and the distance from the origin to home (d_{ih}). The model specification effort is summarized in Table 6, and the final model is given in Table 7.

The set of four accessibility indices evaluated according to Equation 3 for the respective activity types is first added to the previously developed base model (2). The indices as a group have a chi-square value of 7.08, with degrees of freedom (df) of 4, and not significant at $\alpha = 0.05$. Inspection of the individual coefficients indicated that the coefficient of the index for personal business alone was significantly different from zero, but its sign was negative, thus contradicting the hypothesis that higher accessibility induces activity engagement. The final model (Table 7) was developed by eliminating insignificant accessibility indices while adding the origin-to-home travel time variable to the two alternatives--temporary return to home and permanent return to home. These variables are significant as a group ($\chi^2 = 13.04$, with df = 4) and the coefficients of the accessibility indices are positive and lie between 0 and 1 in agreement with the derivation

TABLE 6 Development of Non-Home-Based Activity Choice Models

Model	Log-Likelihood	χ^2 of Added Coefficients	df
Constant terms alone	-872.45		
Base model[a]	-763.43	220.04	29
Base model + $I_a(i,t)$	-759.89	7.08	4
Final model [with d_{ih} and $I_a(i,t)$]	-756.91	13.04	4

[a]See paper by Kitamura and Kermanshah (2).

TABLE 7 Non-Home-Based Activity Choice Model

	Activity Type											
	Personal Business		Social-Recreation		Shopping		Serving Passengers		Temporary Home		Permanent Home	
Variable	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
Constant	-1.1103	-1.95	-4.0188	-4.03	-3.4003	-3.79	-4.0440	-4.01	1.2745	1.50		
PBNS	1.6092	1.78	1.5712	1.89	2.6850	2.50			1.0873	1.34	1.0873	1.34
SREC	1.1335	1.90			1.6292	1.99			0.3576	0.74	0.3576	0.74
SHOP					0.8279	1.15			-0.4479	-1.26	-0.4479	-1.26
t			0.2336	3.85	0.1055	1.96	0.2020	3.19				
exp(-t/10)									0.4610	0.26		
exp(t/10)											0.0561	6.09
CHLDRN			-0.2398	-1.82								
SCHLAG							0.7363	1.86				
CARS									0.0672	0.66	-0.2486	-2.30
PBNS01C	1.1346	2.26			0.3966	1.04						
SREC01C			0.4855	1.12	0.7013	1.79						
SHOP01C					1.4350	3.36						
SRVP01C							1.1578	2.68				
OHTIME											-0.0002	-0.28
CHAINS											-0.1987	-1.60
$I_{srec}(i,t)$[a]			0.2727	1.42								
$I_{srvp}(i,t)$[b]							0.3115	1.69				
d_{ih}									-0.0500	-2.51	-0.0561	-3.14

Note: $L(0) = -985.46$; $L(C) = -872.45$; $L(\beta) = -756.91$; $\rho^2 = 0.132$; N = 550. [Note that L(C) is the log-likelihood with constant terms alone.]

[a]Accessibility index for social-recreation.
[b]Accessibility index for serving passengers.

of the nested logit model (9,10). The result indicates, however that the accessibility variables provide rather marginal improvement to the goodness-of-fit of the model, and the socioeconomic, time-of-day, activity history variables and origin-to-home travel time are the major factors that explain non-home-based activity choices.

The origin-to-home travel time has a significant negative coefficient for both temporary and permanent returns to home. It appears that the variable reflects the sequencing tendency that the locations visited after a completion of nonhome activity tend to be closer to home. Accordingly, the individuals exhibit a higher probability of returning home from a location closer to home. The analysis, which used a large-scale data set from the Detroit metropolitan area (26), showed the same tendency of sequencing. The finding obtained from the two data sets may imply risk-averse planning behavior of the individuals. Locations closer to the home base require less time to visit, and the visits can be arranged with flexibility because they will fit into short time slots available during the day. On the other hand, visiting locations farther from home requires more time and allows less scheduling flexibility. Presumably individuals prefer to make less flexible visits first because of the uncertainty involved in trip making and activity engagement (e.g., it may not be possible to visit farther locations later because of tightened time constraints). A previous study (1) suggested similar planning behavior under uncertainty in sequencing activities in a trip chain. Daily time-use patterns reported in the literature (34) also suggest that less flexible activities tend to be pursued first during the day.

