HIGHWAY RESEARCH RECORD

Number 165

Origin and Destination

Advances in Transportation Planning

8 Reports

Subject Area

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., 1967 Publication 1443

Department of Traffic and Operations

Harold L. Michael, Chairman
Associate Director, Joint Highway Research Project
Purdue University, Lafayette, Indiana

HIGHWAY RESEARCH BOARD STAFF

E. A. Mueller, Engineer of Traffic and Operations

COMMITTEE ON ORIGIN AND DESTINATION
(As of December 31, 1965)

Alan M. Voorhees, Chairman
Alan M. Voorhees and Associates
Washington, D. C.

David K. Witheford, Secretary
Transportation Planning Consultant
West Haven, Connecticut

Willard H. Armstrong, Bureau of Planning Survey, Ohio Department of Highways, Columbus
John E. Baerwald, Department of Civil Engineering, University of Illinois, Urbana
Robert D. Bedwell, Highway Economist, California Division of Highways, Sacramento
Glenn E. Brokke, U. S. Bureau of Public Roads, Washington, D. C.
Nathan Cherniack, Economist, The Port of New York Authority, New York, N. Y.
Donald E. Cleveland, Department of Civil Engineering, University of Michigan, Ann Arbor
Roger L. Creighton, Director, Upstate New York Transportation Studies, Albany
John A. Dearinger, Department of Civil Engineering, University of Kentucky, Lexington
Thomas B. Deen, Alan M. Voorhees and Associates, Washington, D. C.
George Ferguson, Pittsburgh Regional Planning Association, Pittsburgh, Pennsylvania
Thomas J. Fratar, Partner, Tippetts-Abbett-McCarthy-Stratton, New York, N. Y.
W. L. Grecco, Purdue University, Lafayette, Indiana
Harold W. Hansen, Senior Planning Engineer, Portland Cement Association, Chicago, Illinois
Walter G. Hansen, Alan M. Voorhees and Associates, Washington, D. C.
Kevin E. Heanue, Highway Engineer, Urban Planning Division, U. S. Bureau of Public Roads, Washington, D. C.
G. H. Johnston, Assistant Planning Studies Engineer, Traffic and Planning Studies Division, Ontario Department of Highways, Downsview, Canada
Joseph Kates, Traffic Research Corporation, Ltd., Toronto, Ontario, Canada
Louis E. Keefer, Consultant for Transportation Planning and Research, Woodmont, Milford, Connecticut
Norman Kennedy, Professor of Transport Engineering, Institute of Transportation and Traffic Engineering, University of California, Berkeley
Warren B. Lovejoy, The Port of New York Authority, New York, N. Y.
John T. Lynch, McLean, Virginia
Brian V. Martin, Assistant Chief Engineer (Research), Department of Highways and Transportation, Greater London Council, London, England
James J. McDonnell, U. S. Bureau of Public Roads, Washington, D. C.
W. L. Mertz, Technical Director, Tri-State Transportation Committee, New York, N. Y.
John K. Mladinov, Director, Puget Sound Regional Transportation Study, Seattle, Washington

Robbie W. Parker, Vogt, Ivers and Associates, Consulting Engineers, Jeffersonville, Indiana
William S. Pollard, Jr., Memphis, Tennessee
Lloyd A. Rivard, Planning and Programming Engineer, Automotive Safety Foundation, Washington, D. C.
James J. Schuster, Civil Engineering Department, Villanova University, Villanova, Pennsylvania
Arthur Schwartz, Mass Transportation Planner, Tri-State Transportation Committee, New York, N. Y.
Paul W. Shuldiner, Consultant to the Office of the Under Secretary for Transportation, Northeast Corridor Study, U. S. Department of Commerce, Washington, D. C.
Bob L. Smith, Associate Professor of Civil Engineering, Kansas State University, Manhattan
Anthony R. Tomazinis, Chief, Transportation Flows Simulation, Penn-Jersey Transportation Study, Philadelphia, Pennsylvania
George V. Wickstrom, Program Director, New Castle County Program, Wilmington, Delaware
Martin Wohl, National Science Foundation Fellow, University of California, Berkeley
J. E. Wright, Traffic Manager, Planning Survey Division, Texas Highway Department, Austin
F. Houston Wynn, Wilbur Smith and Associates, New Haven, Connecticut

Foreword

Urban transportation studies produce a great deal of literature dealing with perplexing problems of planning adequate transportation facilities for ever-changing urban environments. Findings as portrayed in this RECORD are useful to those engaged in the performance of transportation studies as well as urban planners, mathematicians and highway engineers who must work with study data in arriving at comprehensive transportation plans that satisfy the demands of a mobile public.

This RECORD has eight papers concerned with uses of O-D data; four of these are directly concerned with mathematical modeling problems and the others discuss the study of freight movements, land-use linkages, estimating traffic volume and manufacturing trip-generation data.

