HIGHWAY RESEARCH RECORD

Number 472

Transportation Forecasting

9 reports
prepared for the
52nd Annual Meeting

Subject Areas

55	Traffic Measurements
84	Urban Transportation Systems

HIGHWAY RESEARCH BOARD
DIVISION OF ENGINEERING NATIONAL RESEARCH COUNCIL
NATIONAL ACADEMY OF SCIENCES—NATIONAL ACADEMY OF ENGINEERING

Washington, D.C. 1973

NOTICE

These papers report research work of the authors that was done at institutions named by the authors. The papers were offered to the Highway Research Board of the National Research Council for publication and are published here in the interest of the dissemination of information from research, one of the major functions of the Highway Research Board.

Before publication, each paper was reviewed by members of the HRB committee named as its sponsor and accepted as objective, useful, and suitable for publication by the National Research Council. The members of the review committee were chosen for recognized scholarly competence and with due consideration for the balance of disciplines appropriate to the subject concerned.

Responsibility for the publication of these reports rests with the sponsoring committee. However, the opinions and conclusions expressed in the reports are those of the individual authors and not necessarily those of the sponsoring committee, the Highway Research Board, or the National Research Council.

Each report is reviewed and processed according to the procedures established and monitored by the Report Review Committee of the National Academy of Sciences. Distribution of the report is approved by the President of the Academy upon satisfactory completion of the review process.

ISBN 0-309-02257-6
Library of Congress Catalog Card No. 74-53
Price: $3.80

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

CONTENTS

FOREWORD

The papers presented in this RECORD focus primarily on the development of various analytical tools to be used in the transportation planning process.

Jones and Grecco report on a research project to develop a simplified procedure for major thoroughfare planning in small urban areas. Various growth factor models for developing future estimates of internal traffic in small urban areas were tested, modified, and refined using Lafayette, Indiana, data. The complete procedure was tested and demonstrated in Columbus, Indiana.

Kannel and Heathington discuss the findings of their evaluation of trip generation models based on household data rather than on zonal aggregate data. The major objectives of their research were to examine the form of household travel relations, determine the stability of these relations over time, and evaluate the ability of household models to estimate future travel. Their results indicate that household models could successfully predict household travel.

Pigman, Deen, and Deacon discuss the results of their research project to examine the characteristics of travel to outdoor recreational areas in Kentucky. Data were obtained by means of a license plate, origin-destination survey at 160 sites within 42 recreational areas and by means of a continuous vehicle-counting program at 8 of these sites. Findings and conclusions on vehicle occupancy, types of vehicles, trip length, distribution of travel over time, and hourly volumes are reported.

Deacon, Pigman, Kaltenbach, and Deen report on their evaluation of models of travel flow from population centers throughout the United States to outdoor recreational areas in Kentucky. Attempts to simulate distributed travel flows concentrated on various single-equation models, a cross-classification model, and gravity and intervening opportunities models. Analysis and evaluation of the aforementioned models are given.

Dunphy in his paper states that automobile ownership is generally accepted as the most important determinant of trips made by residents from a traffic zone. The author analyzes the relation between transit accessibility and automobile ownership by eliminating variations in family size and income through household analyses. The findings show that there is a significant correlation between automobile ownership and transit accessibility for almost every category of household. According to the author, this indicates that it may be possible to reduce the level of automobile ownership in an area by improving transit accessibility.

Harris discusses the long-term nature of transportation planning, some recent developments that have cast doubt on the utility of transportation planning, and the need for more explicit procedures for transportation planning.

Goodknight discusses a technique for estimating air passenger travel demands on a statewide basis. The author reports on the model developed through research that uses empirically determined relations between total travel and regional socioeconomic activity as a basis for estimating the pattern of intercity travel by all modes.

Waltz and Grecco report on the results of their research of the evaluation of mailed planning surveys to obtain survey data for transportation planning purposes. The authors state that, on the basis of cost versus information obtained, the results indicate that the combined use of the mail approach with mail, telephone, and personal follow-ups could be comparable to the other methods for planning surveys having an informative purpose.

Hill, Tittemore, and Gendell report on the results of an extensive effort to analyze the temporal distribution of vehicular traffic in eight U.S. urban areas ranging in population from 100,000 to 350,000. Graphical models were developed during the analyses, and travel data in an example urban area, St. Louis, are presented in detail along with tabular and graphical outputs of data for each of the other cities.

SIMPLIFIED PROCEDURE FOR MAJOR THOROUGHFARE PLANNING IN SMALL URBAN AREAS

A. D. Jones, California Polytechnic State University, San Luis Obispo; and
W. L. Grecco, University of Tennessee, Knoxville

The purpose of the research project was to develop a simplified procedure for major thoroughfare planning in small urban areas. Previously developed corridor growth factor models for developing future estimates of internal traffic in small urban areas were tested, modified, and refined. Lafayette, Indiana, data were used. Regression models to provide data usually obtained by use of external cordon surveys were developed. External survey reports from 36 cities in 14 states, ranging in population from 10,000 to 90,000, provided most of the data used. Alternate procedures for providing the external survey information, based on use of historical data from the subject city, were also developed. The completed procedure provides forecast traffic volumes within the accuracy necessary for major thoroughfare planning in small urban areas at low cost and with a level of sophistication that will permit application by personnel usually available in small communities. The feasibility of the complete procedure for providing the required traffic volume for major thoroughfare planning in small urban areas was demonstrated in Columbus, Indiana.

•THE continuing planning process developed as provided by the Federal-Aid Highway Act of 1962 was somewhat precise and detailed in the stated requirements. A complete land-use inventory, an inventory of existing physical facilities, an inventory of population and economic information, a review of existing zoning and subdivision regulation ordinances, an inventory of parking facilities and use, and a complete inventory and study of all other aspects pertaining to or connected with existing traffic were required. The manner in which the detailed inventories would be obtained was carefully outlined. A home-interview origin-destination survey and an external cordon survey were specifically required. With the information from the external survey and other collected data, forecasts of future traffic volumes were made, and total future trip generation by traffic zone within the study area was determined. Through the use of computers, these future trips are distributed among the various traffic zones, and finally the total trips between zones are assigned to a mathematical representation of the major arterial network. Through this process, the planners are able to determine the segments of the transportation system requiring either improvement or further planning for development of complete new segments to handle forecast traffic for the target year. This is accomplished by comparing the assigned traffic to the existing capacity of the individual segments of the system.

Techniques specifically designed to accomplish the same study objectives for the small urban areas as for the large areas have not been developed and tested. In general, especially in studies of areas having a population range of 25,000 to 50,000, it has been the practice to use the same procedures as used in the large-area studies. This, of course, means that, in small urban areas, a much higher overall sample percentage is necessary for the home-interview origin-destination survey. In addition, a highly qualified staff, consisting of professional and technical personnel, is needed to successfully complete a transportation study. The required professional staff for a small urban area will be almost as large as a staff for a large urban area. The end result is that, even if a competent staff were available to complete a transportation study for a small urban area, the cost will be much higher on a per capita basis than for a large urban area.

An additional major disadvantage of the detailed procedures followed in large areas is the length of time required to complete such a comprehensive study. The initial data collection phase of a study will usually require a minimum of 2 years when using established procedures in any urban area. In addition, the maximum benefits are derived from such studies only if they are reviewed and updated every 5 years as a minimum.

In small urban areas of 5,000 to 50,000 population, there are always many existing minor transportation problems. These will become greater as the area grows and automobile registrations increase. These problems can be alleviated by proper planning in the majority of cases even though such planning is not currently required by federal legislation.

For the purposes of this study, a small urban area will be defined as a geographically separate urban area. The size will be limited to less than 100,000 population. Public transportation in such areas usually is nonexistent or, at best, accommodates only a small portion of the population; therefore, study procedures for this area will not be included. This is not to imply that small urban areas can ignore the need to plan for a proper transit system.

A simplified planning procedure for major thoroughfare planning, developed and designed specifically to satisfy the requirements of small urban areas, is needed. The procedure must be easily applied by the type of personnel usually available at the municipal level of government, require a small budget, and produce results with the degree of accuracy necessary for sound thoroughfare planning.

Utilization of existing city personnel would accomplish a multipurpose objective. First, the completion of the study by local personnel would enhance the possibility of developing the all-important continual planning process. Second, this type of procedure would permit maximum efficiency and economy, allowing city personnel to complete the required data collection during normal slack periods in their regular routine. Third, and possibly most important, the involved personnel will gain an overall knowledge of the community and its traffic and transportation problems. The problem was approached with the preceding criteria as a guide.

French's (3) recently completed research study utilized data that are readily available in small areas in order to develop simple models for deriving internal-internal traffic volumes on arterial streets in small urban areas. In his models, French used only three independent variables that required collection of data: dwelling units, retail employment, and total employment per corridor. The procedure also required delineating traffic corridors, establishing the limits of the central business district (CBD) and its environments, and determining an external cordon line. The existing traffic volumes at each intersection of a major arterial with the corridor boundary were also required. All of the required information for the preceding models is readily available in all areas or easily obtainable from uncontrolled aerial photography.

CONCEPTUALIZATION OF PROBLEM

In the large urban areas, the sheer magnitude of the component parts of the transportation system and interactions of the many attracting forces or traffic generators make the problem impossible for the planners to mentally visualize. As a result, a step-by-step planning process is required that includes computerized data collection, assimilation, and summarization. In this manner, the individual components may each be carefully analyzed separately and then combined in any selected manner for analysis of alternate systems. However, this procedure is costly and is not necessary for transportation planning in smaller urban areas.

A majority of small urban areas in this country have experienced a modified sector and concentric circle growth pattern. The CBD remains the major traffic generator in the community, and traffic corridors radiate outward from this center. If major shopping centers exist, they are usually located on the radial arterial streets and are small compared to the size of centers in the central area; therefore, their influence is subsidiary to that of the central area.

The traffic corridor concept of thoroughfare planning is not new. In fact, many transportation planners still feel that this approach to the solution of the problem of providing an adequate transportation system is superior to a zone-by-zone analysis—even in this day of third-generation computers. Using the corridor technique requires that both the capacity of the available thoroughfares and the forecast traffic volumes be determined by corridor. The traffic corridor concept is accepted in principle by all planners who utilize computer capacity-restrained-traffic assignment packages in transportation planning. This particular theory of assignment provides for a reduction in link speed when the assigned volume reaches a predetermined level, with this level being based on the level-of-service concept of capacity. The reduction in link speed forces computation of new zone-to-zone minimum paths or trees and new assignment of trips. This effectively distributes trips over a number of arterial streets serving the same basic traffic movement and provides in essence a corridor assignment. In many cases, the reasons for traveler preference for one arterial over another in a corridor may be a slight travel-time difference or some other factor that can be readily rectified or that no longer exists as volumes increase.

Future demand traffic volumes are necessary for each corridor to permit planning for improvements to handle the demand within the time constraint established. French proposed that this future traffic volume could be obtained by multiplying the existing traffic volumes by a growth factor (3). The growth factor was based on growth of the "activities" in the corridor. Comparison of the street capacities to forecast volumes can then provide an estimate of system deficiencies.

STUDY PROCEDURE

The entire procedure is based on the assumption that the existing travel patterns in the community will remain stable over time. This is considered a reasonable assumption. It can be noted that, even in the very large cities, the basic travel patterns remain substantially the same except for circuitous travel over routes provided by controlled-access facilities that tend to encourage such travel. In the small cities, the growth is usually an extension along present patterns. To disrupt or change the basic travel patterns in a small city requires the elimination of a large portion of the existing street network. This is not likely to occur.

The procedure developed considers external and internal traffic separately, and each will be discussed separately here.

CORRIDOR IDENTIFICATION

A corridor may be defined as an area between traffic divides. It represents the area producing trips served by the one or more basically parallel major streets in the area. The orientation of the corridor in small urban areas would be basically oriented toward the central area because of its predominance as a generator. With a knowledge of the local travel habits, supplemented by aerial photographs, street classifications, land-use maps, and a traffic volume map, the corridor limits may be determined. The corridor boundary should be equidistant between arterials unless physical constraints dictate otherwise. Corridors may overlap with separate corridors identified on circumferential or cross routes.

To select corridors first requires delineation of the central area. This central area would include the CBD "core" and would generally include the "frame" of the CBD. Specifically, the central area would begin at the point where radial corridors and the arterial streets serving the corridors merge and lose their individual identity. Usually the merging movement would be served by cross routes bordering the CBD, providing for dispersal of traffic to the scattered destinations.

Traffic entering small cities is composed of varying percentages of external-internal and external-external traffic. Generally, the composition and magnitude of this traffic are determined by an external cordon survey.

It was determined that there are five different items in external traffic information that require procedures to be developed for application in the planning. These are as follows:

1. Method of forecasting total external traffic volumes,
2. Method of allocating the total external traffic to each external cordon crossing,
3. Method of determining the amount of the forecast total external traffic that is external-internal traffic volume,
4. Method of allocating the external-internal traffic volumes to each external cordon station, and
5. Method of determining the amount of the total external-internal traffic volume that is the external-internal traffic destined to the central area.

There are two separate feasible procedures for determining the total external traffic and the components of external-external and external-internal traffic necessary for the simplified procedure for major thoroughfare planning. The procedure will be selected on the basis of the availability of the following information:

1. Previous external cordon survey study for the area, or
2. Traffic volumes from a past year at each cordon station.

TRAFFIC COMPUTATIONS: EXTERNAL REPORT AVAILABLE

Total External Crossings

With an external survey report available the procedure is greatly simplified. A growth factor based on the increase in vehicle registration should be adequate for forecasting. A calibration period using a growth factor based on 5 to 10 years should provide a check on the accuracy of the procedure. This assumption was tested by using 15 cities in Indiana for a data set (Table 1). Traffic volumes on an external cordon around each city were obtained for two points in time that were not less than 5 and preferably 10 to 20 years apart. The cordon line was established at a point that included the urban area for both years and where an Indiana count station was located. A regression analysis was made using as the independent variable the base year total external traffic volumes multiplied by a growth factor representing the increase in county vehicle registration for the period between the two points. The observed total external traffic volume for the later year was used as the dependent variable. The results of the regression analysis were an R^2 of 0.96 with a standard error of estimate of 3.734.

Allocation of Total External Volumes Among Cordon Stations

A model for allocating total external traffic between stations was developed; eight cities with external survey reports available for two points in time were used for the investigation. Data from a total of 72 external cordon crossings in the eight cities were used in a simple regression procedure. This technique for establishing the correlation between the variables was selected because of its simplicity and adequacy. A statistical test was made that showed that the percentage of the total traffic crossing at each of the external cordon stations remains constant over time. As a test, the percentages of total traffic crossing at each station from the base year reports were used for the dependent or response variable, and the percentages of total traffic crossing at each station in the later reports were used for the independent variable. Values for stations in each city were computed in addition to a regression on a combined sample of all 72 crossings in the eight cities. The results of the analysis were examined to evaluate the comparison. Two null hypotheses were tested in each model as follows: $H:B_0 = 0$ and $H:B_1 = 1.00$.

The first hypothesis tests that the intercept value is equal to zero. The second hypothesis tests that the regression coefficient is equal to 1.00. In all models, neither hypothesis could be rejected at the 5 percent level. The results of the analysis were an R^2 of 0.94 and a standard error of estimate of 2.70.

Split Between External-External and External-Internal Traffic Volumes

When a past external survey is available, the percentages of the total external-internal cordon crossings at each station are probably the best estimate of the percentage of the total external-internal cordon crossings for the study year; however, a procedure to provide this information where a report is not available will be presented later. A regression technique was used to test the assumption that the percentage of the external-internal traffic crossing the external cordon at each cordon station was the same as the percentage of the total external traffic crossing at that same station. Twenty-seven survey reports were used for this phase of the investigation consisting of 232 independent cordon crossing stations. Values were computed for data for each city in addition to the regression on the combined sample. For this regression analysis, the percentage of the total external cordon crossings at each cordon station was used as the dependent or response variable. The percentage of the total external-internal traffic at each cordon crossing was used as the independent variable. Table 2 gives the cities used in this analysis.

External-Internal Volume

The split of the total external traffic volume into the two components, external-external and external-internal traffic volumes, is required for the simplified planning procedure. If a previous external cordon survey has been completed, the percentage of split at each station then provides the best estimate of the present split.

External-Internal Traffic to CBD

The final step necessary to provide a complete package for the simplified planning procedure is to determine a means of establishing the percentage at each external cordon station of the external-internal trips that are destined to the central area or to the screen line where the radial corridors merge and lose identity. Employment has been shown to be a very strong trip indicator in other studies, and a simplified distribution method using employment (but not requiring computer iterations for application) was developed.

Eleven study reports containing detailed employment data and trip information by traffic zones were utilized as the data source for this phase of the investigation. Cities included are given in Table 3. Identification of a central area was initially required. For this study, the central area was defined as the CBD and the contiguous traffic zones where the total employment exceeded the number of residents.

The response variable of the regression analysis was the percentage of external-internal trips destined to the central area. The independent variable was the percentage of the total employment in the central area. The R^2 was 0.75, the regression coefficient was 0.95, the standard error of the estimate was 4.60 percent, and the intercept value was 1.11. The assumption that the percentage of the total external-internal trips with origins and destinations in the central area is the same as the percentage of the total study area employment used in the central area was considered valid.

TRAFFIC COMPUTATIONS: EXTERNAL REPORT NOT AVAILABLE

If an external survey report is not available, but traffic volumes from a past year are available at each cordon station, the following procedure should provide information adequate for planning purposes.

Total External Crossings

The procedure for total external crossings is identical to that previously specified when an external report is available. The growth factor based on the increase in vehicle registration is used, but, once again, a calibration period is used as a check on the procedure. If the calibration is not acceptable, an alternate technique using a regression equation for forecasting the future year's total external cordon crossings is used.

Allocation of Total External Volumes Among Cordon Stations

Distribution of total external volumes among cordon stations follows the procedure previously outlined.

Table 1. Growth factor based on county vehicle registration increase.

City	Population (× 10^3)	Base Year	Base Year Total External Crossing	Vehicle Registration Growth Factor	Present Year Used	Present Year Total External Crossing	Forecast Total External Crossing	Error	Percentage of Error
Kokomo	43.3	1956	31,435	1.72	1968	49,461	54,068	4,607	9.3
Marion	40.0	1958	45,405	1.36	1966	45,375	45,302	-73	-0.2
Elkhart	42.4	1959	50,446	1.33	1967	63,707	67,093	3,386	5.3
Goshen	14.6	1959	28,353	1.33	1967	42,378	37,709	-4,669	11.0
Anderson	69.9	1959	42,378	1.50	1969	63,579	63,567	-12	0.0
Columbus	31.4	1959	31,233	1.64	1971	52,586	51,222	-1,364	-2.6
Bloomington	43.2	1960	32,016	1.76	1970	52,835	56,348	3,513	6.6
Wabash	13.3	1965	27,222	1.11	1970	27,435	30,216	2,781	10.1
Seymour	13.1	1963	27,186	1.27	1968	32,786	34,526	1,740	5.3
Connersville	17.5	1957	18,648	1.37	1970	27,762	25,547	-2,215	-8.0
Lafayette	64.0	1952	41,827	1.90[a]	1970	71,278	76,120	-4,842	6.8
South Bend	122.8	1958	64,500	1.36[a]	1968	88,798	89,400	602	0.7
Muncie	68.1	1957	46,695	1.55	1970	75,798	72,377	-3,421	-4.5
Vincennes	19.7	1953	34,212	1.55	1970	43,947	53,029	9,082	20.7
Logansport	19.1	1952	24,419	1.61	1970	34,786	39,315	4,529	13.0

[a] Growth in vehicle registration for Indiana for cross-state route.

Table 2. External volume and external-internal volume.

City	Sample Size	Intercept Value	Regression Coefficient	Standard Error of Estimate	F-Value	R^2
Independence, Kansas	10	0.02	1.00	1.14	719.54	0.99
Big Rapids, Michigan	6	-4.36	1.26	2.42	206.11	0.98
Richmond, Kentucky	7	-5.28	1.38	3.27	90.08	0.95
Campbellsville, Kentucky	8	-0.81	1.07	0.92	274.06	0.98
Bonham, Texas	6	-1.23	1.07	3.60	36.7	0.90
Center, Texas	8	0.48	0.96	1.83	21.54	0.78
New Castle, Pennsylvania	12	-0.00	1.00	1.57	87.72	0.90
Vincennes, Indiana	13	-0.03	1.01	2.54	113.44	0.91
Bay City, Michigan	6	-3.54	1.21	5.14	9.71	0.71
Ann Arbor, Michigan	5	2.65	0.87	4.71	32.74	0.92
Bowling Green, Kentucky	10	-2.20	1.22	2.48	140.91	0.95
Junction City, Kansas	11	1.48	0.84	3.76	134.93	0.94
Brownwood, Texas	9	-0.98	1.09	1.27	510.60	0.99
Somerset, Kentucky	9	-1.35	1.12	0.91	997.01	0.99
Childress, Texas	7	-3.92	1.27	2.68	165.96	0.97
Bay City, Texas	7	-0.21	1.02	1.27	651.34	0.99
Athens, Texas	9	-0.60	1.05	1.00	380.05	0.98
Caruthersville, Missouri	4	2.56	0.90	2.63	188.59	0.99
Hannibal, Missouri	9	-0.68	1.06	1.36	440.44	0.98
Commerce, Texas	6	-0.88	1.05	2.19	113.80	0.97
Blytheville, Arkansas	8	-0.09	1.01	0.39	4,507.47	1.00
Borger, Texas	6	1.53	0.91	1.18	347.64	0.99
Cynthiana, Kentucky	6	-0.69	1.04	3.09	80.51	0.95
Kinston, North Carolina	13	-0.05	1.01	1.93	85.50	0.89
Charlottesville, Virginia	11	-0.58	1.06	1.21	391.43	0.98
Pulaski, Virginia	10	-0.00	1.00	0.44	5,408.23	1.00
Martinsville, Virginia	16	0.02	1.00	2.37	73.98	0.84
Combined set	232	-0.16	1.01	2.33	4,193.07	0.95

External-External Volume

The total external-external cordon vehicle crossings are forecast using regression modeling and Eq. 1:

$$Y_2 = 4.28 + 0.035\,(X_1) + 0.066\,(X_2) - 0.064\,(X_3) \tag{1}$$

where

Y_2 = total external-external cordon crossings (in thousands),
X_1 = population of the cities larger than subject city within a 25-mile radius of the city center (in thousands),
X_2 = county population density (in population/square mile), and
X_3 = population of the cities smaller than subject city within a 25-mile radius of the city center (in thousands).

The summary table is as follows:

Variable	R	R^2	Increase in R^2	F-Value to Add or Remove
X_1	0.73	0.53	0.53	37.96
X_2	0.82	0.68	0.15	15.27
X_3	0.86	0.74	0.06	7.10

The analysis of variance is as follows:

Source	Degrees of Freedom	Sum of Squares	Mean Square	F-Value
Regression	3	1,339.08	446.36	29.69
Residual	32	481.16	15.04	
Total	35	1,820.24		

The standard error of the estimate was 3.88.

The allocation of the total external-external trips to each cordon station is assumed to be the same percentage as the traffic distribution for the present year.

External-Internal Volume

The methodology is the same as previously specified when an external report is available. The external-internal volume at each cordon station is found through subtraction of the external-external volumes from the station's total external volume.

External-Internal Traffic to CBD

The method used is the same as that previously stated. The percentage of the external-internal trips that are destined to the CBD is the same ratio as CBD employment is to total study area employment. The remaining external-internal traffic is assumed to be distributed to all study zones from each cordon station. The individual flows will normally be small.

ALTERNATE TECHNIQUE

If for the study year the actual traffic volumes at the cordon checkpoints along the corridors do not agree (within reasonable limits) with values provided through this methodology, an alternate technique is proposed. The external trip information can be developed by regression modeling.

Regression analysis, using cross products of certain combinations of the variables, determined from plots, was employed as a means of investigating the interactions. Possible combinations of independent variables versus each of the response variables were plotted. Those combinations of variables that indicated an intersection within the limits of the response variable being investigated, if they met the additional criteria stated previously, were then entered into the stepwise regression program by use of the transgeneration option.

External survey reports from 77 cities in 19 different states were obtained for the original data set. The reports were made available by the state highway departments in each of the states.

A total of 20 independent variables, both quantitative and qualitative, were available in the development of the final regression models for this study. Of these 20, 3 were dummy variables used to represent qualitative factors. The same independent variables were used in developing two models regressing on two different response variables during the model development. The response variables were the total external and total external-external cordon crossings. The cities included in the data set are given in Table 4.

As a result of the regression analysis, a model for predicting total external cordon vehicle crossings (for an inference space of 10,000 to 100,000 population) is given as Eq. 2:

$$Y_1 = 28.55 + 0.068 (X_1) + 0.00009 (X_2) - 369.8 (X_3) + 78.3 (X_4) \tag{2}$$

where

Y_1 = total external cordon crossings (in thousands),
X_1 = county population density (in population/square mile),
X_2 = county area multiplied by population of the cities larger than subject city within a 25-mile radius of the city center (population × square miles),
X_3 = reciprocal of the total study area population, and
X_4 = reciprocal of the total study area employment.

The summary table is as follows:

Variable	R	R^2	Increase in R^2	F-Value to Add or Remove
X_1	0.66	0.43	0.43	25.68
X_2	0.75	0.57	0.14	10.47
X_3	0.81	0.65	0.08	7.39
X_4	0.85	0.72	0.07	8.13

The analysis of variance is as follows:

Source	Degrees of Freedom	Sum of Squares	Mean Square	F-Value
Regression	4	4,519.11	1,129.78	20.10
Residual	31	1,742.47	56.21	
Total	35	6,261.58		

The standard error of the estimate was 7.50.

INTERNAL TRAFFIC

Developing a Growth Factor

A growth factor for each corridor must be established to forecast growth of internal traffic. The factor must adequately represent the growth of all "activities" in the corridor. Corridors commonly contain an agglomeration of land uses, each having a different trip generation rate; therefore, a method of weighting these rates is necessary. The growth factor must reflect the increase, present to target year, of each land use.

The method established to handle the weighting of trip attractiveness of various land uses was to use the percentage of total linked trips by linked-trip purpose. This information was obtained from a review of available origin-destination study reports. Trip purpose may be related to land uses or other parameters to obtain a relative trip attractiveness. Linked-trip purpose percentages for automobile driver trips for various sized cities indicate that the percentages are similar for all sizes of cities.

Parameters that are easily measured and capable of being forecast are needed to indicate trip purposes. The acres of each type of land have been used in many studies for this purpose, but there are problems inherent with this parameter, such as varying densities of development, that make it undesirable.

The total number of employees within the unit of study is a good indicator of work and business trips. This information is available from several sources and is usually listed by business establishments and can be forecast satisfactorily.

Home trips can be determined by using the number of dwelling units per corridor.

Shopping trips may be determined by using the total number of retail employees. The number of retail employees by corridor can be obtained easiest concurrently with collection of the number of total employees necessary for work trips.

Social-recreation trips to clubs, theaters, residential areas, and so forth are difficult to represent with any single parâmeter because of their diversity. The three parameters used for work and business, home, and shopping trips can be assumed to represent these trips without a separate parameter.

Previous research has indicated that these three parameters adequately represent total trips. This project therefore used total employees to represent work and business-linked trips by corridor, the number of retail employees were used for shopping-linked trips, and the number of dwelling units were used for home-linked trips and other trips to residential areas.

The procedure established relative trip production rates for the three parameters in the study area in the following manner. The relative average trip production rate per employee is established by dividing the percentage of the total trips to be represented by that parameter by the total number of employees in the study area. The same procedure would be followed for the remaining parameters. These rates are assumed to remain constant over time and are used in both the base and target years.

The procedure for developing a growth rate by corridor was as follows:

1. The relative trip rates by each parameter are multiplied by the quantity of the parameter in the corridor for the base year, and the products are totaled;
2. The procedure is repeated for the target year using forecast quantities of the parameters;
3. The ratio of the target year sum to the base year sum is the corridor growth factor; and
4. The corridor growth factor multiplied by the base year traffic volume in the corridor gives the forecast or design volume for the corridor.

Plan Evaluation

The evaluation procedures for alternate plans are well documented by many references such as the National Committee on Urban Transportation (1). The evaluation is simplified because the extent to which mass transit vehicles will contribute to congestion or its relief is minor in small urban areas. Furthermore, freeway networks are seldom warranted; therefore, improvements to the existing system are the primary solution to traffic problems. The street-capacity calculations themselves should provide clues as to where additional needed capacity can be provided with minimum expenditures. The Policy and Procedure Memorandum 21-18, U.S. Department of Transportation, Federal Highway Administration, dated May 13, 1971, for the TOPICS program provides a good guide for methods of upgrading existing facilities.

All traffic assignment procedures require engineering judgment. This proposed simplified procedure requires the same judgment in its application. There will be only a few existing parallel arterial streets to handle the corridor traffic, and it can be assumed that traffic can be divided between these streets. Major arterials should always have four moving lanes. This should be sufficient, in most cases, to handle arterial traffic in small urban areas. It should be remembered that reevaluation of the plan will be simple under the proposed procedure; therefore, it can be easily and quickly repeated whenever any substantial variations in forecast traffic volumes are noted.

The usual procedure for major thoroughfare planning is to make an overall 20-year forecast with 5-year step or incremental forecasts to provide information for establishing project construction priorities and for capital improvement programs. The simplified procedure should not be an exception to this procedure; quite the contrary, this is one of the strong points of the method. Simplicity and minimum personnel requirements permit reevaluation on short notice as area development dictates. When unexpected new development occurs, a reevaluation can be quickly accomplished to check proposed plans and provide information for modifications if necessary. The 5-year incremental forecasts will also preclude errors due to large growth factors.

The "best" plan is that plan that satisfies the people of the community and satisfactorily handles the traffic. The simplified procedure described here can be used to develop adequate information to provide direction to those charged with the responsibility of developing a plan, but it is not intended as "the" cookbook solution. Judgment and assistance of those in the area are not only helpful but an absolute necessity when developing a plan that will be acceptable to the community it affects.

Data Collection

Data collection is greatly simplified using this procedure. Dwelling unit data and street inventory information are obtained from aerial photographs for the various time periods. Employment data, obtained from state employment offices, are supplemented by some personal contact. Traffic data are usually available from the state, county, or city. Some additional counts may be required. Automobile registrations are available from the state.

Additional information, such as maps, zoning, and land-use data, are usually available from the city; however, some additional data collection may be necessary.

DEMONSTRATION OF PROCEDURE

Proper evaluation of the growth-factor technique of traffic forecasting requires the establishment of standards for acceptance. The standard of acceptance for this project was established as the point where the predicted volume was within the range of accuracy that would allow a planner or designer to determine the correct number of lanes, proper location of improvements, and proper relative construction priorities for improvements to the major thoroughfare system of a small urban area. In reviewing the described procedure, the reader must not lose sight of the basic advantages and design constraints of this procedure. It is simple and economical to use and can be effectively utilized by personnel possessing a minimum of expertise in transportation planning. In short, it is intended to be as simple and inexpensive as possible while still providing the required information.

Major thoroughfare capacities for planning purposes are based on the 1965 Highway Capacity Manual (4). Using the manual, certain ranges of service volumes for the demonstration city were obtained for a thoroughfare assuming the following: level of service, C; population of city, 75,000; peak-hour factor, 0.85; directional split, 60 to 40; peak-hour volume, 10 percent of ADT; G/C, 0.45; lane width, 10 to 12 ft; no parking; and 20 percent turns. Use of these assumptions gives the following capacity ranges for major thoroughfares: four-lane thoroughfare, 12,000 to 15,000 vehicles per day (vpd); four-lane thoroughfare with left-turn lanes, 15,000 to 19,000 vpd; and six-lane thoroughfare, 19,000 to 23,000 vpd.

This indicates that an estimated volume with an error of approximately 4,000 vpd, for volumes under 19,000, will not change the basic design of the street. If the estimated volume forecast during the planning study is a little less than 15,000 vpd for a four-lane street that is 12 ft wide, no improvements will be recommended. If the volume for the target year actually is between 15,000 and 19,000 vpd, then some widening may be needed at critical intersections to incorporate left-turn lanes. At noncritical intersections, additional green time may be available from the cross street to accommodate the additional volumes. It is apparent from these figures that the underestimation of future traffic by 4,000 vpd of volumes below 19,000 would not create a traffic problem for the target year. Overestimation would not involve significant overdesign unless the estimated volume exceeds approximately 19,000 vpd.

The streets on the major thoroughfare system in any urban area should be designed and constructed to four-lane minimum standards according to the recommendations of the National Committee on Urban Transportation. Volumes in the range below the basic capacity of a four-lane facility therefore do not affect the design in any manner.

FEASIBILITY DEMONSTRATION

The city of Columbus, Indiana, was selected as the site to demonstrate the feasibility of the entire package comprising the simplified procedure for major thoroughfare planning in small urban areas.

Study Area

Columbus is a city of approximately 27,000 population. The city is a typical small city that has experienced a steady growth through the years. Because of the location at the junction point of several rivers and other small streams, the growth has been primarily in the north and east portions of the city rather than concentrically as in many communities.

The completion of I-65, providing a connecting route from Indianapolis to Louisville, Kentucky, and also on the route from Chicago to Florida and other southern states, is the only major change in the highway system in the area during the past 20 years.

A number of major industrial plants such as Cummins Engine and Arvin Industries are located in the city. The city has a higher than average ratio of employment to population because of heavy industrialization. The effect of this factor on trip generation characteristics of the community is to produce a higher percentage of external-internal trips than other comparably sized communities. The city administration and the citizens have long recognized the necessity for sound planning of the future. This progressive attitude is positively indicated by numerous studies and resulting reports on all phases of community development. The abundance of basic material to use for data sources may indicate that the estimated cost for the study should be increased when estimating the cost of application in communities with less basic data.

External Cordon

The study area external cordon was established to include the area expected to become developed by 1990. The I-65 route was utilized as the west cordon limit because of the natural screen line it provides.

Clifty Creek was established as the east boundary. The location of the cordon throughout the study area made possible the use of Indiana State Highway Commission count station locations.

Corridors and Major Thoroughfares

The arterial street plan prepared by De Leuw, Cather and Company (2), currently being used as a guide for Columbus, was used to assist in initial street inventory traffic volume counts and corridor identification.

The identification of corridors for Columbus was accomplished by using the arterial street plan, the existing traffic volume flow map, existing land-use map, and information from personnel familiar with the area.

Seven basically radial corridors were established as shown in Figures 1 and 2. Two of these corridors overlap because of the configuration of the streets. Ind-46 is considered a radial route; however, it causes a 90-deg route change in the approach to the central area screen line, dispersing traffic over five closely spaced streets crossing the screen line. This alignment crossed corridor 4, Central Avenue. This does not create a double count because the procedure uses a growth factor, not trip productions.

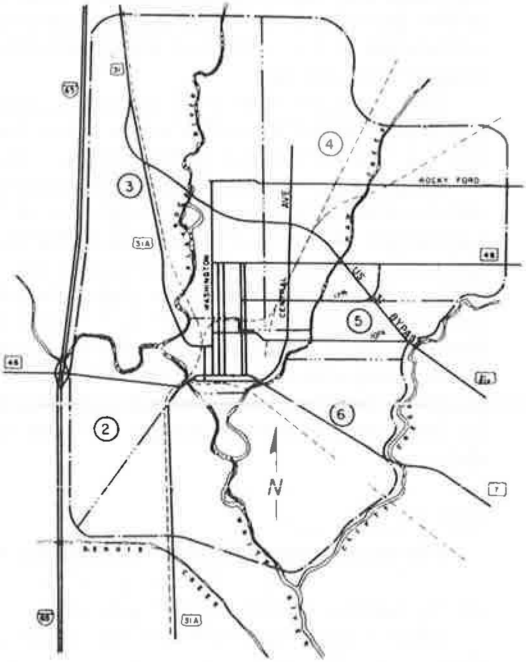
Table 3. External-internal trips and employment.

City	Study Size (× 10^3)	Total Employment	Employment in CBD and Frame Destinations	Percentage of Employment in CBD and Frame Destinations	External-Internal Crossings to CBD and Frame Destinations	Total External-Internal Crossings	Percentage of External-Internal Crossings to CBD and Frame Destinations
Moberly, Missouri	13.4	5,079	2,200	43.3	6,109	14,762	41.6
Henderson, North Carolina	20.4	9,778	2,495	25.5	5,600	20,522	27.3
Lawrenceburg, Tennessee	10.3	5,786	1,546	26.7	2,600	12,322	21.1
Lumberton, North Carolina	20.4	7,166	2,168	30.3	7,400	26,000	28.5
Glasgow, Kentucky	13.0	9,330	2,110	22.6	4,700	15,956	29.5
Franklin, Kentucky	7.3	5,760	1,329	23.1	3,100	10,142	30.6
Cynthiana, Kentucky	6.7	2,900	886	30.6	3,900	11,009	35.4
Frankfort, Kentucky	22.9	16,500	5,415	32.8	5,300	16,287	32.5
Brownsville, Texas	65.0	14,449	5,300	36.7	9,185	25,948	35.4
Champaign-Urbana, Illinois	80.0	33,885	12,898	38.0	8,800	26,714	33.0
Nashville, Tennessee	732.0	142,018	16,800	11.8	8,219	69,199	11.8

Table 4. Populations of cities used in alternate model.

City	City Population (× 10^3)	Study Population (× 10^3)
Moberly, Missouri	13.4	13.4
Lancaster, Ohio	32.9	32.9
Bellefontaine, Ohio	11.3	11.3
Urbana, Ohio	11.2	11.2
Xenia, Ohio	25.4	25.4
Tiffin, Ohio	21.6	21.6
Circleville, Ohio	11.7	11.7
Greenville, Ohio	12.4	12.4
Mount Vernon, Ohio	13.4	13.4
Alpena, Michigan	14.7	17.2
Bay City, Michigan	49.1	82.3
Glasgow, Kentucky	10.9	13.0
Richmond, Kentucky	12.5	12.5
Bowling Green, Kentucky	30.5	33.2
Frankfort, Kentucky	18.0	22.9
Owensboro, Kentucky	42.5	42.5
Elizabethton, Tennessee	10.1	19.7
Henderson, North Carolina	12.7	20.4
Kinston, North Carolina	24.8	49.0
Jacksonville, North Carolina	15.7	30.6
Lumberton, North Carolina	15.3	20.4
Sanford, North Carolina	11.7	16.5
Hays, Kansas	14.0	14.0
Independence, Kansas	11.5	11.5
Pittsburg, Kansas	18.7	18.7
Borger, Texas	13.9	13.9
Bay City, Texas	12.0	12.0
Brownwood, Texas	16.3	16.3
Big Spring, Texas	28.2	28.2
Blytheville, Arkansas	28.3	28.3
Pulaski, Virginia	11.0	13.3
Tuscaloosa, Alabama	61.9	73.2
Gainesville, Georgia	15.4	21.4
Boise, Idaho	75.0	85.3
Billings, Montana	61.6	62.0
Great Falls, Montana	60.0	74.7

Figure 1. Corridors of Columbus.



US-31 Bypass traffic was forecast using a growth factor based on the growth of the entire area from the Tenth Street corridor on the east to the Flatrock River on the west.

Calibration Procedure

The simplified procedure developed recommends the use of two points in time to establish a calibration for the city involved. The project was initiated early in 1971; therefore, 1970 was used for the study year data.

The 1960-to-1970 calibration period for Columbus was selected for several reasons as follows:

1. U.S. Bureau of Census data were available to check dwelling unit counts from aerial photography,
2. Traffic volume counts were available from the Indiana State Highway Commission, and
3. The 10-year period provided a reasonable test of the capabilities of the overall procedure.

A complicating factor that occurred during this period was the construction and opening to traffic of I-65 immediately west of Columbus. The 1963 Columbus arterial plan (2) presented before-and-after volumes throughout the city, providing sufficient information to assess the effect of opening of I-65.

The actual calibration procedure was as follows: 1960 was used as a base year, and growth factors for the external and internal traffic based on the corridors established were computed. After applying the growth factors to the 1960 existing volumes, the resulting forecast 1970 volumes were compared to the observed 1970 volumes. Comparison of the central area screen line volumes within the accuracy necessary for design was considered sufficient to reasonably ensure that the corridors established were satisfactory for planning purposes.

External Traffic

An external traffic growth factor was computed using the increase in total vehicle registration for Bartholomew County for the period 1960 to 1970. The completion and opening to traffic of I-65 just west of the city in late 1962 made direct comparison of 1960-to-1970 external volumes impossible; however, adjustment of the 1960 volumes on US-31 by 30 percent to adjust for the Interstate provides comparable figures. The adjustment factor was provided by information presented in the arterial street plan report (2).

Table 5 gives a comparison of forecast 1970 traffic volumes at external cordon stations (developed by applying a growth factor, based on the increase in total vehicle registration in the county from 1960 to 1970, to the 1960 traffic volumes at each station) and the observed 1970 traffic volumes. The 1960 traffic volumes at the US-31 external station north of the city and US-31 Bypass at Clifty Creek were reduced 30 percent as indicated by the De Leuw, Cather study (2) to adjust for the opening of I-65. The total 1970 forecast external volume was 46,614 vpd as compared to the observed 1970 external volume of 48,683 vpd at the stations. The total error is 2,069 crossings, and, by distributing this among individual stations based on the existing percentages of total external traffic, the maximum error would be 500 crossings. Individual expansions at each station indicate maximum errors of 3,301 vpd at US-31 Bypass at Clifty Creek and 2,790 vpd at US-31 Alternate at Denois Creek. These differences can be attributed to a slight change in traffic patterns occurring subsequent to the opening of I-65. None of the differences was of sufficient magnitude to cause a design change if they were used for a design. However, the future forecast will be based on the 1970 patterns and, therefore, will not reflect these differences because of a slight change in traffic patterns. The growth factor for all external stations was based on the Bartholomew County total vehicle registration increase because I-65 (outside cordon) was considered to be the route selected by cross-state traffic. The comparison thus obtained was considered acceptable, and a growth factor based on county vehicle registration increase was considered acceptable for forecasting external volumes to 1990.

The external-external component of the total external traffic was determined by regression model. The external-external volume thus computed was 9,500 vehicles in 1960 and 10,800 in 1970. This volume was distributed to the external stations using the same percentage as existed for the total external volume.

The percentage of the external-internal traffic to be distributed to the central area was determined by the percentage of total study area employees employed in the central area in 1970. This amounted to 44.3 percent.

Internal Traffic

The internal traffic volume growth factors were computed using the growth of three parameters, dwelling units, total employment, and retail employment, in each of the seven established corridors.

The percentage of the total internal trips to be represented by each of the three parameters of dwelling units, total employment, and retail employment are 50, 35, and 15 respectively. Dwelling unit data by corridor were obtained from aerial photography enlargements (1 in. = 400 ft) for both years. Employment data for both years were obtained from information assimilated and tabulated by the Indiana Employment Securities Division. The base year traffic volumes were obtained primarily from counts made by the Indiana State Highway Commission in 1959, supplemented by information from city files and the arterial street plan report (2). The 1970 counts were from the Commission and City Engineer's Office. Additional counts were provided by city personnel to complete the required information.

The corridor growth factor procedure was used to expand the existing 1960 traffic volumes to 1970 and compared them with the actual observed traffic volumes. The forecast and observed volumes at the central area screen line were compared for each corridor, and additional point volumes were compared on Ind-46 (25th Street) at US-31 Bypass and Washington Street intersections. Table 6 gives the results of this comparison and gives the growth factors used.

The maximum difference between 1970 forecast average daily traffic volumes and 1970 observed volumes was 2,361 in corridor 3, US-31 (N). This is probably due to a slight change in traffic patterns occurring after completion of I-65; however, the difference did not affect the thoroughfare design.

Differences in all other corridors are of such magnitude that designs would have been unaffected. Corridor 7 is one of the major corridors with regard to total traffic magnitude; however, the one-way pairs of Franklin and Lafayette and California and Chestnut, in addition to Washington Street, serve the traffic desiring to enter the central area. Seventeenth Street was not included as an east-west route because of its configuration, which terminates at US-31 Bypass on the east and at a cemetery on the west. It does effectively serve as an overflow or alternate route for Ind-46 and 25th Street for short trips as shown by the existing volumes. The calibration or check period as described here substantiates the corridor identifications and the overall feasibility of the entire procedure with respect to providing adequate accurate design information.

CONCLUSIONS

The completed package for a simplified planning procedure for major thoroughfare planning for small urban areas, using the corridor growth factor technique with synthetically developed external data, provides traffic volumes sufficiently accurate to develop major thoroughfare plans. The methodology fits satisfactorily into the overall planning process, using output from other studies as input to the process. The cost of completing this type of study is a fraction of that required for the home-interview, computer-oriented procedures although the resulting information produced satisfies the same requirements, i.e., design volumes. Detailed cost and time figures were compiled during the feasibility demonstration. After making upward adjustments in these costs to convert from a research environment, the best estimate of the total cost if the study is conducted by city personnel is $15,389.

Specifically, the following conclusions can be drawn from the research:

Figure 2. Columbus radial corridors.

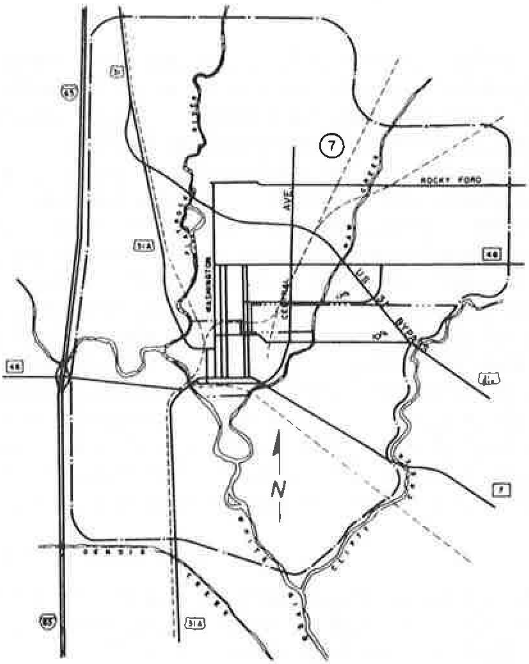


Table 5. External cordon station check for Columbus.