Home-Based Choice Models

Unlike the case of the non-home-based model, the time-of-day variables played less important roles in the home-based destination choice models and the model for personal business alone included the variable. Accordingly, the models gave the appearance of the traditional destination choice models. Inclusion of the accessibility indices in the home-based activity choice model resulted in a small improvement of the log-likelihood value and the indices as a group were not significant at $\alpha = 0.05$. The final model excluded the index for shopping because its sign was negative and insignificant. The other three indices had coefficient values between 0 and 1. However, as in the non-home-based activity choice model, these spatial variables played only marginal roles. It can be concluded that the choice of activity types, whether home based or nonhome based, is largely determined by factors other than the accessibility to opportunities [the estimation results of the home-based choice models can be found elsewhere (26)].

Residual Analysis

Underlying the use of the system of the logit models in this study is the assumption that the random disturbance terms associated with respective alternatives are statistically independent across the alternatives in a choice and also across the choices made by an individual. It appears appropriate to adopt this assumption for the destination choice models when they are formulated by activity types. Also note that the logit model is the only choice model that has been applied successfully to empirical destination choice analysis. The assumption, however, may be less appropriate when applied to a series of activity choices. For example, an individual may have a positive or negative preference for certain activities throughout a day, which can be represented only by disturbance terms that are correlated across choices. Inferring from the known results of linear-regression analysis (35), this by itself does not impose any serious estimation problems. However, the activity choice models of this study contain the history variables that may be viewed as a class of the lagged dependent variable. Presence of the correlation then may lead to inconsistent estimates when the ordinary logit estimation procedure is applied. Although it is beyond the scope of this study to develop an improved estimation procedure, an analysis was carried out to examine possible correlations of the residuals of the choice models. The results are summarized in the following paragraph [further discussions can be found elsewhere (26)].

Presence of correlations among the random disturbance terms across choices were examined by using weighted residuals (36). The residuals were evaluated for up to the sixth activity choice for each individual in the sample with more than one out-of-home activity record. The residuals were then regressed on the set of preceding residuals in order to examine the existence of correlations. The results indicated that the correlations were overall weak and were at the level that would have been expected with independent residuals. The result supports the model development effort of the study and indicates that interrelated choices can be adequately modeled by introducing variables that represent the history of the choices without assuming a complex distributional structure for the disturbance terms of a series of choices.

DISCUSSION OF RESULTS

The results of this study can be discussed from two different perspectives. One is concerned with the improvement of destination choice models toward more appropriate representation of spatial travel patterns of urban residents. The other is concerned with the development of a model system that is capable of evaluating the daily travel pattern as a whole rather than as a collection of isolated and unrelated trip segments.

The empirical analysis of this study has clearly shown that there exists a modification of destination choice models that will lead to better depiction of complex travel patterns. By introducing into the model formulation the travel time between a destination alternative and the home base, it becomes possible to represent the patterns in sequencing activity locations in a trip chain and also to better describe individuals' movement patterns that center around their residence locations. Representation of interrelated destination choices involved in a trip chain can be made by applying the destination choice models in a sequential manner.

The destination-to-home travel time is an important factor that influences non-home-based destination choice as much as the traditionally used origin-destination travel time. Judging from the statistical significance of this variable, its inclusion in the model should contribute to its predictive accuracy. Moreover, this improvement does not require any additional information to be supplied; the model can be estimated by using the standard logit estimation procedure with small-scale survey results. The study results warrant the evaluation of the predictive capability of the proposed non-home-based destination choice model in comparison with that of the conventional model, and further the development of the procedure for model application.