The so-called opportunity model used to distribute the various systems of traffic to a transportation network is examined by Ruiter in the first paper. A promising calibration method was developed which seems freer from errors than other methods. Its application to future work seems warranted. Caswell's paper also deals with the opportunity model and presents formulas and tables designed to reduce errors in its application. Errors in the model are inherent in the use of zone size, trip density, and zonal interaction and the paper gives further insight into these complexities.

A major problem in the urban transportation planning process is the determination of how many users will choose a particular mode of travel. Further development and application of a "modal-split" model on a large scale is discussed by Tomazinis. Using some of the relationships found, better testing of alternative transportation studies is possible. Boyce has used analysis of variance to test some assumptions used with trip-distribution models. A test for trip direction effect is developed and applied.

Studies of freight movement are usually not made or are made on a reduced scale. Wood's paper describes the monumental work under way in the Metropolitan New York area to measure and describe all local and intercity freight traffic by mode. The result of the work will be to develop a base for future projection of needs.

Horton and Shuldiner applied Markov models to analysis of O-D data and found that they offer good potential for adding insight on travel behavior for multipurpose trips. The research indicates that more exploration of the application of this type of model would lead to better understanding and application of trip data.

Mathematical modeling is employed for estimating traffic volume at a point in Schneider's paper. The practical need for an easy and reliable way to estimate average volumes at any particular point is felt by all who are involved in the process of traffic estimation. The research showed considerable promise after preliminary testing and is being carried forward.

Manufacturing trip generation data have been extensively investigated by Kolifrath and Shuldiner and differences in precision obtained by varying the basis of analysis are set forth.

Contents

Improvements in Understanding, Calibrating, and Applying the Opportunity Model

EARL R. RUITER, Traffic Planning Analyst, Chicago Area Transportation Study, Chicago, Illinois

•THE Opportunity Model is the name given to the mathematical procedure developed by Morton Schneider to distribute, over all possible destinations, the actual destinations of all trips having a stated origin. The distinguishing feature of the model is its unique independent variable, intervening opportunities. Although new to the urban transportation field when proposed, this variable has been a feature of earlier models of human behavior in the fields of population migration and intercity travel (4).

Since the completion of its 1980 transportation plan for the Chicago area, the Chicago Area Transportation Study (CATS) has been involved in a number of projects in which detailed traffic forecasts are needed for relatively small areas within the large Chicago metropolitan region. Two of these areas, the Fox River Valley and Lake County, are shown in Figure 1 along with the original CATS region to indicate the types of areas with which CATS has been dealing. The application of the Opportunity Model of the trip distribution process to these areas was more difficult than the application to the entire metropolitan area. Calibration of the model to actual vehicle-miles of travel and to screenline counts was impossible with the two-parameter (two L values) model used in the earlier large area applications. These problems indicated a need for improved methods of applying the model. A study of the theoretical bases of the model and of calibration methods indicated that improved methods using multiple L values could be developed, two for each zone or group of zones, as parameters in the Opportunity Model.

Because of computer size limitations, it is felt that trip distribution applications to small areas within large metropolitan regions will become more important in the years to come, as such regions increase in size. These regions are becoming too large for complete inclusion in a single assignment run. An example is the Chicago region, in which the Chicago-Northwestern Indiana Standard Consolidated Area, as defined by the 1960 Census, includes eight counties (Fig. 1). By comparison, the 1956 CATS area includes parts of only four counties.

This paper discusses the improved understanding of the Opportunity Model which has resulted from the application of the model to small areas. It also explains the calibration methods developed and the results obtained with these methods. The Opportunity Model itself, rather than the CATS assignment system, which uses the model, is the major subject of the paper. Computer-oriented documentation of the assignment system in its entirety is provided in two CATS publications (8, 9).