Route	1960 Volume	Growth Factor	Estimated 1970 Volume	Actual 1970 Volume	Error
US-31 (N)[a]	6,858	1.64	11,247	10,399	848
Ind-46 (E)	3,956	1.64	6,488	6,816	-328
US-31 Bypass at Clifty Creek[a]	5,325	1.64	8,733	12,034	-3,301
Ind-7 at Clifty Creek	4,522	1.64	7,416	7,371	45
US-31 Alternate at Denois Creek	3,562	1.64	5,842	3,052	2,790
Ind-46 (W)	4,200	1.64	6,888	9,011	-2,123
			46,614	48,683	

[a] Existing volumes reduced 30 percent to adjust for opening of I-65 (2).

Table 6. Radial corridors of Columbus.

Corridor	Street or Highway	1960 Volume	Internal Volume	Internal Growth Factor	External Volume	External Growth Factor	Estimated 1970 Volume	Actual 1970 Volume	Corridor Error
Direct Route									
1	US-31 Alternate (S)	7,400	5,170	1.63	2,230	1.64	12,084	10,240	1,844
2	Ind-46 (W)	4,793	2,193	2.98	2,600	1.64	10,799	12,861	-2,062
3	US-31 (N)	5,220	0	2.38	5,220	1.64	8,561	6,200	2,361
4	Central Avenue	10,000	10,000	1.31	—	—	13,100	14,495	-1,395
5	Tenth Street	4,100	4,100	1.34	—	—	5,494	5,400	94
6	Ind-7	14,608	11,898	1.21	2,710	1.64	18,841	16,708	2,133
7	Washington	10,800	8,435	1.54	2,365	1.64	16,869	15,974	
	Franklin	2,500	2,500	1.54	—	—	3,850	3,800	
	Lafayette	1,500	1,500	1.54	—	—	2,310	1,700	
	California	2,300	2,300	1.54	—	—	3,542	2,900	
	Chestnut	1,300	1,300	1.54	—	—	2,002	2,500	
		18,400	16,035	1.54	2,365	1.64	28,573	26,874	1,699
Circumferential Route									
	US-31 Bypass at 25th Street	7,700[a]	6,770	1.46	930	1.64	11,025	12,317	-1,292

[a] Reduced volume by 30 percent (2).

1. The corridor growth factor procedure, in combination with synthetically produced external information, can be used as a complete package to determine future traffic demand within the accuracy necessary for major thoroughfare planning in small urban areas. Tests in Columbus substantiated this fact.
2. The three parameters used for the corridor growth factors determination are adequate to indicate corridor traffic volume growth. The data are easy to obtain and easy to forecast. Aerial photography can be used satisfactorily to obtain dwelling unit information, discern growth patterns, and so forth for use with the simplified procedure.
3. With regard to external traffic in small urban areas, the existing distribution of total external cordon traffic volumes among stations may be used as the best estimate of future distribution of the forecast total external volumes. For small urban areas, a growth factor developed by using the county total vehicle registration increase is sufficiently accurate for thoroughfare planning. Regression modeling can be used to provide the total external-external traffic volumes in a small urban area with sufficient accuracy for thoroughfare planning, and this computation does not require use of computers. The best estimate for distribution of external-external traffic volumes among external cordon stations is the existing percentages of total external volumes.

ACKNOWLEDGMENTS

The authors wish to acknowledge the financial support of the Joint Highway Research Project of Purdue University and the Indiana State Highway Commission (5, 6, 7, 8) and the Federal Highway Administration. Also, the data assistance provided by many state highway departments is greatly appreciated. It should be noted that the final report was accepted by the Federal Highway Administration as documentation of the research performed but with a number of comments and reservations and a request that some reevaluation be done. This is being done, but, as of this date, the Federal Highway Administration does not subscribe to the findings and conclusions.

REFERENCES

1. Better Transportation for Your City. National Committee on Urban Transportation, Public Administration Service, Chicago, 1958.
2. Arterial Street Plan for Columbus, Indiana. De Leuw, Cather and Company, Chicago, April 1963.
3. French, D. K. A Simplified Procedure for Major Thoroughfare Planning in Small Urban Areas. Joint Highway Research Project, Rept. 29, November 1967.
4. Highway Capacity Manual—1965. HRB Special Rept. 87, 1965.
5. Jones, A. D. A Simplified Procedure for Major Thoroughfare Planning in Small Urban Areas. Joint Highway Research Project, Rept. 18, July 1972.
6. Jones, A. D. Study Design for Columbus, Indiana, Major Thoroughfare and Plan Development Study. Joint Highway Research Project, Rept. 19, July 1972.
7. Jones, A. D. Major Thoroughfare Study and 1990 Plan for Columbus, Indiana. Joint Highway Research Project, Rept. 20, July 1972.
8. Jones, A. D., and Grecco, W. L. Procedure Manual for Determining Traffic Patterns for a Simplified Procedure for Major Thoroughfare Planning in Small Urban Areas. Joint Highway Research Project, Rept. 21, July 1972.

TEMPORAL STABILITY OF TRIP GENERATION RELATIONS

Edward J. Kannel, Purdue University; and
Kenneth W. Heathington, University of Tennessee

Trip generation models based on household data rather than zonal aggregate data are evaluated. It has been suggested that analysis of household travel characteristics should precede aggregation so that home-interview data can be used more efficiently. Relations identified at the decision level of travel should have greater causal validity and should be more temporally and spatially stable. The major objectives of this research are to examine the form of household travel relations, to determine the stability of these relations over time, and to evaluate the ability of household models to estimate future travel. The potentials for reduced sample sizes and greater applicability of disaggregate models in different urban areas are also examined. Household travel data were obtained from home-interview surveys in 1964 and 1971. Single-family households interviewed in the 1971 survey represented the identical families that were interviewed in 1964. This unique sampling design permitted the analysis of the effects that changes in the households' socioeconomic characteristics during the 7-year period had on trip production. The results indicated that the household models based on the 1964 data could successfully predict household travel reported by the same households in 1971. The household models from both time periods could be expanded to adequately estimate 1964 reported zonal area travel. Parameters of the disaggregate models also appear more consistent among geographical areas and could be developed with considerably fewer data than comparable zonal models.

•THE methodology of trip generation modeling used in most current urban transportation planning studies is referred to as a zonal analysis concept. The enormous body of data obtained in the home-interview portion of the origin-destination study is aggregated and summarized in larger units of the total study area, the traffic zone. These zones are the smallest areas considered in all further analyses and projections. The aggregated data are used to calibrate generation models that estimate trip production occurring under present economic, social, and physical conditions. Future travel is then estimated assuming that the true causal relations have been identified and that the model parameters will remain stable over time.

This traditional trip generation modeling approach, which is based on aggregated socioeconomic and land-use data, is subject to critical review. The modeling approach has been challenged from at least two major viewpoints:

1. The modeling approach does not allow full consideration of the continuous nature of the travel decision process. Trip generation is only the first stage in the total urban transportation planning process that also consists of trip distribution, modal split, and trip assignment. In current practice, each of the models is normally developed independently of the others. As a result, there is no general assurance that an internally consistent network equilibrium will be achieved. The modeling process acts as though there is a given level of demand irrespective of the transportation system that is available.
2. The use of spatially aggregated data assumes that the relations derived represent the true relations occurring in the units that compose the aggregate total. Further, the aggregate descriptions are assumed to remain stable temporally and thus serve as a basis for prediction of future travel.

Inconsistencies may arise because the transportation system is not explicitly allowed to affect all stages of the model development; therefore, attempts have been made to develop explicit demand models that combine the functions of generation, distribution, and modal split (2, 8, 11). Other research efforts have been directed to the development of a stochastic modeling approach that would retain the sequential nature of current planning models but would incorporate principles of economic utility theory to include more policy-sensitive variables in the model framework (12, 13, 15). These latter approaches also recognize the efficacy of using disaggregate data to estimate model parameters.

Research related to the second major point of discussion, data aggregation effects, has also shown the shortcomings in the aggregate planning model concept (1, 3, 6, 9, 10). Review of the assumptions of the aggregate models has shown that the zonal means are not adequately representative of the individual units composing the mean (9). The reasons for inadequate representation are that the zone sampling distributions are skewed rather than normal so that the sample mean is not the central value and considerable heterogeneity exists within zones with respect to household travel characteristics and socioeconomic traits.

Further, aggregation of the behavioral units to a zonal description "washes out" much of the total variation that exists in the data. The aggregate data may mask the true relations and the causal nature of the explanatory variables. Investigations have shown that aggregation changes the strength of the associations among variables, and that the model parameters are dependent on the size of the area unit selected in the analysis (1, 3, 9). As a result, the calibrated model is applicable only at the macrolevel of analysis and in the geographical area for which it was calibrated. This has further important implications in the continuing phase of the urban transportation planning study as one needs to be concerned with measurement of changing conditions. When the analyst is interested in measurement of changes, particular care must be exercised to carefully identify the explanatory variables to be used and the parameters associated with those variables. Because the aggregate models are based on large volumes of data that are averaged together, the models are not sensitive to subtle changes that occur at the basic decision level of travel. Further, the data measure habits for a single time frame. Because it is financially impractical to obtain the large quantities of data that would be necessary to revise the zonal estimates, the relations observed in the original time frame are generally assumed to be held constant throughout the planning period. Logic suggests that changes in social and cultural patterns and changes in the physical environment will have an effect on urban travel. To be sure, comparisons of aggregated relations within an urban area have been made for different time periods, but these aggregate relations are dependent on the type and level of activity within the area units. Because the size of some zones or the level of activity within the zones may change over the study period, it is difficult to separate the effect of changing area descriptions from the effect of changing relations of the variables in the model. The subtle changes in urban structure and the status and life-style of the individual cannot be detected at the macrolevel of analysis.

DISAGGREGATE TRIP GENERATION ANALYSIS

Although there is a recognized need for considering concepts that incorporate the interacting effects of the total travel decision process, in this paper attention is directed only toward obtaining a better understanding of travel behavior by evaluating trip generation relations at a disaggregate, behavioral level of analysis. The household is taken as the basic decision-making unit for evaluating travel behavior.

Other researchers have pointed out the shortcomings of models based on spatially aggregated data and have indicated the desirability of identifying the more basic relations between the socioeconomic and travel characteristics that occur at the household level of analysis. Analysis at the disaggregate level appears to provide a means of overcoming several of the shortcomings mentioned previously and provides several advantages to the transportation analyst. First, because the analysis is conducted at the household level, the basic relations are not averaged out by aggregation or clouded by the analyst's selection of the area boundaries. Because the parameters of the model are not tied to a particular aggregation scheme, the model can be developed and then applied to whatever aggregation scheme is necessary for the following series of models that are employed. This is of importance in the continuing phase of the transportation study as the size and shape of the planning area change. This greater flexibility in application would allow the analyst to more effectively use data from other public records (e.g., census data) that are summarized in area units that do not conform to the boundary scheme of the transportation study.

A second advantage proposed for the use of household level analysis is that the household provides a common base for comparing travel characteristics in different urban centers. Unlike the artificial aggregate unit such as the traffic zone, the household is basically of the same size and internal consistency in different geographical areas. Because of the common nature of the household unit, one might expect household model parameters to be more consistent from area to area.

A third advantage is that the household relations represent the basic relations at the decision level and therefore are assumed to possess greater causal validity. These causal relations are more likely to remain stable over time, thus forming a more valid basis for the prediction of future trip generations.

Finally, because all the data that are collected in the home-interview surveys are analyzed prior to aggregation, the data are used more completely and effectively. As a consequence, the possibility exists for using smaller sample surveys in the continuing study to measure the changes that occur in the basic relations.

The purposes of the research reported in this paper were to evaluate the form of the relations that occur in household trip generation models and to evaluate the stability of these relations. In addition, the ability of models based on reduced sample size to estimate total area travel is examined. The hypothesis that the parameters of disaggregate generation models are more consistent from one geographic area to another is studied.

STUDY DESIGN

The data used in this study to evaluate the causal validity and stability of household trip generation relations were obtained from home-interview surveys in Indianapolis, Indiana. In this metropolitan area of 800,000 population, a basic transportation study was conducted with the home-interview data collection taking place in 1964. The Indianapolis Regional Transportation and Development Study (IRTADS) is a typical example of a transportation study in which the trip generation formulations are based on aggregate zonal totals. A 5 percent home-interview sample was taken representing over 10,000 interviews. The data were aggregated into 395 zones defined for the study area and factored to represent total travel volumes for the area.

In this research, a second home interview was conducted in 1971 to study changes in household socioeconomic characteristics and travel behavior and to evaluate the stability of household trip generation models over a 7-year period. The latter survey obtained measures of the socioeconomic and travel characteristics of some of the identical families that were interviewed in 1964. Earlier research had suggested several variables that could be evaluated at the household level including family size, automobile ownership, stage in the family life cycle, occupational status, income, type of dwelling unit, and location within the urban structure (7, 10, 16). Although simultaneous evaluation of all levels of all factors would have been desirable, such a design would have required a prohibitively large sample to obtain a sufficient number of cases for statistical stability in all possible levels and combinations of the variables. Instead, a sample was selected that represented all levels of three principal socioeconomic variables: family size, automobile ownership, and income. In the experimental design, the confounding influence of other variables was controlled to the greatest possible extent by careful selection of the 1971 sample. Differences in travel behavior that were caused by differences in life-style of families living in different types of dwelling units (and not by changes in the principal variables being considered) were controlled in the 1971 study by selecting only single-family homeowners. The 1971 survey interviewed the identical families that were interviewed in 1964. Further, only those families were selected that remained at the same dwelling unit from 1964 to 1971. In this way, differences in travel behavior that may have been caused by changes in the living environment could be controlled to a greater degree. Finally, travel variation that may be attributed to seasonal or daily variation was controlled by obtaining both data sets in the fall of the year and by scheduling the 1971 interview schedule such that each household recorded travel in 1971 on the same day of the week as in 1964.

Elimination of all families who were not single-family homeowners or who did not provide complete information in 1964 provided a final list of 4,300 households from which the 1971 sample was selected. Table 1 gives a list of average household and travel characteristics of the 4,300 households and of the 357 households from which completed interviews were obtained in 1971.

HOUSEHOLD TRAVEL CHARACTERISTICS

Several household variables such as family size, automobile ownership, income, labor force, and occupational status were examined to determine their effect on household trip generation rates. All of the variables showed a statistically significant effect when considered alone; however, because of the large intercorrelation among the variables, the effect of any one variable is not independent of the others. For example, although income would be a significant explanatory variable for estimating trip production, the research indicated that income had a greater effect on automobile ownership, which in turn affects household travel (5). Because automobile ownership appeared as a more direct cause of travel, it was selected for use in the prediction models along with the family size variable.

Graphical summaries of the relations between family size and automobile ownership and household travel are shown in Figure 1 and Figure 2 respectively. Figure 1 shows that the relation between family size and home-based trip production is nearly linear for family sizes of four or less, but, as family size increases, the rate of trip production increases at a decreasing rate. This overall nonlinear trend agrees with the findings reported by Oi and Shuldiner (10). Because trip generation models generally assume linear relations, large departures from linearity could have important effects on these prediction models. The analyst must recognize where the assumptions of the model are not met and the consequences of using the variable or model formulation in spite of these irregularities. This will be examined later in connection with evaluation of the predictive ability of the models.

The other significant observation to be made from Figure 1 is the relatively good agreement of the curves for the two data sets. Although changes in the family composition and age structure have occurred over the years, the average trip production for families of similar size for the two periods is relatively stable. This again has important implications in developing models of travel behavior. The stability of the form of the relation indicates that the variable should be useful in forecasting models.

Figure 2 shows the corresponding curve for automobile ownership. The curve exhibits strong linear trends with greater fluctuation from linearity exhibited in the 1971 data. However, the slope and intercept (and thus the effect of automobile ownership on trip production) appear to have shifted in these households over the years. Such a shift in the relations could again have special significance in the planning study. Unless a shift in the value of model parameters is detected by observation at intervals less than the planning period for which forecasts are made, the final estimate could yield considerable error. Because the disaggregate modeling approach is able to detect the subtle changes that occur, it is felt that these relations could be monitored with smaller sample sizes and perhaps with greater frequency to detect these changes.

MODEL DEVELOPMENT

Stability of the disaggregate trip generation relations was evaluated in three stages. First, standard linear regression models were developed from the 1964 and 1971 data sets, and the parameters of these models were compared. Next, the 1964 model was used to predict the volume of travel that should be expected in 1971 if the model relations are sufficiently stable to predict future travel in the households. The regular planning process was then essentially reversed in that the 1971 model was taken back in time to estimate the total zonal movement reported by the 4,300 single-family households in 1964. The 1964 and 1971 models were compared as to their ability to measure the 1964 aggregate home-based trip production. Finally, a disaggregate model from all types of dwelling units was used to estimate total area travel. The consequences of using data that do not appear to meet the theoretical requirements of the model formulation are discussed.

As indicated, the relation between family size and trip production did not appear to be linear through the entire range of the independent variable. Preliminary investigations also showed heteroscedasticity of household trip production variances for all levels of the family size and automobile ownership variables. Further, the sampling distribution of the dependent variable is not a true normal distribution. Although this does not preclude the use of linear regression analysis to estimate parameters for the model, one may not be able to make probabilistic statements about the accuracy of the model parameters with the degree of confidence that is usually associated with the statistical model. Because of these limitations, linear regression techniques were used to evaluate the disaggregate models.

Table 2 gives the results of the linear models for estimating home-based trip productions. As was expected, the parameters of the model have shifted somewhat over time. The degree of change is in agreement with observations made from Figures 1 and 2. That is, the parameter for family size is very similar over the period, whereas automobile ownership has greater variability. Two-way analysis of variance models (ANOVA) with unequal cell sizes were evaluated to test the stability of the relation over time (17). The time factor may be labeled simply as a years' effect, but years is considered only as a surrogate for the effect of changes in other possible pertinent variables such as income and stage in the family life cycle. The statistical analysis indicated that there was not a significant change in the effect of family size over the time period, but the effect of automobile ownership had changed. From Figure 2, one could speculate that the change occurred in the zero- and three-car families. Indeed, when only one- and two-car households were considered, there was not a significant variation due to time changes.

The coefficient of determination, R^2, and the standard error of the estimate provide other measures for comparing the two models. Both models give similar statistics for these measures, but, for the analyst who is accustomed to observing R^2 values of about 0.90 for zonal data, they are unimpressive. However, these values were not unexpected because the models are attempting to explain all of the variation in trip production—not just the variation between zones. Within any household, the number of trips reported may be two to three times the average rate of trip production of all households with similar characteristics. The household model formulated here cannot hope to predict these large variations for each household. The measure of usefulness of the household model for forecasting trip production must be based on its ability to predict average travel for some higher level of aggregation. If the model is successful in accomplishing this task, then model development at the disaggregate level would be of value to the researcher as a means of evaluating causal relations at a behavioral level and to the practitioner for developing area travel forecasts.

Estimation of 1971 Household Travel

The 1964 trip generation model given in Table 2 was first used to estimate home-based trips for the 357 families in 1971. The total estimated home-based travel was 2,542 trips compared to the survey total of 2,498 trips, i.e., an error of less than 2 percent.

Sufficient data were available in one- and two-car households and all family size levels to statistically evaluate discrepancies in the estimated and observed trips using a chi-square contingency analysis (14). The null hypothesis of no difference between the estimated and surveyed trips could not be rejected at the 0.01 significance level. Visual inspection of zero- and three-car families also did not show any major discrepancies.

The household equation was remarkably successful in estimating trip production for these households. Of course the independent variables for the prediction of 1971 trips were known exactly at each household. This is a luxury that is not available in the operational study, but it does exhibit the faithfulness of the model for estimation even though all theoretical considerations of linear regression were not met. In particular, the nonlinear trend for the family size variable did not significantly reduce the effectiveness of the linear model to estimate future travel from the surveyed households.

Estimation of 1964 Single-Family Zonal Trips

Because the independent variables of the household model are linear in form, zonal area trips may be efficiently estimated from the following relation:

$$Y_j = na + b_1X_{1j} + b_2X_{2j} + \ldots + b_mX_{mj}$$

where

Y_j = the number of trips in zone j,
X_{kj} = the zonal total of variable k in zone j,
n_j = the number of households in zone j,
a = the regression constant, and
b_k = the regression parameter for variable k.

Table 3 gives the results of expanding the 1964 and 1971 household equations to obtain estimates of the home-based trips reported by the 4,300 single-family households. Two prominent elements of these statistics deserve attention here. First, when the household equations are expanded to obtain zonal estimates, the percentage of variation is increased from about 35 (Table 2) to 96, whereas the percentage of standard error of the mean is reduced from approximately 60 to 20. The adjusted values are similar to values observed in zonal regression analyses.

The second and most important point to be drawn from the data in Table 3 is the comparability of estimates obtained from the two data sets. The 1971 model estimated the zonal trip productions reported in 1964 with the same statistical efficiency as was possible with the 1964 household data sets. This supports the basic hypothesis of this research; i.e., analysis at the household level should provide relations that are more meaningful, and these relations should remain stable over time. In this study, the disaggregate analysis did detect a shift in the effect of automobile ownership for the families selected, but the overall relation was sufficiently accurate for estimating zonal travel at a second point in time.

GENERATION MODELS FOR ALL TYPES OF DWELLING UNITS

It is recognized that the 357 single-family units selected for the first part of this research represent a limited inference space in that they represent only a portion of the total population. The models developed for this sample can be expanded to give acceptable estimates of travel for the households from which they were selected. Would the same be true if one were to use a sample of all household and family characteristics? Further, these models have been expanded to obtain estimates of reported trips of the households from which the sample was drawn. Can these models be expanded to determine the factored trip volumes that represent the trips of the total population in the study area?

These questions are evaluated here by developing a household travel model using the entire 1964 IRTADS interview data set. In addition, the models based on all types of dwelling units are used to examine the possibility of data reduction in the continuing study and the geographical transferability of the model relations.

Estimation of IRTADS Total Urban Travel

The variables used in the home-based trip production model developed by IRTADS were total zonal population and total automobiles in the zone. The household model developed in this study used household family size and total automobiles in the household

Table 1. Average household socioeconomic and trip production characteristics.

	1964, 4,300 Single-Family Households		1964, 357 Households Sampled		1971, 357 Households Sampled	
Characteristic	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Family size	3.63	1.83	3.64	1.78	3.18	1.60
Persons age 5 and over	3.23	1.58	3.32	1.62	3.09	1.51
Labor force	1.31	0.73	1.39	0.71	1.26	0.79
Automobiles owned	1.39	0.73	1.48	0.75	1.67	0.88
Total trips	9.16	7.47	9.80	7.03	9.31	7.10
Home-based trips	6.98	5.38	7.49	5.24	7.00	5.18
Home-based work	2.00	1.49	2.12	1.34	1.99	1.51
Home-based shop	1.33	2.03	1.43	2.13	1.04	1.73
Home-based school	1.03	1.97	1.08	1.97	0.97	1.08
Home-based other	2.62	3.46	2.85	3.38	2.99	3.31
Mean income (dollars)	8,000	3,900	8,400	4,000	13,000	6,800
Median income (dollars)	7,300	–	7,200	–	14,000	–

Figure 1. Household travel rates for varying levels of family size.

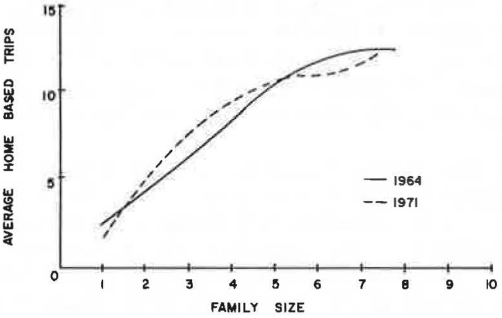


Figure 2. Household travel rates for varying levels of automobile ownership.

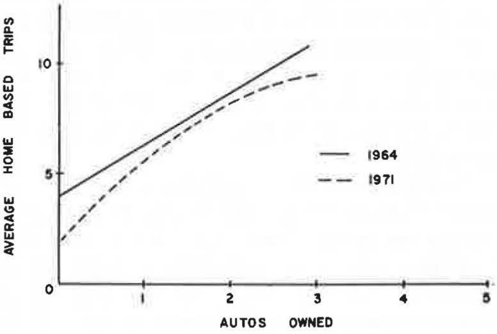


Table 2. Household prediction equations for home-based trips.

	1964[a]		1971[b]	
Variable	Regression Coefficient	Standard Error	Regression Coefficient	Standard Error
Constant	-0.45	–	-0.19	–
Family size	1.40	0.13	1.46	0.15
Automobiles	1.92	0.31	1.52	0.27

[a] $R^2 = 0.34$, standard error of estimate = 4.31, and $\bar{Y} = 7.46$.
[b] $R^2 = 0.36$, standard error of estimate = 4.20, and $\bar{Y} = 7.00$.

Table 3. Summary statistics of single-family household equations expanded to obtain zonal travel estimates.

Statistic	1964	1971
Adjusted R^2	0.96	0.96
Adjusted standard error of estimate	18.4	19.0
Mean of zonal trips[a]	95.8	95.8
Mean of residuals	-4.1	-4.1
Slope ($Y_{actual}/Y_{predicted}$)	0.98	0.98

[a] Dependent variable is 1964 zonal home-based trips, number of zones = 313, and household models based on 357 observations.

to define the equivalent relations. The relations obtained for each of the models were as follows:

1. IRTADS model (4)—Home-based trips/zone = 10.776 + 0.149 (population) + 1.257 (automobiles), and
2. Household model—Home-based trips/HH = -0.232 + 1.015 (family size) + 2.148 (automobiles).

The household equation was expanded and compared with the zonal estimates made by IRTADS. The statistics for the two models are given in Table 4. A special comment is necessary when comparing the data given in Table 4. The IRTADS model is based on data from 389 zones, whereas the household equation is expanded to represent travel from only 326 zones. As a consequence, the mean number of zonal trips is not identical for each model. The reduction in the number of zones is due to elimination in this study of all zones in which there were no reported dwelling units or labor force.

The ability of household equations to estimate zonal travel is shown in Figure 3. This is a plot of predicted home-based trips against factored zonal trip estimates provided by IRTADS. If the model predicted perfectly, all points should fit a 45-deg line passing through the origin. The actual regression line exhibited a slope coefficient of 1.00 and a constant term of -45. This constant is only 1 percent of the mean zonal trips; therefore, the model was accepted as a good fit of the data.

The residuals were examined by plotting the travel volumes against the residuals. This plot exhibited a random scatter of points. Further, Figure 4 shows a histogram of the residual distribution. This plot closely approximates the ideal normal distribution with a mean value of zero. Thus, in this study, it was found that, even though the household data did not meet all the assumptions for linear regression at the household level, residuals from the expanded equation did meet the criteria of independence and normality.

Comparison of the predictive ability of the zonal totals model and the household model indicates that the latter produces estimates with somewhat greater variation. It must be noted, however, that parameters of the IRTADS zonal equations are estimated to produce the minimum error in the zonal productions. By definition, the sum of the residuals must be zero. On the other hand, the parameters of the household model are estimated to produce minimum error at the household level. The mean of the residuals at the household level must be zero, but generally there can be no assurance that the residual sum will be zero when the model is used to estimate larger area travel. The degree to which the mean residual approaches zero provides another measure of the applicability of the expanded equation. The mean residual represents less than 1 percent of the zonal mean.

Potential for Data Reduction in Continuing Study

It was demonstrated previously that the disaggregate household model could be expanded to produce total area travel. However, because the aggregate and disaggregate models were both formulated from a data base that includes more than 10,000 home interviews, there has been no indication that the household modeling approach would save data collection expenditures. It would be necessary to conduct a full-scale analysis of sampling variability and expected errors to estimate potential savings. From this analysis, the ideal sample size necessary to obtain estimates within desired confidence limits could be determined. In this research, a single subsample was drawn to determine the order of magnitude of sample size reduction that might be possible. This subsample was equivalent to a 1 percent sampling rate, whereas the IRTADS sample was designed as a 5 percent sample. Table 5 gives the adjusted household equations given in Table 4 for the 5 percent sample and provides the comparable statistics for the 1 percent sample (2,240 cases). The ability of the two household equations to predict total travel is very similar. The standard error of the estimate is actually somewhat smaller for the smaller sample size, but, on the other hand, the mean residual is larger. Additional research is required to obtain more complete knowledge concerning the full extent of possible data reductions for trip generation as well as the other

Figure 3. Estimated and actual trip productions.

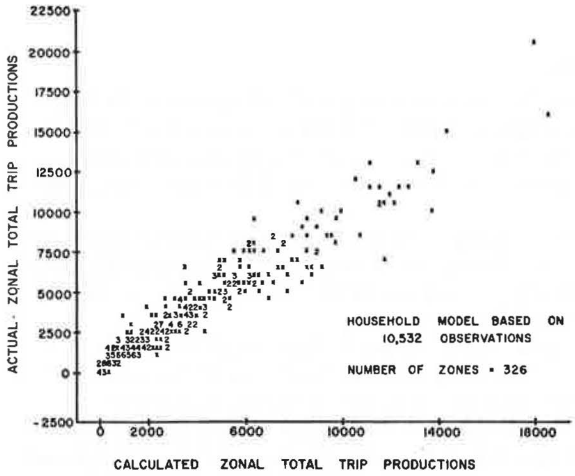


Figure 4. Frequency distribution of zonal residuals determined from expanded household equations.

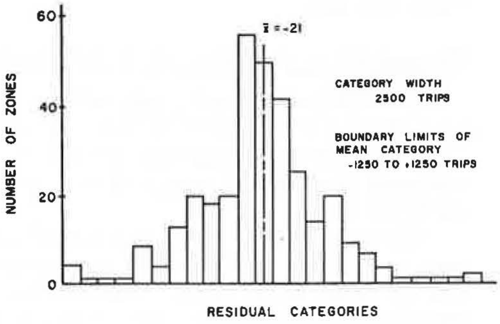


Table 4. Trip generation model statistics for estimation of total home-based trip productions.

Statistic	IRTADS Zonal Model	Household Model
R^{2a}	0.97	0.92
Standard error of estimate	597	948
Mean of zonal trips	3,287	3,947
Number of zones	389	326
Percent standard error	18.2	24.0
Mean of residuals	0	-21
Slope ($Y_{actual}/Y_{predicted}$)		1.00

[a] Adjusted statistics given for household model. Original household model based on 10,532 observations.

Table 5. Predictive ability of models.

	Sample Size	
Statistic	10,532 Households	2,240 Households
Adjusted R^2	0.92	0.92
Adjusted standard error of the estimate	948	922
Mean of zonal trips	3,947	3,947
Mean residual	-21	+56
Slope ($Y_{actual}/Y_{predicted}$)	1.00	1.02

Note: Household equations expanded to estimate home-based travel in 326 zones.

Table 6. Aggregation effects on trip generation model parameters for two urban areas.

			Indianapolis		Tri-State Area	
Level of Analysis	Dependent Variable (Y)	Independent Variable (X_i)	Number of Observations	Model Parameters	Number of Observations	Model Parameters
Household	Trips per household	X_1 = persons per household; X_2 = automobiles per household	10,532	$Y = 1.146 X_1 + 3.169 X_2 - 0.192$	5,032	$Y = 1.064 X_1 + 3.169 X_2 + 0.292$
Zone	Average trips per household per zone	X_1 = average persons per household per zone; X_2 = average automobiles per household per zone	299	$Y = 1.092 X_1 + 5.139 X_2 - 2.37$	305	$Y = 2.054 X_1 + 3.458 X_2 - 2.94$

phases of travel forecasting. Certainly though, the contention that sample size requirements may be reduced for estimation of household trip generation appears to be substantiated.

Geographical Variation in Model Parameters

The final advantage proposed for disaggregate analysis was that observed relations should be more consistent from area to area because the analysis unit is not tied to an artificial area description, and the household unit is of the same basic internal consistency in different geographical areas. A limited examination of this aspect is given in Table 6.

The parameters given in Table 6 provide a measure of the degree to which household and zonal model parameters are comparable for two study areas, i.e., Indianapolis and the tri-state area, which includes New York City. The tri-state area equations were developed in the research by Kassoff and Deutschman (6).

The magnitudes of the household model parameters for the independent variables are strikingly similar for the two study areas, even though the areas themselves would not be considered as comparable in nature. The largest variation is in the magnitude of the constant term. One might reflect that the constant term of the model is the geographic factor that explains differences in household travel in the two areas. Of course, other differences in average trip rates in the areas would be reflected by differences in the average value of the independent variables.

On the other hand, there are substantial differences in the parameters of the zonal-based models. Although this comparison is only for two study areas, the basic premise that household parameters measure a more stable, basic relation appears to be substantiated.

SUMMARY AND CONCLUSIONS

Analysis of travel behavior using the household as the basic unit provides a method of evaluating the changing relations that occur over time. A disaggregate trip generation model developed from data obtained from 357 single-family households in 1964 was able to predict the home-based trips produced by the same families in 1971 with an average error of less than 2 percent. The household models from 1964 and 1971 also exhibited the same degree of statistical efficiency when expanded to estimate total zonal trips reported in 1964 by the single-family households from which the 1971 sample was selected.

The disaggregate model for estimating total home-based travel from all dwelling units was judged to be nearly comparable with the zonal model for estimating present travel. However, because the disaggregate model is sensitive to measurement of change in the behavioral unit, the household model is preferred. Indications are that the data set may be reduced by as much as 80 percent for estimating trip generation parameters at the disaggregate analysis level. Further, because the household is the basic unit in all urban areas, analysis at the household level can help the planner understand true travel variation among geographical areas rather than apparent differences that are a function of the size of area unit selected within the study area.

Certainly, additional research is required to determine the limits of the sample size necessary for estimating travel and the degree of geographical biases that exist. Also, consideration must be given to the data requirements of other aspects of travel forecasting, i.e., trip attraction, distribution, modal split, and assignment. In the continuing study, the analyst must determine the degree to which the existing calibrated models can simulate changing travel patterns. Will the sample size that provides adequate information about changes in trip generation rates also provide sufficient data to evaluate changing attitudes and patterns of spatial distribution? Behavioral model research for the other planning models may also indicate increased efficiency. Careful planning of the survey design may provide information adequate for development of all disaggregate models. If knowledge of the complexities of travel behavior can be attained at this disaggregate level, the analysis could be conducted at this level, and then aggregation may proceed to whatever level is necessary. The important item to emphasize is that the disaggregate model approach is sensitive to changes that occur at the behavioral level and, therefore, provides a means to measure changes. This is an essential consideration as the transportation analyst considers the changing conditions that occur during the continuing planning process. After evaluation of these changes at the behavioral level, aggregation may proceed to whatever analysis unit is necessary.

REFERENCES

1. Bureau of Public Roads. Guidelines for Trip Generation Analysis. U. S. Government Printing Office, 1967.
2. Domencich, T. A., Kraft, G., and Valette, J. P. Estimation of Urban Passenger Travel Behavior: An Economic Demand Model. Highway Research Record 238, 1968.
3. Fleet, C. R., and Robertson, S. R. Trip Generation in the Transportation Planning Process. Highway Research Record 240, 1968.
4. Trip Generation Analysis. Indianapolis Regional Transportation and Development Study, 1966.
5. Kannel, E. J. A Disaggregate Analysis of Urban Travel Behavior. Purdue Univ., PhD dissertation, June 1972.
6. Kassoff, H., and Deutschman, H. Trip Generation: A Critical Appraisal. Highway Research Record 297, 1969.
7. Lansing, J. B., and Hendricks, G. Automobile Ownership and Residential Density. Institute for Social Research, Univ. of Michigan, Ann Arbor, 1967.
8. Manheim, M. Practical Implications of Some Fundamental Properties of Travel Demand Models. Presented at HRB 51st Annual Meeting, January 1972.
9. McCarthy, G. M. Multiple Regression Analysis of Household Trip Generation—A Critique. Highway Research Record 297, 1969.
10. Oi, W., and Shuldiner, P. Analysis of Urban Travel Demand. Northwestern Univ. Press, 1962.
11. Quandt, R. E., and Baumol, W. Abstract Mode Model: Theory and Measurement. Journal of Regional Science, Vol. 6, No. 2, 1966.
12. Reichman, S., and Stopher, P. Disaggregate Stochastic Models of Travel Mode Choice. Highway Research Record 369, 1971.
13. A Disaggregate Behavioral Model of Urban Travel Demand. Charles River Associates, Inc., March 1972.
14. Siegel, S. Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill, New York, 1956.
15. Stopher, P. R., and Lisco, T. Modelling Travel Demand: A Disaggregate Behavioral Approach—Issues and Applications. Transportation Research Forum, 1970.
16. Stowers, J., and Kanwit, E. The Use of Behavioral Surveys in Forecasting Transportation Requirements. Highway Research Record 106, 1966.
17. UNEQUAL. Purdue Univ., Computer Library, 1970.

CHARACTERISTICS OF OUTDOOR RECREATIONAL TRAVEL

Jerry G. Pigman and Robert C. Deen, Kentucky Department of Highways; and
John A. Deacon, University of Kentucky

The purpose of this investigation was to examine the characteristics of travel to outdoor recreational areas in Kentucky. Data were obtained by means of a license-plate, origin-destination survey at 160 sites within 42 recreational areas and by means of a continuous vehicle-counting program at eight of these sites. A computer algorithm was developed for error detection and subsequent adjustment of the volume data as necessitated by occasional malfunction of the traffic recorders and by vandalism. Vehicle occupancy was found to depend on the type of recreational area, distance traveled, and vehicle type. Occupancy increased with increasing distance and was greatest for those vehicles pulling camping trailers. Percentages of the various vehicle types were also influenced by the type of recreational area and the distance traveled. The proportion of camping units in the traffic stream increased with increasing distance of travel. In general, trip lengths were quite short as evidenced by the fact that 60 percent of all vehicles traveled less than 50 miles. However, trip-length distribution was highly dependent on type and location of the recreational area. It is highly recommended that future data collection programs be concentrated on the average summer Sunday so that the maximum amount of usable traffic data can be collected with a minimum of effort.

•IN 1970, the Kentucky Department of Highways initiated a study to examine the characteristics of travel to outdoor recreational areas in Kentucky and to develop a model for simulating these flows. Results of the modeling efforts have been reported elsewhere (3, 6). The purpose of this paper is to describe many of the characteristics of recreational travel that are of interest to highway engineers including vehicle occupancy, type of vehicle, trip-length distribution, and distribution of flows over time. Knowledge of these characteristics is necessary for the efficient design of highways and parking facilities to accommodate recreational travel.

SELECTION OF RECREATIONAL AREAS

A total of 42 recreational areas, encompassing a major part of outdoor recreational activity in Kentucky, were chosen for detailed study. These areas (Table 1) represent a variety of facilities from small fishing lakes to major scenic attractions, a broad geographic distribution within the state, and a wide variety of operating agencies.

Characteristics of the 42 areas are also summarized in Table 1. The characteristic termed "regional impact" was evaluated from two measures of travel obtained from an origin-destination (O-D) survey: the coefficient of variation of the actual number of trips produced by 190 origin zones located throughout the United States and the percentage of trips having lengths greater than 50 miles. Coefficients of variation for those areas having large, medium, and small regional impact averaged 280, 480, and 720 percent respectively. Corresponding average percentages of trips having lengths greater than 50 miles were 66.7, 35.7, and 23.7 respectively.

DATA ACQUISITION

Two surveys were undertaken to provide data for characterizing outdoor recreational travel. One, a traffic volume survey, provided data concerning the fluctuations of traffic volumes over time. The other, an O-D survey, provided information on vehicle occupancies, types of vehicles, trip lengths, and so forth.

Traffic volume data were obtained from continuous automatic traffic recorders located at eight sites considered to be most representative of Kentucky outdoor recreational areas. The punched-tape counters, employing inductive loops for vehicle detection, recorded two-way flows continuously from July 1970 through June 1971. In each case, the recorder was located on a major access road to the recreational area in such a manner as to intercept only recreation-oriented travel. A total of 3,039,403 vehicles were counted at the eight sites during the 1-year survey. This represented an average of about 380,000 vehicles annually per site.

The license-plate O-D survey was conducted at 160 sites, similarly located to intercept only recreation-oriented travel, during the summer of 1970. Each of these sites was associated with 1 of the 42 recreational areas. The sites were carefully selected so that the sum of the flows passing all the sites associated with a given recreational area accurately represented the total flow to that area. The O-D survey at each site was conducted during a 10-hour period of normal peak flow, namely 10 a.m. to 8 p.m. on summer Sundays. No data were collected during holiday weekends. Recorded for each observed vehicle were the direction of movement (arriving or departing), type of vehicle, vehicle occupancy, and license-plate identification. The license-plate identification was used to approximate the zone of origin of the vehicle.

A total of 130,653 vehicles were observed as a part of the O-D survey. Considering those small intervals during each 10-hour period when the surveyors were otherwise occupied, it was estimated that a total of 147,000 vehicles actually passed the 160 sites during the survey period. A further adjustment was made to account for the few instances in which inclement weather prevailed, bringing the total estimated flow to 151,300 vehicles.

TRIP ORIGINS

Of the vehicles observed in the O-D survey, approximately 73.0 percent were licensed in Kentucky. This percentage was sensitive to the type of recreational area, however, and varied from a low of 36.6 percent at the two national parks (Mammoth Cave and Cumberland Gap) to a high of 85.2 percent at facilities administered by the Corps of Engineers, which are predominantly day-use oriented. Table 2 gives the percentages of vehicles from different origins as a function of type of facility. The origins are arranged in Table 2 in approximate order of increasing distance from Kentucky. The effect of geographic proximity is most pronounced. It was also found that about 96.3 percent of all vehicles came from Kentucky and seven nearby states including, in order of highest to lowest visitation, Ohio, Indiana, Illinois, Tennessee, Michigan, Missouri, and West Virginia.

VEHICLE OCCUPANCY

The O-D survey provided information with which to evaluate average vehicle occupancy, that is, the average number of persons in each vehicle. The average occupancy rate for all vehicles was found to be 3.06 persons per vehicle. However, occupancy rate was a function of the type of recreational area, distance traveled, and type of vehicle.

Table 3 gives the effect of type of recreational area on average vehicle occupancy. Lowest occupancy rates of 2.87 to 2.88 persons per vehicle occurred at predominantly day-use, water-oriented facilities. Intermediate rates of 3.13 to 3.26 persons per vehicle occurred at multiple-use facilities, and the highest rates of 3.36 to 3.41 persons per vehicle occurred at scenic areas catering to families and having nationwide interest.

Table 3 also indicates that location of origin affects vehicle occupancy. The average occupancy rate for Kentucky vehicles was 2.94 persons per vehicle and that for the seven primary states outside of Kentucky was 3.41 persons per vehicle. This suggests that occupancy rates may be related to distance traveled, a hypothesis that seems plausible considering that many out-of-state vehicles carry vacationing families.

Table 4 gives the effects of both distance and type of vehicle on occupancy rate.

Table 1. Recreational areas.

Number	Area Name	Regional Impact	Scenic Attractiveness[a]	Lake[b]	Day-Use Facilities[c]	Overnight Accommodations[d]	Other[e]
1	Columbus-Belmont S.P.	S	N	N	M	M	
2[f]	Kentucky Lake-Barkley Lake	L	H	L	L	L	G, OD, SP, SB
3	Lake Beshear-Pennyrile Forest	S	N	S	L	L	G, SP, SB
4	Audubon S.P.	S	N	S	L	M	G, SB
5	Lake Malone S.P.	S	N	L	L	M	SB
6[f]	Rough River Reservoir	S	N	L	L	L	G, SP, SB
7	Doe Valley Lake	S	N	S	S	S	SB
8	Otter Creek Park	S	N	N	L	L	SP
9	Nolin Reservoir	S	N	L	M	M	
10[f]	Mammoth Cave N.P.	L	P	N	M	L	
11	Shanty Hollow Lake	S	N	S	S	S	
12	Barren River Reservoir	S	N	L	L	L	G, SB
13	My Old Kentucky Home S.P.	M	P	N	L	M	G, OD
14	Green River Reservoir	S	N	L	M	S	
15	Dale Hollow Reservoir	M	N	L	M	M	SB
16	Lake Cumberland	M	N	L	L	L	G, SP, SB
17	Natural Arch and Rockcastle areas	M	P	N	M	M	
18	Cumberland Falls S.P.	L	P	N	M	L	SP
19	Wilgreen Lake	S	N	S	S	S	
20	Herrington Lake	M	N	L	S	L	
21	Old Fort Harrod S.P.	L	P	N	M	S	OD
22[f]	Beaver Lake	S	N	S	S	S	
23	Guist Creek Lake	S	N	S	S	S	
24	General Butler S.P.	M	N	S	L	L	G, SP, SB
25	Elmer Davis Lake	S	N	S	S	S	
26	Lake Boltz	S	N	S	S	S	
27	Big Bone Lick S.P.	S	N	S	M	L	
28	Williamstown Lake	S	N	S	S	S	
29	Blue Licks Battlefield S.P.	M	H	N	M	S	SP
30[f]	Fort Boonesboro S.P.	M	H	N	M	M	SB
31[f]	Levi Jackson S.P.	S	N	N	L	L	SP
32	Pine Mountain S.P.	S	N	S	L	M	G, OD, SP
33	Cumberland Gap N.P.	L	P	N	L	L	OD
34	Natural Bridge S.P.	L	P	S	M	L	SP
35	Sky Bridge and Koomer Ridge	L	P	N	M	M	
36[f]	Carter Caves S.P.	M	H	S	L	S	G, SB
37	Greenbo Lakes S.P.	S	N	S	L	L	SB
38	Grayson Reservoir	S	N	L	M	S	
39	Buckhorn Lake	S	N	L	M	L	SB
40[f]	Jenny Wiley S.P.	S	N	L	L	L	G, OD, SP
41	Kingdom Come S.P.	S	N	S	M	S	
42	Fishtrap Reservoir	S	N	L	M	S	

[a] P = primary attractiveness of a scenic or historic nature, H = high scenic or historic attractiveness with a balance of other recreational activities, and N = normal scenic or historic attractiveness.