Another result of the non-home-based model estimation is that the valuation of travel time varies depending on the time of day, presumably because of the tightening time budget constraint toward the end of the day. This constraint on destination choice can be expressed conveniently in destination choice models.

The difficulty of developing a model system of daily travel patterns is perhaps proportional to the complexity of the behavior itself, especially the magnitudes of interdependencies among the choices. This study, together with the previous effort (1,2), has shown that the dependencies can be incorporated into the model system by use of appropriately developed variables that represent the past history of activity. The significance of the variables suggests that their omission will result in serious errors. The endogenous nature of the history variables, however, may create estimation problems when the random disturbance terms of the choice models are correlated across choices. The residual analysis conducted in connection with this study (26) indicated that such correlations are not significant. Although the effort to develop and apply improved and more versatile estimation procedures should continue, it may be appropriate to conclude that the logit model can be used to represent a series of choices and that each choice model can be separately estimated. These results and also the finding from the previous studies (1,2)--that the activity history can be represented in a simple and convenient manner--all suggest that the model structure can be kept simple and that the model system can be applied in a practical manner.

The study findings also suggest that activity and destination choices are influenced by different types of factors, with only a few affecting both. Activity choice is influenced largely by time-of-day, activity history, and socioeconomic attributes of the individuals, whereas spatial factors play only minor roles. On the other hand, the socioeconomic and history variables influence destination choice behavior to a rather limited extent.

The sequential model system developed here, with further extensions and modifications, can be used in several ways. Daily travel patterns can be reconstructed by the system by using the stochastic simulation technique, and impacts of transportation planning options can be evaluated. This reconstruction is more realistic than one by the conventional procedure because the model system accounts for the interdependencies among choices and continuity of trips. The separability of the explanatory variables, together with the previous findings (2) that socioeconomic attributes play only small roles in non-home-based activity choices, may make possible aggregative treatment of individuals when simulating their non-home-based choices; model application may be able to avoid the bookkeeping difficulties that may otherwise arise. The model can also be used to evaluate the likelihood of alternative daily travel patterns that a person may take in response to changes in various elements in the travel environment. The model will serve as a useful supplementary tool to the in-depth game-simulation technique (37) used to evaluate such responses.

The model system, as it is formulated now, is sensitive to travel time, land use variables, as well as socioeconomic variables. The inclusion of time of day offered the possibility of evaluating the effects on travel patterns of the changes in time-related factors such as store hours. The destination attraction measures that are formulated as time-of-day dependent make the model system sensitive to such changes. The estimation results, however, did not show that the accessibility indices, which are also time-of-day dependent, have an important effect on activity choice. This may be caused by the physiological rhythms inherent in human activity patterns and also to the habitual, routine time-use patterns that may be insensitive to changes in the environment. It is quite conceivable that the temporal variations in the supply of opportunities are closely correlated with the time-use patterns, making it difficult to evaluate the sensitivity of activity choices to changes in the availability of opportunities over the 1-day period.

Although it is believed that the proposed sequential model system will resolve many problems of the conventional forecasting procedure, it is of course not devoid of limitations. The model system assumes the structure of (past) history dependency. As a result, the activity and travel patterns predicted by the system may not necessarily agree with the patterns that individuals, who conscientiously plan ahead and schedule future activities, would exhibit in a different travel environment. Theoretically speaking, a future-dependent model system can be obtained from a history-dependent system (1), but practical difficulties involved therein call for other solutions. One possibility is to model the respective model components such that they reflect the individuals' planning effort. An example of such a model can be found in a recent destination choice analysis (25). The activity choice models may be made future dependent by extending the accessibility index among the time dimension to reflect the availability of opportunities during the rest of the day. Note that the system structure can be kept as history dependent after these modifications. Another task that remains to be completed is the development of activity duration models. This is being undertaken while focusing on the relationship between activity durations and their locations, time of day, and history (38). Interrelationships among activity duration, activity choice, and activity sequencing also remain as a subject of future investigation.