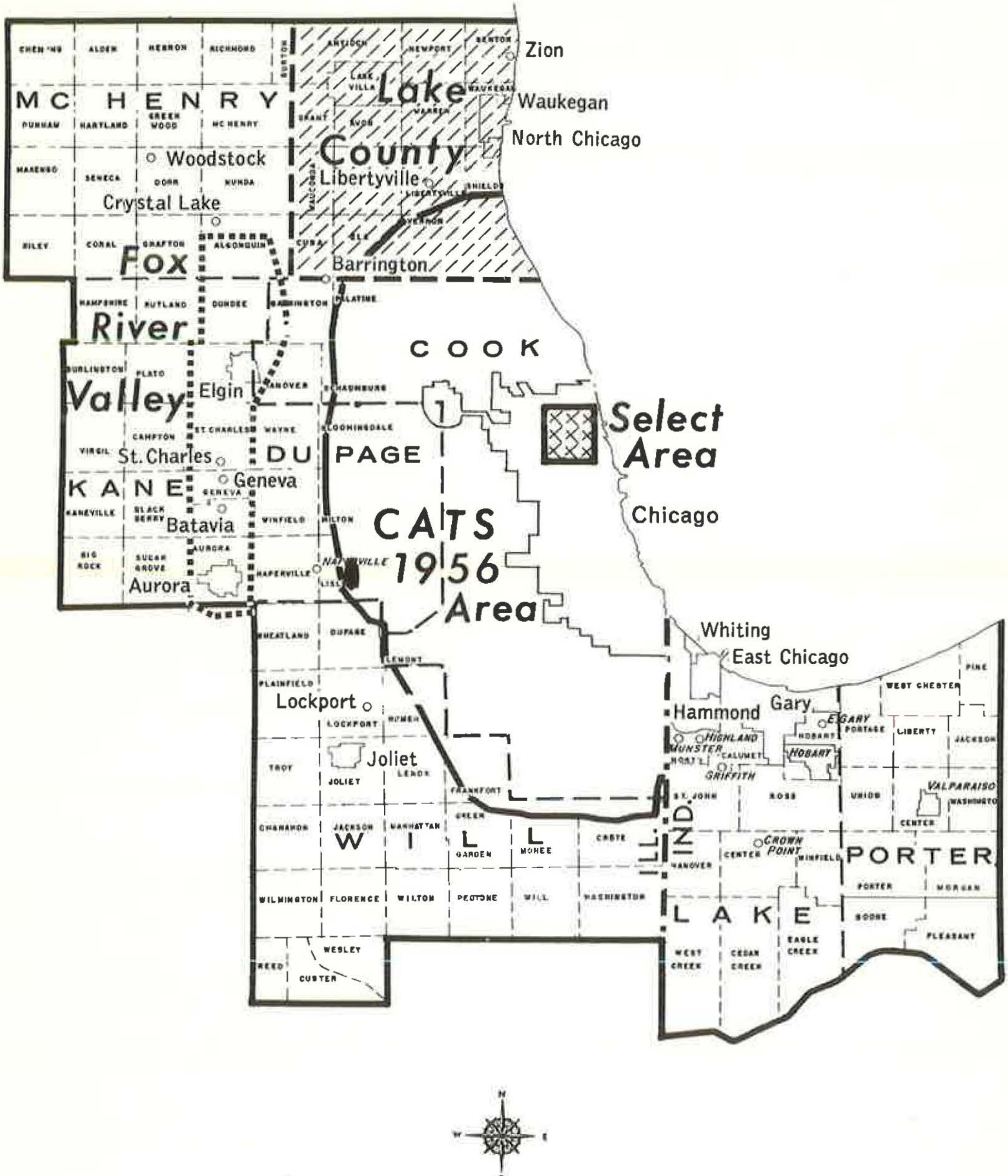
UNDERSTANDING THE OPPORTUNITY MODEL

Hypotheses and Mathematics

The hypotheses and mathematics underlying the Opportunity Model are given briefly as a starting point before the discussion of interpretations of the model and its parameter, and the presentation of relationships between the model and trip parameters (1).

The Opportunity Model is based on the hypotheses that (a) total travel time from a point is minimized, subject to the condition that every destination point has a stated probability of being accepted if it is considered; and (b) the probability of a

Paper sponsored by Committee on Origin and Destination.



Map Prepared By
CHICAGO AREA TRANSPORTATION STUDY

Figure 1. Assignment areas within Chicago-Northwestern Indiana standard consolidated area.

destination being accepted, if it is considered, is a constant, independent of the order in which destinations are considered.

The hypotheses lead to the following mathematical formulation, in terms of limitingly small quantities:

$$dP = L \left[1 - P(V) \right] dV \qquad (1)$$

where

dP = probability that a trip will terminate when considering dV possible destinations;
$P(V)$ = total probability that a trip will terminate by the time V possible destinations are considered;
V = possible destinations already considered, or subtended volume; and
L = constant probability of a possible destination being accepted if it is considered.

The solution of the differential Eq. 1 is

$$P(V) = 1 - e^{-LV} \tag{2}$$

The expected interchange from zone i to zone j (T_{ij}) is the volume of trip origins at zone i (O_i) multiplied by the probability of a trip terminating in j:

$$T_{ij} = O_i \left[P(V_{j+1}) - P(V_j) \right] \tag{3}$$

or

$$T_{ij} = O_i \left[e^{-LV_j} - e^{-LV_{j+1}} \right] \tag{4}$$

The subtended volumes (V's) are the sums of the possible destinations considered before reaching a given zone. As it can be assumed that a zone's trip origins equal its trip destinations over a 24-hr period, V_j can be defined in terms of the trip origins reached before reaching zone j:

$$V_j = \sum_{k=1}^{j-1} O_k \tag{5}$$

where the O_k's are arranged in order of increasing travel time from zone i.