[b] L = lake acreage ⩾ 500, S = lake acreage ⩽ 500, and N = no lake.

[c] L = availability of golf course and/or picnic tables > 150, M = picnic tables ⩽ 150 and no golf course, and S = no picnic tables and no golf course.

[d] L = units (cottages, lodge rooms, and camping sites) ⩾ 90, M = between 15 and 90 units; and S = units < 15.

[e] G = golf, OD = outdoor drama, SP = swimming pool, and SB = swimming beach.

[f] Area at which continuous traffic recorders were operated on major access roads.

Table 2. Percentage of vehicles by origin for different recreational areas.

Origin[a]	National Parks	Land-Between-the-Lakes (TVA)	Daniel Boone National Forest	State Parks	Kentucky Lake (TVA)	Other Areas	Corps of Engineers Facilities	All Areas
Kentucky	36.57	54.05	67.95	70.67	72.08	81.44	85.21	73.02
East North Central states	37.80	23.82	25.10	20.02	18.61	11.68	11.88	18.38
East South Central states[b]	8.14	15.72	1.16	2.69	3.20	1.08	0.82	2.76
South Atlantic states	7.75	1.39	1.28	3.11	1.34	2.17	0.98	2.55
Middle Atlantic states	4.06	0.61	0.78	0.63	0.52	1.11	0.17	0.64
West North Central states	1.76	2.58	0.52	1.74	3.37	0.84	0.42	1.49
West South Central states	1.70	1.48	0.65	0.51	0.63	0.41	0.26	0.53
New England states	0.63	0.09	0.39	0.11	0.03	0.42	0.06	0.12
Mountain states	0.25	0.05	0.26	0.12	0.13	0.29	0.04	0.13
Pacific states	0.77	0.21	0.13	0.29	0.07	0.48	0.08	0.25
Canada	0.52	–	0.25	0.10	0.02	0.08	0.01	0.09
Other	0.05	0.03	1.53	0.01	–	–	0.07	0.04

[a] U.S. Bureau of the Census Divisions. [b] Excluding Kentucky.

Despite large variability in the data, occupancy rate generally increased with increasing distance of travel. The effects were most pronounced for vehicles traveling rather short distances. In addition, sensitivity of occupancy rate to distance was greatest for camping vehicles and least for vehicles with boats.

Highest occupancy rates were observed for cars pulling camper trailers, and lowest rates were observed for the "other" vehicle category, which includes primarily service trucks and motorcycles. The fact that single-unit campers had much lower occupancy rates than cars pulling camper trailers is probably due to erroneous surveys in which some persons riding in the single-unit campers could not be detected by the surveyors and a certain bias caused by rather extensive use of pickup campers by fishermen who usually travel in small groups.

Considerable variation is found in occupancy rates reported by others. To illustrate, an average occupancy rate of 3.2 persons per vehicle has been reported for weekend recreational travel in Kansas (4). Occupancy rates for recreational travel in Arkansas averaged 3.3 persons per vehicle for Arkansas residents and 3.2 persons per vehicle for out-of-state residents (1). Analysis of weekend travel to 10 Kansas reservoirs yielded average occupancy rates ranging from 3.3 to 4.2 persons per vehicle (7). The Kansas data also showed that occupancy rate was affected by trip purpose. Finally, an average occupancy rate of 3.7 persons per vehicle was observed at three parks in Indiana (5). The preceding data together with those reported here substantiate the observation that average occupancy rate for outdoor recreational travel is considerably larger than for other highway travel.

TYPES OF VEHICLES

As anticipated, a large proportion of the vehicles were cars (pickups included) and cars pulling trailers (a total of 96.7 percent). The remainder were single-unit campers (2.1 percent) and motorcycles, trucks, and buses (1.2 percent). Altogether, 3.4 percent of the vehicles had camping units attached and 5.8 percent had boats. Vehicle classification was found to depend on both trip origin and type of recreational area.

The effect of origin can be shown as follows: 2.1 percent of the Kentucky vehicles had camping units and 6.0 percent had boats; respective percentages for Michigan vehicles were 10.4 and 3.9. These and similar data are summarized for the eight primary states contributing to Kentucky recreational travel in Table 5. Origin effects are due in large part to intervening distances (Fig. 1). Decreasing percentage of cars with increasing distance reflected the increasingly greater use of single-unit campers over the longer distances. As distance increased, a greater percentage of recreationists used camping vehicles. Boat use peaked in the distance range of 60 to 90 miles.

The effects of type of recreational facility on vehicle use are quite clear. A high percentage of vehicles with boats was observed at water-based facilities (a high of 12.3 percent at Corps of Engineers facilities compared to a low of 0.6 percent at national parks). The percentage of vehicles with camping units depended in large part on the nature of available camping facilities (a high of 11.2 percent at Land-Between-the-Lakes compared to a low of 3.0 percent at state parks). Table 6 gives these data.

TRIP-LENGTH DISTRIBUTION

Examination of trip origins (Table 2) revealed that most recreationists came from Kentucky. This suggested that most trips to Kentucky outdoor recreational facilities were short-distance trips. The average trip length for all vehicles was found to be 109 miles. However, 60 percent of all vehicles traveled distances less than 50 miles, and 72 percent traveled less than 100 miles. Ungar (8) also showed that outdoor recreational travel is predominantly of the short-distance type. He reported that 50 percent of the recreationists in Indiana traveled distances less than 50 miles and in Kansas less than 40 miles. The corresponding distance for travel in Kentucky was found to be 38 miles.

Trip lengths were found to be a function of type and location of the recreational area. Figure 2 shows trip-length distribution for three state parks representative of large

Table 3. Effect of type of recreation and location of origin on average vehicle occupancy.

Origin	State Parks	National Parks	Corps of Engineers Facilities	Kentucky Lake (TVA)	Land-Between-the-Lakes (TVA)	Daniel Boone National Forest	Other Areas	Total
Kentucky	3.02	3.22	2.84	2.70	3.18	3.44	2.82	2.94
Ohio	3.47	3.37	3.11	3.69	3.61	3.33	3.00	3.37
Indiana	3.34	3.56	3.08	3.23	3.35	3.63	3.16	3.31
Illinois	3.68	3.57	3.43	3.39	3.54		3.38	3.57
Tennessee	3.40	3.29	3.13	3.39	3.23	3.43	3.82	3.32
Michigan	3.50	3.94	3.16	2.97	3.10	4.14	3.31	3.52
Missouri	3.61	3.44	3.14	3.03	3.32	6.00	2.33	3.40
West Virginia	3.60	3.40	3.30	2.86	2.00	6.00	2.40	3.61
All origins	3.13	3.36	2.88	2.87	3.26	3.41	2.87	3.06

Note: Figures in persons per vehicle.

Table 4. Effect of distance and type of vehicle on average vehicle occupancy.

Type of Vehicle	Persons per Vehicle by Distance Interval in Miles											
	1 to 20	21 to 40	41 to 60	61 to 80	81 to 100	101 to 150	151 to 250	251 to 400	401 to 700	701 to 1,300	1,301 to 3,000	Average (all distances)
Car	2.78	3.02	3.28	3.27	3.31	3.29	3.20	3.45	3.39	3.25	3.11	3.07
Car with boat and trailer	3.02	3.14	3.12	3.25	3.13	3.15	3.45	3.19	3.16	3.18	3.60	3.16
Car with boat on top	2.72	3.14	3.05	2.79	3.00	3.09	3.92	3.31	3.00	2.50		3.04
Car with camper trailer	3.06	3.20	3.28	3.45	3.44	3.61	3.63	3.86	4.06	3.60	3.82	3.63
Single-unit camper	2.70	2.55	2.83	3.11	3.06	3.00	2.92	2.99	3.39	3.48	3.36	2.97
Single-unit camper with boat	2.75	2.79	2.71	2.71	2.70	3.27	2.65	3.38	2.94	3.30	4.25	2.96
Other	2.16	1.61	1.92	2.19	5.30	1.63	1.69	4.78	1.57	1.75	20.50	2.67
Average (all vehicles)	2.78	3.02	3.26	3.25	3.30	3.28	3.21	3.45	3.41	3.26	3.28	3.06

Table 5. Effect of location of origin on percentage of type of vehicle.

Origin	Car	Car With Boat and Trailer	Car With Boat on Top	Car With Camper Trailer	Single-Unit Camper	Single-Unit Camper With Boat	Other
Kentucky	90.89	5.27	0.40	0.61	1.08	0.37	1.38
Ohio	86.46	5.34	0.63	3.35	2.62	0.62	0.97
Indiana	87.57	4.51	0.62	2.38	3.15	0.87	0.90
Illinois	88.11	3.36	0.88	3.20	2.72	0.86	0.88
Tennessee	90.99	3.44	0.32	1.59	1.62	1.05	0.99
Michigan	85.74	2.28	0.70	6.08	3.33	0.94	0.94
Missouri	88.67	4.03	0.77	2.82	2.63	0.51	0.58
West Virginia	88.51	2.31	0.79	5.61	1.45	0.46	0.86
All origins	89.95	4.91	0.46	1.36	1.58	0.48	1.26

Figure 1. Effect of distance on percentage of type of vehicle.

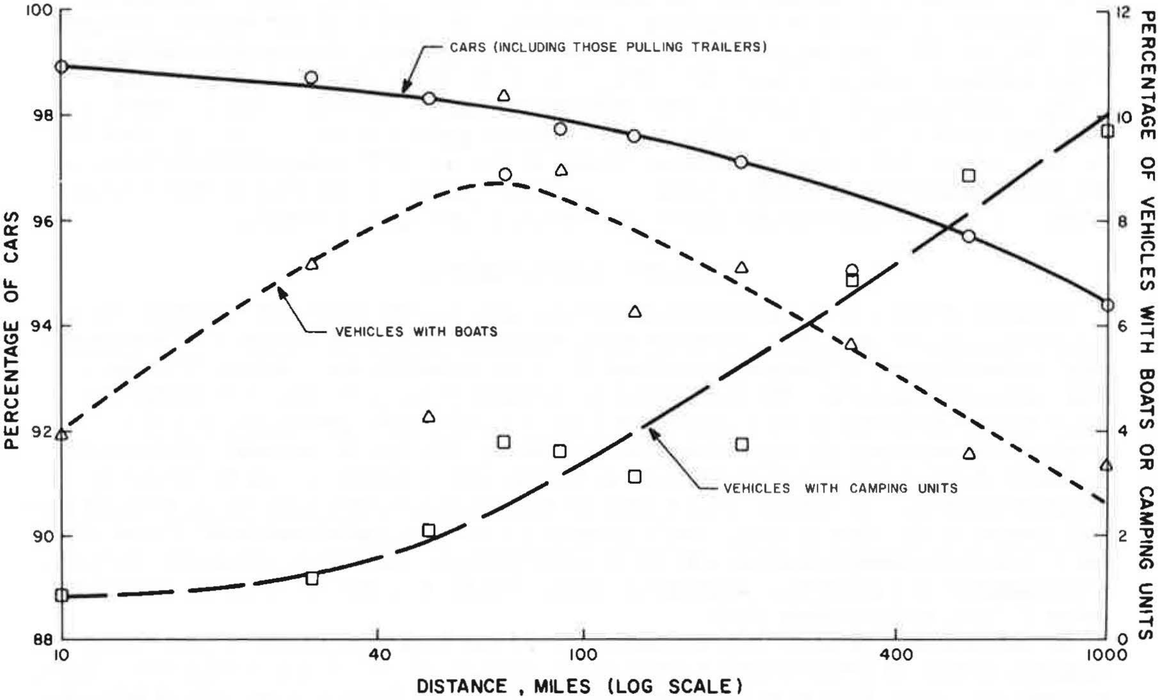


Table 6. Effect of type of recreational area on percentage of type of vehicle.

Type of Facility	Percentage of Cars[a]	Percentage of Camping Vehicles[b]	Percentage of Vehicles With Boats
State parks	97.36	2.95	3.22
National parks	95.56	6.51	0.58
Corps of Engineers facilities	95.71	3.29	12.31
Kentucky Lake (TVA)	96.31	3.81	6.14
Land-Between-the-Lakes (TVA)	90.84	11.24	12.02
Daniel Boone National Forest	96.22	2.99	3.25
Other areas	97.84	2.59	7.15
All areas	96.67	3.42	5.84

[a] Includes cars with boat and camper trailers.

[b] Includes cars with camper trailers and single-unit campers.

regional impact areas (Cumberland Falls), medium regional impact areas (My Old Kentucky Home), and small regional impact areas (Jenny Wiley). Mean trip lengths for the areas classified as having large, medium, and small regional impact averaged 176, 89, and 70 miles respectively (Table 1). Corresponding average percentages of trips having lengths less than 50 miles were 33.3, 64.3, and 76.3 respectively.

Also of considerable interest is the influence of type of vehicle on the distribution of trip lengths (Fig. 3). Cars pulling camper trailers generally traveled the greatest distances. Single-unit campers traveled somewhat shorter distances partially because of the considerable use of single-unit campers by fishermen. Cars without either boats or trailers generally traveled the shortest distances of any type of vehicle.

VOLUME ADJUSTMENTS

Because of vandalism and equipment malfunction, a limited amount of traffic volume data from each of the eight continuous recorders was found to be in error. This necessitated development of computer routines for error detection and subsequent adjustment of erroneous data. These routines were based on the premises that hourly volumes at a given location for a particular hour of the day and a particular day of the week should demonstrate a great deal of consistency throughout the year and that such volumes should reach a minimum in the winter months and a maximum in the spring or summer months. All hourly volume data for a given site were therefore rearranged into 168 groups of 52 volumes each. Each group represented a particular hour of a particular day and was analyzed independently of other groups. Each of the 52 hourly volumes corresponded to a particular week of the year. Figure 4 shows one such group of data taken in Levi Jackson State Park.

Error detection proceeded as follows. Let V_i represent the hourly volume corresponding to the ith week and AV represent the average of the 52 hourly volumes. First, grossly inaccurate data were identified when either of the following two sets of inequalities was satisfied:

$$V_i < 0.05\ AV \text{ and } |V_i - AV| > 80 \tag{1}$$

or

$$V_i > 6.0\ AV \text{ and } |V_i - AV| > 80 \tag{2}$$

Erroneous data so identified were automatically removed from the data set, and seven-item moving averages (MAV_i) were calculated. The second comparisons to detect erroneous data were based on the following two sets of inequalities that compared each hourly volume with the corresponding moving average:

$$V_i < 0.2\ MAV_i \text{ and } |V_i - MAV_i| > 20 \tag{3}$$

or

$$V_i > 2.0\ MAV_i \text{ and } |V_i - MAV_i| > 20 \tag{4}$$

Figure 4 shows, for the group of data at Levi Jackson State Park, four erroneous volumes detected in this way.

Having identified the set of "correct" data, it was necessary to provide more reasonable estimates of the "incorrect" data. This was accomplished by fitting a third-degree polynomial to the correct data and obtaining the desired estimates by interpolation. Figure 4 also shows such a polynomial, which was used to make the required four estimates for this group of data.

The aforementioned procedure for error detection and correction was found to be invaluable to this study even though there was some risk that all erroneous data were not detected and some smaller risk that some correct data were classified as being erroneous. Altogether, 8 percent of the hourly volumes were found to be in error. This

Figure 2. Trip-length distributions for different recreational areas.

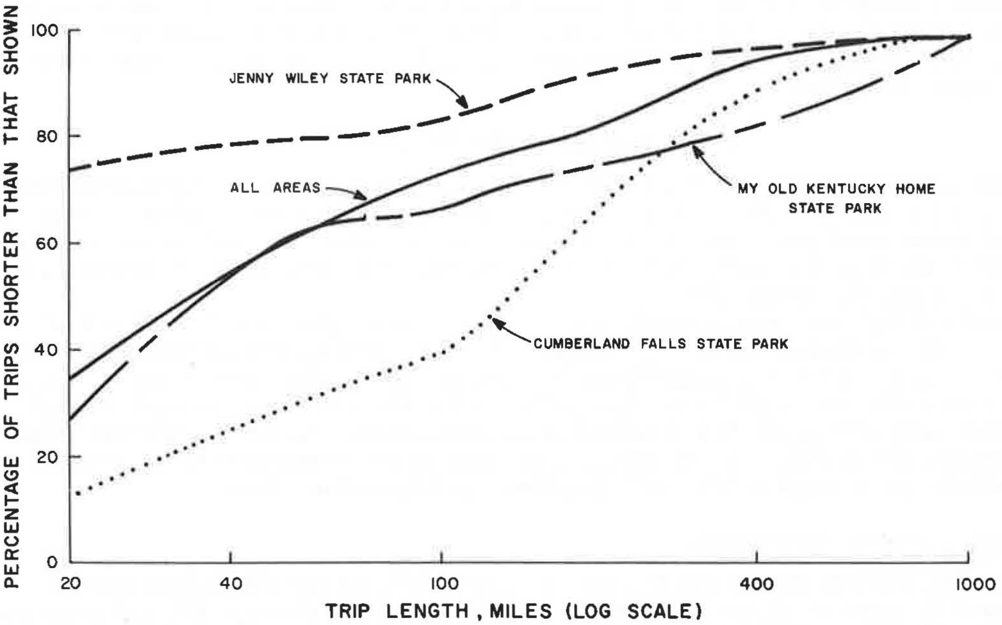
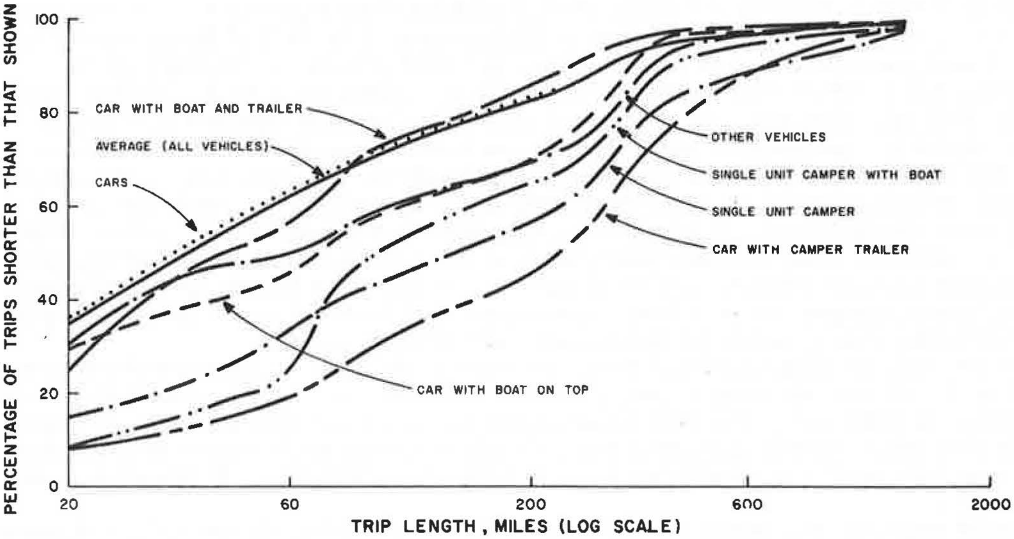


Figure 3. Trip-length distributions for different types of vehicles.



includes data from two locations at which the recorders were known to be inoperative for a cumulative total at each of approximately 2 months. Identical procedures for error detection and correction can be used for other types of hourly volume data collected on an annual basis if suitable modifications are made to the limiting constants in the preceding inequalities.

TIME DISTRIBUTION OF FLOWS

The distribution of recreational traffic volumes over time can be examined in various ways. Data from this study are presented in the following sections to show average and certain highest volumes for different time periods, demonstrate cyclic patterns throughout the year, and allow short-term counts of recreational traffic to be expanded to estimates of certain average flows.

In analyzing these data, a weekend was defined to encompass the 48-hour period from 6 p.m. on Friday to 6 p.m. on Sunday. Seasons were specifically defined as follows: summer (June 20 through September 19), fall (September 20 through December 19), winter (December 20 through March 19), and spring (March 20 through June 19). Average daily traffic (ADT) was defined in the conventional manner as the total annual volume divided by 365. Various summer averages were computed in such a manner as to exclude the summer holidays of Labor Day and Independence Day.

Average and Highest Volumes

Hourly volumes, expressed as a percentage of ADT, for the 200 highest volume hours of the year are shown in Figure 5. Three curves are shown in this and subsequent figures. The upper curve represents the maximum volumes at any of the eight sites, the middle curve represents the eight-site average volumes, and the lower curve represents the minimum volumes at any of the sites. The maximum hourly volume as a percentage of ADT varied from a high of 121.2 percent at Boonesboro to a low of 37.2 percent at Mammoth Cave and averaged 63.2 percent at the eight sites. The 30th highest hourly volumes ranged from a high of 82.9 percent of the ADT at Boonesboro to a low of 24.0 percent at Beaver Lake and averaged 38.8 percent at the eight sites. [Maring has reported 30th highest hourly volumes ranging from 14.5 to 22.3 percent of the ADT (4). However, the locations at which his data were obtained intercepted some travel not specifically destined to outdoor recreational areas.] As anticipated, the 30th highest hourly percentages were considerably greater than those commonly observed for normal urban or rural travel indicating the peaking commonly associated with recreational travel. The highest peaking was observed at Fort Boonesboro State Park, a predominantly day-use facility attracting significant numbers of visitors only during the summer months. Lowest peaking was observed at Mammoth Cave National Park, a scenic attraction of national importance, and Beaver Lake, a small fishing lake attracting fishermen during the spring, summer, and fall months.

In general, the highest volume hours occurred on Sundays. Approximately 83 percent of the 100 highest volume hours occurred on Sundays and approximately 10 percent occurred on Saturdays. The only major exception among the eight sites was at Mammoth Cave where only 38 percent of the 100 highest volume hours occurred on Sundays; the remainder was approximately equally divided among Tuesdays, Wednesdays, and Saturdays.

Daily volumes, expressed as a multiple of ADT, for the 100 highest volume days are shown in Figure 6. The maximum daily volume ranged from a high of 889 percent of the ADT at Boonesboro to a low of 332 percent at Beaver Lake. Matthias and Grecco (5) have reported maximum daily volumes at three parks in Indiana averaging approximately 1,350 percent of the ADT. These data clearly demonstrate the significant daily peaking associated with recreational traffic. The high-volume days shown in Figure 6 were typically associated with summer Sundays. Average summer Sunday volumes ranged from a high of 412 percent of the ADT at Boonesboro to a low of 156 percent of the ADT at Beaver Lake. In general, the average summer Sunday volumes corresponded with the volume associated with the 11th or 12th highest volume day.

Finally, Figure 7 shows the weekly volumes arranged in order of magnitude for the

Figure 4. Fluctuation of hourly volumes throughout the year (12 noon to 1 p.m.).

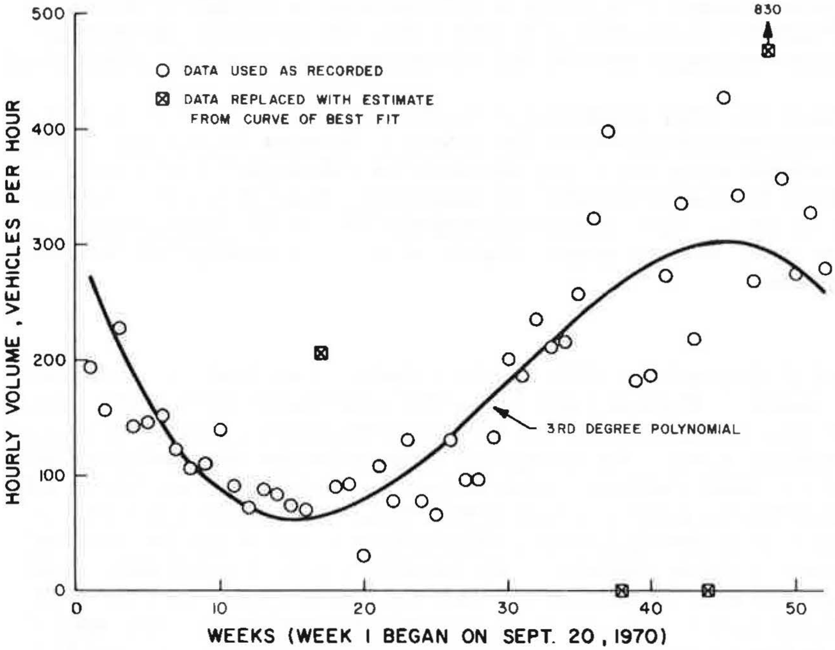
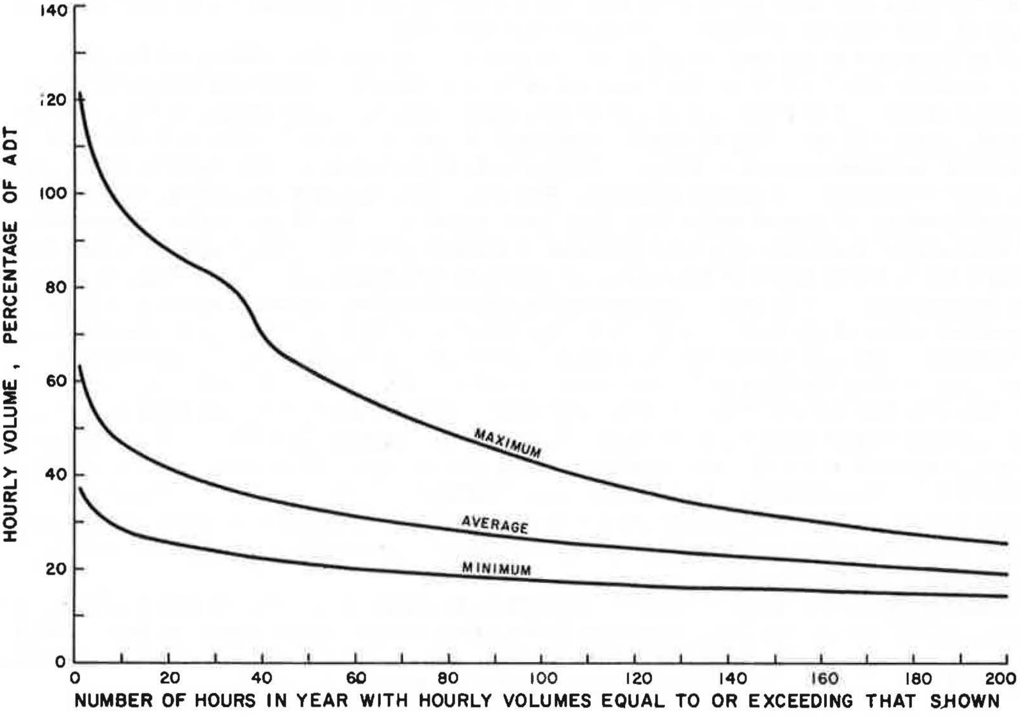


Figure 5. Highest hourly volumes.



52 weeks. A very wide range in weekly volumes is shown by this figure. The average summer weekly volumes ranged from a high of 1,300 percent of the ADT at Mammoth Cave to a low of 800 percent of the ADT at Beaver Lake. On the whole, the average summer weekly volume corresponded with that volume associated with the 10th highest volume week.

A summary of these and other pertinent volume data is given in Table 7. In view of the extreme peaking associated with recreational travel, it seems impractical to design highways serving recreational areas to accommodate the 30th highest hourly volumes. A more practical basis for design would be the peak-hour volume on the average summer Sunday, which on the average corresponds with the 70th to 75th highest hourly volume. Concentration on the average summer Sunday should also greatly facilitate future data collection programs.

Cyclic Patterns

The cyclic nature of recreational travel is also a matter of interest to recreational and highway planners alike. Figures 8 and 9 show the patterns of variation in volumes among seasons and months respectively, and Table 8 summarizes the peak volumes for the individual recreational areas. As anticipated, seasonal peaks occurred in either the spring or summer. Peak seasonal volume ranged from a low of 36 percent of the total annual volume at Beaver Lake to a high of 46 percent at Mammoth Cave and averaged about 40.6 percent at the eight areas. Others have reported similar seasonal peaking. For example, summer visitation, expressed as a percentage of annual visitation, has been reported to be 40 percent at Tennessee and Kentucky reservoirs (2), 62.1 percent at Indiana and Ohio reservoirs (2), and 45.2 percent for several types of recreational areas in Arkansas (1). The differences between the Tennessee and Kentucky reservoir data and the Indiana and Ohio reservoir data may be due in part to climatic influences that, for travel to reservoirs, cause more peaking during the summer months in the colder areas. Data from Arkansas (1) also showed an influence of type of facility with seasonal peaks, varying from a low of 36.3 percent at national parks to a high of 48.8 percent at Corps of Engineers reservoirs.

Monthly peaks at the individual areas occurred in either May, June, or August. May peaking at Beaver Lake and Lake Barkley is probably attributable to large spring fishing activity. The peak monthly volume, expressed as a percentage of total annual volume, ranged from a low of about 15 percent at Beaver Lake to a high of about 24 percent at Boonesboro and averaged 17.6 percent at the eight areas. Others have likewise reported similar monthly peaking. For example, monthly visitation, expressed as a percentage of annual visitation, has been reported to be 14 percent at Tennessee and Kentucky reservoirs (2), 24.1 percent at Indiana and Ohio reservoirs (2), and 16.7 percent for several types of recreational areas in Arkansas (1). The Arkansas study also demonstrated an influence of type of facility with a low monthly peaking of 12.9 percent at national parks to a high of 19.0 percent at Corps of Engineers reservoirs.

Summer daily and hourly cyclic patterns were also investigated. Peak summer flows occurred on Sundays at all recreational areas (Fig. 10). The next highest volume day was Saturday with very little differences among the remaining days of the week. This was somewhat surprising in that it had been anticipated that Friday flows would generally exceed those of other weekdays. On the average, 25 percent of the travel during the typical summer week occurred on Sundays. Smith and Landman also observed notable Sunday peaking in travel to reservoirs in Kansas and, with one exception, reported summer Sunday flows that ranged between 26.5 and 39.0 percent of the corresponding weekly flows (7).

Peaking within the days of summer weekends is shown in Figure 11 and Table 9. The hour of peak flow was typically later on Friday than it was on Saturday; in turn, Saturday peaks occurred later in the day than Sunday peaks. At the same time, Sunday flows were typically more peaked than either Saturday or Friday flows.

Expansion Factors for Short-Term Counts

It is frequently desirable to estimate average traffic volumes based on short-term volume surveys. Table 10 gives a set of factors by which short-term counts taken

Figure 6. Highest daily volumes.

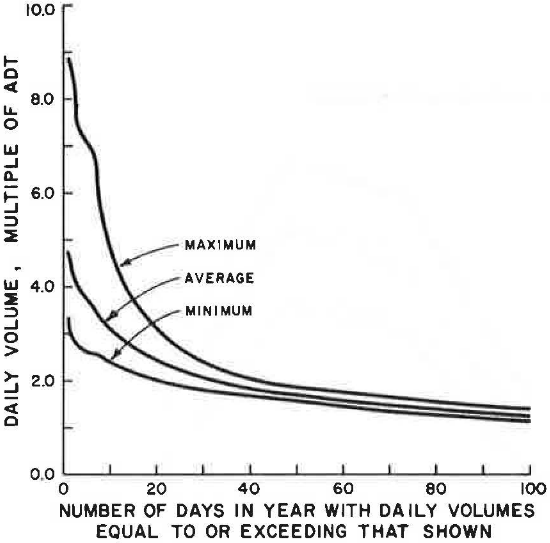


Figure 7. Highest weekly volumes.

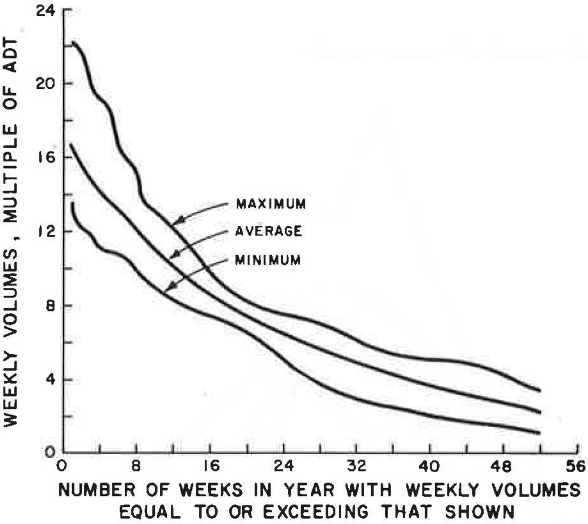


Table 7. Average and highest volumes.

Time Period	Type of Volume	Area 2	Area 6	Area 10	Area 22	Area 30	Area 31	Area 36	Area 40	Average
Week	Maximum	14.3	19.1	17.6	13.4	22.2	15.6	14.7	15.6	16.6
Week	4th highest	13.0	17.1	14.5	10.9	19.0	12.2	13.4	12.8	14.1
Week	8th highest	11.3	14.0	13.1	10.5	15.2	11.1	12.1	9.8	12.1
Week	Summer average	8.3	12.3	13.0	8.0	11.7	10.9	11.4	10.8	10.8
Weekend	Maximum	6.62	11.71	6.26	6.15	14.77	7.10	9.56	6.82	8.62
Weekend	4th highest	5.54	9.89	5.64	5.31	11.28	5.52	6.43	5.75	6.92
Weekend	8th highest	4.66	7.64	4.36	4.51	8.04	5.07	5.70	4.61	5.57
Weekend	Summer average	3.53	6.28	4.43	3.30	6.14	4.94	5.24	4.81	4.83
Weekend	Annual average	2.99	3.49	2.64	3.00	3.93	3.25	3.34	3.06	3.21
Day	Maximum	3.72	6.61	3.50	3.32	8.89	3.47	5.00	3.68	4.77
Day	5th highest	2.91	5.50	2.73	2.60	7.06	2.93	3.69	2.84	3.78
Day	10th highest	2.39	4.08	2.58	2.42	5.03	2.58	3.02	2.47	3.07
Day	20th highest	2.13	2.92	2.34	2.00	3.08	2.11	2.46	2.15	2.40
Day	Summer Sunday average	2.18	3.66	2.18	1.56	4.12	2.54	2.98	2.53	2.72
Hour	Maximum	0.430	0.839	0.372	0.496	1.212	0.612	0.674	0.425	0.632
Hour	15th highest	0.350	0.602	0.264	0.273	0.912	0.310	0.428	0.353	0.436
Hour	30th highest	0.303	0.503	0.254	0.240	0.829	0.284	0.395	0.294	0.388
Hour	50th highest	0.279	0.410	0.242	0.213	0.627	0.262	0.367	0.253	0.332
Hour	100th highest	0.243	0.321	0.222	0.176	0.404	0.217	0.292	0.214	0.261

Note: Volume expressed as multiple of average daily traffic.

Figure 8. Volume variation among seasons.

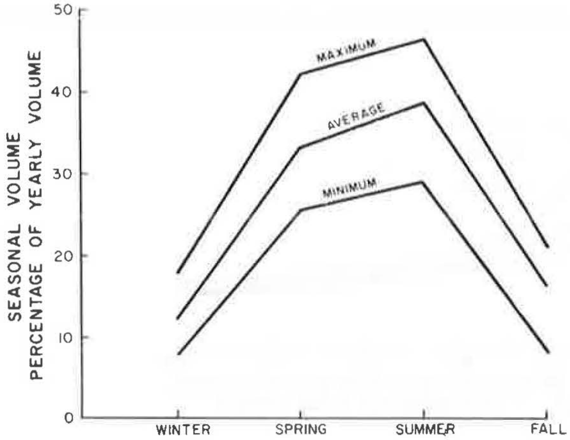


Figure 9. Volume variation among months.

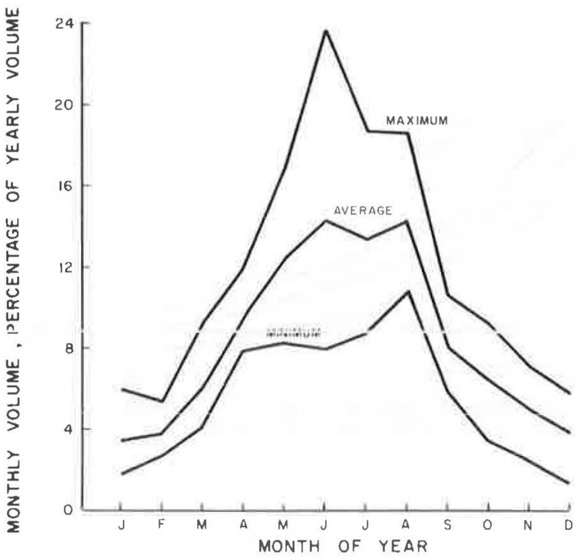


Table 8. Peak volumes.

Area	Period of Peak Volume			Peak Volume		
	Day of Summer Week	Month of Year	Season of Year	Day (percentage of average summer week)	Month (percentage of annual volume)	Season (percentage of annual volume)
2	Sunday	May	Spring	26.46	16.92	38.55
6	Sunday	June	Summer	29.83	19.44	43.84
10	Sunday	August	Summer	16.82	18.51	46.41
22	Sunday	May	Spring	19.41	14.67	36.00
30	Sunday	June	Spring	35.26	23.66	42.16
31	Sunday	August	Summer	23.19	15.64	39.06
36	Sunday	August	Summer	26.30	15.39	40.54
40	Sunday	August	Summer	23.50	16.50	38.52
Average				25.09	17.59	40.64

Figure 10. Volume variation among days throughout average summer week.

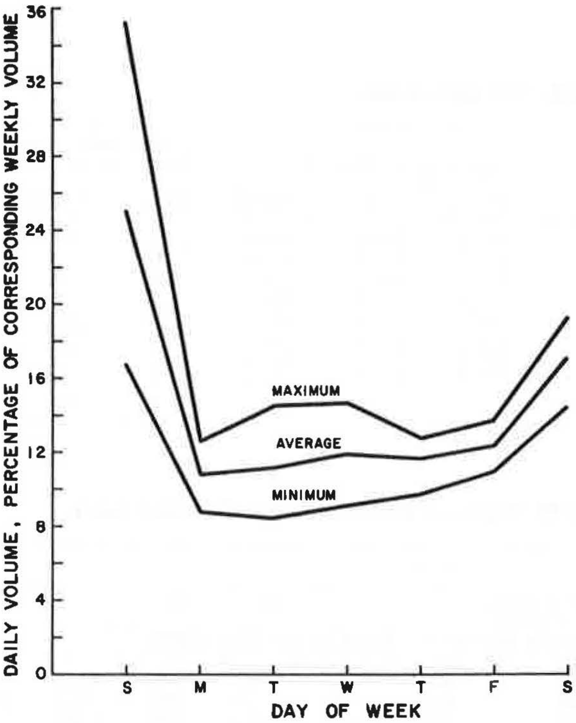


Figure 11. Volume variation throughout average summer weekend days.

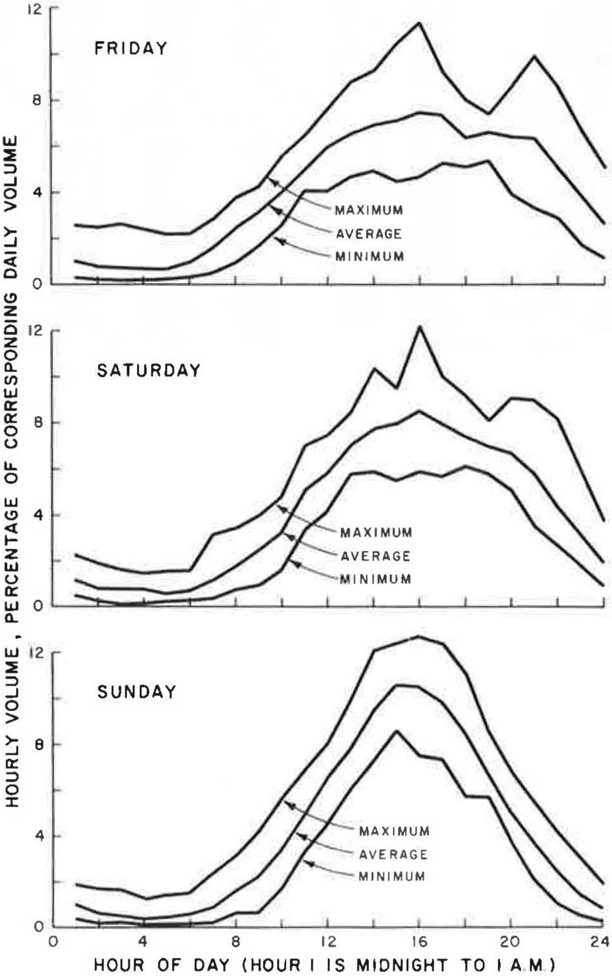


Table 9. Peak-hour volumes.

	Hour of Peak Volume (p.m.)			Peak Hourly Volume (percentage of daily volume)		
Area	Friday	Saturday	Sunday	Friday	Saturday	Sunday
2	3 to 4	3 to 4	3 to 4	11.41	12.17	12.68
6	8 to 9	2 to 3	3 to 4	7.75	8.85	11.14
10	4 to 5	4 to 5	3 to 4	8.44	8.94	9.83
22	7 to 8	3 to 4	2 to 3	6.44	7.06	8.59
30	2 to 3	3 to 4	4 to 5	8.37	10.05	12.39
31	8 to 9	7 to 8	2 to 3	9.90	9.11	11.31
36	4 to 5	2 to 3	1 to 2	7.53	9.20	12.11
40	4 to 5	3 to 4	1 to 2	8.28	7.88	9.72
Average				8.52	9.16	10.97

Table 10. Expansion factors for summer short-term counts.

	Area								
Counting Period	2	6	10	22	30	31	36	40	Average
To Convert Short-Term Count to Average Daily Traffic									
Sunday	0.464	0.292	0.450	0.653	0.255	0.401	0.345	0.388	0.406
Monday	1.049	0.933	0.635	1.067	0.950	0.826	0.861	0.826	0.893
Tuesday	1.056	0.818	0.529	1.097	1.020	0.742	0.873	0.831	0.871
Wednesday	0.950	0.796	0.523	0.928	0.939	0.800	0.753	0.798	0.811
Thursday	1.022	0.786	0.605	0.990	0.883	0.751	0.755	0.790	0.724
Friday	1.070	0.701	0.562	0.971	0.779	0.724	0.705	0.710	0.778
Saturday	0.868	0.446	0.491	0.690	0.516	0.541	0.496	0.509	0.570
Weekend	0.284	0.162	0.219	0.313	0.162	0.206	0.190	0.209	0.218
10-hour Sunday	0.527	0.353	0.564	1.013	0.291	0.520	0.403	0.517	0.524
Week	0.121	0.084	0.075	0.118	0.088	0.087	0.085	0.088	0.093
To Convert Short-Term Count to Average Summer Sunday Traffic									
Sunday	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Monday	2.259	3.196	1.411	1.636	3.729	2.061	2.494	2.130	2.364
Tuesday	2.275	2.803	1.175	1.681	4.004	1.851	2.527	2.141	2.307
Wednesday	2.047	2.727	1.161	1.423	3.686	1.996	2.181	2.058	2.160
Thursday	2.202	2.692	1.344	1.517	3.468	1.875	2.188	2.037	2.165
Friday	2.304	2.401	1.248	1.488	3.061	1.806	2.042	1.830	2.022
Saturday	1.869	1.529	1.090	1.058	2.026	1.350	1.436	1.311	1.459
Weekend	0.611	0.556	0.487	0.480	0.636	0.513	0.549	0.539	0.546
10-hour Sunday	1.134	1.210	1.252	1.552	1.142	1.297	1.167	1.332	1.261
Week	0.260	0.289	0.166	0.181	0.346	0.217	0.247	0.228	0.242

under normal conditions during the summer months can be used to estimate average daily traffic and average summer Sunday traffic. This table, as developed from a complete year of actual data, should prove to be a useful tool for making these conversions. However, the factors vary a great deal among the recreational areas, and their effective use relies on careful study and informed judgment.

SUMMARY AND CONCLUSIONS

The purpose of this study was to examine characteristics of travel to outdoor recreational areas in Kentucky that are of interest to the highway engineer. Recreational travel, like many other types of travel, is highly complex and very much dependent on local conditions. Therefore, many of the specific data assembled here are sensitive to the nature of the recreational area and its location relative to the various origin zones throughout the United States. Some of the principal findings and conclusions of the study follow.

1. Vehicle occupancy, which averaged 3.06 persons per vehicle, is much larger for outdoor recreational travel than for normal highway travel. Occupancy was found to be a function of the type of recreation area, distance traveled, and type of vehicle. Smallest rates were observed at areas having large day-use activity. Among the various types of vehicles, occupancy was highest for cars pulling camping trailers. The sensitivity of occupancy rate to distance traveled was greatest for camping vehicles. However, for all types of vehicles, occupancy rate increased with increasing distance traveled.
2. A large portion of the vehicles were cars (96.7 percent). The remainder were single-unit campers (2.1 percent) and motorcycles, trucks, and buses (1.2 percent). Altogether, 3.4 percent of the vehicles had camping units attached and 5.8 percent had boats. The nature of the recreational facilities had a decided impact on the proportion of camping units and boats. The proportion of camping units also increased significantly as distance of travel increased. Boat use peaked in the distance range of 60 to 90 miles.
3. Trips to outdoor recreational areas of the type found in Kentucky are relatively short as evidenced by the fact that 60 percent of all vehicles traveled less than 50 miles. Trip lengths were definitely dependent on the type and location of the recreational area, and, for areas having a large regional impact, average trip length can be quite large. Vehicles with camping units travel on the average much longer distances than other types of vehicles.
4. The distribution of recreational travel over time can be investigated by conducting long-term, continuous volume surveys. A very effective method was developed and applied here for the detection and correction of erroneous data collected from the long-term operation of continuous traffic recorders. With minor modifications, this method should prove useful in all long-term, continuous vehicle-counting programs.
5. The distribution of recreational traffic over time is highly dependent on the nature of the recreational area, nature of the recreationists, and location of the areas in relation to population centers. In any case, however, recreational travel is much more variable over time than other forms of highway travel. Evidence and documentation of this peaking is presented in terms of highest hourly volume, highest daily volume, and highest weekly volume plots.
6. The maximum hourly volumes averaged 63.2 percent of the ADT, whereas the 30th highest hourly volumes averaged 38.8 percent of the ADT. Design of highway facilities serving recreational travel to accommodate the 30th highest hourly volume appears in many cases to be impractical. A more practical basis for design is the peak-hour volume on the average summer Sunday. This volume on the average corresponded with the 70th to 75th highest hourly volume. It should be emphasized, however, that proper selection of a design-hour volume is a complex task including economic analyses and, of necessity, must vary from situation to situation. Volumes during the summer week averaged 1,080 percent of the ADT, those during the summer weekend averaged 480 percent, and those on summer Sundays averaged 270 percent.
7. Cyclic volume variations for the seasons of the year, months of the year, days of the summer week, and hours of the summer weekend are documented here. The

peak seasonal volume averaged 40.6 percent of the total annual volume and occurred in either the spring or summer seasons. The peak monthly volume averaged 17.6 percent of the total annual volume and occurred in either May, June, or August. Sunday was always the peak day of the summer week except for holidays, and, on the average, 25 percent of the weekly volume was observed on Sunday. The peak hourly volume on Sundays occurred within the interval of 1 to 5 p.m. and averaged 11 percent of the 24-hour Sunday flows.