REFERENCES

1. R. Kitamura. Sequential, History-Dependent Approach to Trip-Chaining Behavior. In Transportation Research Record 944, TRB, National Research Council, Washington, D.C., 1983, pp. 13-22.
2. R. Kitamura and M. Kermanshah. Identifying Time and History Dependencies of Activity Choice. In Transportation Research Record 944, TRB, National Research Council, Washington, D.C., 1983, pp. 22-30.
3. T. Hagerstrand. What About People in Regional Science? Regional Science Association Papers, Vol. 24, 1970, pp. 7-21.
4. I.G. Heggie and P.M. Jones. Defining Domains for Models of Travel Demand. Transportation, Vol. 7, 1978, pp. 119-135.
5. W. Brog and E. Erl. Application of a Model of Individual Behaviour (Situational Approach) to Explain Household Activity Patterns in an Urban Area to Forecast Behavioural Changes. In Recent Advances in Travel Demand Analysis (S. Carpenter and P. Jones, eds.), Gower, Aldershot, England, 1983, pp. 350-370.
6. R. Kitamura, L.P. Kostyniuk, and M.J. Uyeno. Basic Properties of Urban Time-Space Paths: Empirical Tests. In Transportation Research Rec-

ord 794, TRB, National Research Council, Washington, D.C., 1981, pp. 8-19.