Although Eq. 5 could be substituted into Eq. 4 to express T_{ij} completely in terms of trip origins and the L value, it is more convenient to leave the equations as given in Eq. 4.

Initial applications of the Opportunity Model showed that it would be necessary to specify more than one value for L, because of the differing probabilities of acceptance associated with different types of trips. For example, people are more selective in choosing places to work than they are in choosing places to shop for groceries. Three trip subpopulations (short, long residential, and long nonresidential) with two L values (short and long) satisfactorily represented empirical trip data for large regions. A mathematical statement of the Opportunity Model, as used in the CATS assignment system, is, therefore:

$$T_{ij} = \sum_{k=1}^{3} O_{ik} \left[e^{-L_k V_{jk}} - e^{-L_k V_{j+1,k}} \right] \tag{6}$$

where k ranges over the three trip subpopulations.

The discussion of interpretations of the model and the L value which follows is based on just one of the trip subpopulations of Eq. 6. It is assumed that the model holds in the simple form of Eq. 3. When the time comes to speak of operational problems, the necessary trip subpopulations will be reintroduced.

Interpretations of Model

The Opportunity Model can be considered in its broadest sense as an explanation of human behavior, as stated in the two foregoing hypotheses. The fact that the model

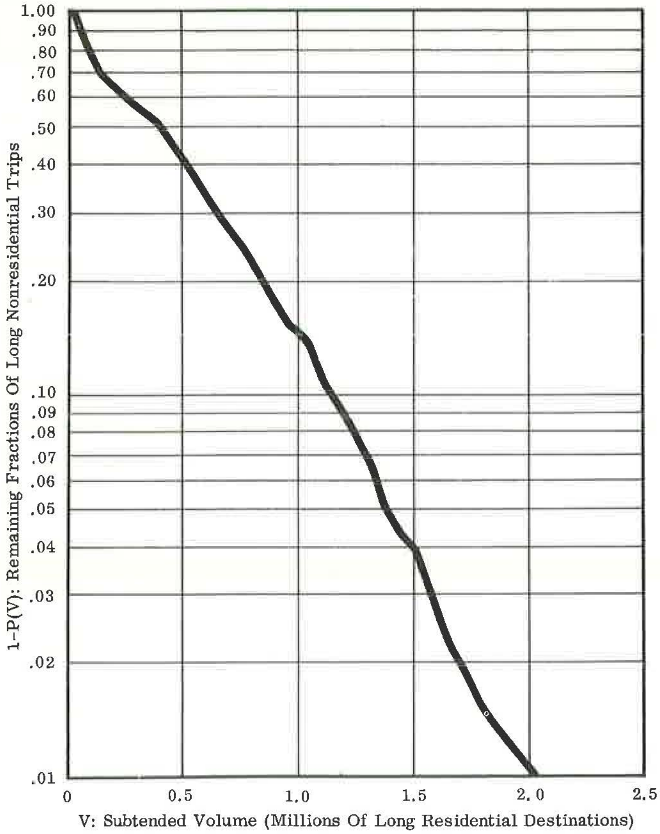


Figure 2. Cumulative distribution of long residential trips from CATS zone 001 according to number of opportunities.

has proved to be satisfactory for metropolitan regions indicates that, when averaged over a large area, people do behave as hypothesized in the model.

The model can be interpreted in a more limited sense by considering the mathematical expression of the model given in Eq. 2 as an equation which is to be fitted to empirical data by adjusting the parameter, L. Equation 2 can be changed to a linear form by rearranging and taking natural logarithms of both sides. This procedure results in the following:

$$-LV = \ln \left[1-P(V) \right] \tag{7}$$

Empirical values of V and 1-P(V) for a given zone can be plotted on semilog graph paper (Fig. 2). Theoretically, this plot will be a straight line for all trip types and all origin zones. The parameter, or L value, can be found by least squares regression. The procedure is very simple once values of V and 1-P(V) have been determined from survey data. These empirical values can be found once actual trip interchanges from a given zone are arranged in travel time order.

This interpretation of the Opportunity Model indicates that once the concepts of subtended volume (V) and probability of trip termination [P(V)] are understood, or simply accepted, the model can be thought of as a statement that a semilog relationship tends to exist between V and 1-P(V). The model expressed in a statement of this type may appeal to those who have difficulty visualizing the basic hypotheses of the model.

Interpretations of L Value

The curve-fitting approach previously discussed leads to a graphic interpretation of the L value. This parameter can be viewed simply as the slope of the straight line which best fits a set of empirical V and ln [1-P(V)] data.

Mathematically, the L value can be expressed:

$$L = \frac{-\ln [1-P(V)]}{V} \tag{8}$$

Two characteristics of the L value are evident in Eq. 8. The sign of L is always plus, because 1-P(V) is always less than one and its natural logarithm is always negative. The units of L are (1/opportunities), or (1/trip ends), as the numerator of Eq. 8 is unitless and the denominator has the unit opportunities, or trip ends. Experience with empirical data indicates that L is always very small, usually of the order of 10^{-5}, and always much less than one.