8. It is practical to estimate ADT and average summer Sunday traffic from the results of short-term counting programs. Factors that permit such estimates have been documented here.

ACKNOWLEDGMENTS

Material presented in this paper was based in part on a planning study conducted by the Kentucky Department of Highways in cooperation with the Federal Highway Administration. The opinions, findings, and conclusions are not necessarily those of the Federal Highway Administration or the Kentucky Department of Highways. The authors wish to acknowledge the assistance of the following agencies in the conduct of the O-D surveys: U.S. Army Corps of Engineers, U.S. Forest Service, National Park Service, Tennessee Valley Authority, Kentucky Department of Natural Resources, and Kentucky Department of Parks. Especial gratitude is expressed to Spindletop Research, Inc., and the Kentucky Program Development Office for their assistance and consultation throughout the study.

REFERENCES

1. Travel in Arkansas—1964. Division of Planning and Research, Arkansas State Highway Department.
2. James, L. D., and Lee, R. R. Economics of Water Resources Planning. McGraw-Hill, New York, 1971.
3. Kaltenbach, K. D. Application of Gravity and Intervening Opportunities Models to Recreational Travel in Kentucky. Kentucky Department of Highways, Research Rept. 336, Aug. 1972.
4. Maring, G. E. Weekend Recreational Travel Patterns. U.S. Department of Transportation, Highway Planning Technical Rept. 18, Feb. 1971.
5. Matthias, J. S., and Grecco, W. L. Simplified Procedure for Estimating Recreational Travel to Multi-Purpose Reservoirs. Highway Research Record 250, 1968, pp. 54-69.
6. Pigman, J. G. Influence of Recreational Areas on the Functional Service of Highways. Kentucky Department of Highways, Research Rept. 310, Aug. 1971.
7. Smith, B. L., and Landman, E. D. Recreational Traffic to Federal Reservoirs in Kansas. Engineering Experiment Station, Kansas State Univ., Special Rept. 70, Aug. 1965.
8. Ungar, A. Traffic Attraction of Rural Outdoor Recreation Areas. NCHRP Rept. 44, 1967.

MODELS OF OUTDOOR RECREATIONAL TRAVEL

John A. Deacon, University of Kentucky; and
Jerry G. Pigman, Kenneth D. Kaltenbach, and Robert C. Deen,
Kentucky Department of Highways

The purpose of this investigation was to evaluate models of travel flow from population centers throughout the United States to outdoor recreational areas in Kentucky. Data were obtained by means of a license-plate, origin-destination survey at 160 sites within 42 recreational areas and by means of a continuous vehicle-counting program at 8 of these sites. Attempts to simulate distributed travel flows concentrated on various single-equation models, a cross-classification model, and gravity and intervening opportunities models. The cross-classification model was found to be an acceptable means for simulating and predicting outdoor recreational travel flows and was decidedly superior to the other models. From the cross-classification model, per capita distributed flows were found to decrease at a decreasing rate with increasing population of the origin zone, increase at a variable rate with increasing attraction of the recreational area, and decrease at a decreasing rate with increasing distance. The intervening opportunities model was found to be unacceptable as a distribution model because it could not effectively accommodate the widely differing sizes of the 42 recreational areas. The gravity model was quite effective in distributing actual productions and attractions. Problems associated with the gravity model were limited to difficulties in accurately estimating trip productions and attractions in the trip generation phase of analysis.

THIS paper describes a comprehensive evaluation of several models of travel flow from population centers throughout the United States to outdoor recreational areas in Kentucky. Particular attention is focused on the information needs of highway planners. They require information such as simulation of the flow of vehicles within a short time period such as a day; simulation of distributed flows, that is, the flow from each origin zone to each recreational area; and consideration of all major recreational areas within the geographic bounds of interest regardless of type, function, or ownership.

NATURE OF PROBLEM

Conceptually, recreational travel flow is related to various factors determining that flow as follows:

$$V_{ij} = f(D_i, S_j, PR_{ij}, T, \overline{S}_{ij}, \overline{D}_{ij}, M) \tag{1}$$

where

V = distributed recreational travel flow from origin zone i to recreational area j,
f = some function,
D_i = recreational demand at zone i,
S_j = recreational supply at area j,
PR_{ij} = average price of the recreational experience,
T = time period,
$\overline{S}_{ij}$ = supply of other recreational areas and facilities that competes with recreational area j for the limited demand at zone i,
$\overline{D}_{ij}$ = demand of other origin zones that competes with origin zone i for the limited recreational supply at area j, and
M = miscellaneous factors.

Thus, recreational flow may be visualized as a delicate equilibrium among the demand for recreational experiences, supply of recreational opportunities, price of recreation as modified by the competitive nature of the system, and other miscellaneous considerations. Two primary tasks of traffic flow modeling are to identify the most relevant, quantifiable, independent variables of Eq. 1 and to select a suitable function or algorithm for relating the dependent with the independent variables. Figure 1 shows many specific factors that have been used by others to quantify the conceptual variables of Eq. 1.

Recreational travel flow models may be classified in either of two distinct categories. The first includes "total flow" models designed to simulate the total flow produced at an origin zone or the total flow attracted to a recreational area. The second includes "distributed flow" models designed to simulate the flow between each origin zone and each recreational area. Output from distributed flow models can be used, through appropriate summation, to produce total flow simulations for both origin zones and recreational areas. The following are some prior studies in which recreational travel models have been developed: total flow models (18, 19, 22) and distributed flow models (1, 5, 6, 8, 10-14, 18-22).

The literature review failed to identify any distributed flow model that was superior to the other types. Therefore, it was decided to investigate four types, including single-equation, cross-classification, gravity, and intervening opportunities models. Single-equation models, used quite successfully by others (11, 14, 21), are particularly easy to calibrate and apply. Cross-classification models, apparently not used for recreational travel, have been successfully used for other travel not only as a simulation model but also as a means for visual examination of data trends (7). Finally, gravity and intervening opportunities models have been used quite successfully not only for recreational travel but also for travel in urban areas (3, 4, 16).

SURVEY PROCEDURES

Data for calibrating and evaluating the various models were collected by means of a license-plate, origin-destination (O-D) survey at 160 recreational sites in Kentucky during the summer of 1970. These data were supplemented by a traffic volume survey using continuous automatic traffic recorders at eight of the sites.

Peak travel to most outdoor recreational facilities in Kentucky occurs on summer Sundays, excluding from consideration certain holiday periods. The O-D survey was, therefore, conducted on Sundays, and modeling efforts were concentrated on average summer Sunday flows, a flow period suitable for planning and design of both recreational and highway facilities. Surveys were conducted at each site from 10 a.m. to 8 p.m. by one to three persons, depending on the level of travel anticipated. Data recorded for each observed vehicle included direction of movement (arriving or departing), type of vehicle, number of persons in the vehicle, and license-plate identification.

The license-plate identification was used to approximate the origin of the vehicle. A total of 190 origin zones were identified—120 counties in Kentucky, 10 zones in Ohio, 8 zones in Indiana, 6 zones in Tennessee, 3 zones in Michigan, and 1 zone for each of the remaining 43 coterminous states.

Each of the 160 survey sites was associated with 1 of 42 recreational areas. The sites were carefully selected such that the sum of flows passing all the sites associated with a given recreational area accurately represented the total flow to that area.

The 42 areas represent a major part of outdoor recreational activity in Kentucky. Specific areas were chosen to represent a variety of facilities from small fishing lakes to major national scenic attractions, a wide geographic distribution within the state, and a wide variety of operating agencies.

Details concerning the study techniques and other related information can be found elsewhere (15). However, it must be noted here that the license-plate, O-D study was found to be a very efficient way to obtain useful flow data even though certain information, such as trip purpose, could not be obtained and some error was introduced by assuming the point of the trip to be identical with the location of vehicle registration. Concentration on the period of normal peak flow, that is, the summer Sunday, proved extremely efficient and completely compatible with data requirements of this study.

Figure 1. Factors influencing outdoor recreation travel flow.

- A. Origin of Recreational Demand
  - 1. Participant
    - a. Family Characteristics
      - (1) Income of family head
      - (2) Education of family head
      - (3) Occupation of family head
      - (4) Leisure of family head (work week and paid vacation)
      - (5) Race
      - (6) National origin
      - (7) Automobile ownership
      - (8) Location of residence (urban or rural)
    - b. Individual characteristics
      - (1) Leisure
      - (2) Age
      - (3) Marital status
      - (4) Sex
      - (5) Education
  - 2. Origin area
    - a. Total population
    - b. Degree of urbanization
    - c. Median family or percapita income
    - d. Median education
    - e. Percentage of blue- or white-collar employees
    - f. Automobile ownership or registration
    - g. Retail sales
    - h. Property value
    - i. Median age
    - j. Median leisure (work week and paid vacation)
    - k. Race ratio
    - l. Nativity ratio
    - m. Unemployment ratio
    - n. Proportion of various types of employment
    - o. Residential density
    - p. Number of dwelling units

- B. Price of Recreational Experience (monetary and non-monetary)
  - 1. Spatial separation characteristics
    - a. Travel route quality
    - b. Travel time
    - c. Out-of-pocket travel costs
    - d. Distance (airline, road, or other)
  - 2. Charges for use of recreational facilities
  - 3. Cost of equipment rental or ownership

- C. Time Characteristics
  - 1. Holidays
  - 2. Cyclic conditions
    - a. Season
    - b. Month
    - c. Day of week
    - d. Time of day

- D. Competition
  - 1. Supply
    - a. Accessibility to closer recreational areas
    - b. Distance ratio (nearest competing area)
    - c. Sum of attractiveness of closer areas
    - d. Other
  - 2. Demand
    - a. Accessibility to closer origin zones
    - b. Sum of population closer
    - c. Other

- E. Miscellaneous Considerations
  - 1. Regional preferences
  - 2. Other

- F. Supply of Recreational Opportunities
  - 1. Water-oriented facilities
    - a. Lake
      - (1) Total acres
      - (2) Water level, temperature, and quality
      - (3) Miles of shoreline
      - (4) Acres for fishing, water skiing, boating, and sail boating
      - (5) Length or acres acres of beach
      - (6) Swimming areas
      - (7) Number of boat-launching ramps
      - (8) Number of rental boats
      - (9) Number of slips (open and closed)
    - b. Swimming pools
      - (1) Number
      - (2) Size
      - (3) Availability of bath house
  - 2. Intensive-use facilities
    - a. Number of golf holes
    - b. Area available for field sports
    - c. Number of tennis courts
    - d. Number and types of playgrounds
    - e. Availability of shooting range
    - f. Availability of archery range
    - g. Availability of bicycle rentals
    - h. Availability of sky lift
    - i. Availability of amusement park
    - j. Availability of skating rink
    - k. Availability of riding stables
  - 3. Extensive-use facilities
    - a. Trails and paths
      - (1) Miles of bicycling paths
      - (2) Miles of hiking and walking paths
      - (3) Miles of horseback-riding paths
    - b. Area available for hunting
  - 4. Composite size of area
    - a. Total undeveloped acreage
    - b. Total developed acreage
    - c. Total water acreage
  - 5. Eating facilities
    - a. Restaurant (number of seats)
    - b. Concessions
    - c. Picnicking
      - (1) Number of tables or area available
      - (2) Number of grills
      - (3) Number or area of shelters
      - (4) Availability of drinking water
    - d. Distance to nearest inn or store
  - 6. Overnight accommodations
    - a. Camping
      - (1) Number of sites and(or) acres
      - (2) Availability of bathhouse
      - (3) Availability of flush or pit toilets
      - (4) Availability of electricity
      - (5) Availability of laundry facilities
      - (6) Availability of firewood
      - (7) Availability of drinking water
    - b. Other
      - (1) Number of cottages
      - (2) Number of lodge rooms
      - (3) Number of motel rooms
      - (4) Total number of overnight accommodations
  - 7. Quality of physical environment
    - a. Terrain
    - b. Vegetation and shade
    - c. Wildlife
    - d. Water and shoreline
    - e. Climate
    - f. Historic and(or) cultural attractions
  - 8. Activities available
    - a. Wildlife exhibits
    - b. Naturalist service
    - c. Number of drama or concert seats
    - d. Museum
    - e. Lectures
  - 9. Other
    - a. Distance to nearest airport
    - b. Capital investment in recreational facilities

DEPENDENT VARIABLE

The number of vehicles departing a recreational area during the 10-hour survey period (10 a.m. to 8 p.m.) on the average summer Sunday was chosen as the dependent variable of the modeling efforts. The 10-hour period was selected to encompass the hours of primary flow in such a way that the endurance of one survey crew would not be exceeded. Departing flows were chosen to avoid a bias toward day users arriving on Sundays. In all cases, the number of vehicles departing during this period was, for all practical purposes, equal to the number of vehicles arriving during the same period. Use of the average summer Sunday avoided extreme peaks associated with summer holidays. At the same time, summer Sunday flows occur with sufficient frequency to justify their use in planning and design.

The 10-hour departing vehicular flow has little direct use in highway planning and design. However, it may be readily factored to yield estimates of more relevant flow variables. For example, the 10-hour departing flows can be multiplied by a factor ranging from 0.25 to 0.29 (average of 0.27) to estimate peak-hour, two-directional flows. Average summer Sunday, 24-hour, two-directional flows can be estimated by applying similar factors of 2.27 to 2.66 (average of 2.44) to the 10-hour departing flows. Finally, 10-hour departing flows can be multiplied by a factor of 0.58 to 1.13 (average of 0.91) to estimate average daily traffic. Average daily traffic, a two-directional 24-hour flow, is defined as the total annual flow divided by 365. The preceding factors were obtained by analyzing continuous traffic-count data obtained at seven sites located in large part at multipurpose state parks. The eighth site, at which volumes were continuously recorded, was excluded because it was not representative of typical recreational travel in Kentucky.

TOTAL FLOW MODELS

The gravity and intervening opportunities models required, as input, estimates of the number of trips produced at each origin zone that are destined to Kentucky outdoor recreational areas and estimates of the number of trips attracted to each recreational area. Such estimates are usually based on total flow models evaluated using regression techniques.

Productions

Kaltenbach (9) has summarized many independent variables used by others in regression equations for estimating productions. These include total population, urban population, number of dwelling units, median age, median family income, retail sales, sex, race, educational level, various measures of accessibility to recreational opportunities, and others. Chosen for evaluation here were total population, motor vehicle registration, total number of dwelling units, number of dwelling units per square mile, average effective buying income per household, and accessibility to recreational opportunities. Unfortunately, when the Kentucky origin zones were analyzed, very large linear correlations were found among the first four of these independent variables. Accordingly, population was chosen to represent this set of variables in order to avoid potential difficulties. Accessibility to recreational opportunities was expressed as

$$AR_i = \sum_j A_j F_{ij} \tag{2}$$

where

AR_i = accessibility of origin zone i to recreational opportunities,
A_j = number of trips attracted to recreational area j, and
F_{ij} = F-factor of the gravity model corresponding to the distance between i and j.

Separate models were developed for out-of-state origin zones and in-state (Kentucky) origin zones to reflect distinctively different patterns in trip production. Several production equations evaluated are given in Table 1. The accuracy of these equations, as measured by the squared correlation coefficient R^2, is somewhat marginal. At the same time, a generalized, second-degree polynomial in the three independent variables yielded little increase in accuracy. Similarly, a cross-classification model showed no improvement.

Therefore, the following models were judged to be the most suitable among those investigated:

For out-of-state zones,

$$P_i = 803.1\ POP_i^{1.05}\ I_i^{4.19}\ AR_i^{1.03} \tag{3}$$

For in-state zones,

$$P_i = 4{,}050.3\ POP_i^{0.93}\ AR_i^{0.54} \tag{4}$$

where

P_i = productions of origin zone destined to Kentucky recreational areas,
POP_i = total population of the zone in millions,
I_i = average effective buying income per household of the zone in tens of thousands of dollars, and
AR_i = accessibility of zone to Kentucky recreational areas in millions of accessibility units.

Population and accessibility were important for both in-state and out-of-state zones, whereas family income significantly improved the accuracy only for out-of-state productions. Equations 3 and 4, combined with projections of future per capita recreational travel (2), allow predictions to be made of future productions of trips destined to Kentucky outdoor recreational areas.

Attractions

Development of a model to accurately simulate attractions was particularly difficult because of the wide variety among the 42 recreational areas. These areas included small fishing lakes such as Beaver Lake, large water-based resort complexes such as Kentucky Lake-Lake Barkley, and national scenic attractions such as Mammoth Cave. Kaltenbach (9) has also summarized many of the independent variables used by others to estimate trip attractions. Based on this summary, it was concluded that independent variables affecting attractions should include measures of the extent of water-oriented facilities, measures of the availability of overnight accommodations, measures of the development of day-use facilities, measures of the accessibility to population centers, and measures of the quality of the physical environment including historic, cultural, and scenic attractions.

The extent of water-oriented facilities was measured in terms of lake acreage (LAKE), linear feet of swimming beach (BEA), and square feet of swimming pools (POOL). Overnight accommodations were expressed as the sum of the numbers of campsites, cottages, and motel or lodge rooms (ON). Number of golf holes (GH), number of picnic tables (PIC), number of drama seats (DRAM), miles of hiking trails (HIK), and miles of horseback trails (HB) were used as appropriate measures of the development of day-use facilities. Accessibility to population centers was defined as

$$AP_j = \sum_i POP_i F_{ij} \tag{5}$$

where AP_j = accessibility of recreational area j to population. Unfortunately, it was impossible to devise suitable measures of the quality of the physical environment, and this factor had to be omitted from the analysis.

Linear regression analysis yielded the following simple equation for estimating attractions:

$$\begin{aligned} A_j = {} & \underset{(0.17)}{10.2\ \mathrm{GH}} + \underset{(2.08)}{3.28\ \mathrm{PIC}} + \underset{(0.14)}{0.324\ \mathrm{ON}} + \underset{(0.10)}{0.0643\ \mathrm{DRAM}} + \underset{(0.15)}{2.24\ \mathrm{HIK}} + \underset{(0.45)}{8.17\ \mathrm{HB}} \\ & + \underset{(0.83)}{0.293\ \mathrm{BEA}} + \underset{(1.92)}{0.227\ \mathrm{POOL}} + \underset{(4.46)}{0.0986\ \mathrm{LAKE}} \end{aligned} \tag{6}$$

The t-ratio for each regression coefficient, defined as the ratio of the value of the coefficient to its standard error, is shown in parentheses. Regression coefficients significantly different from zero at the 95 percent confidence level have t-ratios in excess of about 2.0. Unfortunately, Eq. 6 contains several independent variables not significantly different from zero at the 95 percent confidence level. Development of a similar equation in which all the independent variables are statistically significant yields the following:

$$A_j = \underset{(4.09)}{4.09\ \mathrm{PIC}} + \underset{(2.16)}{0.211\ \mathrm{POOL}} + \underset{(7.26)}{0.1111\ \mathrm{LAKE}} \tag{7}$$

Accuracy obtained with both Eqs. 6 and 7 was reasonably good as evidenced by R^2's of approximately 0.88. The R^2 was increased to 0.92 when the accessibility term, defined by Eq. 5, was included in either an additive or multiplicative form. However, use of this accessibility term was considered unacceptable because of the unreasonable negative coefficient in the additive equation and the similarly unreasonable negative exponent in the multiplicative equation.

Equation 6 or 7, combined with projections of future per capita recreational travel (2), can be used to make suitable predictions of future attractions for most recreational areas. However, attractions will generally be underestimated for recreational areas of high scenic appeal or areas that are very close to large population centers.

DISTRIBUTED FLOW MODELS

Single-Equation Models

Many of the factors in Figure 1 that influence outdoor recreational travel could have been considered as possible candidates for the independent variables of single-equation models. However, it was obvious that, to be manageable, the number of independent variables had to be much less than the number of factors shown in Figure 1. Furthermore, Matthias and Grecco (11) and Tussey (21) have concluded that simpler equations often produce better predictions than more complex ones.

Based on the literature review and the ease of acquiring data, we decided to represent the recreational demand at each origin (D_i of Eq. 1) by the single variable of population. This is certainly the most important of the demand-generating factors and one that is easy to acquire and easy to predict for future time periods.

The supply of recreational facilities (S_j of Eq. 1) was represented by attractions as estimated by Eq. 6. Selection of the estimated attractions to represent supply was based on the desirability for achieving consistency within the data base; a desire to include measures of day-use activity, overnight accommodations, and water-based activity; the necessity for including facilities present at all recreational areas; and an analysis of the importance of the variables based on the literature review.

The final factor to be considered was the price of the recreational experience (PR_{ij} of Eq. 1), represented here by the distance separating the origin zone from the recreational area. To determine the required 7,980 distances, we established a system of nodes including the 190 origin-zone nodes and the 42 recreational-area centroids. Links were then constructed connecting all adjacent nodes. Airline distances were used for the links interconnecting the 120 Kentucky origin zones, the 42 recreational areas, and the zones of Ohio, Indiana, Tennessee, and Michigan. Over-the-road distances were used outside these five designated states. The minimum path distances from each origin zone to each recreational area were determined using ICES TRANSET I (17).

Having selected the independent variables, the form of the expression to be evaluated was

$$V_{ij} = f(DIS_{ij}, POP_i, A_j) \tag{8}$$

where

V_{ij} = 10-hour departing vehicular flow between recreational area j and origin zone i,
f = some function,
DIS_{ij} = distance in miles between the recreational area and the origin zone,
POP_i = population of the origin zone in thousands, and
A_j = estimated attractions of the recreational area as defined by Eq. 6.

The first phase of the analysis was an attempt to simulate flows at individual recreational areas, disregarding effects of varying attractions by treating each area separately. Results of this analysis for three of the recreational areas are given in Table 2. In all cases, the attempt to use linear regression analysis on a transformed nonlinear equation proved futile. Hence, results from only nonlinear regression analyses are reported here. A similar difficulty has been noted previously by Matthias and Grecco (11).

First, the basic linear equation

$$V_{ij} = k_1 + k_2 DIS_{ij} + k_3 POP_i \tag{9}$$

was tested to verify the suspected nonlinearity. Small R^2's for each of the three recreational areas given in Table 2 were evidence of this nonlinearity.

Next, a relation of the type reported and used successfully by Tussey (21) was investigated:

$$V_{ij} = k_1 DIS_{ij}^{k_2} POP_i \tag{10}$$

Table 2 gives the notable improvement in R^2 that Eq. 10 offered as compared with Eq. 9. It was suspected, however, that the simple expression for the effect of distance in Eq. 10 would not be valid for such a wide range in distances as encountered in this study. A simple means for treating such a situation is to use dummy variables as indicated in the following equation:

$$V_{ij} = k_1 DIS_{ij}^{x_1 k_2 + x_2 x_3 + x_3 k_4} POP_i \tag{11}$$

where

x_1 = 1 for $0 < DIS_{ij} \leq 100$ and 0 otherwise,
x_2 = 1 for $100 < DIS_{ij} \leq 300$ and 0 otherwise, and
x_3 = 1 for $DIS_{ij} > 300$ and 0 otherwise.

Little or no improvement in R^2 resulted from the use of Eq. 11. Accordingly, use of dummy variables was dismissed from further consideration.

Concern for the effects of distance persisted, however, and it was decided to separate the data set into three parts based on short-, medium-, and long-range distance intervals and to evaluate Eq. 10 separately for each of these data subsets. Results of this evaluation, also given in Table 2, yielded no significant improvement over Eq. 11 or the first use of Eq. 10. It was concluded, therefore, that the effect of distance on distributed travel flows was adequately expressed by Eq. 10.

Preliminary examination of the O-D data had revealed that the per capita flows seemed to depend on the population of the origin zone, increasing population causing a decreasing per capita flow. This suggested that an equation of the following form might prove beneficial:

$$V_{ij} = k_1 DIS_{ij}^{k_2} POP_i^{k_3} \tag{12}$$

A nonlinear regression analysis was performed using Eq. 12 and data from Columbus-Belmont State Park. A substantial improvement was noted in R^2. However, the exponent on the population term was negative. Such an exponent fails to meet the test of reasonableness and suggests a high collinearity between the population and distance variables. Because of this unreasonableness and operational difficulties encountered in the regression analysis for the other two recreational areas given in Table 2, further attempts to examine Eq. 12 were abandoned.

A final equation of significant interest was reported by Matthias and Grecco (11) and is of the following form:

$$V_{ij} = k_1 e^{k_2 DIS_{ij}} POP_i \tag{13}$$

where e = base of natural logarithms. Equation 13, although producing satisfactory results as noted in Table 2, proved slighty inferior to Eq. 10.

It was next necessary to modify the form of the model to accept attractions (Eq. 6) as an independent variable measuring the supply of recreational opportunities. For these analyses, the data were separated into two subsets—one for distances less than or equal to 100 miles and the other for distances greater than 100 miles—in an attempt to reduce the population-distance collinearity and to recognize the large number of very small distributed flows for the longer distances. Because there were so many zero flows associated with the long-distance subset, cross-classification techniques were selected as the most acceptable means of analysis. The cross-classification matrix consisted of 180 cells representing all possible combinations of six distance groups, five population groups, and six attractiveness groups. Each distributed flow was entered into the appropriate cell as a departing flow per thousand people, and the weighted mean of all flows within each cell was recorded as the representative value.

The first model to be evaluated for the short-distance subset by nonlinear regression represented the following modification of Eq. 10:

For $DIS_{ij} \leq 100$,

$$V_{ij} = k_1 DIS_{ij}^{k_2} \; POP_i A_j^{k_3} \tag{14}$$

The total R^2 resulting from the use of this model was 0.28, and only 17 percent of the individual R^2's for the 42 recreational areas exceeded 0.50. These results were considered to be unsatisfactory, and the following model was suggested as a possible improvement:

For $DIS_{ij} \leq 100$,

$$V_{ij} = k_1 DIS_{ij}^{k_2} \; POP_i^{k_3} \; A_j^{k_4} \tag{15}$$

Unlike prior efforts to raise the population term to a power, this effort succeeded in producing the following acceptable least squares equation:

For $DIS_{ij} \leq 100$,

$$V_{ij} = 1.107 \; DIS_{ij}^{-1.083} \; POP_i^{0.441} \; A_j^{0.868} \tag{16}$$

A total R^2 of 0.40 resulted from the use of this model. Detailed comparison of simulated versus actual flows indicated that the model consistently underestimated the larger flows and overestimated the smaller ones. However, all attempts to develop more accurate nonlinear regression models were unsuccessful.

Cross-Classification Model

Development and application of a cross-classification model is almost a trivial matter once the independent variables have been identified. For the analysis reported here, the same independent variables were used as for the single-equation models. The dependent variable was the 10-hour departing flow per 1,000 population of the origin zone. Figure 2 shows the complete model and identifies the categories into which the independent variables were classified. An R^2 of 0.68 was obtained using this model.

Portions of the model have been plotted (Fig. 3 through 5) to indicate visually the

Figure 2. Distributed vehicle flows per 1,000 people from cross-classification analysis.

POPULATION (THOUSANDS)		0-10	10-100	100-1000	1000-10000	10000-100000
ATTRACTIVENESS INDEX FACTOR GROUP	DISTANCES (MILES)					
0- 100	0- 20	0.95898163	0.37657559	0.16223729	0.0	0.0
	20- 40	0.07621366	0.04382936	0.09810883	0.0	0.0
	40- 60	0.03046736	0.00665962	0.01014474	0.02425961	0.0
	60- 80	0.00447205	0.00213163	0.00075684	0.00793951	0.0
	80- 100	0.00501749	0.00134144	0.00087748	0.00135821	0.0
	100- 150	0.0	0.00209034	0.00086263	0.00042550	0.0
	150- 250	0.00236395	0.00113672	0.0	0.00008044	0.0
	250- 400	0.0	0.00194506	0.0	0.0	0.0
	400- 700	0.0	0.0	0.0	0.00001943	0.0
	700-1300	0.0	0.0	0.0	0.00002829	0.00000457
	1300-3000	0.0	0.0	0.0	0.00001465	0.00000711
100- 250	0- 20	1.13544655	5.72978306	0.0	0.0	0.0
	20- 40	0.50813001	0.64762914	0.16062135	0.0	0.0
	40- 60	0.09077013	0.07542700	0.19504023	0.0	0.0
	60- 80	0.00946701	0.04417120	0.03474323	0.0	0.0
	80- 100	0.00978377	0.02821740	0.01148133	0.0	0.0
	100- 150	0.01283454	0.01465168	0.00521773	0.01120973	0.0
	150- 250	0.0	0.01779335	0.00267404	0.00496950	0.0
	250- 400	0.00974108	0.01038040	0.00106779	0.00049610	0.00026823
	400- 700	0.0	0.0	0.0	0.00023177	0.00013441
	700-1300	0.0	0.0	0.0	0.00014438	0.00004732
	1300-3000	0.0	0.0	0.0	0.00001783	0.00003980
250- 500	0- 20	13.60512066	2.15327835	1.69190311	0.0	0.0
	20- 40	0.45618343	0.84385180	0.0	0.05734091	0.0
	40- 60	0.07118195	0.20437711	0.05361288	0.0	0.0
	60- 80	0.08550048	0.09662765	0.02636402	0.0	0.0
	80- 100	0.08958763	0.07254964	0.11188710	0.00930038	0.0
	100- 150	0.12461966	0.03304999	0.04490374	0.00147254	0.0
	150- 250	0.06225098	0.03363845	0.01334620	0.01168360	0.00223527
	250- 400	0.00599639	0.00172808	0.00508884	0.00435554	0.00706346
	400- 700	0.0	0.0	0.00032421	0.00202752	0.00088352
	700-1300	0.0	0.0	0.00074924	0.00085544	0.00085450
	1300-3000	0.0	0.0	0.00032310	0.00038885	0.00028194
500- 1000	0- 20	17.07408142	14.42647648	4.35972214	0.0	0.0
	20- 40	1.23048592	0.98168427	0.06762052	0.0	0.0
	40- 60	0.34941846	0.26402664	0.11306220	0.0	0.0
	60- 80	0.08732462	0.07660019	0.69479340	0.0	0.0
	80- 100	0.04565455	0.04019441	0.06184201	0.04536866	0.0
	100- 150	0.07295336	0.03790832	0.01202674	0.01365991	0.0
	150- 250	0.02795955	0.02301007	0.00477299	0.00526530	0.0
	250- 400	0.01548490	0.00816158	0.00185738	0.00121748	0.00260782
	400- 700	0.0	0.0	0.0	0.00050416	0.00029665
	700-1300	0.0	0.0	0.00008404	0.00026949	0.00026645
	1300-3000	0.0	0.0	0.0	0.00010759	0.00011941
1000- 2000	0- 20	19.38731934	5.39795589	0.0	0.0	0.0
	20- 40	1.09620857	1.13166714	0.49376857	0.0	0.0
	40- 60	0.22912484	0.44439262	0.34142214	0.0	0.0
	60- 80	0.05523006	0.12133151	0.40397137	0.08435732	0.0
	80- 100	0.06004418	0.04569305	0.04844257	0.09810972	0.0
	100- 150	0.02523594	0.04007056	0.01873372	0.01772470	0.0
	150- 250	0.03705391	0.01600631	0.00513345	0.00448848	0.0
	250- 400	0.01967793	0.00619185	0.00224304	0.00144763	0.00132626
	400- 700	0.0	0.0	0.00062903	0.00060745	0.00077247
	700-1300	0.0	0.0	0.00013438	0.00035780	0.00025655
	1300-3000	0.0	0.0	0.00027983	0.00028034	0.00013434
2000- 4000	0- 20	9.30527592	16.86503601	0.0	0.0	0.0
	20- 40	1.61003971	2.61544514	1.88730049	0.0	0.0
	40- 60	0.24922538	0.68204987	0.00874927	0.0	0.0
	60- 80	0.23705786	0.32020891	0.04441848	0.0	0.0
	80- 100	0.10578489	0.10133439	0.09276676	0.0	0.0
	100- 150	0.18476230	0.10318834	0.05605559	0.05523141	0.0
	150- 250	0.07273781	0.08328956	0.02152548	0.03683314	0.0
	250- 400	0.15019166	0.04602881	0.01176453	0.00443741	0.00067058
	400- 700	0.0	0.0	0.00090592	0.00138012	0.00214037
	700-1300	0.0	0.0	0.00047370	0.00087972	0.00033799
	1300-3000	0.0	0.0	0.00041116	0.00012996	0.00023882
4000-10000	0- 20	4.85889149	21.91233826	0.0	0.0	0.0
	20- 40	0.24458832	27.33007813	0.0	0.0	0.0
	40- 60	0.0	2.28969002	0.0	0.0	0.0
	60- 80	0.93365526	1.73152637	0.00729106	0.0	0.0
	80- 100	0.55362606	0.55689758	1.63489532	0.0	0.0
	100- 150	0.29150647	0.23062080	0.19563627	0.0	0.0
	150- 250	0.31635976	0.20414138	0.07663280	0.16524690	0.0
	250- 400	0.0	0.0	0.04418130	0.00874706	0.00447054
	400- 700	0.0	0.0	0.00181183	0.00149137	0.00092559
	700-1300	0.0	0.0	0.0	0.00086951	0.00094638
	1300-3000	0.0	0.0	0.00088904	0.00042389	0.00044779
10000-20000	0- 20	107.38320923	111.47634888	0.0	0.0	0.0
	20- 40	41.39472961	21.06471257	0.0	0.0	0.0
	40- 60	6.48586330	20.13973999	0.66606885	0.0	0.0
	60- 80	5.39302994	6.34304714	0.0	0.0	0.0
	80- 100	1.95128021	2.99995136	0.0	0.0	0.0
	100- 150	0.19966675	0.72808444	0.64073777	0.08552021	0.0
	150- 250	0.49463910	0.54417735	0.32898664	0.06843203	0.0
	250- 400	0.25180978	0.30700815	0.06867200	0.05273020	0.19035572
	400- 700	0.0	0.0	0.02297622	0.01006312	0.0
	700-1300	0.0	0.0	0.00549426	0.00490166	0.00383404
	1300-3000	0.0	0.0	0.00165747	0.00344839	0.00119410

Table 1. Production equations.

Equation	Squared Correlation Coefficient	
	Kentucky	Out-of-State
$P = a_1 + a_2POP + a_3AR$	0.67	0.10
$P = a_1 + a_2POP^{a_3} + a_4I^{a_5} + a_6AR^{a_7}$	0.71	—
$P = a_1POP^{a_2}I^{a_3}AR^{a_4}$	0.71	0.84
$P = (a_1 + a_2AR)^{a_3}\ (1 - e^{-a_4POP})\ I^{a_5}$	0.74	0.83
$P = a_1POP^{a_2}AR^{a_3}$	0.70	0.71

Note: P = productions of an origin zone, POP = total population of zone, I = average effective buying income per household in zone, AR = accessibility of zone to Kentucky recreational opportunities, a_i = constants, and e = base of natural logarithms.

Table 2. Regression analysis for three recreational areas.

Equation Number (see text)	Squared Correlation Coefficient		
	Columbus-Belmont State Park	Kentucky Lake-Lake Barkley Complex	Lake Beshear-Pennyrile State Park
9	0.01	0.09	0.02
10	0.76	0.66	0.59
11	0.76	0.66	0.60
10[a]	0.76	0.71	0.61
12	0.95	—	—
13	0.71	0.57	0.60

[a] Separate calibrations were made for three data subsets based on distance intervals of 0 to 100 miles, 100 to 300 miles, and greater than 300 miles.

Figure 3. Effect of population on flow rate.

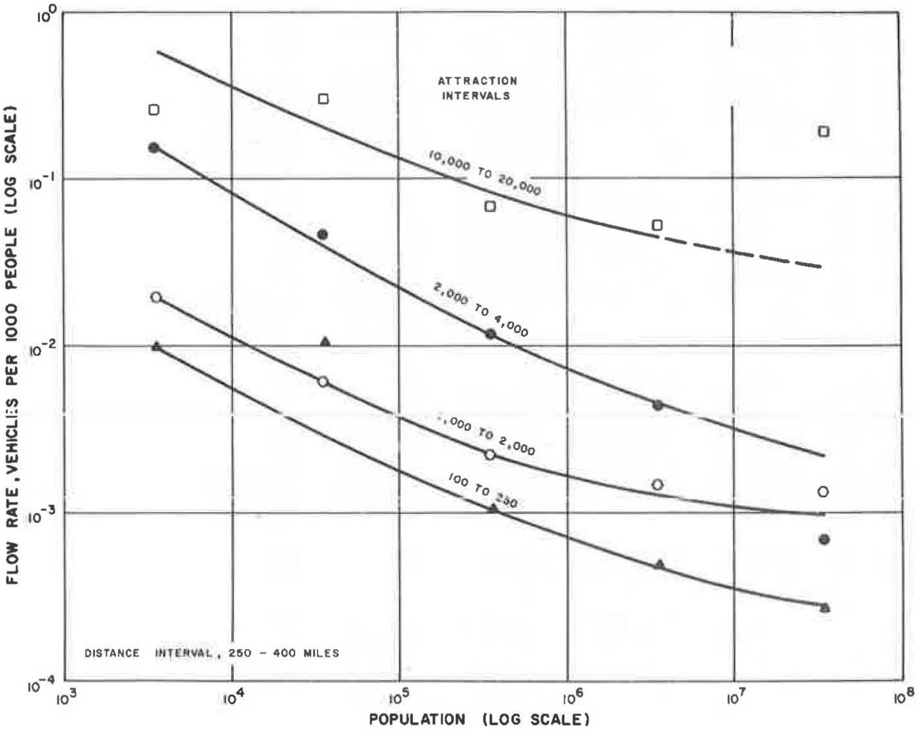


Figure 4. Effect of attractiveness of recreation area on flow rate.

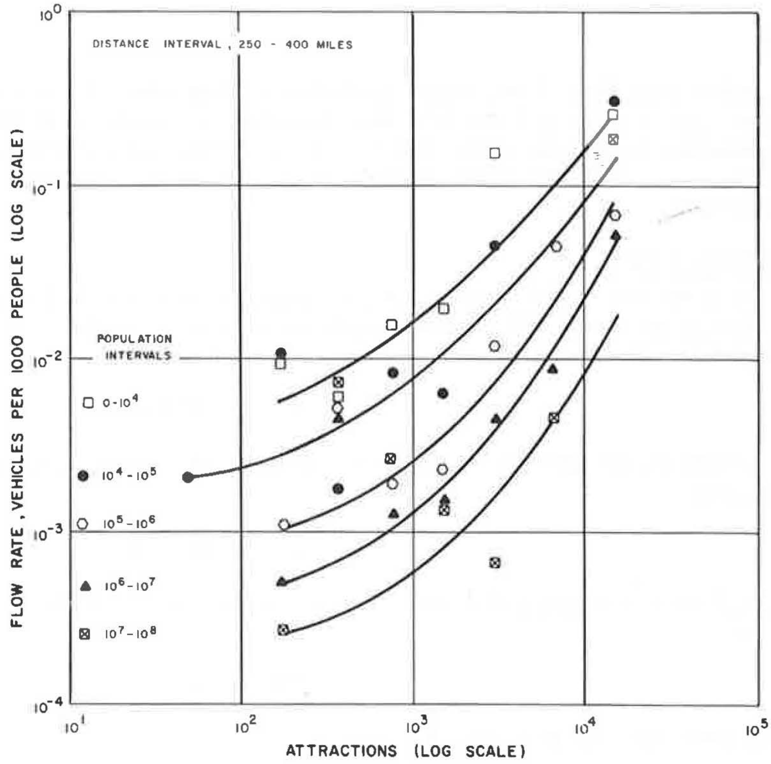
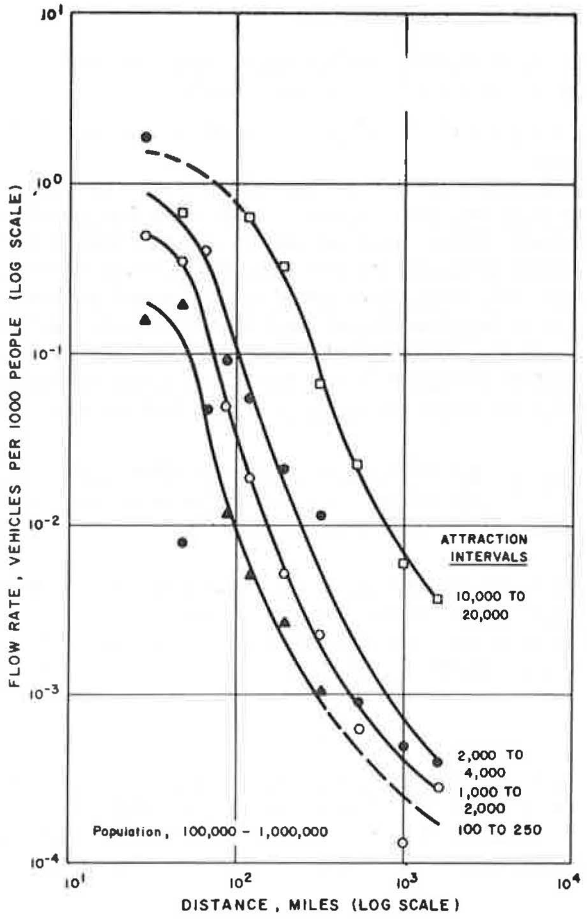


Figure 5. Effect of distance on flow rate.



effects of the three independent variables on flow rate. From the cross-classification model, per capita distrubuted flows were found to decrease at a decreasing rate with increasing population of the origin zone, increase at a variable rate with increasing attractions of the recreational area, and decrease at a decreasing rate with increasing distances.

Gravity Model

The gravity model in all of its varied forms is certainly the most widely used trip distribution model. The model employed here is of a form described by the Federal Highway Administration (3):

$$V_{ij} = P_i A_j F_{ij} / \sum_k A_k F_{ik} \tag{17}$$

In practice, the attractions (A_j) of Eq. 17 are replaced by "adjusted" attractions (AA_j) to yield

$$V_{ij} = P_i AA_j F_{ij} / \sum_k AA_k F_{ik} \tag{18}$$

Equation 18 was applied iteratively until the following constraining equality was satisfied:

$$\sum_i V_{ij} = A_j \tag{19}$$

Adjusted attractions were calculated as

$$AA_j = AA'_j A_j / \sum V'_{ij} \tag{20}$$

where

AA'_j = adjusted attractions from the prior iteration, and
V'_{ij} = distributed flows from the prior iteration.

A maximum of 10 iterations was required in this study to satisfy Eq. 19 and thereby balance the trip ends.

The gravity model must be calibrated before it can be applied; that is, the F-factors are determined as a function of distance. This was also an iterative, numerical procedure. A set of F-factors was first assumed, and the distributed flows (V_{ij}) were estimated using the actual productions and attractions from the O-D survey. During calibration, the average trip length estimated by the model was required to be within 3 percent of the average trip length obtained from the O-D survey. In addition, the percentage of trips occurring within each of 19 distance intervals, as estimated by the model, was required to be within 5 percent of the corresponding value obtained by survey. If these conditions were not satisfied, new factors were estimated as follows:

$$\text{New F} = \text{old F}\ \frac{\text{percentage of trips in distance interval by O-D survey}}{\text{percentage of trips in distance interval by latest gravity model distribution}} \tag{21}$$

The process was then repeated until the convergence criteria based on average trip length and trip-length distribution were satisfied.

F-factors obtained from the calibration phase are given in Table 3. They are approximately related to distance as follows:

$$F_{ij} = k/DIS_{ij}^{2.1} \tag{22}$$

For purposes of comparison, F-factors developed by Smith and Landman (19) and Ungar (22) are also given in Table 3. With the exception of the shorter distances, F-factors developed here compared quite favorably with those of Ungar. However, they showed little similarity to the irregular F-factors developed by Smith and Landman.

The gravity model, using the F-factors of Table 3 and actual O-D survey productions and attractions, simulated trip interchanges quite accurately as evidenced by an R^2 of 0.89. Average trip length and trip-length distribution were also acceptable. However, when using simulated productions (Eqs. 3 and 4) and attractions (Eq. 6), the R^2 decreased to 0.52, indicating that the greater problem in using the gravity model for recreational travel is not the distribution model itself but rather the trip generation phase in which productions and attractions are estimated.

Intervening Opportunities Model

Like the gravity model, the intervening opportunities model is a distribution model requiring trip-end data as input. The model can be stated mathematically (4) as

$$V_{ij} = P_i\left[e^{-LA} - e^{-L(A+A_j)}\right] \tag{23}$$

where

L = probability that a random destination will satisfy the needs of a particular trip, and

A = sum of attractions of all recreational areas closer to origin i than recreational area j.