7. P. Burnett and S. Hanson. Rationale for an Alternative Mathematical Approach to Movement as Complex Human Behavior. In Transportation Research Record 723, TRB, National Research Council, Washington, D.C., 1979, pp. 11-24.
8. R. Kitamura. Serve Passenger Trips as a Determinant of Travel Behaviour. In Recent Advances in Travel Demand Analysis (S. Carpenter and P. Jones, eds.), Gower, Aldershot, England, 1983, pp. 137-162.
9. M. Ben-Akiva and S.R. Lerman. Disaggregate Travel and Mobility-Choice Models of Accessibility. In Behavioural Travel Modelling (D.A. Hensher and P.R. Stopher, eds.), Croom Helm, London, England, 1979, pp. 654-679.
10. H.C.W.L. Williams and M.L. Senior. Accessibility, Spatial Interaction, and the Spatial Benefit Analysis of Land Use Transportation Plans. In Spatial Interaction Theory and Planning Models (A. Karlqvist, L. Lundqvist, F. Snickars, and J.W. Weibull, eds.), North Holland, Amsterdam, 1978, pp. 253-287.
11. A.G. Wilson. Entropy in Urban and Regional Modelling. Pion, London, England, 1970.
12. R.A. Cochrane. A Possible Economic Basis for the Gravity Model. Journal of Transport Economics and Policy, Vol. 9, No. 1, 1975, pp. 34-49.
13. J.H. Niedercorn and B.V. Bechdolt. An Economic Derivation of the "Gravity Law" of Spatial Interaction. Journal of Regional Science, Vol. 9, No. 2, 1969, pp. 273-282.
14. S. Hanson. Urban-Travel Linkages: A Review. In Behavioural Travel Modelling (D.A. Hensher and P.R. Stopher, eds.), Croom Helm, London, England, 1979, pp. 81-100.
15. S. Hanson. Spatial Diversification and Multipurpose Travel: Implications for Choice Theory. Geographical Analysis, Vol. 12, 1980, pp. 245-257.
16. K.P. Burnett and D.C.L. Prestwood. The Development of Disaggregate Trip-Distribution Models. In Behavioral Travel-Demand Models (P.R. Stopher and A.H. Meyburg, eds.), D.C. Heath, Lexington, Mass., 1976, pp. 109-124.
17. S.R. Lerman and T.J. Adler. Development of Disaggregate Trip-Distribution Models. In Behavioral Travel-Demand Models (P.R. Stopher and A.H. Meyburg, eds.), D.C. Heath, Lexington, Mass., 1976, pp. 125-139.
18. P.M. Jones. Destination Choice and Travel Attributes. In Determinants of Travel Choice (D.A. Hensher and Q. Dalvi, eds.), Prager, New York, 1978, pp. 266-311.
19. D. Damm. Parameters of Activity Behavior for Use in Travel Analysis. Transportation Research, Vol. 16A, 1982, pp. 135-148.
20. T. Adler and M.E. Ben-Akiva. A Theoretical and Empirical Model of Trip Chaining Behavior. Transportation Research, Vol. 13B, 1979, pp. 243-257.
21. G.S. Root and W.W. Recker. Toward a Dynamic Model of Individual Activity Pattern Formulation. In Recent Advances in Travel Demand Analysis (S. Carpenter and P. Jones, eds.), Gower, Aldershot, England, 1983, pp. 371-382.
22. F.S. Koppelman and E.I. Pas. Travel-Activity Behavior in Time and Space: Methods for Representation and Analysis. In Measuring the Unmeasurable (P. Nijkamp, ed.), Martinus Nijhoff, The Hague, Netherlands (in preparation).
23. S.R. Lerman. The Use of Disaggregate Choice Models in Semi-Markov Process Models of Trip Chaining Behavior. Transportation Science, Vol. 13, 1979, pp. 273-291.
24. R. Kitamura and T.N. Lam. A Time Dependent Markov Renewal Model of Trip Chaining. In Proceedings of Eighth International Symposium on Transportation and Traffic Theory (V.F. Hurdle, E. Hauer, and G.N. Steuart, eds.), University of Toronto Press, Toronto, Ontario, Canada, 1983, pp. 376-402.
25. R. Kitamura. Incorporating Trip Chaining into Analysis of Destination Choice. Transportation Research, Vol. 18B, No. 1, 1984, pp. 67-81.
26. M. Kermanshah. Time, History, and Location Dependencies of Spatial Choice. Ph.D. dissertation. Department of Civil Engineering, University of California, Davis, 1984.
27. T. van der Hoorn. Experiments with an Activity-Based Travel Model. Transportation, Vol. 12, 1983, pp. 61-77.
28. L.P. Kostyniuk. A Behavioral Choice Model for the Urban Shopping Activity. Ph.D. dissertation. State University of New York, Buffalo, 1975.
29. M.E. Ben-Akiva, L. Sherman, and B. Kullman. Disaggregate Travel Demand Models for the San Francisco Bay Area: Non-Home-Based Models. In Transportation Research Record 673, TRB, National Research Council, Washington, D.C., 1978, pp. 128-133.
30. S.R. Lerman and J.J. Louviere. Using Functional Measurement to Identify the Form of Utility Functions in Travel Demand Models. In Transportation Research Record 673, TRB, National Research Council, Washington, D.C., 1978, pp. 78-86.
31. M. Shapcott and P. Steadman. Rhythms of Urban Activity. In Timing Space and Spacing Time, Volume 2: Human Activity and Time Geography (T. Carlstein et al., eds.), Edward Arnold, London, England, 1978, pp. 49-74.
32. L.P. Kostyniuk and R. Kitamura. Temporal Stability of Urban Travel Patterns. Transport Policy and Decision Making, Vol. 2, 1984, pp. 481-500.
33. G. Argyrakos. The Influence of Time on Trip Distribution--A Short-Term Forecast Approach. Traffic Engineering and Control, Vol. 21, No. 3, 1980, pp. 114-118.
34. R. Maw. Analysing Demand for Leisure Facilities. Built Environment, Vol. 1, 1972, pp. 519-522.
35. H. Theil. Principles of Econometrics. Wiley, New York, 1971.
36. D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers in Econometrics (P. Zarembka, ed.), Academic Press, New York, 1974, pp. 105-142.
37. P.M. Jones, M.C. Dix, M.I. Clarke, and I.G. Heggie. Understanding Travel Behaviour. Transport Studies Unit, Oxford University, Oxford, England, 1980.
38. R. Kitamura. Time-of-Day, Location, and History Dependencies of Nonhome Activity Durations in Daily Activity-Travel Behavior. Presented at 30th North American Meetings of the Regional Science Association, Chicago, Nov. 1983.