These three characteristics all support the interpretation of the L value as a modified probability quantity. Just as for other probability quantities, L lies between zero and one. Unlike more common probabilities, L is not unitless. It can be thought of as the probability, per individual opportunity or trip end, of destination acceptance. A reading of the hypotheses of the model shows that this interpretation of the L value is in agreement with the model's theoretical basis.

The interpretation of the L value leaves room for these parameters to vary from origin zone to origin zone. In fact, the interpretations given here are based on one origin zone and would be seriously limited if L values could not vary from zone to zone. The realization that multiple L values are desirable from an interpretive point of view was the first of two breakthroughs to CATS researchers attempting to apply the Opportunity Model in small areas. The second breakthrough was that the CATS assignment system was understood well enough so that it could be modified to accept multiple L values (9).

Relating L Values to Trip Parameters

The L value has been interpreted in terms of subtended volume and fraction of trips unsatisfied. For the extremely simplified situations in which trip end density is assumed to be constant, the L value can be expressed in terms of average trip length and average trip end density. As trip length and trip end densities are more common trip parameters than subtended volume, the expression obtained, although a simplification, provides insights into the nature of the L value.

In addition to the assumption of uniform trip end density, it is necessary to assume that the time ranking of possible destinations can be replaced by a distance ranking without loss of accuracy. This assumption would be true if the speed in all parts of the transportation system were constant, or nearly so.

Since the Opportunity Model is probabilistic in nature, the probabilistic concepts of mathematical expectation can be used to find desired averages. For example, average trip length ($\bar{r}$) may be found by performing the following integration:

$$\bar{r} = E(d) = \int_a^b f(V)\, dP(V) \tag{9}$$

where

d = distance variable (mi);
E(d) = expected value, or average, of distance variable (mi);
f(V) = expression of distance in terms of variable V (subtended volume);
dP(V) = density function of variable V; and
a, b = lower and upper limits on V.

The density function needed can be obtained by differentiating Eq. 2:

$$dP(V) = Le^{-LV}\, dV \tag{10}$$

Because of the assumptions of constant density and distance ranking, subtended volume can be expressed in terms of distance as follows:

$$V = \rho \pi d^2 \tag{11}$$

where ρ = average trip end density (trip ends/mile^2). Equation 11 can be solved for d to obtain f(V):

$$d = f(V) = \left[\frac{V}{\rho \pi}\right]^{1/2} \tag{12}$$

Equations 10 and 12, along with the limits on V (0 and ∞), can be substituted into Eq. 9:

$$\bar{r} = \int_0^{\infty} \left[\frac{V}{\rho \pi}\right]^{1/2} Le^{-LV}\, dV \tag{13}$$

Carrying out the integration and simplifying, the following expression for $\bar{r}$ is obtained:

$$\bar{r} = \frac{1}{2}\left[\frac{1}{\rho L}\right]^{1/2} \tag{14}$$

or, solving for L,

$$L = \frac{1}{4 \rho \bar{r}^2} \tag{15}$$

Although it must be remembered that Eq. 15 is a gross approximation, it does indicate that L tends to be inversely proportional to trip end density and to the square of average trip length. Experience with CATS assignments has confirmed these tendencies.

A dimensional analysis of Eq. 15 indicates that, as ρ is expressed in trip ends per square mile and $\bar{r}$ is expressed in miles, the L value has the units (1/trip ends). Thus, Eq. 15 agrees dimensionally with the theoretical basis of the Opportunity Model.

CALIBRATION

The major benefit of the improved understanding of the Opportunity Model has been the enhanced ability to calibrate individual assignments. Calibration techniques are desired so that empirical data or predictions, such as total vehicle mileage and screenline traffic counts, can be duplicated by the assignment process without running a large number of expensive "trial and error" computer runs.

Criterion

The matching of actual or predicted average trip lengths by the assignment was the criterion which led to the most useful calibration techniques. Since the number of trips in an area must be specified before an assignment can be run, and since total vehicle-miles of travel is the product of total trips and their average trip length, the matching of average trip lengths means that actual or predicted vehicle-miles of travel will be matched by the assignment.