The opportunities model of Eq. 23 does not automatically distribute all of the productions. This potential difficulty can be readily overcome by adding a constant K as follows (16):

$$V_{ij} = KP_i\left[e^{-LA} - e^{-L(A+A_j)}\right] \tag{24}$$

in which

$$K = 1/\left(1 - e^{-L\sum_k A_k}\right) \tag{25}$$

Trip-end balancing is also required with the opportunities model to ensure that

$$\sum_i V_{ij} = A_j \tag{26}$$

This can be done by rewriting Eq. 24 in terms of "adjusted" attractions (AA and AA_j) as

$$V_{ij} = KP_i\left[e^{-LAA} - e^{-L(AA+AA_j)}\right] \tag{27}$$

Equation 27 was applied iteratively until the trip ends were balanced; that is, Eq. 26 was satisfied. Adjusted attractions were computed following each iteration using Eq. 20.

Calibration of the opportunities model entails selection of the value of the probability parameter L that yields the best simulation of the actual O-D trip interchanges. Smith and Landman (19) suggested an iterative process whereby an initially assumed value of L is adjusted so that the simulated average trip length is nearly equal to the actual average trip length. For each iteration, a new L is calculated as follows:

$$\text{New L} = \text{old L}\ \frac{\text{calculated average grip length from prior iteration}}{\text{actual average trip length}} \tag{28}$$

This method of determining L was originally attempted here, but convergence was extremely slow. Therefore, a new method was used whereby the initially assumed estimate was modified by a given increment in successive iterations and the optimum L selected as that that maximized R^2. This incremental method proved much more effective than the method suggested by Smith and Landman.

The best value of L was found to be 0.00033. This compared with a value of 0.00069 as reported by Smith and Landman (19). The large difference between these two L-values was due in part to the large difference in the total number of attractions between the two studies.

Using actual attractions and productions, the calibrated model simulated trip interchanges with an R^2 of 0.70. This was considerably less than that achieved with the gravity model. A second evaluation was made using the opportunities model in which trip ends were not forced to balance. This yielded an improved R^2 of 0.79 but, of course, violated the constraint of Eq. 26. It was concluded that the low accuracy achieved with this model was probably due to the fact that the 42 recreational areas demonstrated such a wide range in attractions (from a low of 45 to a high of 18,220). Pyers (16) has reported a similar problem and suggested that it might be overcome by using two different values of L—one for small generators and one for large generators. This possibility was not investigated here.

When simulated productions and attractions were used with the opportunities model, the accuracy with which trip interchanges were simulated, as measured by R^2, was 0.40. The large reduction in R^2 from 0.70 when actual productions and attractions were used further indicated that trip generation was a greater problem in recreational travel modeling than trip distribution.

COMPARISON OF MODELS

Adequacy of the four distributed flow models can be evaluated in many ways. Perhaps the best way is to compare the accuracy with which the 7,980 trip interchanges of the O-D survey can be simulated by each of the models. The R^2, a measure of this accuracy, is summarized for each of the types of models given in Table 4. The cross-classification model, which explained approximately 68 percent of the observed variance, was definitely the most accurate of the four models. A similar measure of accuracy is the percentage of the 42 recreational areas for which the models can simulate trips with an R^2 of at least 0.50. Based on this measure, the superiority of the cross-classification model is again given in Table 4.

Good distributed flow models will likewise accurately simulate average trip length and trip-length distribution. Table 4 indicates that, with the exception of the opportunities model, all models were satisfactory in simulating average trip length. A comparison of the actual and simulated trip-length distributions is shown in Figure 6. The cross-classification model was superior for simulating trip-length distribution, and the gravity model was adequate. However, the single-equation and opportunities models produced distributions that significantly departed from the actual both in position and in shape.

All models were calibrated essentially on the basis of average conditions. The degree to which the flows at any particular recreational area could be accurately simulated depended to a significant degree on how much the area deviated from average. Thus, for recreational areas that had significant day-use activity commonly associated with shorter trips, such as Lake Cumberland and Lake Barkley, the models predicted a longer than actual average trip length. On the other hand, for areas of primarily national interest, such as Mammoth Cave, the models predicted a shorter than actual average trip length. The manner in which this difficulty can be overcome is not readily apparent unless a stratification based on trip purpose can be used. This is obviously impossible with data obtained from a license-plate, O-D survey such as reported here.

Other factors useful in comparing model types are simplicity and ease of application. However, the single-equation and cross-classification models offered certain advantages over the gravity and opportunities models. These included more limited input data requirements and the possibility for making predictions without the use of a computer. In addition, they allowed less restrained use of independent judgment and permitted a single recreational area to be examined by itself.

In comparing only the two distribution models, the gravity model was considerably more accurate than the opprotunities model and simulated the actual trip-length distribution much better. It was also considerably less costly to calibrate and apply. In general, computer cost for the opportunities model was found to be three or four times more than that for the gravity model. The gravity model was able to handle the wide variety of sizes of the recreational areas, whereas the opportunities model was not able to do so.

Table 3. F-factors for gravity model.

Distance Interval (miles)	F-Factor[a]		
	Developed Here	Smith and Landman (19)	Ungar (22)
0 to 10	10,735.62		1,545
11 to 20	3,400.18	4,290	1,267
21 to 30	917.27	4,090	750
31 to 40	483.68	2,540	376
41 to 60	162.22	2,790	180
61 to 80	90.21	90.2	90.2
81 to 100	36.09	22.9	54.4
101 to 125	21.01	11.5	34.6
126 to 150	11.60	4.69	22.9
151 to 200	8.86	0.70	13.6
201 to 250	5.07	0.00	6.2
251 to 325	3.11		
326 to 400	1.40		
401 to 550	0.65		
551 to 700	0.29		
701 to 1,000	0.20		
1,001 to 1,300	0.12		
1,301 to 1,700	0.08		
1,701 to 3,000	0.05		

[a] F-factors of Smith and Landman and Ungar were modified by factoring to achieve conformity at a distance of about 70 miles.

Table 4. Model evaluation.

Model	Total R^2 [a]	Percentage of Recreational Areas With $R^2 \geq 0.50$ [b]	Average Trip Length[c] (miles)
Cross classification	0.679	45	113.7
Gravity	0.519	31	115.9
Single equation[d]	0.403	19	110.3
Opportunities	0.396	10	126.1

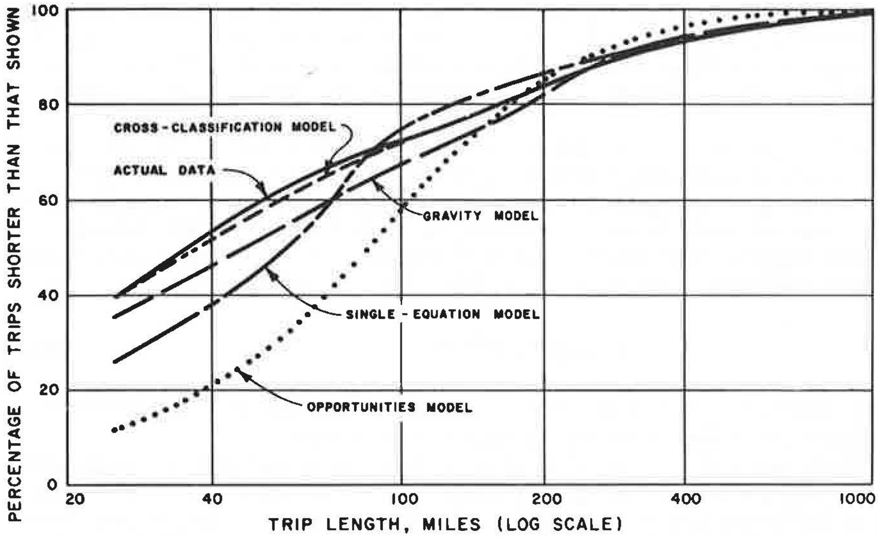
[a] Determined on basis of 7,980 distributed flows.

[b] Percentage of the 42 recreational areas having individual $R^2 \geqslant 0.50$.

[c] Actual average trip length was 109.0 miles.

[d] Equation 16 for distances less than or equal to 100 miles and a cross-classification model for greater distances.

Figure 6. Trip-length distributions.



Based on the preceding evaluations, the cross-classification model was certainly the best of the four models investigated here. Development of this model makes available for the first time an acceptable technique for simulating travel flows to outdoor recreational facilities in Kentucky. When coupled with projections of trends in per capita recreational activity (2), the cross-classification model should prove most effective in predicting future flows to either existing or proposed recreational facilities. Any type of outdoor recreational area can be considered so long as it is possible to estimate its attractions either by comparison with existing facilities or by the use of Eq. 6 or 7. The specific Kentucky model may have limited potential for use outside the state because recreational demand, the mix of available recreational facilities and activities, and consumer preferences vary regionally.

SUMMARY AND CONCLUSIONS

The purpose of this study was to evaluate different models for simulating average summer Sunday flows to outdoor recreational areas in Kentucky from population centers throughout the United States. The primary findings and conclusions of the study are as follows:

1. The impact of recreational travel can be evaluated in a way that is beneficial to highway planners by estimating distributed vehicular flows among all origin zones and all recreational areas during a short time period such as a day. The average summer Sunday is the day of most intense interest because outdoor recreational travel typically peaks on summer Sunday afternoons.

2. Overall results indicate the license-plate, O-D survey is a most satisfactory way to gather O-D data of the type required here, particularly because it enables maximum utilization of personnel without requiring voluntary participation of the traveler and because it allows a very large sampling rate. The time selected for the O-D survey, 10 a.m. to 8 p.m. on summer Sundays, proved to be completely acceptable. However, to be most useful, the O-D survey must be supplemented by a continuous traffic-counting program.

3. The pattern of trip production to outdoor recreational areas in Kentucky differed between in-state and out-of-state origin zones. For in-state zones, population (POP) and accessibility to recreational opportunities (AR) were the most significant indicators of productions. For out-of-state zones, population, average income (I), and accessibility to recreational opportunities were found to be significant. The best equation for simulating productions (P) was found to be of the following general form:

$$P = k_1 POP^{k_2} AR^{k_3} I^{k_4} \tag{29}$$

However, such an equation explains only about 70 percent of the variance for in-state zones and about 84 percent of the variance for out-of-state zones.

4. Attractions (A) to recreational areas of varying types and sizes can be reasonably approximated by a linear equation involving the nature and extent of recreational facilities. The following facilities, listed in the order of highest to lowest significance, were identified as having important effects on attractions and were judged essential for encompassing the wide range of recreational areas studied: water area, picnic tables, swimming pools, horseback trails, beach, golf, hiking trails, overnight accommodations, and outdoor drama. The linear equation utilizing these variables explained about 89 percent of the variance in attractions. However, this equation proved unsuitable for simulating attractions at areas deviating significantly from the average, such as those of high scenic interest and those highly accessible to large population centers.

5. Four types of travel models, including single-equation, cross-classification, gravity, and intervening opportunities models, were evaluated here. The cross-classification model was found to be the most acceptable means for simulating and predicting distributed outdoor recreational travel flows. In virtually any travel modeling effort, cross-classification analysis can be gainfully employed if only for the purpose of visually depicting the effects of various independent variables.

6. The cross-classification model demonstrated that per capita distributed flows decrease at a decreasing rate with increasing population of the origin zone, increase at a variable rate with increasing attractions of the recreational area, and decrease at a decreasing rate with increasing distance.

7. The best single-equation model for simulating distributed flows (V_{ij}) for short-range travel was of the form

$$V_{ij} = k_1 DIS_{ij}^{k_2} \, POP_i^{k_3} \, A_j^{k_4} \tag{30}$$

where DIS_{ij} = distance between origin zone i and recreational area j. This nonlinear flow equation, as others investigated here, had to be evaluated using nonlinear regression analysis. Linear regression using transformed (linearized) equations proved totally unsuitable.

8. The gravity model is a simple and effective model for distributing recreational trips. Accuracy of the trips so distributed depends in large part on the accuracy of estimating productions and attractions. F-factors developed in the gravity model calibration are a convenient and useful means for explaining the effects of distance on travel impedance.

9. The intervening opportunities model can be calibrated very effectively by incrementing the probability parameter L in such a way as to maximize the accuracy of the trip-interchange simulation. However, the opportunities model was found to be decidedly inferior to the gravity model. The intervening opportunities model cannot produce satisfactory results with only one value of L if recreational areas of widely differing attractiveness are present in the study area.

10. For flow models using distinct trip generation and distribution phases, trip generation was found to be the most critical problem in outdoor recreational travel modeling.

ACKNOWLEDGMENTS

Material presented in this paper was based on a planning study conducted by the Kentucky Department of Highways in cooperation with the Federal Highway Administration. The opinions, findings, and conclusions are not necessarily those of the Federal Highway Administration or the Kentucky Department of Highways. The authors wish to acknowledge the assistance of the following agencies in the conduct of the O-D surveys: U.S. Army Corps of Engineers, U.S. Forest Service, National Park Service, Tennessee Valley Authority, Kentucky Department of Natural Resources, and Kentucky Department of Parks. Especial gratitude is expressed to Spindletop Research, Inc., and the Kentucky Program Development Office for their assistance and consultation during all phases of the study. Use of the University of Kentucky Computer Center is also acknowledged.

REFERENCES

1. Boyet, W. E., and Tolley, G. S. Recreation Projection Based on Demand Analysis. Jour. of Farm Economics, Vol. 48, November 1966, pp. 984-1001.
2. Bureau of Outdoor Recreation. 1965 Survey of Outdoor Recreation Activities. U.S. Government Printing Office, 1967.
3. Bureau of Public Roads. Calibrating and Testing a Gravity Model for Any Size Urban Area. U.S. Government Printing Office, 1965.
4. Chicago Area Transportation Study. Vol. 2, 1960.
5. Ellis, J. B. A System Model for Recreational Travel in Ontario: Further Results. Ontario Department of Highways, Rept. RR148, July 1969.
6. Ellis, J. B., and Van Doren, C. S. A Comparative Evaluation of Gravity and System Theory Models for Statewide Recreational Traffic. Jour. of Regional Science, Vol. 6, No. 2, 1966, pp. 57-70.
7. Federal Highway Administration. Guidelines for Trip Generation Analysis. U.S. Government Printing Office, June 1967.

8. Gyamfi, P. A Model for Allocating Recreational Travel Demand to National Forests. Highway Research Record 408, 1972, pp. 50-61.
9. Kaltenbach, K. D. Application of Gravity and Intervening Opportunities Models to Recreational Travel in Kentucky. Kentucky Department of Highways, Research Rept. 336, August 1972.
10. Knetsch, J. L. Outdoor Recreation Demands and Benefits. Land Economics, Vol. 39, November 1963, pp. 387-396.
11. Matthias, J. S., and Grecco, W. L. Simplified Procedure for Estimating Recreational Travel to Multi-Purpose Reservoirs. Highway Research Record 250, 1968, pp. 54-69.
12. Merewitz, L. Recreational Benefits of Water Resources Development. Water Resources Research, Vol. 2, 1966, pp. 625-640.
13. Milstein, D. N., et al. Michigan Outdoor Recreation Demand Study—Volume 1: Methods and Models. Michigan Department of Conservation, Technical Rept. 6, June 1966.
14. Pankey, V. S., and Johnston, W. E. Analysis of Recreational Use of Selected Reservoirs in California. Office of the Chief of Engineers, U.S. Department of the Army, July 1969.
15. Pigman, J. G. Influence of Recreational Area on the Functional Service of Highways. Kentucky Department of Highways, Research Rept. 310, August 1971.
16. Pyers, C. P. Evaluation of Intervening Opportunities Trip Distribution Model. Highway Research Record 114, 1966, pp. 71-98.
17. Ruiter, E. R. ICES TRANSET I—Transportation Network Analysis: Engineer's Users Manual. Massachusetts Institute of Technology, Research Rept. R68-10, 1968.
18. Schulman, L. L. Traffic Generation and Distribution of Weekend Recreational Trips. Joint Highway Research Project, Purdue Univ., Rept. 8, June 1964.
19. Smith, B. L., and Landman, E. D. Recreational Travel to Federal Reservoirs in Kansas. Engineering Experiment Station, Kansas State Univ., Special Rept. 70, August 1965.
20. Thompson, B. Recreational Travel: A Review and Pilot Study. Traffic Quarterly, Vol. 21, No. 4, October 1967, pp. 527-542.
21. Tussey, R. C., Jr. Analysis of Reservoir Recreation Benefits. Water Resources Institute, Univ. of Kentucky, Research Rept. 2, 1967.
22. Ungar, A. Traffic Attraction of Rural Outdoor Recreation Areas. NCHRP Rept. 44, 1967.

TRANSIT ACCESSIBILITY AS A DETERMINANT OF AUTOMOBILE OWNERSHIP

Robert T. Dunphy, Metropolitan Washington Council of Governments

Automobile ownership is generally accepted as the most important determinant of the number of trips made by residents of a traffic zone. For this reason, the way in which it is forecast can have a dominant influence on a regional travel forecast. Because most automobile ownership forecasts have been independent of the transportation alternatives being tested, a major portion of the regional travel demand was set prior to the distribution of travel and the allocation among modes. This paper analyzes the relation between transit accessibility and automobile ownership by eliminating variations in family size and income through a household analysis. The findings show that there is a significant correlation between automobile ownership and transit accessibility for almost every category of automobile ownership in an area by improving transit accessibility. Such a finding could have a major effect on estimates of regional travel demand in areas where major transit improvements are made. Furthermore, reductions in future regional automobile ownership levels that would accrue from a major transit improvement could be considered as a benefit of the transit improvement. These findings could tend to make transit investment slightly more favorable than when the only benefits considered are improved ridership for existing transit users and some diversion of trips from automobile to transit.

•RESEARCH into the demand for urban transportation has shown that automobile ownership is the variable that exhibits the closest association with reported trip generation rates (1). Moreover, when used as a variable to estimate the relative use of different travel modes, automobile ownership rates are much more important than other variables (2). Because of the major importance of automobile ownership in forecasting travel demand, the way in which automobile ownership rates are forecast can be the primary determinant of the amount of travel demand on a future network.

This research has investigated the relation between the transportation system and automobile ownership. Specifically, the effect of variations in transit accessibility on levels of automobile ownership has been analyzed.

PREVIOUS RESEARCH

In a review of previous research of methods of estimating automobile ownership, Deutschman found that the three variables most commonly used were family size, income, and residential density (3). The size of the family largely determines the amount of travel that would be made by that household in the absence of any financial constraints. The income of the family determines the extent to which travel demands can be satisfied through the ownership of one or more automobiles. Residential density determines the percentage of travel desires that can be satisfied by walking trips, which are not counted in the traditional travel survey. In very high-density areas, it is possible for a large percentage of people to walk to shops, schools, recreation, and even work. In low-density areas, only persons living adjacent to shopping centers generally walk to them. Residential density may be considered to be a location variable because it affects the number of opportunities that can be reached from a location in a given amount of time.

It is normally greatest near the central business district (CBD) and declines regularly with increasing distance from the center of the region.

From the perspective of the transportation systems planner, the use of only these independent variables in predicting car ownership, and eventually the number of trips, makes the resulting forecast independent of the transportation system. In testing transit alternatives, this means that the number of transit trips is largely predetermined regardless of the type of system tested. In an attempt to overcome this analysis difficulty, Ferrari and Shindler (2) found that automobile ownership rates varied with the relative level of service provided between the transit and highway systems. This relative transit accessibility is actually a location variable similar to residential density because it indicates the number of opportunities that can be reached from an origin in different time intervals. It tends to be highest in the core of the region and to decline with increasing distance from the CBD. This relation is caused by the centralized orientation of most transit systems. Because of this orientation, transit service tends to be best in the downtown area and progressively worse with increasing distance from the core.

Besides affecting automobile ownership, transit accessibility was shown to be related to those other factors that are accepted as determinants of automobile ownership—family size, income, and residential density. For this reason, some questions have been raised about whether transit accessibility can actually affect automobile ownership or whether it is simply correlated with other factors that are more causative in nature. For example, it has been suggested that transit accessibility is an effect rather than a cause of car ownership. If transit service were provided to a greater extent only in low-income or high-density residential areas, which were assumed to generate the patronage needed to support transit service, then the relation between car ownership and transit accessibility would be meaningless for affecting total transit demand. One of the problems with this type of analysis is that it deals with aggregates of car ownership, income, and family size for an area. However, recent work in trip generation analysis identifies a need for household analysis, in which the basic unit is not an average rate for a traffic zone but rather the average rate for an individual household. Such disaggregate analysis might solve some of the problems previously mentioned (4).

APPLICATION OF HOUSEHOLD ANALYSIS

A major criticism of Ferrari and Shindler's work related to the use of average household characteristics for a traffic zone as the independent variables. Because both family size and income were correlated with transit accessibility, it was difficult to determine the exact relation between car ownership and transit accessibility with all other characteristics held constant. One way to control variations in household characteristics is to perform a household analysis in which the basic observation is an individual interview rather than a zonal average. This type of analysis has been recommended for studies of trip generation, primarily because it attempts to explain more of the basic variation in trip-making (5).

The latter type of analysis was used in this study. The basic data were developed from a home-interview survey conducted by the National Capital Region Transportation Planning Board (TPB) and Metropolitan Washington Council of Governments (COG) in 1968. The data were disaggregated by type of household, which made it possible to formulate for each type of household a simple linear regression.

$$Y_{is} = a + bX$$

where

Y_{is} = the number of automobiles owned by a household of a given income and size,
X = the transit accessibility available to that household, and
a, b = regression coefficients calculated by standard least squares techniques.

Household Categories

Disaggregation of households into individual categories for this analysis makes it possible to eliminate the effect of correlation between the independent variables. Furthermore, it permits a test of whether automobile ownership can be affected by transit accessibility for some types of households but not for others. For example, it could show that those households that would be most willing to exchange automobile ownership for transit accessibility might be small households (especially those with only one person) and very poor households. It would also seem logical that there would be less of a possibility to reduce car ownership through increased transit accessibility for larger, higher income families.

Because of the need to preserve as much of the original variation in the data as possible, grouping of types of households was kept to a minimum. Plotting the relation between car ownership and income for different sizes of households showed that, for any income group, a plateau seemed to be reached for car ownership in families with more than three persons. As shown in Figure 1, it seemed that larger households did not have higher automobile ownership rates. To maintain a larger sample size, all households with more than four persons were combined with the four-person category. Each of the original 10 income groups coded in the survey was used. Table 1 gives the number of samples in each category.

Transit Accessibility

It was mentioned previously that transit accessibility, like residential density, is a location variable. It reflects not only the spatial distribution of opportunities about a point but also the relative speed with which these opportunities may be reached by a given transportation system. Because the distribution of opportunities about a point is a result of development density and location, transportation accessibility is actually a density measure that also incorporates network speed. The measure of transit accessibility currently being used at COG for work travel is the percentage of jobs reached from an area in 45 min by transit (6). It was determined that, in 1968, the average worker could reach three-quarters of the regional employment in 45 min by the fastest mode (usually highway). Because this time boundary accounted for 9 out of every 10 work trips, it was felt to be a representative boundary for commuting travel.

FINDINGS

As given in Table 2, the calculated **F**-ratios between automobile ownership and transit accessibility were significant at the 99 percent level of confidence for all but three categories. These three categories were one-person households with annual incomes of $4,000 to $5,999; one-person households with annual incomes of $20,000 to $24,999; and households having four or more persons with annual incomes of more than $25,000.

Not only does this analysis show that there is generally a significant correlation between car ownership and transit accessibility to employment, but also it shows that the three exceptions do not follow a clear pattern. Except for the highest income households of four persons or more, transit accessibility to employment appears to have a significant impact on the number of cars owned. Moreover, as shown in Figures 2 through 5, the effect of transit accessibility on car ownership is approximately the same at all income levels for a given family size. Although increasing income results in a higher level of car ownership with a constant family size and accessibility for almost all of the regression equations, the slopes of the curves are very similar for different income categories within a particular household size. The exact equations are given in Table 3.

SUMMARY

This study has shown that there is a statistically significant relation between automobile ownership and transit accessibility, even when the other significant household characteristics of family size and income are held constant. It appears that a high level of accessibility to employment by transit may reduce the need to own cars. This

Figure 1. Relation between average car ownership and household income.

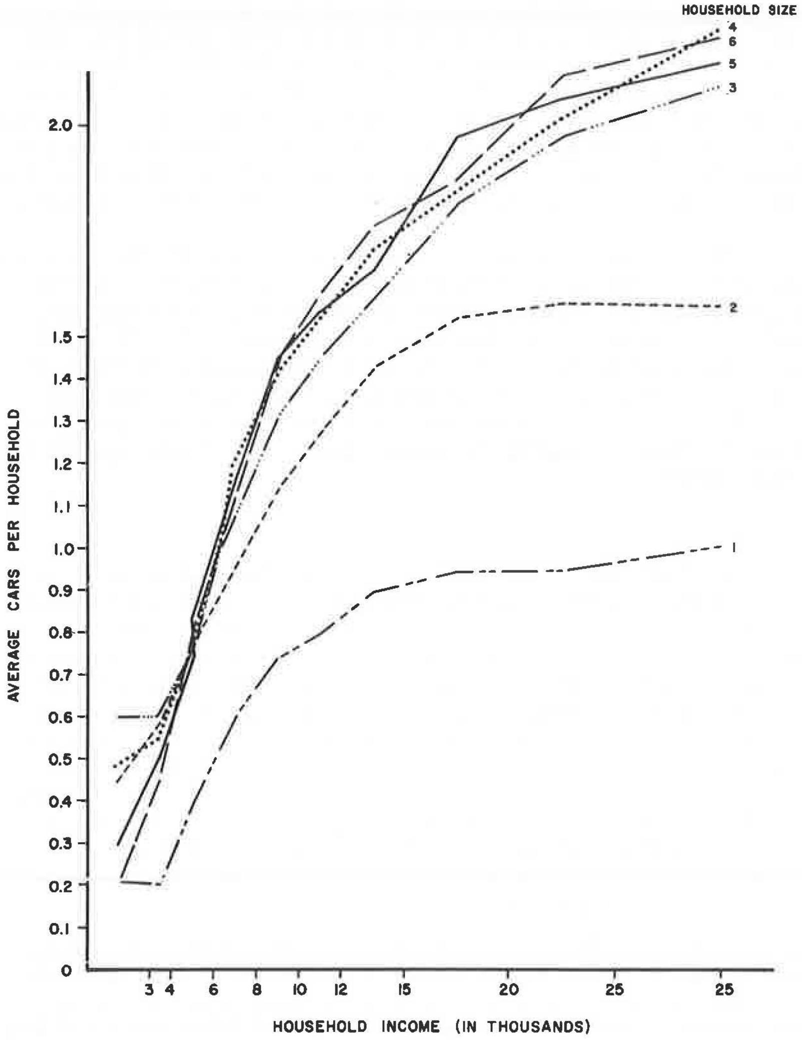


Table 1. Distribution of samples by household characteristics.

1968 Annual Household Income (thousands of dollars)	Number of Persons in Households 1	2	3	≥4	Total
0 to 3	1,871	300	110	126	2,407
3 to 4	593	393	107	127	1,220
4 to 6	784	939	318	410	2,451
6 to 8	882	1,437	567	788	3,674
8 to 10	536	1,505	730	1,130	3,901
10 to 12	360	1,397	723	1,307	3,787
12 to 15	228	1,266	722	1,425	3,641
15 to 20	129	921	602	1,356	3,008
20 to 25	38	378	306	683	1,405
>25	36	312	217	486	1,051
Total	5,457	8,848	4,402	7,838	26,545

Table 2. Calculated F-ratio for regression equations.

1968 Annual Household Income (thousands of dollars)	Number of Persons in Household			
	1	2	3	≥4
0 to 3	32.13	75.67	19.37	26.56
3 to 4	58.45	62.02	21.22	43.94
4 to 6	5.04[a]	184.34	80.24	92.55
6 to 8	139.58	273.33	95.35	200.00
8 to 10	90.12	303.71	108.40	178.92
10 to 12	72.34	194.74	71.31	159.48
12 to 15	38.14	226.74	61.25	121.08
15 to 20	15.52	120.12	51.04	176.98
20 to 25	6.86[a]	25.07	22.04	31.42
>25	9.12	55.61	16.91	1.75[a]

[a] F-ratio is not statistically significant at the 99 percent level of confidence.

Figure 2. Relation between car ownership and transit accessibility to employment (one-person household).

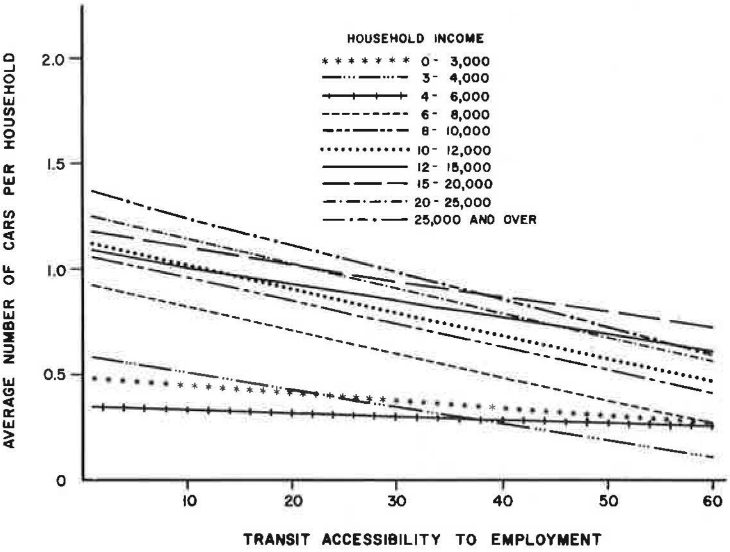


Figure 3. Relation between car ownership and transit accessibility to employment (two-person household).

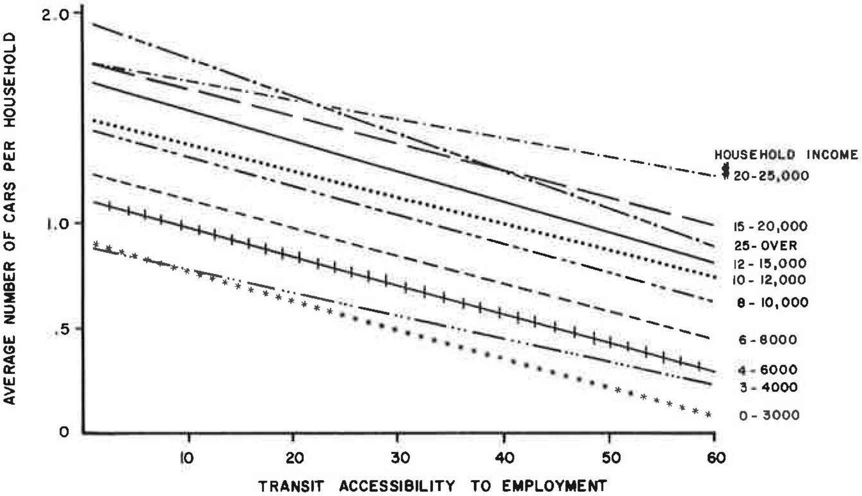


Table 3. Summary of car ownership regression equations.

Income (thousands of dollars)	Number of Persons in Household	Regression Equation[a]	r	s_y	$s^2_{y.x}$
0 to 3	1	Y = 0.4835 - 0.3621 x	-0.13	0.57	0.32
	2	0.9155 - 1.4155 x	-0.45	0.59	0.28
	3	1.2202 - 1.9705 x	-0.39	0.87	0.64
	≥ 4	1.0234 - 1.8983 x	-0.42	0.74	0.44
3 to 4	1	Y = 0.5727 - 0.7619 x	-0.30	0.48	0.21
	2	0.8775 - 1.0898 x	-0.37	0.60	0.31
	3	1.0461 - 1.5119 x	-0.41	0.73	0.45
	≥ 4	1.1950 - 2.0075 x	-0.51	0.69	0.35
4 to 6	1	Y = 0.3512 - 0.1581 x	-0.08	0.48	0.23
	2	1.1163 - 1.378 x	-0.41	0.65	0.36
	3	1.2547 - 1.7369 x	-0.45	0.78	0.49
	≥ 4	1.2465 - 1.7048	-0.43	0.77	0.48
6 to 8	1	Y = 0.9189 - 1.0924 x	-0.37	0.55	0.27
	2	1.2572 - 1.363 x	-0.40	0.65	0.36
	3	1.3507 - 1.4926 x	-0.38	0.70	0.42
	≥ 4	1.5059 - 1.8386 x	-0.45	0.76	0.46
8 to 10	1	Y = 1.0758 - 1.1074 x	-0.38	0.54	0.24
	2	1.4552 - 1.4696 x	-0.41	0.68	0.39
	3	1.5276 - 1.4741 x	-0.36	0.70	0.43
	≥ 4	1.681 - 1.6601 x	-0.37	0.74	0.48
10 to 12	1	Y = 1.1282 - 1.1184 x	-0.41	0.50	0.21
	2	1.4973 - 1.2614 x	-0.35	0.66	0.38
	3	1.6601 - 1.3139 x	-0.30	0.74	0.50
	≥ 4	1.7705 - 1.5363 x	-0.33	0.75	0.51
12 to 15	1	Y = 1.108 - 0.8452 x	-0.38	0.41	0.14
	2	1.6801 - 1.4404 x	-0.39	0.65	0.36
	3	1.7991 - 1.1838 x	-0.28	0.72	0.48
	≥ 4	1.8834 - 1.3035 x	-0.28	0.74	0.50
15 to 20	1	Y = 1.1858 - 0.8889 x	-0.33	0.48	0.21
	2	1.7611 - 1.3001 x	-0.34	0.65	0.38
	3	1.9945 - 1.2122 x	-0.28	0.75	0.52
	≥ 4	2.1207 - 1.6754 x	-0.34	0.76	0.51
20 to 25	1	Y = 1.2577 - 1.1307 x	-0.40	0.53	0.24
	2	1.7546 - 0.8933 x	-0.25	0.62	0.36
	3	2.1536 - 1.1953 x	-0.26	0.76	0.54
	≥ 4	2.2535 - 1.1827 x	-0.21	0.81	0.63
>25	1	Y = 1.372 - 1.2995 x	-0.46	0.51	0.20
	2	1.9547 - 1.7736 x	-0.39	0.74	0.46
	3	2.324 - 1.3818 x	-0.27	0.82	0.62
	≥ 4	2.3994 - 0.4185 x	-0.06	0.95	0.90

[a] Y = average cars per household, and x = ratio of regional employment reached in 45 min by transit.

Figure 4. Relation between car ownership and transit accessibility to employment (three-person household).

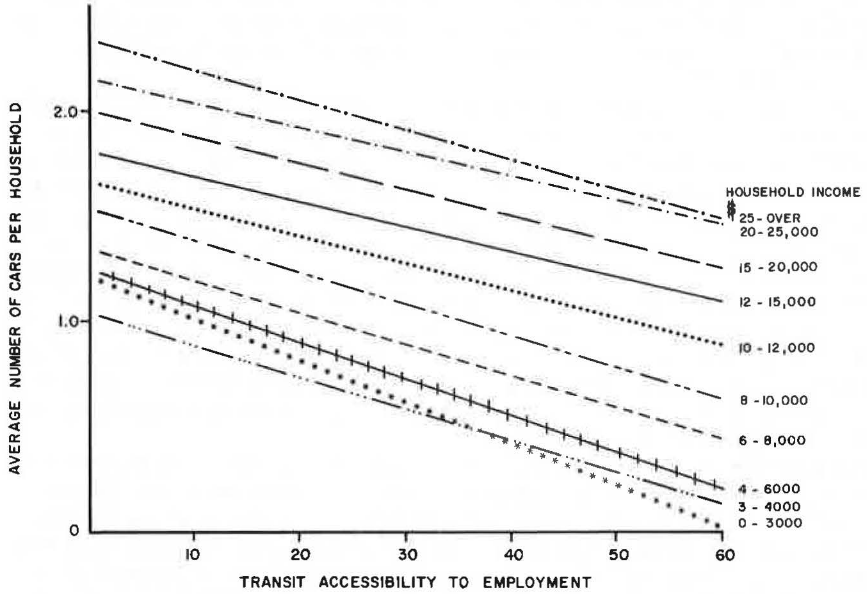
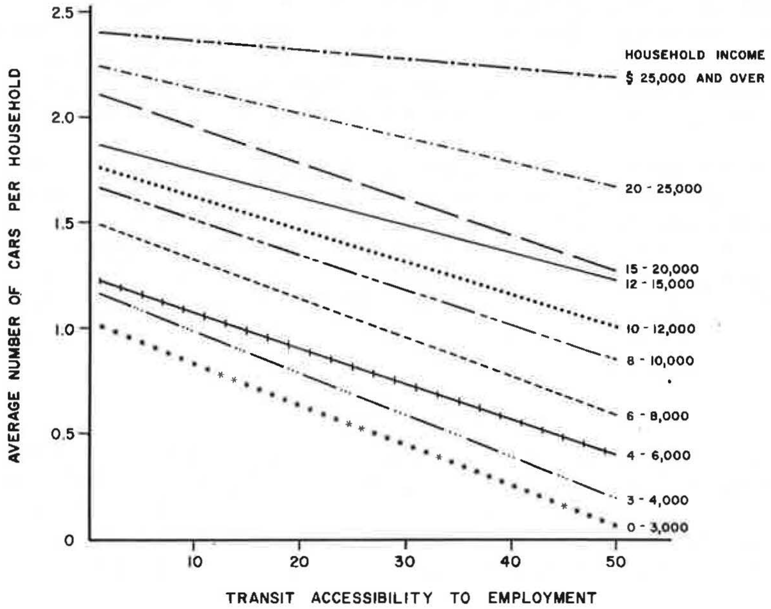


Figure 5. Relation between car ownership and transit accessibility to employment (household of four or more persons).



effect might be to eliminate the need for a second car in two-car families or perhaps to make it possible for some households to exist without any car. In both cases, the family is simply trading off its expenditure for automobile travel for a public expenditure for transit services. In fact, in some cases this may not even be a complete shift of travel from automobile to transit, but simply an awareness that the transit system is there if it were ever needed.

This analysis has dealt entirely with employment-related transit accessibility. This does not mean, however, that those households having a lower level of car ownership because of high transit accessibility to jobs are not also able to use the transit system for other purposes. Although the measurement of accessibility for nonwork purposes is much more complex than that for work purposes, it is likely that this type of accessibility may also affect automobile ownership. In fact, those areas with the highest level of transit accessibility for work trips also have the highest level of accessibility for nonwork travel. The most important effect of transit accessibility may be to eliminate the need for an extra car to go to work. However, a high level of accessibility to nonwork destinations may make it possible for some of the other trips that would have been made in the car to be completed by transit. Innovative transit services such as dial-a-ride may be able to generate sufficient accessibility to nonwork destinations to reduce the need for multiple-car ownership.

Although this analysis deals with a single point in time, it is possible to assess the effect of policy changes over time. For a given point in time, a significant relation has been determined between automobile ownership of different households and transit accessibility to employment of those households. It is not unreasonable to suggest that, for a given household or group of households, a vastly improved transit accessibility may reduce the number of automobiles owned. In fact, a survey of riders on a special commuter bus service in Reston, Virginia, showed that many riders had already reduced the number of cars owned by their families as a result of the service (7).

This analysis has given further support to the theory that provision of good transit service can affect the automobile ownership rate in an area. The magnitude of this effect was shown in an evaluation of a proposal for a new town in the Washington area that included a special transit system. Given the forecast of resident income and family size characteristics, it was found that the level of automobile ownership would be 26 percent below that that would be expected in a similar suburban community with average transit service. A comparable reduction in automobile trip generation could also be expected.

Reductions in automobile travel constitute a public sector benefit, especially if they result in a reduction in highway construction or operation costs. However, the benefits of reducing automobile ownership can be much more significant to the individual. Because the cost to own and operate an automobile can average more than $1,300 per year (8), provision of transit services that eliminate the need for a second car could be a measurable benefit to multicar households. If such benefits are included in the evaluation of proposed transit improvements, it may be possible to justify a higher level of transit service than that that currently exists in many suburban areas.

ACKNOWLEDGMENTS

The preparation of this report has been financed in part through grants from the District of Columbia Department of Highways and Traffic, Maryland Department of Transportation, Virginia Department of Highways, and the U.S. Department of Transportation, Federal Highway Administration and Urban Mass Transportation Administration.

The author would like to thank Vergil G. Stover for his continuing assistance.

REFERENCES

1. Oi, W., and Shuldiner, P. An Analysis of Urban Travel Demands. Northwestern Univ. Press, Evanston, 1962.
2. Ferrari, M. G., and Shindler, R. Auto Ownership as Affected by Transportation System Alternatives. Traffic Engineering, Oct. 1967, pp. 24-28.

3. Deutschman, H. D. Auto Ownership Revisited: A Review of Methods Used in Estimating and Distributing Auto Ownership. Highway Research Record 205, 1967, pp. 31-49.
4. Federal Highway Administration. Guidelines for Trip Generation Analysis. U.S. Government Printing Office, June 1967.
5. McCarthy, G. M. Multiple-Regression Analysis of Household Trip Generation—A Critique. Highway Research Record 297, 1969, pp. 31-43.
6. Wickstrom, G. V. Defining Balanced Transportation—A Question of Opportunity. Traffic Quarterly, July 1971, pp. 337-350.
7. Morin, D. A., Lindsey, L. C., and Riener, R. A Report on the Reston Bus Passenger Survey. Federal Highway Administration, unpublished rept., Jan. 1972, p. 13.
8. Liston, L. L., and Gauthior, C. L. Cost of Operating an Automobile. Federal Highway Administration, U.S. Department of Transportation, April 1972.

NEED FOR EXPLICIT TRANSPORTATION PLANNING PROCEDURES

Britton Harris, University of Pennsylvania

This paper develops the thesis that the major investment in formal methods in transportation planning studies has been in the field of analysis. This analysis covers three broad areas: data collection, statistical analysis, and model construction. The relation among these areas of investigation and theories of behavior and transportation systems performance are explored briefly. These questions are then enlarged to indicate the relation between the transportation system and the social and economic system (spatially distributed) that the transportation system serves, influences, and is influenced by. A transition to the second part of the paper deals with the optimizing nature of planning in general, as suggested by cost-benefit analysis and various decision procedures for the allocation of public resources. The mathematical program as a paradigm of the planning process is also briefly explored, and those elements that are to be discarded in the subsequent discussion are summarized briefly. The second part develops the contention that most actual transportation planning, as distinct from analysis, is conducted on an intuitive basis and according to professional practices that may or may not be well suited to the problem. A tentative sketch of the elements of these traditional methods as related to mathematical programming concepts is then developed. Several typical transportation planning problems are discussed. Finally, the utility and the nature of heuristic methods for finding optimal solutions to these and similar problems are sketched, and suggestions are made as to probable fruitful means of developing better planning systems.

•TRANSPORTATION planning is an activity that has long-term results. The facilities that are put in place now will still be operational to a large extent in the year 2020. In fact, a review of the history of many metropolitan areas shows that trails laid out by Indians and early settlers are still main channels of communication and transportation. Only in recent years, with major investments in the Interstate System, have some of these long-standing patterns been destroyed or modified, and it seems likely that new patterns established by new modes of transportation will have the same permanent effects as the early establishment of primitive trails.

One of the principal reasons for the persistence of channels of movement is what might be called the intensification effect of the interaction between land use and transportation. Principal transportation routes attract activities, and the growth of activities requires the improvement and expansion of transportation routes or the provision of supplemental and parallel facilities. This positive feedback guarantees the persistence of some patterns of activity and provides a major problem in planning.

Despite the long-term nature of transportation planning in principle, some recent developments have cast doubt on the utility of transportation planning as it has been practiced in recent years. These doubts arise from a number of sources: citizen opposition to the environmental impacts of the automobile, local resistance to the disruption caused by new facility construction, and growing uncertainty as to the technical future of power production, propulsion systems, and transportation technology in general. These three factors and some others suggest caution in the development of long-term plans. On the other hand, such long-term planning cannot be abandoned because of the anticipated size and durability of investments and because future provisions for transportation, except by air or tunnel, require the establishment and preservation of long and continuous corridors.

Federal and state governments and transportation planners have implicitly realized that these considerations are not to be taken lightly. Twenty percent of the U. S. gross national product is devoted to transportation. It has been fashionable to attempt to build a Machiavellian theory of transportation difficulties based on the magnitude of the supplier interests—the oil companies, motor companies, highway construction lobbies, and state highway officials. Although these vested interests do exist and undoubtedly contribute to the institutional inflexibilities with which transportation planning has to deal, there is another side to the problem. The fact that so much is spent on transportation directly implies that transportation plays a very important and positive role in the organization of our economy and our private lives. Such powerful economic forces, viewed from the consumer side, cannot be manipulated in either the short or the long run without serious large-scale investments in the planning and operation of a variety of facilities. It is also obvious that a very substantial intellectual and organizational apparatus has been developed for dealing with these problems of planning and providing facilities.

There are several major shortcomings of the planning process. The relative effort invested in detailed planning and engineering design of facilities as compared with the overall design of systems is disproportionately large. The emphasis in analysis and planning has been too responsive to the popularity of the automobile and has not until recently given adequate attention to other modes of transportation. There is an institutional anti-urban or at least pro-rural bias in the United States that has influenced the provision of transportation facilities. In transportation planning for urban areas, inadequate means have been developed for joint planning of transportation and nontransportation facilities, and the impact of transportation on land use has not been adequately accounted for.

I have taken the position for some time that the transportation planning process as currently conducted is up to a point a phenomenally successful and well-conceived enterprise. Let me define some of the best elements of this conception before defining some of the points where a shift of emphasis is needed to ensure adequate further progress. In this discussion, I shall focus principally on metropolitan-area transportation planning.

The origin-destination (O-D) survey and the approach to transportation analysis that arises out of it are remarkable examples of a type of behavioral social study. The uniqueness of such studies is especially remarkable because their principal development came from engineering.

The use and manipulation of the masses of data produced by O-D surveys and other parts of major transportation studies required the development of substantial competence in data management and manipulation. These competencies expanded into the field of computer utilization, from which has come computerized data manipulation and computerized models.