A case could conceivably be made for using the criterion of matching actual or predicted trip travel times which are more directly related to the Opportunity Model. The main reason for rejecting this criterion is that experience indicates travel time data obtained in travel surveys is much less reliable than travel distance data. Travel time

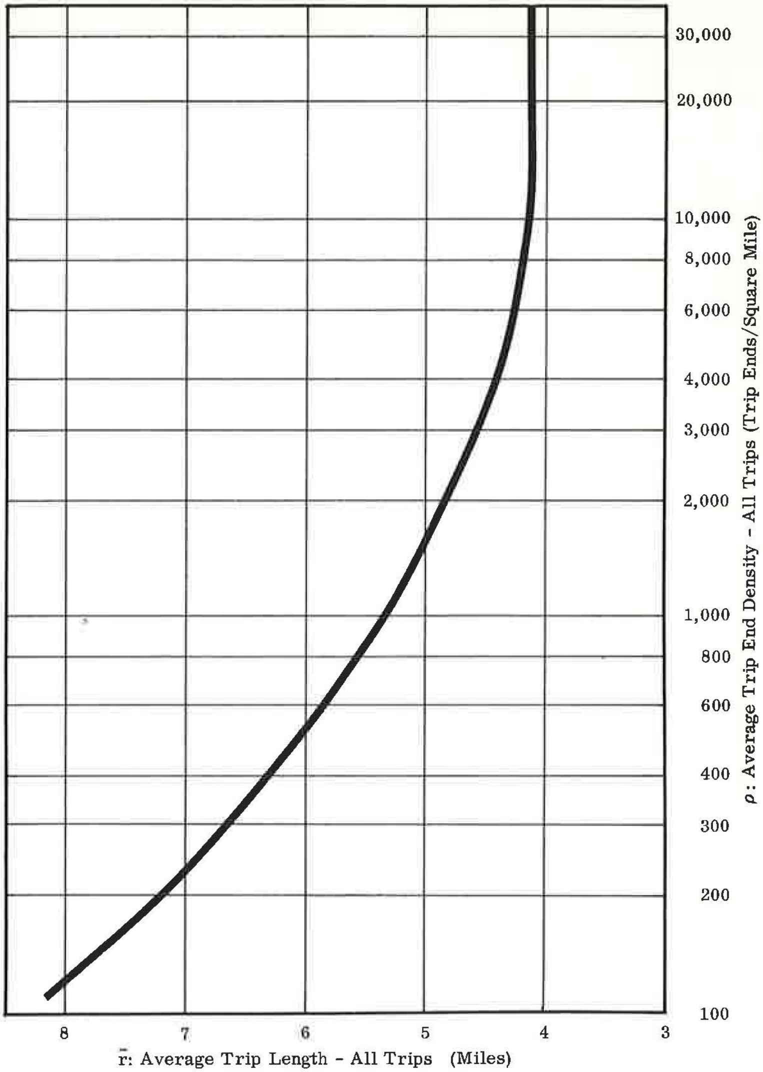


Figure 3. Relationship between total average trip length and total trip end density.

data must be estimated by the trip maker, and this estimation is often quite gross. Travel distances, on the other hand, can be calculated from origin and destination information, which typically is much more accurately reported in travel surveys.

The prediction of future average trip lengths must, of course, be accomplished when assignments serving as predictions of the future are to be calibrated. Although the problem of predicting future average trip lengths is far from completely solved, it is being dealt with. A National Cooperative Highway Research Project has as its goal the determination of trends in average trip lengths which can be observed over time on an area-wide basis (3, 7).

An investigation of zone-by-zone variations in average trip lengths, as they existed at the time of CATS 1956 travel surveys, has been conducted. One result of the investigation has been the discovery that an important source of zone-to-zone variation in average trip lengths is zonal trip end density. A hand-fitted plot of zonal average trip length vs zonal trip end density is shown in Figure 3. The fact that zones with few trips tend to have a long average trip length is apparent. This tendency is an affirmation of the hypothesis of the Opportunity Model which states that the satisfaction of trips and, therefore, average trip length, is affected by the number of available destinations.

As more sources of variation in zonal average trip lengths which can be applied to a future situation are discovered, it will be possible to predict zonal average trip lengths, once area-wide trends have been determined. For example, if it is assumed that the curve of Figure 3 will remain constant over time, future average trip lengths can be determined easily by reading values from the curve once future trips have been generated for each zone.

Single L Value Calibration Method

The first attempt to develop a calibration method was the application of Eq. 14 to the case in which only one L value is desired per trip population. This method has been attributed to Morton Schneider (5). Inasmuch as Eq. 14 is approximate, and the "constant" term is not exactly 0.25 in each case, the equation has been modified to a ratio form so that the results of an already completed assignment can be used as added information when planning a new assignment. The ratio form is

$$\frac{\bar{r}_1}{\bar{r}_2} = \frac{\sqrt{L_2 \rho_2}}{\sqrt{L_1 \rho_1}} \tag{16}$$

where the average trip lengths ($\bar{r}_1$ and $\bar{r}_2$) and average trip end densities (ρ_1 and ρ_2) are obtained by considering the entire assignment area. The subscripts 1 and 2 refer to two particular times, places, or assignment runs.

The L value obtained by using Eq. 16 is only a first approximation. The results of the assignment run using this first L value must be investigated to determine the direction in which the second L value must differ from the first. This trial and error process must continue until the desired accuracy is achieved.

Multiple L Value Calibration Methods

The use of single L values for each trip population in an assignment run for a small area does not satisfy other criteria even when the total vehicle mileage criterion is met. For this reason, research at CATS has shifted to the calibration of multiple L value assignments. The remainder of the methods discussed in this paper, all multiple L value calibration methods, are labeled the empirical, statistical, and iterative calibration methods for convenience. Actually, all three are empirical, statistical, and iterative in some sense of the meaning of these words.

Empirical Calibration Method—Short Trips

The first multiple L value method developed at CATS for short trip L values is essentially an extension of the use of Eq. 16 to a number of groups of zones having similar average trip lengths and trip end densities. Individual L values then are obtained for each group of zones, termed a density class. The method is summarized briefly here; a complete description is given by Muranyi and Miller (6).

The method was developed before the assignment program had been modified to accept multiple L values automatically. Costly and error-prone manual stopping of the program and inserting of new L values was necessary; therefore, the number of density classes was held to five or six. The method is not limited to a small number of cases, however, and could be applied to each zone individually if desired.

One necessary revision of Eq. 16 was a change of the density variable. The use of an area-wide average would defeat the purpose of treating each density class individually. By using the average density within a three-mile radius of each origin zone, the area in which nearly all of the short trips from each zone would end was included. The densities used in the equation were, therefore, the average of these modified densities for all zones included in the density class.

The empirical nature of the method arises from the use of an initial, approximate assignment to a known situation to obtain the data corresponding to the subscript 1 in Eq. 16. In this respect, the method is similar to the single L value method. These data are used to obtain an empirical constant for each density class defined as follows:

$$C_i = \hat{r}_{1i} \sqrt{L_{1i} \hat{\rho}_i} \tag{17}$$

where

$\hat{r}_{1i}$ = average trip length for density class i, obtained from initial assignment;
L_{1i} = L value for density class i used in initial assignment; and
$\hat{\rho}_i$ = average trip end density for density class i.

Once values of C_i have been determined and values of $\bar{r}_i$ are obtained from given data or curves such as those in Figure 3, short L values for use in assignments to known situations can be found by using the following relationship:

$$L_i = \frac{C_i^2}{\hat{\rho}_i \bar{r}_i^2} \tag{18}$$

Future short L values can be obtained by assuming that the C_i vs $\hat{\rho}_i$ relationship found to hold for present assignments will continue to hold in the future. Eq. 18 then can be applied by using new densities and average trip lengths.

Empirical Calibration Method—Long Trips

It was not possible to define a meaningful average long trip end density corresponding to the density within a three-mile radius area used for short trips, so an adaptation of Eq. 14 could not be used for present assignments of long trips. Instead, Eq. 8 was used to insure that the correct number of trips from each density class would go farther than ten miles. The determination of the L value in each case is straightforward once V and P(V) have been determined from the available empirical data.

Future long L values can be determined by using Eq. 16, the ratio form of the approximate L value, the trip parameter relationship.

Statistical Calibration Method

The statistical calibration method was developed by Emilio Casetti at CATS when it became evident that Eq. 15 had serious deficiencies which limited its applicability to the prediction of zonal L values. The method uses multiple regression statistical techniques to determine relationships for test zones in a given area and with given trip ends, between arbitrarily chosen L values, resulting average trip lengths, and trip end densities determined within various cutoff points. Cutoff points indicate the truncation of the allocation of trips from a given origin at a given percent level. For example, a 60 percent cutoff point corresponds to the point at which 60 percent of the trips available in a given zone of origin have been allocated. The density within a given cutoff point varies depending on the L value used.

The equation to which a least squares fit is obtained is the following:

$$\log L = \log a_0 + a_1 \log \rho_1 + a_2 \log \rho_2 + \ldots + a_n \log \rho_n + a_{n+1} \log \bar{r} \tag{19}$$

where

ρ_i = trip end density within cutoff point i, using an arbitrarily selected L value;
a_i = coefficient obtained using multiple regression techniques; and
$\bar{r}$ = average trip length of all trips assigned with a given L value, using an arbitrarily selected cutoff point.

Because this equation is linear in the logarithmic transformations of the variables, standard multiple regression methods can be used to find the a_i coefficients. A more compact form of the equation is obtained by taking antilogarithms of both sides:

$$L = a_0 \cdot \rho_1^{a_1} \cdot \rho_2^{a_2} \cdots \rho_n^{a_n} \cdot \bar{r}^{-a_{n+1}} \tag{20}$$

The family of curves represented by Eq. 20 was selected for use because Eq. 15 indicates a power relationship may be expected to exist between L, $\bar{r}$, and ρ. Eq. 20 represents the generalization of this power relationship to a large family of curves from which the best curve can be found.