The models of transportation behavior that are a well-standardized part of most metropolitan transportation studies are to a large extent more complex than a great many models of related types that have been devised or suggested by economists and sociologists. I maintain that these models are essentially behavioral and that they are frequently superior to the substitutes proposed by critics. I am not aware of any other field in which such massive detailed projections can be made by reasonable and systematic means. In spite of this overall endorsement of the general package of transportation demand projection, I still have many reservations. One such reservation is that the models appear to be lacking in generality. If this were not so, a single model package could be applied without new surveys to almost any metropolitan area in the country. This lack of generality is of utmost importance in considering long-term projections making use of new transportation technologies and new land-use arrangements; if a predictive procedure cannot be generalized in 1970, it is very difficult to see how it can be applied to 2000 or 2020.

The evolution of transportation demand modeling has followed a familiar inductive paradigm. In the earlier stages of O-D surveys (and continuing to some extent to the present) relations were established through simple statistical models (for example, the classical gravity model). The computer has greatly expanded the capability of transportation studies to deal with large masses of data and substantially expanded numbers of variables in this most simple statistical framework. Such statistical analyses following on data collection constitute, of course, an important and possibly essential step in the development of a predictive capability. This step, however, is not complete without the use of models in a more sophisticated sense. This sophistication takes two distinct but related forms. The first of these is computer simulation of large systems. The prototypical example is Morton Schneider's original program for trip distribution and assignment. Typically, such a large system simulation makes very heavy use of computers and permits the manipulation of very large numbers of elements. I have purposely not tried to make a narrow definition of this system simulation concept because it exists independently of the content of problems in the mind of the model builder, although its specific form is in each case ultimately determined by the nature of the problem that is being solved. We now know that there are perhaps three major formal questions influencing the structure of these models, each of which must be answered in a larger context. These questions are whether a dynamic model is needed, what types and levels of aggregation are to be permitted, and what importance is to be given to stochastic events.

The second aspect of modeling has to do with content. A typical example in this area might be a generalized model of modal choice. This example illustrates the fact that, after 30 years of experience with transportation modeling, 15 of them quite intensive, there are many aspects of behavior in transportation demand that are still inadequately understood by transportation planners. The example also shows that there is a tendency for basic research in the transportation field to be driven from the level of aggregated and descriptive models to the level of individual and household behavior. Finally, these models will in all probability turn out to be not generally susceptible to the naive statistical methods that were in vogue 15 years ago; they will probably require concepts and methods having to do with nonlinearities, discontinuities, and other troublesome aspects of realistic models of individual behavior.

The partial solution of all of these problems in the prediction of transportation demand (and the difficulties that arise in trying to extend these successes) gives impetus to the development of new models in the field of urban land use, the delivery of urban services, and some aspects of urban social interaction. This field is deeply indebted to transportation planning for data, statistical methods, computer systems, and the initial steps in understanding spatial processes. The need for solving some of these problems has risen in transportation planning from at least two sides. First, it is now quite clear that transportation is an intermediate service that meets defined social and economic needs and, as such, cannot be considered in isolation from these needs. The original transportation study land-use projections recognize this interaction in an elementary way, but the need for detailed knowledge of the functioning of the system has increased as problems of equity have come to the fore. On the other hand, the development of land uses in response to the provision of transportation had unanticipated consequences on the performance of transportation plans. Plainly, it is beginning to be recognized that the general purpose of planning is to improve jointly the system of transportation services and land uses and that each may be used as an instrument to influence the behavior of the other. All of these considerations have led to the development of locational models that are partly related to and partly independent of transportation planning and transportation analysis.

If we take a broad view of all of this work, we can be reasonably well satisfied with the extent to which such planning is widely understood and widely disseminated through the highway engineering and highway planning profession, partly through the efforts of the Highway Research Board. We must be disappointed that the land-use modeling effort has not received the same systematic development and dissemination. We must still be dissatisfied with the nature and limitations of some of the models currently in use, but most particularly we must define and acknowledge a specific limitation of this work with very far-reaching consequences. Almost without exception, data collection, analysis, and model building serve two important planning purposes that are necessary but not sufficient for a successful planning process. The first and perhaps minor purpose is to establish a baseline description of the status of the system and the metropolis at the beginnning of the planning period. The second and dominating role of these models is to predict the performance of plans or proposals. Subject to the many qualifications mentioned previously, transportation demand models in particular can now project the response of the system to major changes in the system itself and in the environment. This is done at a scale and level of detail that is remarkable for social science modeling; however, the entire apparatus stops at the point of making predictions. The planned changes that are the object of policy-making are entirely outside the modeling system. It is now appropriate to turn to the source of plans and to discuss the process by which they could ideally be generated.

Two different major views of the objectives of transportation and land-use planning may be developed depending on personal predilections and roles within the planning process. A short-range view of the planning process emphasizes the main constraints that have been previously mentioned. In the light of these constraints, it is sometimes difficult to find a plan that can feasibly be applied with any hope of using available funds or meeting a subset of local needs or both. In a variety of ways, transportation planning viewed in this way is very constrained, and the problem to be solved is only that of finding a feasible solution.

The difficulty with this type of planning is that its continued exercise may lead the total system in undesired directions. I therefore lean to the second view, which maintains that, over the long run, major changes can be made in the total system. In effecting these changes, dealing properly with the constraints is an important activity. If necessary, redefining them or removing them can be accomplished. Viewed in its totality, long-term planning attempts to approach an optimal solution to the problems with which it is designed to deal—in this case, transportation and land use. Such an effort has to take into account resource and social constraints and the costs of actually searching for an optimal solution. There are many indications that the principal thrust of public policy is in the direction of optimality rather than feasibility. Stylized procedures such as benefit-cost analysis, cost effectiveness, and program evaluation are all designed to focus public action on the most effective use of resources. A similar result is also achieved through emphasis on "balanced programs," in which no more efficient allocation resources can be found by transferring expenditures from one item to another. In what follows, therefore, despite many important qualifications, I will treat the problem of planning as if it were a problem in optimization.

The principal paradigm for optimization (and a most useful one for discussing the structure of the planning process) is mathematical programming. At a later point, I shall suggest that planning as it is and should be practiced cannot conform with this paradigm, but, at this point in the discussion, it is necessary to develop and fix ideas with respect to the nature of optimization. Every mathematical program has a handful of principal features whose analogs are in most cases easily recognized in the planning process.

Each program has an objective function or measure of performance that must be maximized or minimized. In planning parlance, this represents a weighted set of goals or, in more sophisticated terms, some sort of social welfare function. There are many difficulties in composing such an objective function, and these are especially acute in a pluralistic society and in times of relatively intense social conflict.

Mathematical programs are subject to some set of constraints. These constraints may be social, political, economic, or natural. Very frequently the constraints represent social goals that are established outside of the program and for which, beyond certain levels, no trade-offs are permitted. In most cases, the imposition of constraints makes it easier to solve a mathematical programming problem, but, at the same time, these constraints foreclose choices that might be important in the planning process.

In addition to providing an objective function and constraints we must frequently structure the problem in some particular manner. These structures have two different roles. In the first instance, they may be purely definitional or mathematical and serve the purpose of framing the problem so that it is more easily solved. In the second place, they may involve some correspondence with the real world, for instance by expressing the hierarchical nature of a metropolitan governmental organization or of a highway system. Frequently, the natural structure provides a basis for solution simplifications, as when the hierarchical nature of a problem permits a technical decomposition into interacting subproblems.

Next, every mathematical program has a systematic procedure for searching the solution space, by which it is guaranteed that the optimum will be found. We define solution space as all possible combinations of decisions that do not violate the constraints. One of the principal objectives of mathematical programming is to specify a means by which this optimum may be found by eliminating many solutions on logical grounds rather than examining every individual one.

Finally, and most important for purposes of this discussion, every mathematical program has to include an evaluation process by which, as the successive solutions are examined, their value or performance is established and a basis is laid for searching for the next step in the improvement process. Ordinarily, in mathematical programming, calculation of this new objective function is very simple. In linear programming, for example, it arises automatically out of the selection of each successive improvement of the solution.

It is the nature of large combinatorial problems that the number of possible solutions is extensive and that considerable attention must be given to all of the foregoing aspects of the problem of finding an optimum solution. In transportation and land-use planning, the number of variables and interactions is very large, and even the simplest possible formalism—that of linear programming—can readily generate complex problems. If in addition we add other conditions that generally exist in these types of problems, the number of steps in a solution becomes still larger. These complications include nonlinear objective functions, nonlinear constraints, zero-one or integer values for the variables, and multiple local optima. It may be categorically asserted that, for the overwhelming bulk of these problems and even with the simplest possible calculation of the objective function, it is impossible to explore all local optima and to find the optimum optimorum.

There is, however, one main and related subsidiary point of overwhelming importance when we consider the relation of the foregoing paradigm to transportation planning. The principal point is that the evaluation of the worth of a transportation and land-use plan is a cumbersome and extended process. For even a simple number of evaluations using currently existing techniques, scores of thousands of dollars worth of computer time and scores of man-years of staff time are necessary to specify elaborate plans, predict and tabulate the results, and evaluate these predictions according to some standards of decision-making. The subsidiary aspect of this problem is that the current techniques for predicting impacts on transportation and land-use plans do not lend themselves well to generalizations and simplifications. Thus, if we ask what the relative impact of two different levels of capital budgeting for transit systems would be on the city of Philadelphia, we would probably receive an answer that this requires the complete evaluation of selected proposed plans embodying these levels of expenditure. Some procedures of plan-making urgently require the ability to make decisions at a high level of generality to eliminate or "bound out" certain lines of development. In the absence of generalized evaluation techniques, the entire planning process becomes even more difficult.

We may now express one of the most difficult aspects of the urban metropolitan planning process in terms of a rather straightforward contradiction. On the one hand, our present tools for the analysis of proposed plans are quite accurate, but they are elaborate and cumbersome. We have no easy way of analyzing the impacts of either small changes in plans or decisions at the most general level. The available resources therefore permit the exploration of only a very few well-developed cases. On the other hand, a complete optimizing process involves very extensive explorations of possible solutions. Even in those numerous and quite general cases in which a complete implicit exploration is impossible, ordinary prudence would dictate that we explore a substantial number of useful plans, including some rather "far-out" solutions, before developing a limited number of final schemes in detail. A failure to follow this procedure most probably results in overlooking important and innovative solutions to problems that might usefully receive more consideration.

In general, my conclusion is that there is an imbalance in effort between the improvement of plan evaluation methods (including the prediction of demand) and the improvement of planning methods themselves. I would not recommend any cutback in the first effort because the total resources devoted to these two developmental activities still fall far short of a desirable level. There is a great deal more room for improvement in the design of our systems than present analysis and design or planning techniques can achieve. My general suggestion therefore would be that what is needed is a moderate augmentation of research in prediction and plan evaluation and a considerable increase in the investigation of planning methods. In the remainder of the paper, I will discuss some of the more salient aspects of planning methodology and possible steps toward its improvement.

It is obvious that, when confronted with the paradox just discussed, the average transportation planning study has a number of systematic methods for reducing the contradiction to manageable proportions. One such method is to use simplified models of prediction and evaluation, but this option is not openly available although we will see that it appears to be implied by some other simplifications. Most of the reduction in effort in exploring a wide range of possible plans is done by paring down the choices that are believed to be useful. It is apparent, therefore, that transportation planners have a hidden agenda by which planning choices are narrowed down and a final limited number of sketch plans are arrived at. The principal difficulties with this hidden procedure are the following. First, because the plans are not publicly known, they cannot be criticized by those interested in the outcome of the transportation planning process. Second, because such plans are arrived at in private, it is impossible for interested members of the public to intervene at the early stages. Third, because the process is somewhat personal and individualistic, it cannot easily be replicated. Thus, different planners might achieve basically different results. Fourth, because the procedure is not explicit and well-defined, it cannot be validated or usefully employed by others to vary the starting assumptions and achieve differential results in a systematic way. Fifth, as in all of the preceding cases, it is difficult to systematically transmit knowledge about such hidden planning methods, and the instruction and training of good planners are extremely difficult. All of the foregoing argues for the idea that planning should be conducted by a process that is well-defined, publicly known, open to examination and intervention at various points, and reproducible and that has a clear separation between those parts that depend on individual judgment and those parts that may be considered automated or computerized.

There are two principal forces driving transportation planning in the direction of a more completely specified procedure along the foregoing lines. The first of these is the increasing public concern over the way in which transportation plans are developed and over their impacts on neighborhoods and on the environment, and the second is the increasing difficulty and complexity of transportation planning. Such difficulty and complexity arise out of the increased number of choices that can be made in an affluent society and out of the technological uncertainty regarding the future of transportation itself. In order to understand how such a policy might be more specifically articulated, we can compare some of the things that planners actually do with some of the processes that arise in the formulation of mathematical programming solutions to the problem of optimization.

The formulation of the objective function is equivalent to the definition of social goals and is receiving increasing attention in many aspects of governmental planning. The advance formulation of goals proceeding from the abstraction of general social welfare down to concrete operational policies is an exceedingly difficult process, precisely because it is approached in the abstract. Fortunately, planning is a cyclical process, and the actual procedure of articulating plans and submitting them to public discussion tends to clarify the nature of the goals held by the planners, decision-makers, and public at large. This particular aspect of feedback in the planning process deserves substantial strengthening in transportation planning. It is quite true that there is a large-scale public desire for improved highway transportation that has been recognized by transportation planners in the Federal Highway Administration. At the same time, however, the attention to public thinking in the content of transportation plans and concern with alternatives both within the automotive system and between the automotive and other systems have been totally inadequate. The formulation of goals and objective functions is not, however, the principal part of the process with which I am now concerned.

Planners customarily develop constraints that, in one or another sense, reduce the number of possible solutions to their problems and in all likelihood simplify the solution in other ways. In connection with transportation, these constraints may be budgetary, legal, customary, or physical. All of these constraints are subject to change in one way or another, and, if the costs of the changes could be specified, they could be removed from the constraint set and placed in the objective function. This would permit greater flexibility in planning so that a wider range of choice might become available. The formulation of constraints therefore represents an advance decision by the transportation planner that, outside of certain bounds, the costs or discontinuities of selected policies are excessively burdensome. For example, the idea of congestion pricing of highway facilities is ordinarily ruled out of plan formulation and testing because it is currently not legal in most aspects of federal highway construction. In addition, this legal provision is based on a long-standing customary tradition, and changing it might involve considerable political difficulties. Finally, the technical problems of charging and enforcing congestion pricing are considerable. In the short run, all of the reasons for maintaining a particular constraint on transportation planning are valid. Many constraints of this type also gradually arise as standards of engineering practice and are applied almost without thinking by transportation planners. In most cases, these professional standards are probably well justified, but in some they may have outlived their usefulness. A constant flexibility as to the possibility of changing standards and constraints should be a part of the transportation planner's operating rules, and every effort should be made to specify both implicit and explicit constraints so that the concerned public may understand the rationale behind some aspects of transportation planning.

The most troublesome part of transportation planning involves the development and testing of an adequate variety of alternatives. This difficulty may be said to arise at every level in the planning process, from the smallest elements of facility location to the largest aspects of total system design. To suggest that there are various levels in the process already anticipates the suggestion that it is probably possible, at least in certain respects, to break down the planning of the transportation system in a hierarchical fashion. It also appears likely that a hierarchical breakdown corresponds in its structure to certain large-scale engineering aspects of the problem. We may point out that this is not necessarily the case, although its logic is embedded in a great deal of transportation planning and analysis. The decomposition could be hierarchical by political jurisdiction or in some other fashion by type of movement such as people versus goods and trip purpose.

Decomposition principles for solving large mathematical problems are gradually becoming more important and can often be implicitly related to the practical decomposition of problems both in the real world and in the planning process. Three important features of this decomposition must be borne in mind. First, the system that is being decomposed should itself be adequate in size for dealing with the total problem, properly defined. Second, the decomposition should facilitate rather than confuse or complicate the solution of the problems. Third, there must be a reciprocal iterative relation among the different levels of the decomposition. The last provision means that we cannot plan lower level systems once and for all without referring back to the larger context in which they are embedded and evaluating the larger system. This evaluation may impose changes on our previous plans for the lower level systems. It seems probable that one source of public dissatisfaction with transportation planning has been inadequate attention to the interaction among subsystems. The decomposition occurs at a very high level in the federal government, and what might be called "recomposition" at the local level, where the systems interact, is very difficult. In addition, we should note that, because the federal government has very little responsibility for land-use planning, this aspect of the system is not automatically included in the total problem subject to decomposition.

A hierarchical approach to decision-making facilitates the process known in mathematical programming as "branch and bound," by means of which large classes of solutions are ruled out. If it can be readily shown that certain combinations of high-level decisions are impractical or have a very low benefit-cost ratio, all the subsequent decisions that might depend on these can be aborted. Thus, for instance, a large-area metropolitan transportation plan that calls for all transit or all automobile facilities could automatically be excluded. The difficulty with these large-scale exclusions is that they depend very substantially on planning intuition and not on a direct evaluation of their implications. We urgently need predictive methods that can evaluate a partial statement of a plan rather than a fully developed and articulated plan. Such evaluations ought to be scientifically based and open to public inspection. Obviously also, as with all other prediction methods to be discussed, speed is an essential element in guaranteeing the capability of exploring a large number of possibilities.

Some principal large-scale options in urban transportation planning are configurational in nature. A typical example of this general approach is the year 2000 exploration for the Washington area. In these explorations, the gross interaction between land use and transportation was made perfectly apparent and was to some extent systematically explored. We need, for each particular case of configurational planning of this type, a method for specifying different configurations in a meaningful way that facilitates systematic explorations. In giving a related illustration of the difficulties in this matter, Marvin Manheim offered a hierarchical approach to highway route location that started at the highest level with the assignment of broad bands of location for every exploration. The possible number of these bands is infinite in continuous space, and no systematic procedure was proposed for exploring them without either major duplication of effort or major omissions of likely potentialities. In general, these are the twin dangers of any ill-defined exploratory procedure.

Even better definition will not completely eliminate the possibility of missed combinations. At some level of decomposition of a general planning problem, a level of detail may be encountered where there is some hope of actual optimization. I specify that this is largely a hope because, in the overwhelming majority of practical cases, the hope cannot be fully realized. Nevertheless, subject to the conditions established by higher level planning assumptions, certain problems can be examined in some detail, and fairly firm plans can be developed. What is too often forgotten is that these detailed plans depend in very large measure on the assumptions of the decomposition. As the planning problem is reexplored with a different combination of high-level assumptions, the subsystem optimization should produce different results.

A simple illustration that provides very many interesting sidelights is the problem of route location that constantly arises in highway and transit planning and that has generated many of the most difficult current political problems in plan implementation. This problem was explored graphically by Alexander and Manheim in a manner somewhat different from the more systematic treatment by Manheim mentioned previously, but these graphic methods have been used in a number of other situations including some criticisms of route location decisions mounted by citizen groups. If the sole problem is to connect two separated points by a facility, the graphical methods involve using a set of overlays that show impediments to route location at various levels of intensity. These may be natural physical features, cost of land acquisition, environmental damage, destruction of historical monuments, concentrated political opposition, and so forth. These graphical representations can be overlaid and "eyeballed" to select what may be believed to be a superior or even optimal location. In this simple form, the problem is easily converted to a dynamic programming minimum-path problem that can be solved very rapidly with current computer techniques. It would be quite possible to vary the weighting of the different impediments to route location so as to express the different value systems of participants in disputes. These might then produce a variety of different route locations that could be examined and discussed much more intelligently than has frequently been the case.

It is very rare that a complete optimal solution of the type just described can be found and implemented. Even the simple route location problem rapidly becomes more complicated in a real-world situation. First, there are many hidden and complicated features that may be overlooked in generating data for a model of the type described. Second, there are impacts, such as community disruption, that we are not yet able adequately to measure and model. Third, the problem as defined neglects, for example, the important aspect of service to intermediate points. The real nature of the problem therefore rapidly escalates to one that must be solved by so-called heuristic methods. It is at this point that we need a more active effort to specify what can be done by computer and what involves human intervention. Also a more precise specification is necessary of what form the intervention will take.

A similar example is the problem of network optimization. Here again, there are no completely successful optimizing models. Branch and bound techniques have been found to be excessively time-consuming on all but the smallest problems. The very interesting optimal spacing suggestions of the Chicago area transportation study are not deterministic with regard to the actual location of network links. They provide a general concept of how a system may be brought to a balanced state where the benefit-cost ratios are uniform throughout the system. Here again, heuristic techniques are urgently needed.

Probably the most important single element of heuristic search is a means for improving given solutions systematically. In the more complicated route location problem, this might be a systematic means for making small displacements of different parts of the route that would cumulatively lead to a locally optimal solution. In the optimal network problem, such an improvement method would most likely be swapping, or the deletion and addition of links to the system, once again leading to a local optimum. The essential problem in each such case is to formulate the problem correctly: first so that a systematic improvement may actually be hoped for and second so that the computation of these improvements is extremely rapid. Exactly this form of systematic improvement is used, for example, in linear programming, but it is not heuristic because it is guaranteed to find an optimal solution.

This observation leads me once again to reemphasize the local nature of optima achieved by stepwise improvements of plans. As a simple example, if the route location problem is being solved by incremental adjustments and the route has been located on the wrong side of a mountain, it will probably never be moved to the right side. It is thus highly probable that, even for relatively low-level optimization problems, the difficulty of exploring distinctively different alternatives still exists and can be very troublesome.

It may also be well to reemphasize at this point the fact that the optima achieved in solving a lower level problem depend very much on the terms in which those problems are framed, that is, on higher level policy decisions that may be involved in a decomposition procedure. The optimal network problem obviously depends on land-use and locational decisions, overall level of spending, various constraints, and the way in which the objective function is formulated at the high level and disaggregated for application to the particular subproblem. Similar observations could be made about the route location problem. In the simple problem, the objective function is probably the principal determinant of route location. As the problem is made more complex, all of the other features that have been discussed may gradually enter into its solution.

If it were desired to optimize land uses and location with the transportation system fixed, similar decomposition problems arise. It seems likely, for example, that, except for social externalities or social preferences, residential location patterns by themselves could be optimized by a linear programming approach. On the other hand, the industrial assignment problem of locating interacting industries is a quadratic programming problem that becomes very difficult to solve for large numbers of locators. This quadratic programming problem could be extended jointly to include the location of residences and workplaces, once again given a fixed transportation system. These problems are in their own way every bit as intractable and difficult for land-use planners as the problems that I have discussed previously are for transportation planners.

I mention these land-use optimization problems because, for large urban metropolitan areas, it seems likely that the real planning problem is not to optimize either transportation or land use but to optimize them jointly. Curiously enough, this idea was proposed in a more limited context in a memorandum from Robert Murray Haig to the New York Regional Plan Association almost 50 years ago. He suggested that the best urban plan would be the one that minimized the total of transportation costs and land rents. This simple concept was proposed as a subject of study for the Association, but was not what Haig himself was able to carry out. The roughly described solution method corresponds in a general way with linear programming, which was not developed until 20 years later. Haig was principally concerned with land use and did in fact assume that the transportation system was fixed. This, however, is by no means a necessary assumption, and, for sufficiently drastic changes in land uses, it is obviously untenable. I am not aware that anyone has proposed a practical means for systematically tackling this combined problem, let alone a rigorous one that would produce truly optimal results.

I hardly need emphasize further the fact that the size and complexity of metropolitan planning problems, together with their nonlinearities and discontinuities, make mathematical programming solutions of the global problems largely infeasible. This unfortunate fact greatly magnifies the importance of heuristic methods. In the context of this discussion, these heuristic methods introduce two more or less distinct acts of an artistic or creative nature into the planning process. Neither of these creative activities can be made into an explicit and reproducible planning process. The best that can be done is that they may be justified after they have been completed on the basis of general acceptance.

The first of these acts is the design of the heuristics themselves. They can be rationalized or sketched out in the terms that I have used in the introductory portions of this paper, and their justification may be more firmly established. Heuristic methods will ordinarily contain in the second place steps of human or planning intervention where the inputs are also creative and where the final justification can only be on the basis of results. Here, however, there is a subsidiary point of very great magnitude. Not only are some suggestions bad, but a preoccupation with mediocre suggestions may prevent finding a really first-rate solution. For this reason, the importance of brainstorming and counterplanning should probably be enhanced so that alternatives may be generated outside of the planning process itself. A greater openness on the part of transportation planners to this kind of intellectual and popular input should be a final and most important ingredient in a new style of transportation planning.

MODEL FOR ESTIMATING REGIONAL AIR PASSENGER TRAVEL DEMANDS

John C. Goodknight*, Florida Department of Transportation

The model developed by this research uses empirically determined relations between total travel and regional socioeconomic activity as the basis for estimating the pattern of intercity travel by all modes. Origin-destination surveys of both intercity highway traffic and air passenger traffic together with census data provide the basis for determining these relations. A comparison of costs for air travel versus surface travel between each pair of communities within the study area provides the basis for estimating the modal split. Comparisons of estimated air travel and observed air travel for all Texas cities with commercial air passenger service in 1967 indicate that the model provides reasonable estimates of air passenger travel generated by individual communities. Although discrepancies between estimated and observed volumes do exist and future refinement of the model may be warranted, it is recognized that the potential demand for air travel is not the only factor used in resource allocation decisions. Social, political, and economic factors will unquestionably continue to influence decisions concerning the development of the air transportation system.

•UNTIL recently, there has been little air transportation planning activity at the state or regional level. However, recent developments in the aviation industry, including increased federal emphasis on regional air system planning and rapid growth of the third-level carrier systems, have brought the need for more refined and powerful tools for estimating regional air travel demands sharply into focus.

Techniques that have been applied at the local level and at the national or international level have been oriented to the large cities and do not provide suitable information for decisions concerning services and facilities at the small communities. A recent draft of a planning document prepared by the Federal Aviation Administration emphasizes the need for more refinement in the techniques for air transportation planning. In discussing methods for the estimation of regional demand for air passenger service, this document suggests that, "This is a fertile area for research...which remains as a future effort" (1).

PURPOSE AND SCOPE OF THIS STUDY

The purpose of this research is to develop and test the feasibility of a technique or model for estimating the magnitude and geographical distribution of demand for commercial air passenger travel. The term "demand" is frequently associated with the price-quantity relation. As applied to the model developed in this study, however, demand simply refers to the volume of traffic that would be generated under a specified set of relative prices, transportation system configuration, and pattern of regional socioeconomic activity. Because the nonhub airports and their connecting routes constitute vital components of the statewide air transportation system, this model is intended particularly for estimating potential demand at communities that currently have

---

*This research was performed while the author was affiliated with the Texas Transportation Institute.

no commercial air service or have only limited service. However, it is also intended to apply to communities with well-established service.

Particular attention has been given to developing the model so that it is sensitive to the influence of the quality of air service on the volume of air travel generated, particularly for smaller communities with no air service or with only limited air service; the impact of changes in service to one community on the demand for and economic viability of service to nearby communities; and the complementarity between air transportation and other modes.

DESCRIPTION OF THE MODEL

Summary of Operational Characteristics

Figure 1 shows the general operation of the model. Essentially, it operates by first estimating magnitude and distribution of all intercity travel (irrespective of mode) generated by the region being studied. It then estimates the modal split for each city pair on the basis of comparative costs (which include both direct out-of-pocket costs and time costs). Travel cost calculation allows for variations in the traveler's income, trip purpose, and number of persons traveling together. Finally, the model tabulates the total number of air trips on each link of the air network.

Network Delineation

In delineating zones for this analysis, the entire United States is subdivided into a series of zones. The state or region for which the travel estimates are desired (e.g., Texas) is subdivided in greatest detail. The smallest geographic unit used in testing the model was the county. At greater distances, larger zones were used as shown in Figure 2.

The air transportation network, represented for this analysis as a series of links and nodes, is based on the route descriptions given elsewhere (2). It is, of course, simplified and shows detailed linkages only within the area of interest (Texas) but also includes linkages between Texas and other major cities.

In developing the model, it was assumed that intercity passenger travel is limited to two modes: commercial airline and private automobile. Because of the relatively small percentage of passengers carried by bus, rail, and water, this assumption is appropriate for areas such as Texas. (Trip generation factors developed for this model exclude general aviation travel.) Because of the ubiquity of the highway system, it was further assumed that the highway network is continuous (i.e., it is not described by a series of links and nodes but by only the coordinates of the cities representing each zone). The coordinates provide the basis for calculating the mileage between any pair of cities and for estimating the travel costs.

Estimation of Total Intercity Travel

Three principal sources of data provide the basis for estimating the pattern of total intercity travel: the U.S. Census of Transportation (3), the annual Origin-Destination Survey of Airline Passenger Traffic (4), and various origin-destination surveys from the urban transportation studies for the study region.

The principal application of the census data is in describing the characteristics of the trips and trip-makers (i.e., trip purpose, income, and so forth) although the census does provide information relating to the amount of travel by individuals. The origin-destination data, on the other hand, provide more relevant information on the total magnitude and the spatial distribution of travel. Table 1 gives the trip generation characteristics determined for the study area.

The technique employed here for estimating the distribution of travel consists of enumerating all possible destinations (or origins) for trips produced at a given base zone (in this case, a county). Each of these possible destinations (or origins) is assigned a factor that indicates its attractiveness to trips to or from the base zone. These factors, multiplied by the total number of trips produced at the base zone, give the number of trips between the base zone and each other zone. By repeating this

Figure 1. Analysis of commercial air passenger demand.

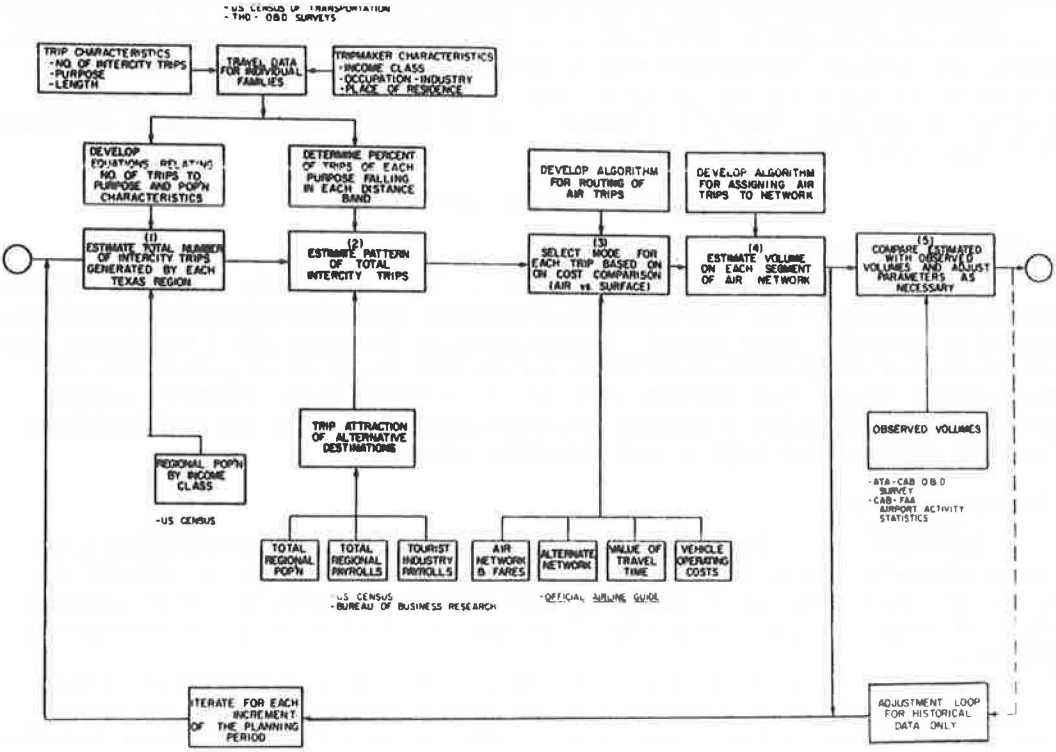


Table 1. Intercity travel by air and highway for selected Texas communities.

City	Estimated Population (1967)	100 to 199 Miles		200 to 499 Miles		500 to 999 Miles		>1,000 Miles	
		Total Trips	Trips/ Person	Total Trips	Trips/ Person	Total Trips	Trips/ Person	Total Trips	Trips/ Person
Population of More Than 100,000									
El Paso	334,134	1,428,808	3.94	754,259	2.07	965,474	2.67	340,311	0.94
Population of 20,000 to 100,000									
McAllen	79,006	221,678	2.75	340,230	4.03	46,355	0.59	55,974	0.71
Sherman-Dennison	51,576	398,653	7.75	241,995	4.68	184,581	3.58	14,637	0.28
Victoria	42,645	714,520	16.75	82,100	1.93	11,050	2.59	16,560	0.39
Borger	25,286	32,110	1.27	123,090	4.88	70,850	2.81	3,425	0.13
Temple	34,300	226,320	6.60	118,630	3.46	33,800	0.99	9,720	0.28
Killeen	26,500	125,800	4.75	57,500	2.17	39,100	1.48	47,900	1.81
Paris	25,200	67,670	2.68	182,210	7.22	10,320	0.41	350	0.01
Total	284,513	1,786,751	6.28	1,145,755	4.03	396,066	1.39	148,566	0.52
Population Less Than 20,000									
Port Lavaca	11,950	115,000	9.63	17,350	1.45	3,340	0.28	3,340	0.28
Childress	6,560	125,500	19.15	38,800	5.93	11,180	1.71	7,220	1.10
Athens	8,046	340,000	28.60	84,500	10.50	19,280	2.40	1,330	0.17
Mineral Wells	12,451	126,400	10.20	209,000	16.85	39,300	3.17	20,350	1.64
Stamford	5,418	291,500	53.90	166,200	30.75	10,800	1.99	1,350	0.25
Levelland	11,700	72,300	6.19	28,300	2.42	16,220	1.39	0	0
San Marcos	14,700	243,000	16.53	130,200	8.84	2,090	1.42	6,270	0.43
Total	70,825	1,203,700	17.00	674,350	9.53	102,210	1.44	39,860	0.56
Grand totals	689,472	3,419,259	4.97	2,574,364	3.74	1,463,750	2.13	528,747	0.77

analysis and using each zone within the study area as the base, the pattern of total travel is obtained. This procedure is shown in Figure 3.

Travel patterns for different trip purposes can be expected to differ. For example, on a nationwide basis, business travel will be strongly oriented to the concentration of business activity in the Northeast. Trips for recreation and entertainment will be more concentrated along routes leading to vacation centers such as Las Vegas or Miami, and the distribution of trips made to visit family or relatives can be expected to closely parallel the national distribution of population.

In synthesizing the travel patterns, the model provides for separate estimates for three trip purposes. These are given in Table 2 together with the socioeconomic activity measures used to indicate the relative attractiveness of a region for trips of each purpose. Table 2 also gives the distribution of travel by purpose for each of three income classes used in developing the multipurpose trip table.

Modal Choice Analysis

A traveler's choice of mode for a given intercity trip can be approximated by comparing the costs of the trip via the alternative modes available and then selecting the mode or combination of modes for which the perceived cost is the least. In the strictest sense, the true travel costs include calculable costs such as vehicle operating costs, air fare, and travel time costs and psychic factors such as convenience, security, safety, and personal preference.

The effects of psychic factors on choice of mode cannot be readily evaluated but, to some degree, can be expected to offset each other (i.e., one person's preference for air travel will be offset by another's prejudice against air travel). Furthermore, for business travel, which accounts for a large segment of the commercial air travel market, the choice of mode can be expected to be much more sensitive to calculable costs than to personal preferences. Indeed, the choice will frequently be made by the employer rather than the traveler. It would, therefore, appear that the error resulting from assuming that the net effect of these psychic factors is zero will not seriously affect the validity of the analysis.

The cost of travel by automobile between two points consists of two principal components: vehicle operating cost and value of passenger travel time. Vehicle operating cost is calculated from the distance between the two cities or nodes being considered. An average perceived operating cost of $0.05 per mile is used in this analysis because it represents approximately the fuel, maintenance, and repair costs. For business travel, a higher rate would be more appropriate; however, the value of travel time associated with business travel is relatively high. Thus, the time cost represents a larger fraction of the total cost, and the choice of mode is relatively insensitive to the rate used in calculating vehicle operating cost for business travel. The allocation of this cost among all passengers making a particular trip can be accounted for by considering the distribution of person-trips by size of the travel party in estimating the surface travel cost. Data from the U.S. Census of Transportation (3) provide the basis for this allocation.

In calculating the passenger time costs, this analysis assumes that value of travel time is directly related to the traveler's annual salary and to the purpose of the trip. Factors used to estimate these values as as follows:

Trip Purpose	Value of Travel Time
Business and conventions	Twice hourly salary rate
Personal business	Hourly salary rate
Recreation and entertainment	None

Average annual income for each class is taken as follows:

Income Range	Assumed Average Income
Less than $5,999	$ 3,000
$6,000 to $9,999	$ 8,000
More than $9,999	$15,000

In allowing for the fact that the value of travel time may differ for different persons traveling together (i.e., value of time for children is very low), the model calculates travel time costs on the basis of the following assumptions:

1. For business travel, all persons traveling together value their time equally; and
2. For all other purposes, the full value of time applies for the first person, one-half of this value for the second person, and one-fourth of the full value for all other persons.

Air Travel Costs

Figure 4 shows the scheme used for calculating the cost of an air trip. Any trip may consist of all or a portion of these segments. Thus, the cost of the trip is a function of the routing. On the other hand, the routing selected for a particular trip is a function of costs; therefore, both must be determined iteratively. In calculating these costs, appropriate factors are included to allow for terminal impedances and other costs encountered in traveling by air. These include both time and out-of-pocket costs.

Trip Routing

Figure 5 shows the operation of the algorithm used to determine the least cost routing through the transportation network. This describes the trip in terms of both automobile and airline travel. It operates iteratively and is an integral part of the modal choice analysis.

Assignment of Trips to the Network

The output of this analysis is a tabulation of air trips along each link of the network. The output also includes the total number of air trips generated at each city within the analysis area. Highway trips, however, are not identified.

MODEL TEST

Data Base Used for Testing Model

In testing the model, 1967 was used as the base year for comparison. Estimates of 1967 air passenger volumes were compared with data from the 1967 survey (4). The 1967 estimates represent the output from the demand analysis based on socioeconomic data for 1967 and the approximate configuration of the air transportation system at that time.

Comparison of Results

In general, it is to be expected that a generalized model would be more reliable for estimating the pattern of long-distance air travel than for short-haul travel. Beyond a certain distance, the relative attractiveness of alternative destinations seems to be little affected by distance. Linkages between specific cities or regions are of a general nature and do not describe specific ties such as between large trading centers and the outlying communities.

On the other hand, factors affecting the pattern of travel and the traveler's choice of mode for the short-haul market are more varied and difficult to completely describe in a generalized model. State capitals or other major governmental or institutional centers generate significant amounts of air commuter traffic. Similarly, major financial centers appear to be the focal point for "single-day" air travel from the surrounding areas. In certain circumstances, where topographic constraints impose major discontinuities on the highway system, there is also a greater tendency for short-distance travelers to use air.

Figures 6 and 7 show the comparisons between observed and estimated air travel volumes for long-haul (interstate) and short-haul (intrastate) air travel respectively. (These comparisons are displayed on a logarithmic scale for convenience.)

Figure 2. Delineation of external regions.

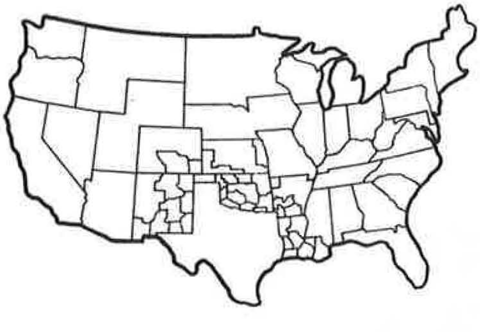


Table 2. Distribution of travel (percent) for each income level.

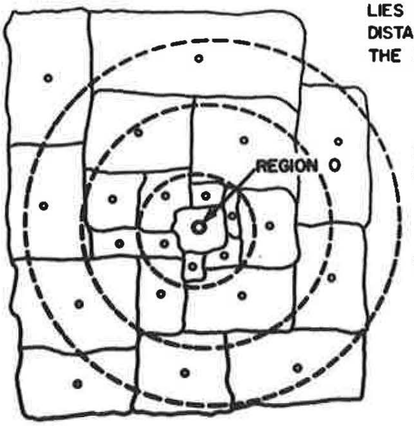
Trip Purpose	Associated Attractiveness Factor	Family Income Level: Less Than $6,000	Family Income Level: $6,000 to $9,999	Family Income Level: More Than $9,999
Business and conventions	Total taxable payrolls[a]	9.7	13.4	24.7
Visit friends and relatives and personal business	Total population	55.3	46.0	33.4
Recreation and entertainment	Hotel and motel payrolls[a]	35.0	40.6	41.9

[a]Payroll data refer to payrolls subject to social security taxation during the first quarter of the year. These data are taken from County Business Patterns, U.S. Bureau of the Census, 1967.

Figure 3. Method for estimating distribution of total travel.

NOTE:

IT IS ASSUMED THAT THE ENTIRE REGION LIES WITH THE SAME DISTANCE BAND AS THE NODE



The number of trips of purpose k generated by the base zone 0 and attracted to any other zone d is given by

$$T_{odk} = (R_{jk})\left(\frac{A_{dk}}{\Sigma A_{ijk}}\right)P_o$$

where

R_{jk} = trip generation rate for distance interval j and purpose k (Tables 1 and 2),
A_{dk} = socioeconomic descriptor representing the attractiveness of zone d for trips of purpose k (Table 2),
A_{ijk} = socioeconomic descriptor for purpose k and zone i within the distance band j, and
P_o = population of base zone 0.

Figure 4. Components of cost for air trip.

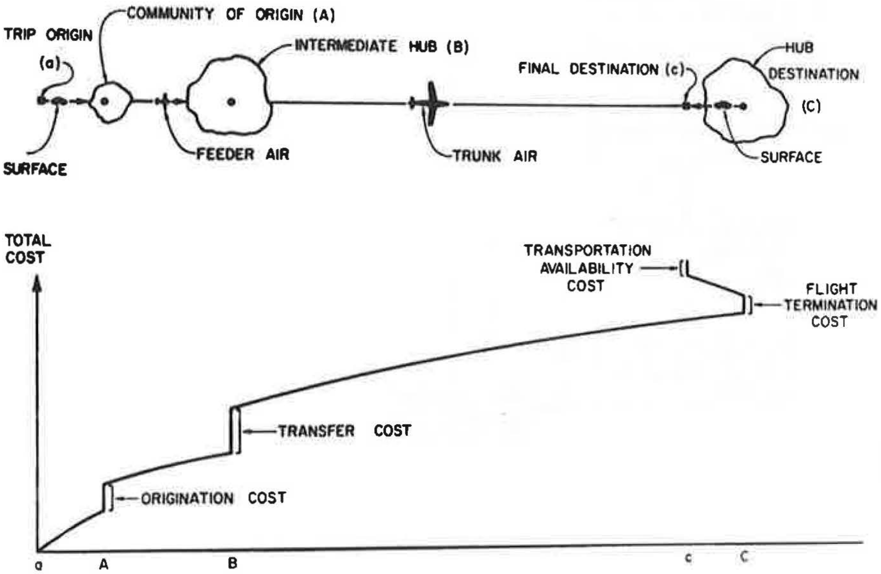


Figure 5. Algorithm for selecting least cost routing through air network.

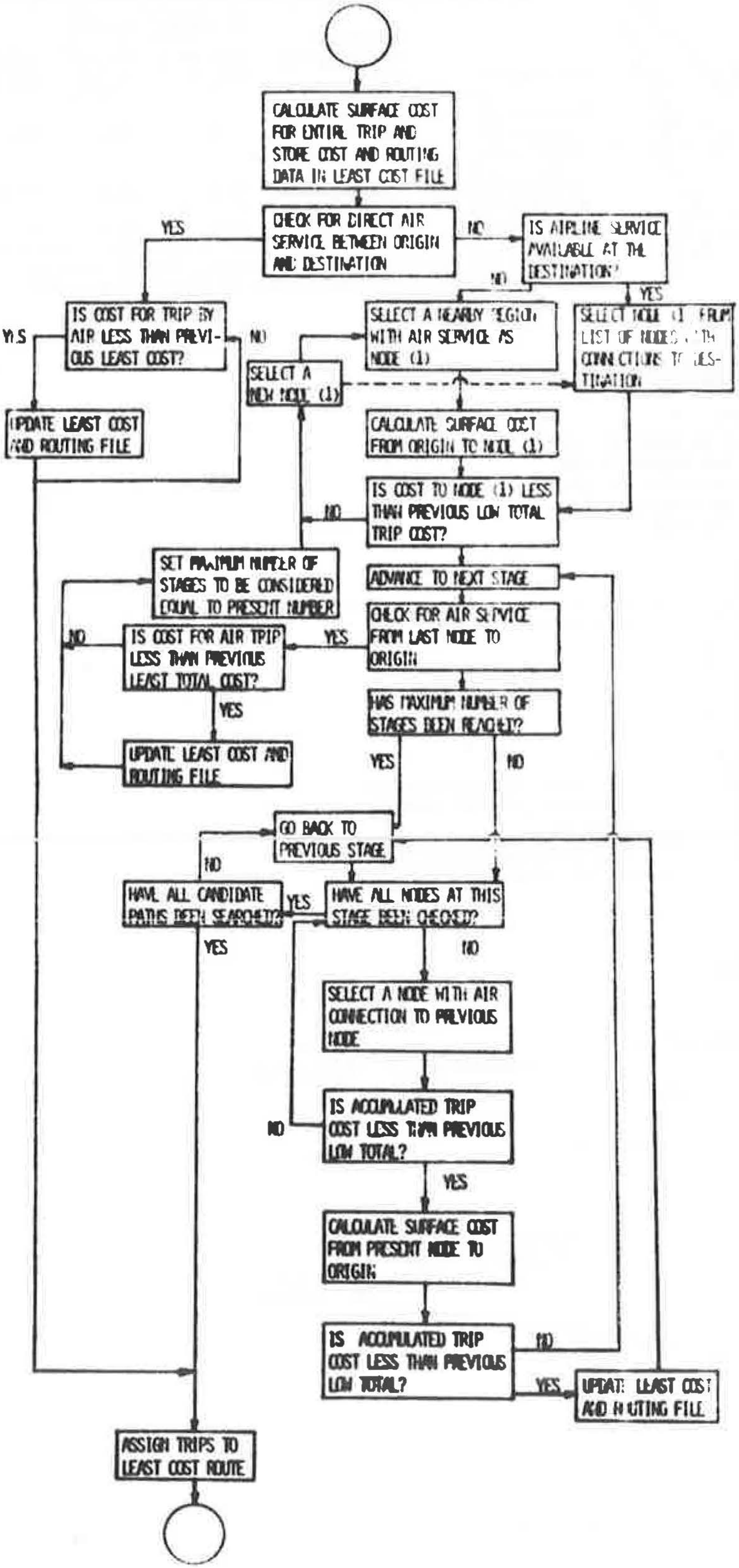


Figure 6. Comparison of estimated and observed interstate air passenger trips for 1967.

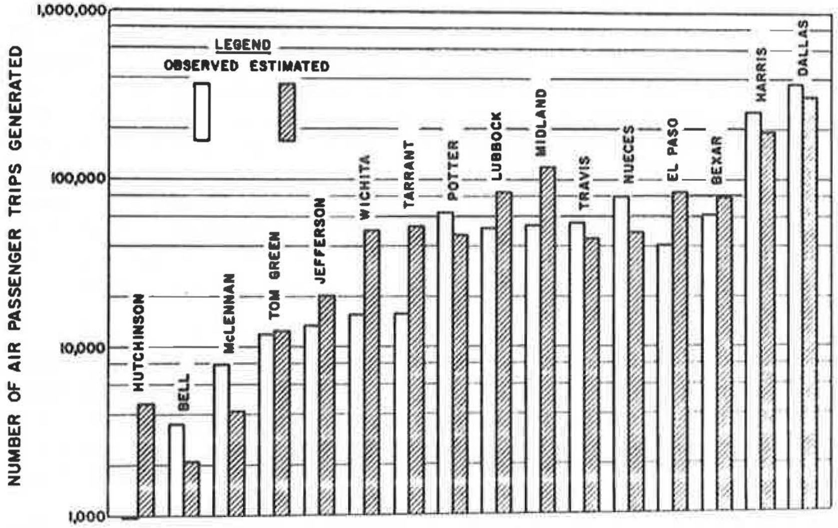
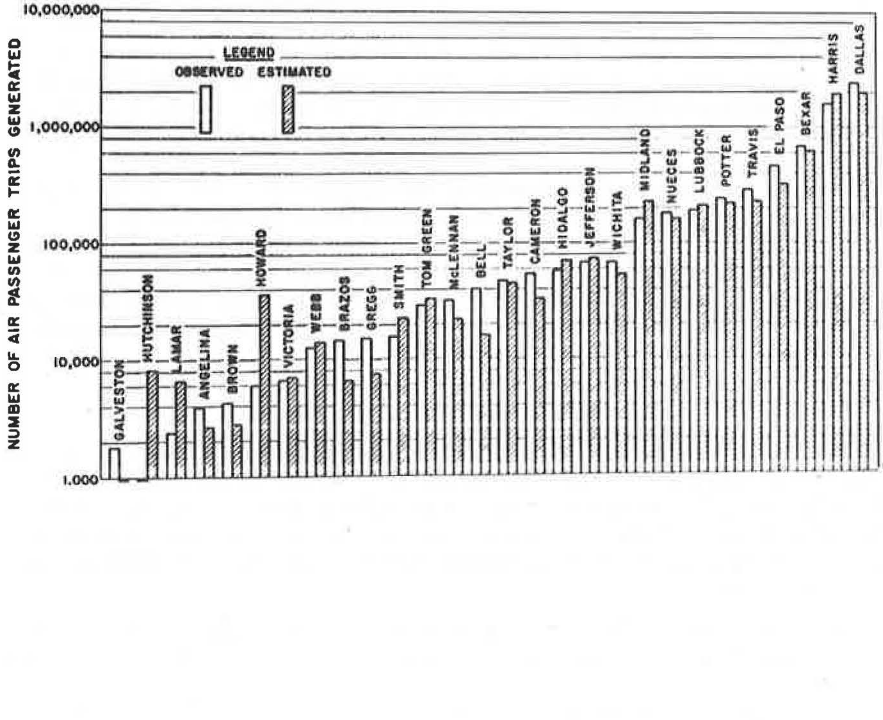


Figure 7. Comparison of estimated and observed intrastate air passenger trips for 1967.



Examination of the comparisons for interstate travel indicates relatively close agreement between observed and estimated travel volumes for the larger urban areas. For smaller areas, the relative variation is greater but is generally explainable. For example, the fact that observed travel is higher than the estimated levels for Wichita and Bell counties can be attributed largely to the major military installations in these areas. Similarly, the state capital and a major university in Travis County and another major university in Brazos County explain the low estimated volumes for thses areas.

Close proximity to other airports of greater or comparable size accounts for estimating variation for Hutchinson, Cameran, and Hidulgo counties. Significant underestimation of traffic at Galveston occurs because Galveston is separated from the nearby Houston Intercontinental Airport by Galveston Bay, which increases the relative attractiveness of the Galveston-Houston air connection for trips originating in Galveston.

For several counties in the West Texas plains area (i.e., Lubbock, Midland, Tom Green, and Howard), per capita income is relatively high, and there is generally a high propensity to travel. However, general aviation activity in this area is also considerably above average. It would, therefore, appear that many candidate air travelers in this area substitute general aviation for commercial air travel. (Wide separation of communities and lack of concentration of travel between them inhibit development of a viable commercial air service that can satisfy the existing demand.)

The comparison for intrastate air travel indicates generally similar patterns as for the interstate travel, but the relative variation appears to be somewhat greater. However, the fact that the short-haul (intrastate) travel accounts for only about one-fourth of the total air travel generated by this area lessens the significance of this greater variation.

Errors in the origin-destination data used as the basis of comparison represent another potential source of variation between the observed and estimated volumes. These observed air travel volumes are actually estimated from a 10 percent sample of ticket coupons. In addition to sampling errors, the fact that information contained on the tickets frequently does not exactly describe the traveler's actual trip introduces an indeterminate bias into the observed data.

Much of the variation previously identified could be significantly reduced by "fine tuning" of the model. Possible improvements include the development of trip generation relations that more accurately reflect the effect of character and level of economic activity on a region's trip generating potential and the refinement of the network description to more precisely account for discontinuities in the highway system in regions of detailed study. Even in its present form, however, the estimates appear adequate for providing general system development criteria.

SUMMARY

This research has developed a model for estimating the magnitude and geographic distribution of air travel demands at the regional level. The model represents, to a large degree, a synthesis of the basic concepts and relations used by previous techniques. However, it permits examination of the following factors that are especially important for planning a regional air transportation system and have not generally been integrated in previous models:

1. The influence of the level of air service provided to a community on the volume of air travel generated, particularly for small communities with no air service or with only limited service;
2. The impact of changes in service to one community on the demand for and economic viability of service to nearby communities; and
3. The complementarity between air transportation and other modes.

Comparison of estimated air travel volumes and observed volumes indicates that the method provides reasonable estimates of air passenger demand. Although there are discrepancies between the observed and estimated air passenger volumes, these discrepancies are not generally serious, and the likely sources of such errors are apparent. Further refinement of the model may ultimately be desirable to reduce these discrepancies, but the procedure in its present form provides useful information for the planning and development of a statewide or regional air transportation system.

REFERENCES

1. Transportation Demand Area Concept. Federal Aviation Administration, U.S. Department of Transportation, January 1970.
2. Official Airline Guide, North America Edition. Reuben H. Donnelley, Oak Brook, Illinois.
3. Census of Transportation—1967. National Travel Survey TC67-N1, U.S. Government Printing Office, 1969.
4. Origin-Destination Survey of Airline Passenger Traffic. Civil Aeronautics Board.

EVALUATION OF A MAILED PLANNING SURVEY

Eugene W. Waltz, Indiana Department of Commerce; and
W. L. Grecco, University of Tennessee

A methodological evaluation of a mailed planning survey was made to obtain information that would help assess the application of mail surveys for providing planning information. The scope of the study was limited to evaluating the combined application of mail and nonmail follow-up procedures for reducing nonresponse and total survey costs and to assessing how critical nonresponse might be to planning surveys. A mailed planning survey of a small community population was conducted using mail, telephone, and personal follow-ups. Survey respondents were determined by selected sociodemographic characteristics, and the cost and contribution of the follow-up procedures were also determined. The survey obtained residents' opinions and suggestions for improvements in community facilities, services, and conditions over 12 major categories of community concern. Survey response was found to be more strongly associated with the resident time of the respondents than it was with their age, sex, socioeconomic status, family status or size, tenure, or type of dwelling unit. Respondents were found more likely than nonrespondents to be old, long-time residents, and owners of single-unit dwellings. Nonmail follow-ups were found to be effective in reducing the typical socioeconomic bias found in the response to the mail-out portions. On the basis of cost versus information obtained, the results indicated that the combined use of the mail approach with mail, telephone, and personal follow-ups could be comparable to the use of other methods for planning surveys having an informative purpose.

SURVEYS of economic base, land use, transportation, and population predominate among those considered essential to urban planning. Recently, planners have come to be increasingly interested in using attitude and opinion survey data in the planning process. Although experience in conducting social surveys for planning purposes is still limited, initial experiences in surveying urban residents' reactions to urban problems indicate a wide range of application. For example, planning surveys have been conducted to obtain residents' general evaluation of their community environments (1), preferences for accessibility to selected neighborhood services (2), suggestions for needed community facilities and programs (3), attitudes about the relative importance of the livability features of their community (4), and comments about subjects related to community objectives (5).

The orientation of this study might best be indicated with the aid of a conceptual framework suggested by Gans (6). Gans defined two conceptual environments: the potential environment (i.e., as seen by the planner) and the effective environment (i.e., the version of the potential environment that is manifestly or latently adopted by users). The viewpoint of this study was that survey data of perceptions and reactions to community facilities, services, or conditions are, in part, descriptive information about the effective environment. This descriptive kind of information is not considered sufficient within itself for explaining why specific groups have particular preferences or for predicting what the effective environment will be. These latter purposes are more of an analytical or explanatory nature that presume a solid base of descriptive information that is currently not available (7). This study was limited to the informative survey purpose on the premise that descriptive survey data can perform the role of giving the planner more information and insight about the effective environment.

In the past, planning agencies have relied on public hearings and discussions with organized reference groups to obtain informative inputs from the public. Sample surveys offer an additional approach to obtain information from a wider cross section of the general public. There are various survey approaches that could be taken. The mail and interview survey methods are predominantly used. In part, the typical problems associated with mail and interview surveys still prevent most small community (less than 100,000 population) planning agencies from undertaking surveys more frequently. Both methods have their advantages and disadvantages depending on the specific survey situation and purpose. The response problem of mail surveys often precludes their consideration as a possible alternative, and this may have contributed to their limited application in planning studies.

In many situations, the type of data sought dictates the use of the interview method. In other instances, either the mail or interview technique could be used when closer attention is given to the data sought and the actual use to which they will be put. For such situations, the economic advantage of using mail surveys with follow-up procedures makes their application in the planning context attractive both for periodic data collection and for situations where limited funds preclude the interview method. For example, one of the surveys done by Barnes secured 71 percent return for a personally delivered questionnaire with three follow-ups, two by mail and one by telephone (8). Such mail survey applications in planning, however, have been limited. Associated with this limited experience is a lack of information that would help evaluate the combined use of the mail approach with follow-up methods in various planning situations to reduce overall survey costs and nonresponse.

STUDY PURPOSE AND SCOPE

The general objective of this study was to make a methodological evaluation of a mailed planning survey for the purpose of obtaining further information that would help evaluate the application of mail surveys in the planning context. The scope of the research was limited to estimating the suitability of the mail approach for planning surveys of the general population and its combined use with different follow-up procedures to reduce nonresponse and total survey costs. A mailed planning questionnaire with both mail and nonmail follow-up procedures was used to achieve the primary study objectives: to determine the respondents and nonrespondents on the basis of selected individual and household characteristics and to determine the cost and contribution of follow-up procedures for reducing nonresponse.

STUDY DESIGN

The general design of this study was to determine the characteristics of the respondents to a mail survey and to evaluate the cost and contribution of follow-up procedures. The approach was similar to that used by Hochstim and Athanasopoulos (9). Empirically, individual and household characteristics would be determined for the respondents and household characteristics for the nonrespondents. Follow-up procedures would be considered for their cost per return, number of returns, and how much they improved the sample estimates of the population on selected characteristics. The analysis would consist basically of determining any characteristic differences between respondents and nonrespondents.

The population for this study was the resident households within the corporate city limits of Lafayette and West Lafayette, Indiana. At the time this study was initiated, a home-interview travel survey was being conducted that covered both the cities and the surrounding county. Available from this survey was an accurate sampling frame of households that could be utilized to reduce research costs and to allow a complete enumeration of the study sample on household characteristics. Furthermore, the substantive opinion data could be made available to the transportation and development study.

Selected Procedures

Four follow-ups were chosen for use after the questionnaire was mailed. Two mail follow-ups and two nonmail follow-ups were selected. The two mail follow-ups were a reminder postcard and an additional mailing of the questionnaire, cover letter, and return envelope. The other two follow-ups were a telephone call reminder and a simplified personal contact.

A two-stage design was selected to evaluate the mail and nonmail follow-ups. The basic reason for selecting this type of design was to evaluate the mail follow-ups in the same way that they would be used in practice and still allow a separate comparison of the telephone and personal follow-ups. The procedure was chosen to be as follows. After the initial mailing, the mail follow-ups would be successively sent to the nonrespondent households. After the mail-out portion was completed, the remaining nonrespondent households would be divided into two groups—one to receive a telephone call follow-up and the other a personal follow-up. By using two treatment groups, a comparative evaluation could then be made of using either the telephone or the personal follow-up directly after the mail portion of a survey.

Several considerations were made in selecting what individual and household characteristics were to be obtained. For comparative purposes, characteristics that had been used in past mail research were desired. Of these characteristics, those having a possible association with the interest in the survey subject matter were chosen for descriptive comparisons of the opinion data. The household characteristics were also selected on the basis of their availability for the study population from other data sources.

The individual characteristics selected were age, sex, education, socioeconomic status (SES), and resident time. The household characteristics chosen were city location, occupation of household "head," family composition, number of persons living in the household, home ownership, and type of dwelling unit structure.

Questionnaire Design

There exist a wide range of community subject areas about which planners would be interested in obtaining residents' opinions. Answers to the question of community needs were considered to be of basic informative use. For this reason, the questionnaire design was focused on the question of community improvements. A choice was made to obtain residents' suggestions for improvements in facilities, services, or conditions over some major categories of community concern such as health care, housing, education, transportation, and recreation.

The preceding choice was based on several considerations. Planning is directly or indirectly concerned with the provision of most community facilities and services. Residents' opinions about the same could be useful in locating problem situations needing further study. Also, by using enough major categories to cover most subject areas of community concern, the possible response biases from the variation of public interest and awareness in different subject areas could be attenuated. The scope of these categories would, however, prevent an in-depth coverage with a short questionnaire.

Three basic types of questions at the community level were selected to obtain residents' opinions about improving the community:

1. In which major categories does the community need the most improvement?
2. Should improvement of specific facilities, services, or conditions be given priority?
3. What is the relative importance of several community projects that were then under consideration?

The type of questions used in a mail questionnaire can affect response rates. It is generally recommended in self-administered questionnaires to use mostly closed-form questions such as checklists, rating scales, or inventories to make responding easier (10). In this case, a closed-form structure could have resulted in a questionnaire composed of several "omnibus" checklists of facilities, services, and conditions for evaluative ratings. Such a design was not generally recommended either because it could produce superficial responses and respondent boredom. Furthermore, it would have required prejudgment of what particular facilities, services, or conditions should be listed. These problems were alleviated by choosing one open and two closed forms. For the second question given previously, the freedom of an open-ended form would obtain more information, and, for the first and third questions, a closed form was considered adequate.

Shown in the Appendix are the four pages of the final questionnaire design. These were printed on one 8½- by 14-in. sheet of white paper with black ink. The page was then folded in booklet form to give the questionnaire a "shorter" look. For the final design, several changes were made in the questions and their wording, based on both pilot test and pretest. The wording of the postcard reminder used for the first mail follow-up was similar to that used by Nichols and Meyer (11).

With the study objectives and the practical constraints in mind, an initial sample size of approximately 500 dwelling units for this study was considered large enough to keep the sample estimates of the population proportions on most characteristics within 10 percent at the 95 percent confidence level (12).

Research costs were reduced by taking the sample for this study from the list of dwelling units selected for interviewing for the Greater Lafayette Transportation and Development Study. The population for that study was all the dwelling units in Tippecanoe County. From an updated field listing of all dwelling units compiled during a land-use inventory 1 year prior, the study selected a systematic sample of every eighth dwelling unit.

Using the aforementioned list, with a random start, every fifth dwelling unit address was selected resulting in 886 dwelling units. Because of fraternities and out-of-the-city addresses, the resulting sample size for the initial mailing was 489 dwelling units.

The initial mailing to this sample would solicit any adult member of the dwelling unit to be the respondent. Although this would present a sampling bias with respect to the population of individuals, it was still the practical approach that a planning agency might use in a mail survey of the general public.

The study was conducted under the name of the Greater Lafayette Community Improvement Study with no organizational sponsor stated.

DATA COLLECTION

A survey in accordance with the study design was taken during October and November. Accurate accounts of the material and labor costs for each procedural stage were kept. The execution of each procedural stage will be briefly discussed.

Mail-Out Questionnaire

The questionnaire, cover letter, and return envelope were mailed with hand-stamped, first-class postage to the 489 selected dwelling unit addresses. For several reasons, only 454 questionnaires were delivered.

Six days after the initial mailing, postcard reminders were mailed to all dwelling units that had not responded. Eleven days after the postcard reminder, a second complete mailing of the same questionnaire with cover letter and return envelope was made to the nonresponding households. Ten days after the second mailing of the questionnaire, the mail portion of the survey was ended.

The returns were then examined for their usability. A return was classified as usable if the respondent answered at least one of the substantive opinion questions. Only five of the questionnaires received were unusable. Conservatively, these were classified as nonresponse.

After the mail-out portion of the survey, there remained 209 nonrespondent cases. These nonrespondent cases were put in numerical order of their case numbers. A systematic sampling of the cases into two groups was then made. The follow-up treatments were then arbitrarily assigned.

Telephone Follow-Up

The telephone and city directories of Lafayette and West Lafayette were used to obtain the telephone numbers of the case addresses. The telephone reminder calls were started 3 days after the cutoff date of the mail portion of the survey. Not all of the households for which a number was listed could be contacted by telephone. Thus, only 65 percent of the subsample was reached by telephone. This percentage was lower than what had been expected. The use of the telephone, however, has to be considered with this associated limitation.

Personal Follow-Up

The canvass of the subsample of households was started 5 days after the cutoff date of the mail-out portion of the survey. All 105 dwelling unit addresses in the subsample were visited one time. Originally, it had been planned to give the household member the option of either completing the questionnaire in the presence of the collectors or completing it at her convenience and returning it by mail. It became apparent, however, after a few contacts that trained personnel would be required to tactfully induce a household member to interrupt her activity and complete the questionnaire on the spot. Pursuing this optional approach would have made the use of untrained personnel for this type of follow-up questionable. For this reason, the approach at the remaining households was to only ask the household member to complete the questionnaire at her earliest convenience and return it by mail. At those households where a personal contact was made, the conversational approach was similar to the approach used with the telephone calls.

At households where no one was at home, a reminder was left. A cover letter, questionnaire, and return envelope were left at the door.

Nonrespondent Household Characteristics

Five characteristics of the nonresponding households were obtained: resident time, occupation of the head of the household, city location, type of dwelling unit, and whether the dwelling unit was owned or being rented. The information for the latter three characteristics was obtained from the data collected in the transportation and development study. The resident time and occupation of the head of the household were determined by using both the preceding study data and the 1970 and 1971 city directories. The occupation listed in the city directory was taken as that of the head of the household. In those cases where the occupation was not reported in the city directory and where the resident time was 1 year or more, the occupational data were taken from the transportation study data. The occupations were coded in the following categories:

1. High SES—professionals, technicians business managers, owners, officials;
2. Middle SES—clerical workers, salesmen, craftsmen, foremen, etc.;
3. Low SES—operatives, unskilled workers, service workers, domestics, etc.;
4. College students; and
5. Retired.

METHODOLOGICAL RESULTS

Initial Sample

The initial dwelling unit sample was checked for any serious bias with respect to the study population. In Table 1, the initial sample proportions on dwelling unit location, type of structure, and tenure are compared with those reported for the study population in the 1970 Census of Housing. The survey sample was proportionally about the same as the population on city location and slightly overrepresentative on single-unit and rented dwellings (Table 1). Even though these latter biases are very small, all comparisons were made with the enumerated sample values to account for these slight differences. The data were compiled and analyzed using a specialized (13) computer program.

Returns

The returns for each procedural stage are given in Table 2. The overall return rate for the survey was 67 percent. The return to the mail-out portion was about what had been expected, 52.7 percent. The percentage of return after the postcard reminder was higher than what had been expected from the pretest results. The return rate for the first two waves was more than 11 percent higher than what was found in the pretest, 42.5 versus 31.0 percent. Some of the improvements could have resulted from differences between the final survey and the pretest, such as improved questionnaire design, timing, or sponsor. The final questionnaire was also shorter and had a better appearance than the pretest form.

After the mail-out portion of the survey was finished, 209 sample households had not responded. These households were separated into two groups for the telephone and personal follow-ups as previously explained. The dwelling unit characteristics of the two groups are given in Table 3 for comparison. Very small differences existed between the two groups on the characteristics shown.

The telephone follow-up obtained 24 additional returns, or 23 percent of the subsample. The low percentage for the subsample is partially attributable to the fact that only 65 percent of the subsample households could be reached by telephone. On the basis of the number contacted, the return rate was 36 percent. This return rate was twice that obtained for the second mailing (18 percent). For those contacted, the response rate was still lower than what had been expected. Voiced intention of cooperation by household members over the telephone proved to be an unreliable criterion.

The use of telephone reminder calls must be considered in the context of the study limitations (e.g., some households not having telephones). The use of a third mailing to those households that cannot be reached by telephone could be an effective supplement to this approach. Omitting the use of this third mailing was an oversight of the study.

The simplified personal follow-up obtained 40 additional returns, or 38 percent of the subsample. Fifty-two percent of the subsample households were personally contacted with the remainder having a reminder letter, questionnaire, and return envelope left at their door. The return rate for those households personally contacted was 40 percent; for those not at home, it was 36 percent. Unexpectedly, both treatments were comparably effective. The impressions given by household members personally contacted caused an overexpectation of likely returns. On the other hand, the returns from those households receiving the notice of a visit and a questionnaire was not expected to have, as it did, a return rate higher than the second mailing of the mail-out portion of the survey.

Costs

Accurate accounts were made of all labor and material costs associated with each procedural stage of the data collection. Table 4 gives the cost accounts for each stage by items of expense. The initial sampling and listing of case addresses were charged to the initial mailing stage. As noted, labor time was converted at the rate of $3.00 per hour.

The overall survey cost for data collection was $541 with an average cost per return of $1.78. As shown, the cost if only the personal follow-up had been used was $1.91 on the basis of a projected overall return of 71 percent. If a telephone follow-up, supplemented by a third mailing, had been used, the overall return rate would likely have been comparable but somewhat lower in cost. On the basis of the cost and return data obtained, a similar survey combining the use of all these procedures for economy and effectiveness could be conducted as follows: initial mailing; postcard reminder; second mailing; postcard reminder; telephone call reminders supplemented by a third mailing of the questionnaire, cover letter, and return envelope; and a simplified personal follow-up.

Whatever combined approach is selected for following up a mailed survey, a certain degree of nonresponse can be expected even when interview follow-ups are used. For example, the mail surveys conducted by the U.S. Bureau of Census were followed up by

Table 1. Dwelling unit characteristics of initial sample.

Dwelling Unit	Survey Sample (N = 454)	1970 Housing Census (N = 22,188)
Lafayette	71.8	70.1
West Lafayette	28.2	29.9
1-unit structure	69.8	65.5
2 or more units	30.2	34.5
Owned	59.7	62.2
Rented	39.2	37.8
Undetermined	1.1	—

Table 2. Survey response by procedural stage.

Number	Procedural Stage	Number	Number of Returns	Return Rate (percent)	Percentage of Total
1	Initial questionnaire mailing	454	115	25.4	37.8
2	Postcard reminder	339	78	23.0	25.6
3	Second questionnaire mailing	261	47	18.0	15.5
4A	Personal follow-up	105	40	38.1	13.2
4B	Telephone follow-up	104	24	23.1	7.9
—	All stages	—	304	67.0	100.0

Table 3. Dwelling unit characteristics of telephone and personal follow-up groups.

Dwelling Unit	Telephone Follow-Up	Personal Follow-Up
Total	104	105
Lafayette	75	75
West Lafayette	29	29
1-unit structure	70	66
2 or more units	34	39
Owned	51	50
Rented	50	53
Undetermined	3	2

Table 4. Data collection costs by stage.

Item	Initial Mailing	Postcard	Second Mailing	Telephone Calls	Personal Visits	All Stages
Materials	41.00	7.00	21.00	1.50	5.50	76.00
Mailing expenses	81.00	23.00	42.00	3.00	6.00	155.00
Office work[a]	42.00	9.00	18.00	36.00	—	105.00
Collectors[a]	—	—	—	—	75.00	75.00
Travel expenses	—	—	—	—	10.00	10.00
Supervision[a]	60.00	12.00	15.00	9.00	24.00	120.00
Total Cost	224.00	51.00	96.00	49.50	120.50	541.00
Returns	115	78	47	24	40	304
Return rate (percent)	25.4	23.0	18.0	23.1	38.1	67.0
Cost per return	1.95	0.66	2.04	2.06	3.01	1.78
Cumulative cost per return	1.95	1.50	1.58	1.63[b]	1.91[b]	—

[a] Time accounts converted at the rate of $3.00 per hour.

[b] Based on projected return.

both telephone and personal interviews (14). The nonresponse to these surveys ranged from 17 to 24 percent. Similarly, the Hochstim and Athanasopoulos study still had 14 percent nonresponse after an interview follow-up (9). This same degree of nonresponse is typical of that expected in complete interview surveys that do not use substitution (15). In most cases, approximately 15 percent nonresponse could be expected when typical follow-up techniques are used. If a mailed survey obtained 50 percent return and a telephone reminder obtains another 15 to 20 percent, approximately 40 to 50 percent of the remaining sample is still not likely to respond. The decrease in the expected return rates at the later follow-up stages makes the cost of the follow-up a more determining factor in its use at these stages. In some cases, combining the telephone call with the more economical simplified personal follow-up might be a more acceptable alternative than an interview follow-up in view of the expectedly low return rate and the high cost of interviews.

Respondents and Nonrespondents

Comparison of the survey respondents and nonrespondents was made basically to determine how the survey respondents, after each procedural stage, compared with the sample enumeration on selected characteristics; what significant differences in selected characteristics existed between the respondents and nonrespondents; and whether any of the selected characteristics were associated with the wave of return. Table 5 gives a summary of the response after each procedural stage of the selected individual and household characteristics. Also, available enumeration values of the characteristics for either the sample or the study population are given for comparison. Some of the values given in Table 5 are proportionally different from the enumeration by less than 10 percent [(percent difference/enumerated percent) × 100 percent < 10 percent]. Before making more detailed statistical comparisons, this 10 percent criterion will be used for cursory comparisons.

The total survey returns were reasonably comparable to the enumeration values on the variables of sex, city, occupational SES, and type of dwelling unit structure. Sex and city were the only variables within 10 percent after each stage. The bias on age and homeownership was consistent over all stages. The categorical distributions of returns on household composition and number of persons were also similar for each wave of return. The differences between early and late respondents were reflected in characteristics such as sex, education, resident time, and occupational SES.

For considering survey response and the selected characteristics more specifically, two statistical analyses were performed with the data. These were a comparison of survey respondents and nonrespondents on selected characteristics, and a test of association between the wave of return and the selected characteristics.

Chi-square (χ^2) was used as the test statistic for significant differences from what would be expected from the hypothesis of equal proportionality. The level of confidence chosen for rejecting the equal proportionality hypothesis was the 0.10 probability level.

The strength of associated differences was measured by using the nonparametric statistic, Cramer's V, which is defined as

$$V^2 = \frac{\chi^2}{N \min \begin{pmatrix} r - 1 \\ c - 1 \end{pmatrix}}$$

where $\min \begin{pmatrix} r - 1 \\ c - 1 \end{pmatrix}$ is the minimum value of either the rows or columns minus one. This statistic takes on values ranging from 0 to 1, for no association to a perfect association respectively and accounts for unequal rows and columns. Even though values of Cramer's V between 0 and 1 do not have much intuitive meaning, the statistic does serve as a comparative indicator of the strength of different associations.

Table 6 gives the survey respondents and nonrespondents by the characteristic variables enumerated for the sample. As shown, the respondent group had proportionally a larger number of long-time residents, homeowners, and persons living in single-unit

Table 5. Cumulative survey response by selected characteristic.

Characteristic	Initial Mailing (N = 115)	Postcard Reminder (N = 193)	Second Mailing (N = 240)	Telephone, Personal Follow-Up (N = 304)	Enumeration (N = 454)
Individual					
Age (years)					
21 to 34	34.0	33.2	35.8	34.9	40.8[a]
35 to 54	33.0[b]	31.1[b]	29.6	31.3[b]	33.4[a]
≥55	33.0	35.7	34.6	33.9	25.8[a]
Sex					
Male	53.0[b]	45.3[b]	45.8[b]	45.7[b]	50.2[a]
Female	47.0[b]	54.7[b]	54.2[b]	54.3[b]	49.8[a]
Education (years)					
<12	12.2	13.4	16.1	14.5	–
12 to 15	40.0	44.7	43.4	47.4	–
≥16	44.3	39.4	38.0	36.1	–
Not reported	3.5	2.5	2.5	2.0	–
Resident time (years)					
≤10	44.4[b]	35.9[b]	36.5	33.7	47.1[c]
11 to 29	26.0[b]	27.2[b]	29.1[b]	32.3	27.8
≥30	29.6	34.8	34.3	34.0	25.1
City					
Lafayette	67.8[b]	71.5[b]	72.1[b]	72.0[b]	71.8
West Lafayette	32.2	28.5[b]	27.9[b]	28.0[b]	28.2
Household					
Occupation (head)					
High SES	40.9	34.7	32.9	30.7[b]	27.8
Middle SES	30.4[b]	33.6[b]	33.7[b]	33.9[b]	33.9
Low SES	14.7	15.5	17.1	18.7[b]	20.7
College student	6.1	6.7	7.5	7.8[b]	8.4
Retired	7.0[b]	8.3	7.9	8.2	6.8
Undetermined	0.9	1.0	0.8	0.7	2.4
Composition					
Single	11.5	12.6	13.9	13.9	–
Married, no children	29.2	30.9	31.1	31.5	–
Married, with children	53.1	48.7	47.9	47.0	–
Other	6.2	7.9	7.1	7.6	–
Number of persons					
1	13.0	13.5	14.6	14.5	–
2	30.4	35.2	34.6	34.9	–
≥3	56.5	51.3	50.8	50.7	–
Dwelling unit					
Owned	66.1	70.5	70.0	65.8	59.7
Rented	33.9	29.5	30.0	33.9	39.2
Undetermined	–	–	–	0.3	1.1
1-unit structure	75.7[b]	76.2[b]	75.0[b]	72.7[b]	69.8
2 or more units	24.3	23.8	25.0	27.3[b]	30.2

[a] Values are from the 1970 Census of the Population for the study cities. All other values are from the transportation study data.

[b] Values are proportionally different from the enumeration by less than 10 percent.

[c] Includes the resident time of the head of the nonresponding households.

Table 6. Respondents and nonrespondents by enumerated characteristics.

Characteristic	Total Respondents (N_R = 304)	Non-Respondents (N = 150)	χ^2 Test of Significant Difference (p)	Cramer's V
City				
Lafayette	72.0	71.3	NS[a]	–
West Lafayette	28.0	28.7		
Resident time (years)				
≤ 10	33.7	68.5[b]	0.0001	0.34
11 to 29	32.3	21.5		
≥30	34.0	10.0		
Occupation (head)[c]				
High SES	37.1	26.7	NS	–
Middle SES	40.5	42.5		
Low SES	22.4	30.8		
Dwelling unit				
Owned	66.0	48.6	0.001	0.17
Rented	34.0	51.4		
1-unit structure	72.7	64.0	0.07	0.09
2 or more units	27.3	36.0		

[a] Not significantly different at the 0.10 level of confidence.

[b] Resident time of head of nonresponding households.

[c] Respondents = 254; nonrespondents = 111.

dwellings. The nonrespondents were more likely to be short-time residents, renting, and living in multiple-unit structures. These three characteristics are probably highly correlated with each other. The groups were comparable on city location and showed typical high and low SES differences.

The Cramer's V measure of the strength of the association indicates comparatively that resident time in the community was the most distinguishing characteristic between the respondents and nonrespondents of those considered. The mean resident time of the respondents was 23.4 years; for the nonrespondents, it was 7.35 years. Within the 10-years-or-less category, 74 percent of the nonrespondents had been in the community for only 3 years or less, whereas only 43 percent of the respondents within this category had resident times of 3 years or less.

The association between survey response and resident time was further examined by using the other characteristics as test factors, or controls, to see if the association was conditional on any of these variables. Table 7 gives the survey respondents and nonrespondents by resident time controlling on occupational SES, tenure, type of structure, and city. As shown, the association was still statistically significant for all the subgroups and comparable in strength to the original association. These results reinforce the conclusion that the resident time in the community has a more dominant influence on the response to a mailed community-related questionnaire than any of the other variables considered. Also, as shown by the relative values for Cramer's V given in Tables 7 and 8, target populations low on SES and short on resident time will likely be the most unresponsive group to a mailed community questionnaire.

The association of survey response and resident time is not considered surprising. It merely reinforces the common-sense notion that community interest and awareness are likely to be higher among long-time residents than they are among recent arrivals. Linking longer resident time with increased community awareness and survey response would reinforce the findings of past mail-survey research that the interest in the survey subject matter is the strongest determinant of response.

Although the characteristics given in Table 9 were found to have statistically significant associations with the wave of return, all the associations were comparatively weak as reflected by the values for Cramer's V. The practical significance of these results is only that the bias in mail returns is more likely to be on these characteristics than the others considered, and the use of the nonmail follow-ups helped reduce these biases. For example, the respondents to the nonmail follow-ups were significantly different (χ^2 probability is less than 0.10) from the respondents to the mail-out portion on education, resident time, occupational SES, and homeownership.

In summary, the sample returns were found to be underrepresentative of the younger age group (21 to 24 years old), males, short-time residents, renters, and persons living in multiple dwelling unit structures. The returns were overrepresentative of the older age group (55 years old and more), females, long-time residents, homeowners, and persons living in single-family dwelling units. The differences between respondents and nonrepondents on city and occupational SES were less than those cited previously. The most significant difference found was on resident time with shorter time residents showing the greatest degree of nonresponse of any group considered.

Technically speaking, the bias found in the sample returns on some of the sociodemographic characteristics could be crucial for surveys having an explanatory or analytic purpose of inferring behavioral variables from attitude data. Planning surveys seeking attitude or opinion data about what residents perceive to be the major sources of dissatisfaction within a community subject area have more of an informative than an explanatory purpose. Primary to the consideration of using the mail-survey approach for this informative purpose is assessing how much information is lost because of nonresponse and to what degree the information obtained is peculair to the characteristic nature of the respondents. From the opinion data collected in this study, one cannot determine if, or how strong, a relation might exist between a group's reactions to its community environment and its sociodemographic characteristics. The data gathered, however, are considered of sufficient scope and detail to make some assessment on how crucial the nonresponse bias is to the informative survey purpose.

Table 7. Survey response by occupational ranking.

Resident Time (years)	High SES		Middle SES		Low SES		Owners		Renters	
	R	NR	R	NR	R	NR	R	NR	R	NR
≤ 10	48.3	81.3	28.7	60.8	10.7	66.7	25.6	50.7	50.5	89.2
11 to 29	34.8	15.6	33.7	33.3	46.4	16.7	34.9	39.4	27.5	5.4
≥ 30	16.9	3.1[a]	37.6	5.9	42.9	16.7	39.5	9.9	22.0	5.4
N(100 percent)	89	32	101	51	56	36	159	71	95	74
P(X^2)/Cramer's V	0.005/0.30		0.0001/0.38		0.001/0.58		0.0001/0.30		0.0001/0.41	

[a] Expected cell frequency less than 5.

Table 8. Survey response by dwelling unit and location.

Resident Time (years)	1-Unit Structure		2-Unit Structure		Lafayette		West Lafayette		Total	
	R	NR	R	NR	R	NR	R	NR	R	NR
≤ 10	27.6	61.5	51.4	81.1	25.2	60.4	57.1	88.4	33.7	68.5
11 to 29	35.9	26.0	21.6	13.2	35.5	27.4	23.4	7.0	32.3	21.5
≥ 30	36.4	12.5	27.0	5.7	39.3	12.3	19.5	4.7	34.0	10.0
N(100 percent)	217	96	74	53	214	100	77	43	304	150
P(X^2)/Cramer's V	0.0001/0.33		0.002/0.32		0.0001/0.36		0.002/0.31		0.001/0.34	

Table 9. Wave of return by selected characteristics.

Characteristic	Wave of Return (percent) 1	2	3	4	N	X^2 (p)	Cramer's V
Sex					304	0.06	0.16
Male	53.0	33.0	48.9	45.3			
Female	47.0	66.7	51.1	54.7			
	100.0	100.0	100.0	100.0			
	(115)	(71)	(47)	(64)			
Education (years)					298	0.01	0.17
< 12	12.6	15.4	27.7	8.1			
12 to 15	41.4	52.6	38.3	62.9			
≥ 16	45.9	32.1	34.0	29.0			
	(111)	(78)	(47)	(62)			
Occupation (head)					254	0.10	0.14
High SES	48.0	30.6	30.8	28.3			
Middle SES	35.0	48.4	41.0	41.5			
Low SES	17.0	21.0	28.2	30.2			
	(100)	(62)	(39)	(53)			
Resident time (years)					304	0.10	0.14
≤ 10	44.4	35.9	36.5	33.7			
11 to 29	26.0	27.2	29.1	32.3			
≥ 30	29.6	34.8	34.3	34.0			
	(115)	(78)	(47)	(64)			
Dwelling unit					303	0.02	0.19
Owned	66.1	76.9	68.1	50.8			
Rented	33.9	23.1	31.9	49.2			
	(115)	(78)	(47)	(63)			

Note: Numbers in parentheses are the actual number of returns.

CONCLUSIONS

The conclusions that were drawn from the results of this study are as follows:

1. The combined use of mailed questionnaires with follow-up procedures is an economical approach for obtaining subjective opinion data from the general public.
2. For planning surveys seeking residents' subjective opinions about their community environments for informative uses, the bias due to nonresponse may not result in any serious loss of information if greater than 60 percent return is achieved.
3. A mailed-out community-related survey is not likely to achieve much more than 50 percent response unless nonmail follow-up procedures are used.
4. The combined successive use of a telephone and simplified personal follow-up to a mailed community survey is likely to be comparable to an interview follow-up on the basis of the cost versus the information obtained.
5. Respondents to a community-related mail survey are more likely to be old, longtime residents owning a single-unit dwelling than are nonrespondents.
6. The use of nonmail follow-up procedures in a mail survey can help reduce the typical SES bias found in the response to a mail survey.
7. The response to a community-related mail survey is likely to be more strongly associated with the respondent's resident time than it is with his age, sex, SES, family status or size, tenure, or type of dwelling unit.

ACKNOWLEDGMENT

The authors wish to acknowledge the funding of this project by the General Electric Corporation and the data collection assistance provided by the Greater Lafayette Transportation and Development Study.

REFERENCES

1. Wilson, R. L. Livability of the City: Attitudes and Urban Development. In Urban Growth Dynamics (Chapin, F. S., Jr., and Weiss, S. F., eds.), John Wiley and Sons, New York, 1962.
2. Peterson, G. L., and Worrall, R. D. An Analysis of Individual Preferences for Accessibility to Selected Neighborhood Services. Highway Research Record 305, 1970, pp. 99-111.
3. Aguar, C. E. The Use of Surveys in Planning. Jour. of the American Society of Planning Officials, 1965, p. 106.
4. Lamanna, R. A. Value Consensus Among Urban Residents. Jour. of the American Institute of Planners, Vol. 30, No. 4, November 1964, pp. 317-323.
5. Community Directions for the City and County of Denver: A Resume of Opinions of Certain Denver Citizens on Subjects Relating to Denver Community Objectives. Denver Planning Office, Comprehensive Planning Bull. 1-2, October 1964.
6. Gans, H. J. People and Plans. Basic Books, New York, 1968, pp. 1-11.
7. Saroff, J. R., and Levitan, A. Z. Survey Manual for Comprehensive Urban Planning. Development Research Associates, Inc., Anchorage, 1969.
8. Barnes, C. F., Jr. Living Patterns and Attitude Surveys. Highway Research Record 187, 1967, pp. 43-54.
9. Hochstim, J. R., and Athanasopoulos, D. A. Personal Follow-Up in a Mail Survey: Its Contribution and Its Cost. Public Opinion Quarterly, Vol. 34, 1970, pp. 69-81.
10. Oppenheim, A. N. Questionnaire Design and Attitude Measurement. Basic Books, New York, 1966, p. 81.
11. Nichols, R. C., and Meyer, M. A. Timing Post Card Follow-Ups in Mail Questionnaire Surveys. Public Opinion Quarterly, Vol. 30, No. 2, 1966, p. 306.
12. Household Survey Manual 1969. U.S. Bureau of the Budget, 1969, p. 58.
13. Nie, N., Bent, D., and Hull, C. H. Statistical Package for the Social Sciences. McGraw-Hill, 1970.
14. Census Use Study: Area Travel Survey Report No. 11. U.S. Government Printing Office, 1970.

15. Lazerwitz, B. Sampling Theory and Procedures. In Methodology in Social Research (Blalock, H. M., Jr., and Blalock, A. B., eds.), McGraw-Hill, New York, 1968, p. 314.

APPENDIX

FINAL QUESTIONNAIRE DESIGN

GREATER LAFAYETTE
COMMUNITY IMPROVEMENT STUDY

DIRECTIONS: *Most of the following questions can be quickly checked or filled-in. Others allow you to answer in your own words.*

① First, we would like to know how long you have lived:

a. in the Lafayette area? ___ yrs.

b. at your present address?___ yrs.

② How long do you expect to live in the Lafayette area?

☐ Indefinitely ☐ At most, only a year

☐ Only a few years ☐ Don't Really Know

All cities seem to have their good and bad points.

③ First, what features of the Greater Lafayette area do you *like* the most?

____

____

____

____

☐ No Opinion

④ Listed below are some major categories in which local improvements might be made. Please check(✓) the three (3) categories in which *you* think the Lafayette area needs the *most* improvement.

☐ 1. Community Appearance
☐ 2. Education
☐ 3. Environmental Protection
☐ 4. Health Care
☐ 5. Housing
☐ 6. Local Economy
☐ 7. Local Government
☐ 8. Public Assistance
☐ 9. Public Safety
☐10. Public Utilities
☐11. Recreation
☐12. Transportation
☐ No Opinion

⑤ Are there particular facilities, services, or conditions you would like to see improved within *any* of the categories above?

☐ Yes ☐ No ☐ Don't Know

a. *If "Yes",* what improvements would you like to see made? *(Please write your answer(s) in the spaces below)*

1. In Category No. ___, *I would like to see* ________________

____

____

2. In Category No. ___, *I would like to see* ________________

____

____

3. In Category No. ___, *I would like to see* ________________

____

____

4. In Category No. ___, *I would like to see* ________________

____

____

b. *If* you suggested *more than one* improvement above, which *one* would you like to see done *first?*

Suggestion No. ___

⑥ Listed below are some specific items of local concern.

How *important* do you think each of these items would be for *improving the Lafayette area?*

(Please circle your answer for each item)

	Very Important	Somewhat Important	Not Very Important	Not Important At All	Don't Really Know
a. Public parking garages downtown	1	2	3	4	DK
b. Combining City-County services: parks, police, fire, sewage, etc.	1	2	3	4	DK
c. Expanding the County park system	1	2	3	4	DK
d. Establishing an area-wide vocational high school	1	2	3	4	DK
e. Expanding and improving the bus service	1	2	3	4	DK
f. Increase the supply of public housing	1	2	3	4	DK
g. Developing the Lafayette riverfront as a park area	1	2	3	4	DK
h. Relocating the downtown railroads	1	2	3	4	DK
i. Downtown urban renewal	1	2	3	4	DK
j. Building wildcat reservoir	1	2	3	4	DK
k. Other ____________________	1	2	3	4	DK

⑦ Now, think for a moment about your part of town.

If the local city government could spend alot of money on a new program to improve your neighborhood, what do *you* think they should spend it on?

____

____

In a community-wide survey, a statistical check must be made to insure that all kinds of people have participated. We ask you to complete the following questions to make such a check possible.

- Your age bracket is:

  ☐ Under 25 yrs. old ☐ 35 to 44 ☐ 55 to 64
  ☐ 25 to 34 ☐ 45 to 54 ☐ 65 or over

- You are a: ☐ Male ☐ Female

- The years of education you have completed:

  (Please Circle One)

Grade School	High School	College, Business, or Trade
6 or less 7 8	9 10 11 12	13 14 15 16 17 or more

- The number of persons living in your household is: ____

- Their relationship to you is:

  ____

  (e.g., wife, husband, son, daughter, brother, uncle, etc.)

- The occupation of the head of your household is:

  ____

  (e.g., clerk, machinist, typist, sales manager, fireman, etc.)

If you have any further suggestions for improving the Lafayette area, please write them below.

____

____

We Thank You For Your Help

ANALYSIS OF URBAN AREA TRAVEL BY TIME OF DAY

Donald M. Hill and Larry Tittemore, Peat, Marwick, Mitchell and Company, Washington, D.C.;
David Gendell, Federal Highway Administration

This research project is a thorough analysis of the temporal distribution of vehicular travel in eight U.S. urban areas having populations of 100,000 to 3,500,000. Graphical models were developed during the analysis, and travel data from St. Louis are presented in detail along with tabular and graphical outputs of data for each of the other cities. Programs were developed that summed the total number of persons in motion and the total vehicle-miles of travel occurring every tenth of an hour for a 24-hour period. Five submodels were used to aggregate hourly travel into similar time groups: wee hours, morning, midday, afternoon, and evening. Areawide traffic count data were used to determine total vehicle-miles of travel occurring on various classes of highway, and at varying distances and orientations from the central core city, on directional and nondirectional bases.

•HIGHWAY congestion is a generally recognized problem of the U.S. transportation system in urban areas. Every major metropolitan area suffers some form of roadway congestion during peak use periods. The levels of congestion are directly related to the fact that total vehicular travel is not uniform throughout the 24 hours of the day; yet, for the most part, the supply of transportation service is uniform.

Traffic congestion and the associated problems of limited highway capacity and travel delays occur during the peak periods of the day, most noticeably in the afternoon hours of 4 to 7 p.m. when 40 to 42 percent of the daily vehicular traffic occurs. The second, and often very pronounced, short-term peak occurs in the morning hours between 7 and 9 a.m. In these two periods of relatively short duration, the capacity of the highway system is often approached. However, for most of the day, the capacity of the transportation system far exceeds the level of highway traffic.

The cost of highway investment is directly related to this peaking phenomenon. Indeed, if it were not for this peaking in traffic demand, the required highway investment could be considerably less than it is now or than it is anticipated to be in future years. Consequently, urban planners acknowledge the need to develop a complete understanding of urban travel behavior and associated temporal characteristics so as to analyze and evaluate alternative levels of transportation investments.

RESEARCH APPROACH

City Selection

An extensive effort was undertaken to contact state highway departments and local planning agencies of more than 50 urban areas in 26 states. Care was exercised in this selection of areas to ensure that the cities that were selected were representative of U.S. urban areas. Extensive data collection was undertaken in 20 of these cities, and a final choice was narrowed to the following cities: Boston, Massachusetts; Louisville, Kentucky; St. Louis, Missouri; Seattle, Washington; Oklahoma City, Oklahoma; Stockton, California; Fall River, Massachusetts; and Colorado Springs, Colorado. Data were also obtained for Manchester, New Hampshire, for checking the analysis. The geographical distribution of these eight cities is shown in Figure 1, and their descriptive characteristics are given in Table 1. The selection process resulted in a good cross section of large, medium, and small urban areas with a reasonable representation of high and low population density, high and low 24-hour modal split, geographic distribution, and age of central city.

Phase A: Area-Wide Analysis of Travel by Time of Day

Phase A involved the analysis of the hourly distribution of person-trips and travel by trip purpose and mode as reported in the base-year origin-destination (O-D) survey of each of the urban areas. Standard survey record files numbers 2, 3, 4, and 5 of the home-interview survey, external cordon line interview survey, truck survey, taxicab survey, and interzonal skim distances over minimum time paths for the base-year highway network all were utilized.

Phase A data processing included the tabulation for the survey files of trips in motion by time of day and of vehicle-miles of travel by time of day. Only the distributions of total vehicle-miles of travel were fully analyzed and graphical relations or models researched and calibrated. Five periods of the day were selected: wee hours (midnight to 5 a.m.), morning (5 to 9 a.m.), midday (9 a.m. to 2 p.m.), afternoon (2 to 8 p.m.), and evening (8 p.m. to midnight). After some experimentation, these groups were established, and models for each hour, or combination of hours in each period, were developed that related percentage of vehicle-miles of travel to the following socioeconomic characteristics: population, degree of compactness, and volume-capacity (V-C) ratio.

Phase B: Analysis of Travel by Time of Day for Specific Facility Types and Locations

The selection of the study areas for the phase B analysis was conducted in parallel with the phase A selection process. Traffic count data were collected for nine areas, of which three were later dropped, leaving the six locations of Boston, St. Louis, Louisville, Seattle, Stockton, and Fall River. Data for Manchester were collected also for use in checking the analysis with the six cities. Data assembled consisted of hourly traffic counts from throughout each urban area for the same year as the study area's O-D survey and the preceding for following years, if available. The data obtained were nondirectional vehicle counts, directional vehicle counts, and classified counts depending on availability. Overall, nondirectional counts for approximately 2,000 locations were obtained for the six study areas and subsequently processed into a common format for analysis purposes. Stratification of highway facility by type, location, and orientation to city center is given in Table 2.

The data obtained and processed are representative of April, May, September, or October traffic and are generally typical of an average weekday. The traffic data assembled consisted of approximately 35 to 50 percent of the total traffic in the selected urban area.

PHASE A: AREA-WIDE ANALYSIS OF TRAVEL BY TIME OF DAY

The hourly distribution of vehicle-miles of travel in a typical urban area, St. Louis, is shown in Figure 2 for internal automobile driver travel by purpose (home-based work, nonwork, and non-home-based). Internal automobile driver, taxicab, trucks, and total internal and external vehicle distribution are shown in Figure 3.

The peak hour for total vehicle-miles of travel is from 4 to 5 p.m. Both total automobile driver and total truck travel also peak in this hour. Internal home-based work automobile driver trips peak between 7 and 8 a.m. and between 4 and 5 p.m., but home-based nonwork automobile driver trips peak between 7 and 8 p.m. Internal taxicab trips peak between 7 and 8 a.m. In conclusion, although 4 to 5 p.m. is the peak travel period, considerable variation in the travel distribution occurs depending on the purpose of travel (work, nonwork, non-home-based) and mode (automobile, taxicab, and truck) of vehicular travel.

The distribution of travel varies among cities. The explanation for this variation is described in the results of the models developed for total vehicular travel. In

Figure 1. Geographical distribution of cities analyzed in time-of-day study.

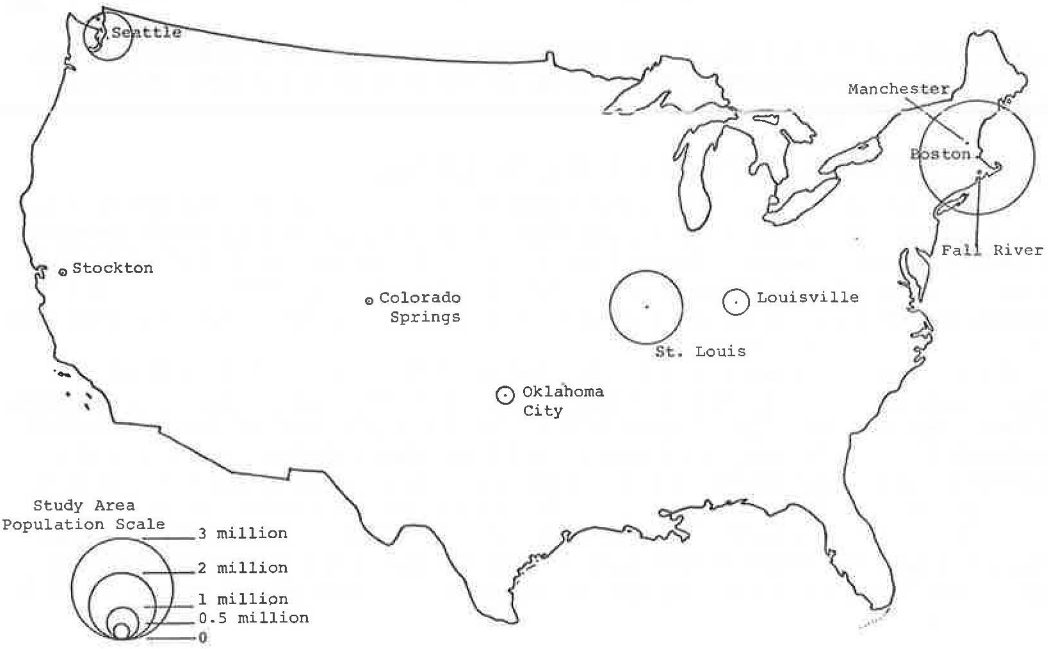


Table 1. Study area descriptive statistics.

Urban Study Area and Year	Total Population	Total Employment	Automobiles Owned	Gross Land Area (square miles)	24-Hour V-C Ratio[a]	24-Hour Modal Split (percent)	Employment Compactness Ratio[b]	Population Compactness Ratio[c]	Density Ratio[d]
Boston (1963)	3,541,000	1,297,000	1,066,000	2,500	0.04	11.9	0.45	0.30	10.0
St. Louis (1965)	2,175,000	878,000	758,000	1,640	0.48	5.1	0.48	0.30	7.9
Seattle (1961)	1,373,000	465,000	520,000	1,000	0.49	5.6	0.43	0.40	5.0
Louisville (1964)	752,000	310,000	249,000	910	0.44	6.0	0.69	0.70	10.0
Oklahoma City (1965)	564,000	229,000	231,000	1,250	0.30	0.8	0.81	0.75	1.5
Colorado Springs (1964)	174,000	48,000	68,000	290	0.23	1.4	0.60	0.55	3.5
Fall River (1963)	138,000	46,000	49,000	110	0.59	2.0	0.90	0.75	2.1
Stockton (1967)	170,000	56,000	66,000	190	0.26	1.1	0.66	0.56	4.2
Manchester (1964)	113,000	38,000	35,000	220	0.30	3.2	0.93	0.80	5.3

[a] V-C ratio = 24-hour vehicle-miles of travel (excluding school trips) divided by 10 times the hourly capacity.
[b] Employment compactness ratio = central city employment divided by study area employment.
[c] Population compactness ratio = central city population divided by study area population.
[d] Density ratio = central city population density divided by study area population density.

Table 2. Stratification of highway facilities.

Cell Number	Description of Facility		
	Type of Facility	Facility Location	Orientation to Center of Study Area[a]
1	Freeway	Central core	All orientations
2	Expressway	Central city	Radial
3	Expressway	Central city	Circumferential-crosstown
4	Expressway	Suburb	Radial
5	Expressway	Suburb	Circumferential-crosstown
6	Expressway	Rural	Radial
7	Expressway	Rural	Circumferential-crosstown
26	Expressway	Other subcenter	Radial
27	Expressway	Other subcenter	Circumferential-crosstown
8	Arterial	Central core	All orientations
9	Arterial	Central city	Radial
10	Arterial	Central city	Circumferential-crosstown
11	Arterial	Central city	Feeder to expressway
12	Arterial	Suburb	Radial
13	Arterial	Suburb	Circumferential-crosstown
14	Arterial	Suburb	Feeder to expressway
15	Arterial	Rural	Radial
16	Arterial	Rural	Circumferential-crosstown
17	Arterial	Rural	Feeder to expressway
18	Arterial	Other subcenter	Radial
19	Arterial	Other subcenter	Circumferential-crosstown
20	Arterial	Other subcenter	Feeder to expressway
21	Collector	Central core	All orientations
22	Collector	Central city	All orientations
23	Collector	Suburb	All orientations
24	Collector	Rural	All orientations
25	Collector	Other subcenter	All orientations

[a] Or to center of subcenter, as applicable.

modeling, the distribution of total vehicle travel was considered. Thorough analysis and modeling of the five component distributions (i.e., internal automobile, taxicab, and truck and external automobile and truck) would probably improve understanding of the total vehicle distribution and should be undertaken when the opportunity presents itself. In the present analysis, for example, knowledge of the distribution of home-based work internal automobile driver trips was very useful in the interpretation of the peak-period portions of the total vehicle distribution.

After some preliminary comparison of the study areas' travel on a strictly chronological basis (i.e., the same time period for all study areas), it became obvious that comparisons were better made between time periods of comparable functional significance. Analysis of the distributions, expressed in the standardized hour periods led to the conclusion that the best comparisons would be obtained by assembling groups of hours of similar character (Table 3).

The cumulative percentage of travel in each of these groups of hours for all eight study areas is shown in Figure 4. After rank-ordering the hours by percentage of daily travel within these groups, models were developed for all 24 1-hour periods. Individual attention was paid to the 3 highest hours of the afternoon and 4 of the morning. The remaining hours were treated primarily as groups or were derived in proportion to other hours. This approach allowed some interesting detailed analysis of the most significant hours, although aggregating the lesser hours of diverse character at a tractable level.

The characteristics of the study areas that proved most definitive in this analysis were the study area size, as measured by population, and the level of congestion on highway facilities, as measured by the 24-hour ratio of volume to capacity for the urbanized portion of the study area. Congestion levels were obtained from the overall 24-hour modal split, which proved useful in some instances. In the morning peak period, a measure of population centralization proved most significant. This measure was taken as the ratio of central city population to study area population. A similar ratio of employment was actually preferred, but the two ratios were very highly correlated with each other, and population was held to be the more readily obtainable of the two statistics.

With only eight study areas, it was virtually impossible to include more than three variables in the development of any model, and generally only two were useful. It is entirely possible that inclusion of more study areas in this investigation could result in a revised shape of the models and perhaps allow use of other secondary variables to account for some of the situations that did not model well with present variables. Particular attention was given therefore to the reasonableness and internal consistency of the models developed, for the greatest confidence in the shape of the curves as currently modeled. A description of the individual models developed is as follows.

Wee Hours Period

Travel during this period is of very little consequence, amounting to approximately 2.4 percent of the total daily travel. An attempt was made to correlate the variations with several descriptive variables, but it was not successful. Therefore, 2.4 ± 0.6 percent of the average value for these eight study areas is recommended for the total amount of travel during the wee hours. The average breakdown by hours is as follows:

Hour	Average Percentage of Total Daily Travel
12 p.m. to 1 a.m.	0.75
1 a.m. to 2 a.m.	0.50
2 a.m. to 3 a.m.	0.35
3 a.m. to 4 a.m.	0.30
4 a.m. to 5 a.m.	0.50

Morning Period

The 3 highest hours were modeled by lumping together the 2 highest hours, splitting

Figure 2. Hourly distribution of vehicle-miles of internal auto driver travel by purpose.

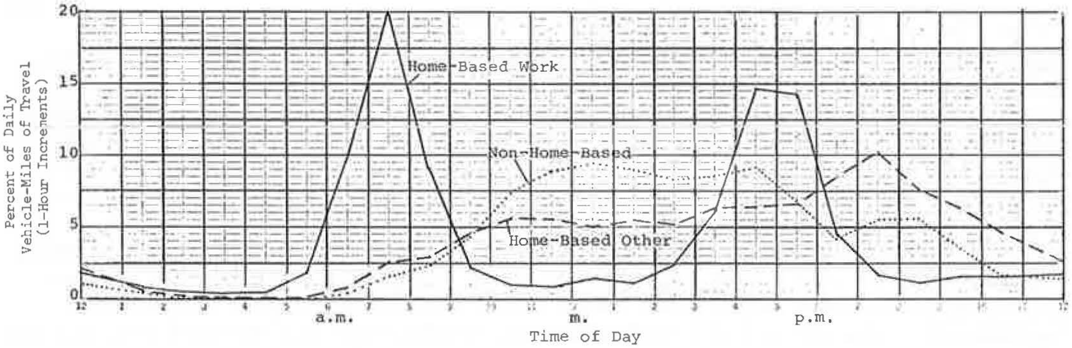


Figure 3. Hourly distribution of vehicle-miles of internal automobile driver, taxicabs, trucks, and total internal and external vehicles.

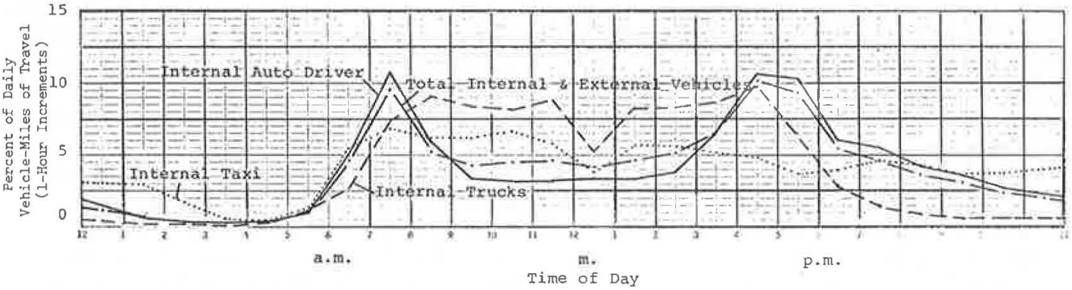


Table 3. Total vehicular travel from O-D surveys summarized in standardized hour periods.

	Percentage of Daily Travel Occurring in Each Time Period							
Standardized Time Period	Boston (+0.5)[a]	St. Louis (+0.2)	Seattle (+0.4)	Louisville (+0.5)	Oklahoma City (+0.6)	Colorado Springs (+0.1)	Stockton (+0.9)	Fall River (+0.3)
Morning								
5 to 6 a.m.	1.81	1.80	1.76	2.83	1.65	1.49	3.18	1.83
6 to 7 a.m.	6.52	6.04	6.23	7.03	6.61	4.77	6.57	5.62
7 to 8 a.m.	8.80	9.52	8.58	8.01	8.75	6.71	6.52	8.67
8 to 9 a.m.	5.47	5.38	4.98	4.97	5.26	4.64	5.03	5.22
Midday								
9 to 10 a.m.	4.72	4.37	4.53	4.82	5.05	4.84	5.28	4.69
10 to 11 a.m.	4.75	4.53	4.95	4.67	5.14	5.34	5.65	4.89
11 a.m. to noon	4.85	4.51	5.08	4.93	5.41	5.85	5.89	5.21
12 to 1 p.m.	4.81	4.33	4.95	4.89	5.58	6.27	5.82	5.51
1 to 2 p.m.	5.13	4.72	5.17	5.13	5.64	5.80	6.09	5.38
Afternoon								
2 to 3 p.m.	6.08	5.35	6.39	6.14	6.47	6.15	7.29	5.45
3 to 4 p.m.	7.64	7.29	8.09	8.52	7.98	7.27	8.47	7.03
4 to 5 p.m.	10.52	10.79	10.62	9.67	10.73	10.26	9.67	9.09
5 to 6 p.m.	7.44	8.10	7.17	6.94	7.38	7.92	6.39	6.78
6 to 7 p.m.	5.30	5.19	4.96	4.81	5.35	5.36	4.95	6.02
7 to 8 p.m.	4.40	4.48	4.19	4.22	3.60	4.93	2.99	5.45
Evening								
8 to 9 p.m.	3.36	3.49	3.28	3.53	3.07	3.70	2.48	4.36
9 to 10 p.m.	2.69	2.91	2.66	2.69	2.32	2.60	2.03	3.40
10 to 11 p.m.	2.29	2.33	2.13	1.80	1.40	2.09	1.67	2.41
11 p.m. to midnight	1.43	1.87	1.62	1.50	0.81	1.59	0.99	1.40
Wee hours								
12 to 1 a.m.	0.72	1.16	1.10	0.87	0.52	0.89	0.69	0.54
1 to 2 a.m.	0.39	0.62	0.51	0.49	0.36	0.47	0.48	0.27
2 to 3 a.m.	0.24	0.41	0.29	0.43	0.26	0.28	0.39	0.20
3 to 4 a.m.	0.22	0.33	0.20	0.41	0.21	0.24	0.52	0.18
4 to 5 a.m.	0.46	0.51	0.42	0.75	0.41	0.50	0.94	0.46
Total	100.04	100.04	99.86	100.00	99.96	99.96	99.98	100.06
Period subtotals								
Wee hours	2.03	3.03	2.52	2.95	1.76	2.38	3.02	1.65
Morning	22.60	22.74	21.55	22.85	22.27	17.61	21.30	21.23
Midday	24.25	22.47	24.68	24.39	26.82	28.10	28.73	25.68
Afternoon	41.38	41.20	41.42	40.30	41.51	41.89	39.76	39.82
Evening	9.77	10.60	9.69	9.52	7.60	9.98	7.17	11.57
Offset morning period								
Peak minus 2 hours	2.24	2.01	1.93	2.84	2.27	1.74	3.63	2.00
Peak minus 1 hour	5.01	5.43	5.29	6.13	4.78	4.39	4.59	4.65
a.m. peak hour	9.11	9.60	8.93	8.31	9.47	6.75	7.46	9.09
Peak plus 1 hour	6.24	5.70	5.40	5.57	5.75	4.73	5.62	5.59

[a] Offsets from actual p.m. clock hours shown in parentheses below each location.

that sum into two parts, and separately modeling the third hour. The model for the top 2 hours is shown in Figure 5 with the portion allocable to the higher hour shown in Figure 6. Increasing population size is seen to cause an increase in the percentage of daily travel in this 2-hour period, a consequence of more extensive home-to-work trip-making and longer trip lengths in the larger metropolitan areas. The impact of greater diffusion of the population base (and also the employment base) was also noted and used to account for the relatively low level of travel in Stockton and Colorado Springs as compared to Fall River, all of which are of comparable size.

The fraction attributable to the higher hour of these two is seen to be higher for the smaller study areas, decreasing as population increases (Fig. 6). This corresponds to the concept of the occurrence of rather broad peak periods in large metropolitan areas and narrower, sharper peaks in small study areas. This suggests a number of effects: One is congestion, which is usually worse in large areas; another is the greater diversification of activities in a large study area, promoting diffusion of trip-making away from a specific peak hour; and a third is that travel develops earlier in large areas because of the longer time required by many commuters to travel to work in large metropolitan areas as compared to travel time in small areas. It is very interesting to note that, in the home-based work travel distributions for Stockton and Louisville, there are two distinct start times for work shifts in these two areas, separated by 1 hour. This split of starting times had a marked effect in reducing the peak-hour percentage of travel, making the 2 top hours more closely equal.

The percentage of daily travel in the third highest a.m. hour is very nearly a constant 4.5 to 5 percent for all study areas, increasing slightly for larger study areas.

The fourth hour, split before and after the 3 high hours, is quite low in volume and proved to be difficult to relate to any meaningful region-wide descriptive variable. A constant value of 2.2 percent was determined as the appropriate average value to assign for this hour. It was applied with reasonable accuracy in most cases.

Midday Period

Travel in this time period is basically for non-work-related trip purposes and was found to be a function of population and V-C ratio. The population impact was the reverse of the situation in the peak periods, as might be expected. Small cities may be characterized as having less diverse travel patterns, more to-and-from-home-for-lunch trips, and travel is more restricted to daytime hours, whereas large metropolitan areas continue to show activity and, hence, travel in the evening hours. Thus, the midday percentage of daily travel decreases as population becomes larger. Although the fraction of daily travel may be less in larger areas, the amount of vehicular travel remains significant because the daily total is quite large. Thus, the impact of increasing daily congestion levels continues to force a reduction in the percentage of travel during the midday period. The instantaneous V-C ratio may not be as high during the midday as during the peak periods, but it is still larger than during the evening period.

The model for the aggregate percentage in this 5-hour period is shown in Figure 7. There is so little meaningful variation among the hours in this group that it is unimportant to model them explicitly. Dividing the aggregate percentage by five yields an average hourly percentage that may be taken as within ±10 percent of all hourly values for the period.

Afternoon Period

The 3 highest hours of this period constitute the p.m. peak period. Each of these 3 hours had been successfully modeled independently. Total study area population and the ratio of area-wide daily volume to capacity are the major variables. Modal split is slightly noticeable in the 2 highest hours but not in the third. This is an acceptable finding because it is only in the most highly congested times (peak hour) that significant diversion of trips to public transportation takes place. It is important to note here that a large modal split may occur as a consequence of high congestion, as represented by V-C ratio.

Figure 4. Cumulative percentage of total travel occurring in each group of hours for each city.

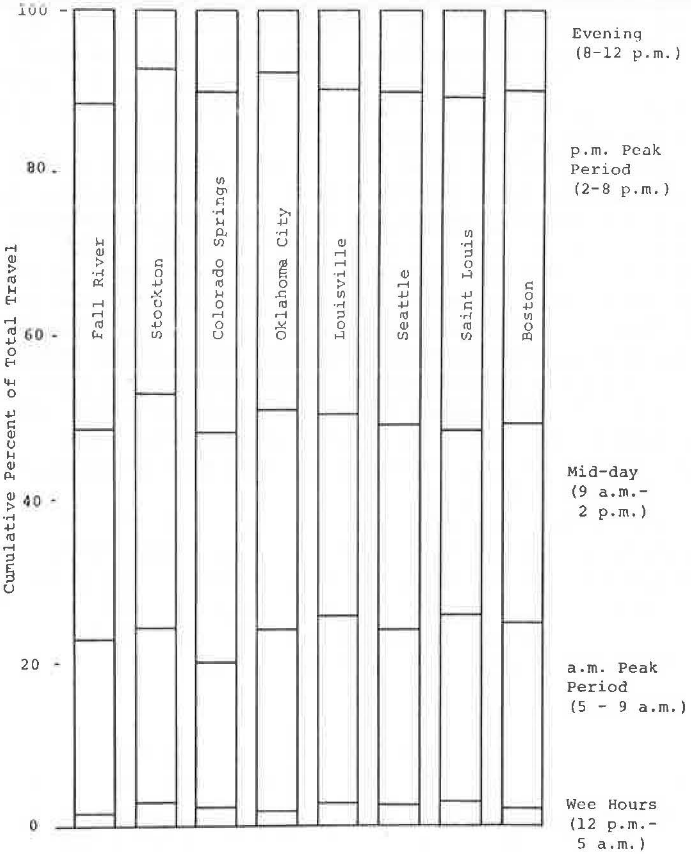
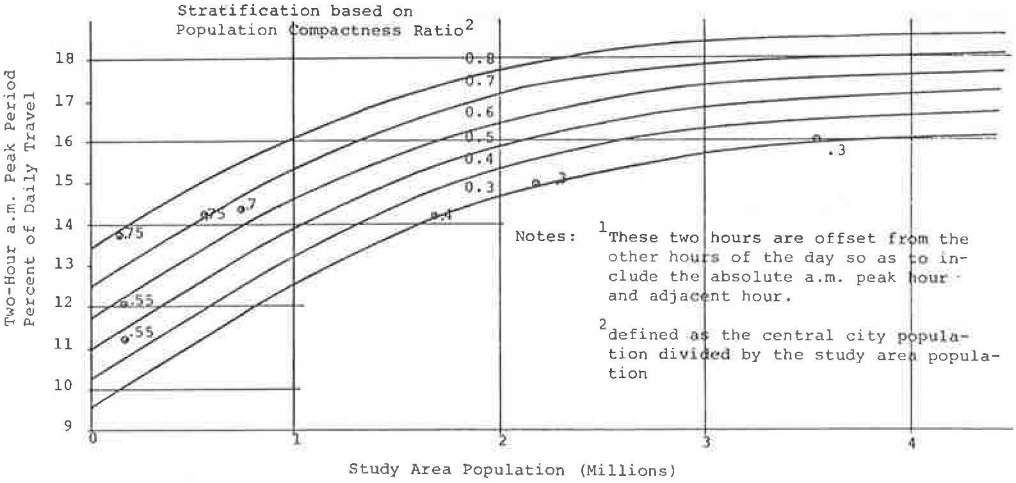


Figure 5. Percentage of travel in morning 2-hour period.



Notes: [1] These two hours are offset from the other hours of the day so as to include the absolute a.m. peak hour and adjacent hour.

[2] defined as the central city population divided by the study area population

The relations between hourly travel percentages and the regional descriptive variables are shown in Figures 8, 9, and 10 for the 3 highest hours. The form is similar to the morning peak period curves, except that the V-C ratio replaces population compactness ratio. The percentage generally increases with increasing population, again reflecting diversity of travel purposes and patterns in larger areas, whereas increasing congestion lowers the percentage ostensibly forcing some travel to occur in hours that the drivers might not have freely chosen.

The lower 3 hours of this period happen to occur in the early evening, for the most part, and include some of the travel presumed to be deferred from the peak period. Figure 11 shows how each of these 3 hours is derived as a percentage of the highest hour in the group (p.m. peak hour). Population enters mildly, but otherwise the factors are nearly constant from all study areas. Note that the first of these decreases with increasing population, representing the immediate reaction to the previous peak hours, and that the next 2 hours gradually shift back to the familiar positive trend.

As checks on overall accuracy of the modeling, the sum of these latter 3 hours should work out to be approximately 15 percent. The eight study areas all fall in the range of 14 to 16½ percent. Similarly, the range for the sum of all 6 hours was found to be from 39½ to 42 percent, which can also be used as a check.

Noteworthy phenomena in this 6-hour group include the split-peak aspect of the Lousiville and Stockton distributions and the fact that Stockton and Colorado Springs are frequently quite different in their distributions although they are practically identical with respect to most study area descriptive variables.

It has been assumed in developing the factors for the lower 3 hours that an erroneously high percentage would be modeled for the peak hour of Louisville and Stockton, and thus the lower hours are modeled to be factored from this value. The second highest hour, as predicted by this model, is as low in such cases as the first hour is high.

The difficulties in matching up the data from Colorado Springs and Stockton emphasize the fact that either the data contain errors or there are other as yet unknown variables that could differentiate between these cities, given more intensive research in this area.

Evening Period

All attempts to model this period accurately were fruitless. In every case, two or three of the eight study areas were misrepresented by 30 to 50 percent, whereas the others were well represented. In consequence, it is proposed that this 4-hour period be assigned a flat value of 9.5 percent. Some of the smaller study areas had values as much as 2 percent above or below this level, but no variable was found that could describe these variations. Attempts were made to correlate this period to the V-C ratio, midday period's percentages, morning period's percentages, and evening period's percentages. All proved particularly incapable of satisfactorily representing some of the small urban areas.

Given the 4-hour total as allocated by the preceding method, it is possible to distribute accurately this percentage among the 4 hours. A rather linear decline was noted from the highest to the lowest of these hours in all cases. Only the slope, or rate of decay, varied among the study areas. As would be expected, travel diminishes most rapidly for small areas and most slowly for the large areas where there is much more late-night activity. The decrease per hour, δ, is to be used as follows:

$$\text{Highest hour} = \frac{\text{total percentage}}{4} + \tfrac{3}{2}\,\delta$$

$$\text{Each lower hour} = \text{preceding hour} - \delta$$

These rank-ordered hours were in most cases also in chronological order from 8 p.m. to midnight.

Figure 6. Morning peak-hour fraction of 2 highest morning hours.

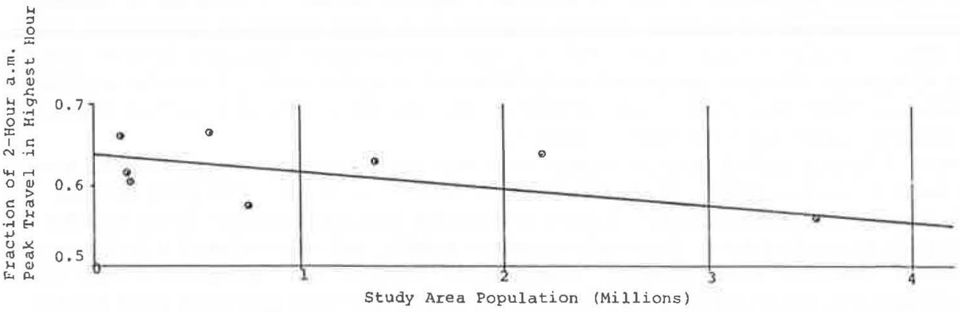


Figure 7. Percentage of total travel in 5-hour midday period.

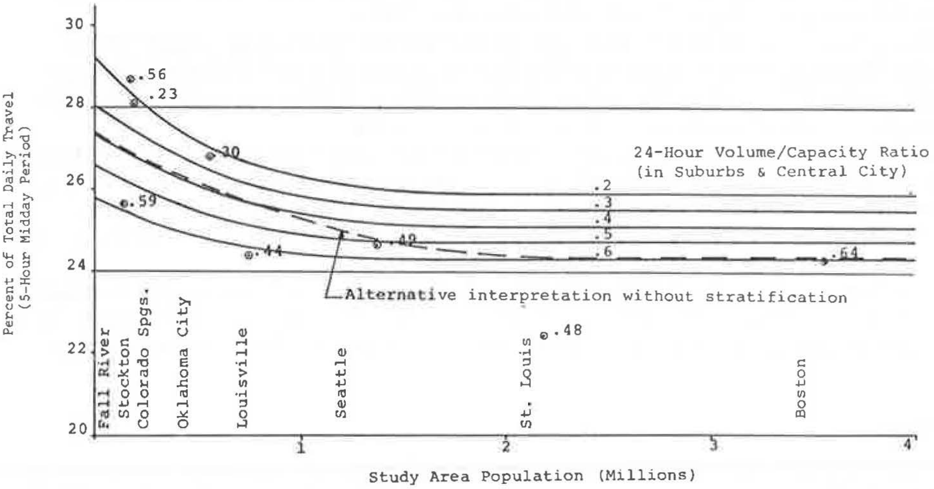
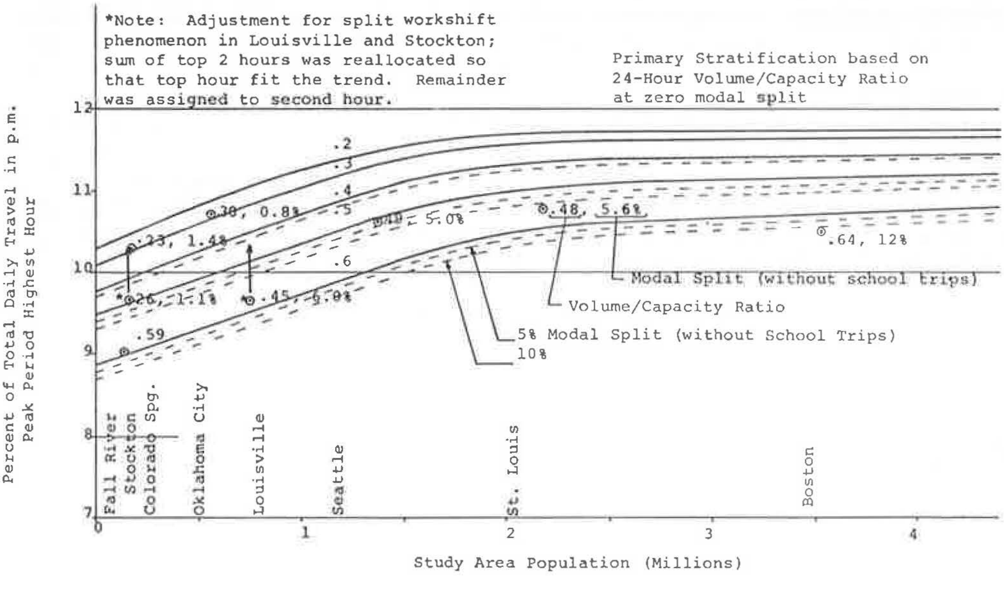


Figure 8. Percentage of total travel in p.m., highest hour.


*Note: Adjustment for split workshift phenomenon in Louisville and Stockton; sum of top 2 hours was reallocated so that top hour fit the trend. Remainder was assigned to second hour.

Primary Stratification based on 24-Hour Volume/Capacity Ratio at zero modal split

Figure 9. Percentage of travel in p.m., second highest hour.

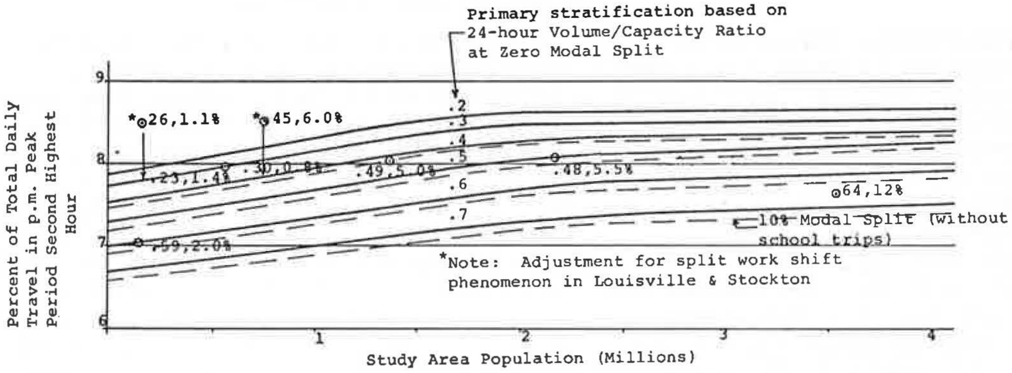


Figure 10. Percentage of total travel in p.m., third highest hour.

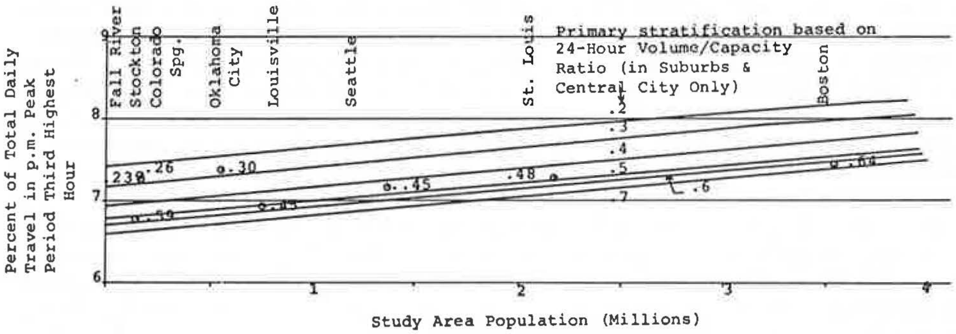
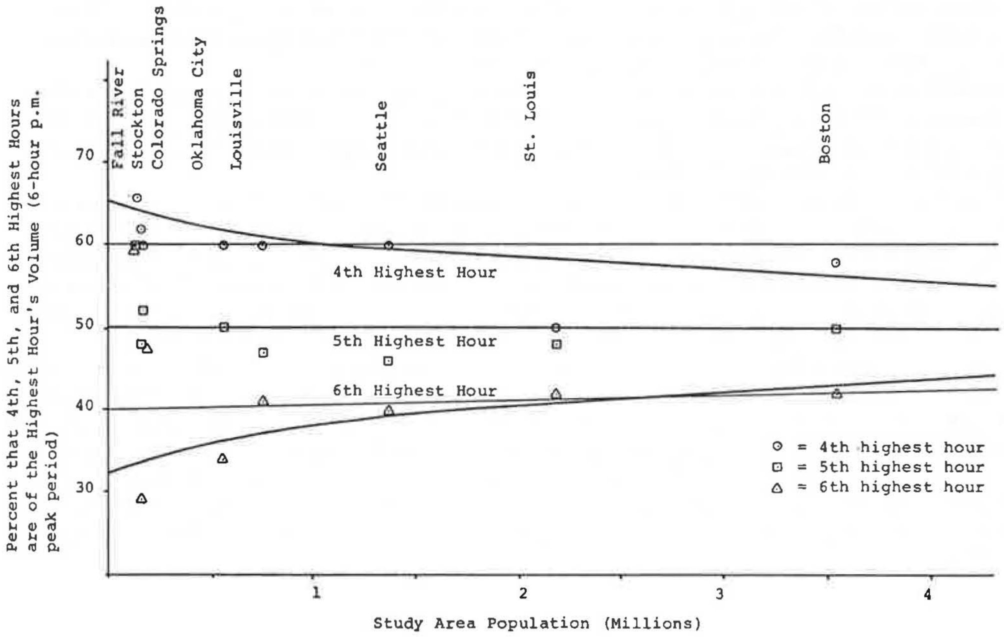


Figure 11. Percentage of travel in p.m., 3 lowest hours.



PHASE B: ANALYSIS OF TRAVEL BY TIME OF DAY FOR SPECIFIC FACILITY TYPES AND LOCATIONS

The temporal distribution of total vehicular travel, as measured from area-wide coverage traffic counts, was conducted in two separate steps. The first was to disaggregate the 24-hour area-wide total vehicular travel into the 27 classes of highway facilities used in this research investigation. The second step involved the development of hourly distributions for each of the 27 classes (where data existed) for 2,000 urban area locations and the further disaggregation of the distributions according to predominant direction of travel within each cell, again subject to the availability of data (Table 2).

The modeling procedure used to allocate the 24-hour total vehicular travel to the 27 classes of highway facilities was structured such that the total vehicular travel by classifications that were outside the current urban-in-fact area (e.g., the rural area and other subcenters within the urban transportation area) were removed at the beginning. The total vehicular travel on collector facilities within the urban-in-fact area was removed next.

Series of submodels were developed for the temporal distributions of travel developed for each of the 27 classes of highway facilities on a nondirectional basis. The initial analysis of the distributions developed for each of the study sites indicated that major differences in temporal distribution of travel occurred within a given facility class depending on whether it was located within a small or large study area. Therefore, the submodels were further stratified into small urban areas (less than 250,000 population) and large urban areas (more than 250,000 population). This resulted in the final 41 nondirectional submodels. Another observation was from the comparison of cities, which showed that St. Louis' distribution differs from Boston's because of a lower congestion factor.

The next series of submodels took the nondirectional temporal distribution of total vehicular travel and split it in the two directions of travel. Again, as with the nondirectional submodels, the classifications were further stratified, dependent on urban size, with the stratifications the same as before. This led to the creation of 39 directional submodels. Travel in the (morning) peak direction accounted for a high of 70 percent between 7 and 8 a.m. Boston was generally higher than St. Louis, again because of the higher congestion factor.

Examples of the temporal distributions of nondirectional travel were prepared in particular for the following types of facilities: freeway-expressway, central city—radial; freeway-expressway, suburb—radial; arterial, central city—circumferential-crosstown; and arterial, rural—circumferential-crosstown.

The distributions are consistent for all the cells, and the findings generally match distributions of the area-wide analysis. Travel is low in the wee hours, peaks in the morning peak hours, falls off in the midday hours, peaks again in the afternoon hours, and then falls off in the evening hours.

The most noteworthy difference occurs between Boston and St. Louis. St. Louis' travel has a higher peak in the morning and afternoon hours and peaks about 1 hour sooner, but a lower distribution occurs midday and in the evening hours. This finding parallels closely the results of area-wide analysis because of the lower V-C ratio in St. Louis. It is recognized that Boston has higher overall congestion, and, hence, the temporal distribution tends to be more evenly distributed throughout the day.

Plots of the percentage of travel in morning peak direction distributed to time of day were prepared for directional travel. As in the case of nondirectional travel, the most significant difference in the temporal distributions is between large-city and small-city groupings. The percentage of travel in the morning peak direction is lowest in the wee hours (42 percent), falls off and then peaks again at 55 percent between 6 and 7 p.m., and drops off to 42 percent from 11 p.m. to midnight. St. Louis shows a generally lower percentage (approximately 5 percent) than Boston because of the lower level of traffic congestion.

CONCLUSION AND RECOMMENDATIONS

In general, the results of the research achieved the objectives of the program. There are, however, several areas where further research could well lead to increased knowledge and improved modeling techniques.

The first of these recommendations is to use the total vehicular travel data from the same sites as was used in this modeling effort and to expand the number of independent (causal) variables that could be used in the modeling process. For example, urban area characteristics could be disaggregated by subarea and subclass. Also, V-C ratio and perhaps modal split to transit could be calculated for time periods consistent with the modeling periods rather than on just a 24-hour basis.

The second recommendation is to include more flexibility through the use of more urban area studies. The number of study areas (eight and six used in phase A and phase B respectively) are at best the very minimum number acceptable. As it is, there are still a number of urban area types and sizes that are not represented in the data used to create the models. Also, the limited amount of information available did not allow for the independent checking of the models. For these reasons, it seems that the addition of several more sites would be appropriate.

The next recommendation would be to carry out this investigation for two or more points in time using the same study site. This effort had been intended for the original research investigation, but it was found that the time and effort required to locate and collect data in a compatible format from the older (pre-1960) studies were markedly greater than permitted by the time constraints of the research project. It would be most interesting to carry out this time-series analysis for both phases A and B. However, based on the results obtained from this study, it might be almost as interesting to carry out the analysis using only phase B data, which are considerably more available and would, therefore, be much easier to obtain than data for phases A and B together.

The final recommendation deals with attempting to obtain phase B data for a shorter time period than the 1-hour basis used, particularly during the morning and afternoon time periods. These data (perhaps on a 15-min interval basis) would allow for the identification of absolute peak hours and periods of travel as was the case in phase A. Although the differences between clock hours and absolute peak hours were not too great for area-wide data, they were acceptable. It would be expected that these differences are perhaps somewhat greater when individual facilities in the highway network are considered.

The four recommendations for further research listed previously are only a few of the possible ones growing out of this research investigation.

ACKNOWLEDGMENTS

The authors wish to acknowledge the helpful cooperation of FHWA personnel. The assistance of Michael Birdsall and Robert Hammond is greatly appreciated for assembling and processing the statistics used in this paper.

The authors wish to point out that the conclusions presented in this paper do not necessarily represent the opinions of their agencies.

SPONSORSHIP OF THIS RECORD

GROUP 1—TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Charles V. Wootan, Texas A&M University, chairman

Committee on Transportation Forecasting
George V. Wickstrom, Metropolitan Washington Council of Governments, chairman
Mark M. Akins, Daniel Brand, Austin E. Brant, Jr., Henry W. Bruck, Francis E. Coleman, Thomas B. Deen, Robert B. Dial, William L. Grecco, Charles H. Groves, John R. Hamburg, Lawrence V. Hammel, Donald M. Hill, Wolfgang S. Homburger, Frank E. Horton, Gerald H. Johnston, Louis E. Keefer, Robert Kochanowski, Herbert S. Levinson, Brian V. Martin, James J. McDonnell, Eugene R. Mills, Kenneth W. Shiatte, Paul W. Shuldiner, Vergil G. Stover, Anthony R. Tomazinis, Robert T. Wood, Richard D. Worrall

James A. Scott, Highway Research Board staff

The organizational units and the chairman and members are as of December 31, 1972.