The various ρ_i variables were introduced into the model to provide some measure of the density variations ignored in Eq. 15. Tests indicated that six density variables, corresponding to cutoff points ranging from 0. 60 to 0. 98 using an L value of 80 × 10^{-6}, resulted in a satisfactory predictive equation.

The cutoff point to be used in determining $\bar{r}$ is 0. 80, as this value of $\bar{r}$ is most closely correlated with L.

The step-by-step procedure recommended for collecting the test data needed, and for using the resulting equation, is as follows. Inasmuch as a major feature of the method is its recognition of variable trip end density and the pattern of this variation is unique for each assignment area, it is not recommended that a_i coefficients found for one area be used to predict L values in another. Therefore, the test data must be collected for each new assignment area.

1. Select a workable number of test zones representing ranges of trip end densities and average trip lengths.
2. Select a reasonable number of test L values representing the probable range of actual L values.
3. Distribute trips using each L value from each test zone.
4. Calculate the trip end densities within each cutoff point for each test zone, using an L value of 80 × 10^{-6}.
5. Calculate the average trip length within the 0. 80 cutoff point for each L value and for each test zone.
6. Use a standard multiple regression computer program to transform all data using a logarithmic transformation and to determine the regression coefficients, a_i.
7. For each zone or group of zones for which an L is to be determined, calculate the trip end densities within each cutoff point, using an L value of 80 × 10^{-6}. Also, determine the actual average trip length within the 0. 80 cutoff point from survey data.
8. Use Eq. 19 or 20 to determine log L or L.

Obviously, the method is neither fast nor simple, but it is likely to be more accurate than the empirical method, because variable densities are recognized and included in the method. A computer program has been written to simplify the execution of steps 4 and 5. An indication of the accuracy of the method is that the multiple correlation coefficient obtained in the test case was 0. 815.

Iterative Calibration Method

Both of the previously discussed multiple L value calibration methods were lacking in ease of application. The search was continued, therefore, for an easy and accurate calibration method which would use the fewest possible data not needed as input for an actual assignment run. A number of approximations on the order of Eq. 15 but using some sort of varying trip end densities were found no better than available methods.

It finally was necessary to use the same relationship between trip ends, average trip lengths, and L values as that existing in the assignment program. This relationship is simply the discrete version of Eq. 9:

$$\bar{r}_o = \frac{\sum_{j=1}^{n} d_{oj} \left[e^{-LV_j} - e^{-LV_{j+1}} \right]}{1 - e^{-LV_n}} \tag{21}$$

where

$\bar{r}_o$ = average trip length for zone o,
d_{oj} = distance from zone o to zone j, and
n = total number of zones.

Although zones are ranked by time in the assignment system, it is necessary to assume for the calibration method that a distance ranking suffices. Otherwise, it would be necessary to use assignment output to determine assignment inputs.

The data needed to use Eq. 21 to find L values are $\bar{r}_o$'s, d_{oj}'s, and V_j's. However, the V_j's are just summations of O_i's (see Eq. 5), which are assignment inputs. The $\bar{r}_o$'s must be determined externally, just as with any of the calibration methods. The d_{oj}'s can be calculated if each zone is assigned X and Y grid coordinates. In summary, the data needed to solve Eq. 21 can be obtained if trip origins, X and Y coordinates, and average trip length are specified for each zone.

Equation 21 reveals that the L value cannot be isolated algebraically on one side of the equation. This is where iteration comes in. Iterative methods for solving non-linear equations such as Eq. 21 are presented in texts on numerical analysis, such as Hildebrand (2). A common iterative method is the modification of a nonlinear equation $f(x) = 0$ to the form $x = F(x)$ and then by using the recurrence relation of the form:

$$x_{k+1} = F(x_k) \tag{22}$$

The procedure involves choosing an x_0 as an initial approximation, finding $x_1 = F(x_0)$, and continuing until the difference between x_k and x_{k-1} is sufficiently small. Hildebrand points out that the method is guaranteed to converge only if:

$$\left| \frac{dF(a)}{da} \right| < 1 \tag{23}$$

where a is the true value of x.

In applying this method to Eq. 21, a function F(L) can be found by multiplying both sides of the equation by $L/\bar{r}_o$:

$$L = F(L) = \frac{L \sum_{j=1}^{n} d_{oj} \left[e^{-LV_j} - e^{-LV_{j+1}} \right]}{\bar{r}_o \left[1 - e^{-LV_n} \right]} \tag{24}$$

Since F(L) depends on a large number of parameters, it is difficult to check it for convergence in the general case. However, Eq. 14 is approximately true, so a test of the F(L) obtained by multiplying both sides of this simpler equation by $L/\bar{r}$ should indicate whether or not the F(L) of Eq. 24 will converge. Multiplication of Eq. 14 by $L/\bar{r}$ results in:

$$L = F(L) = \frac{1}{2\bar{r}} \left[\frac{L}{\rho} \right]^{1/2} \tag{25}$$

Differentiating,

$$\frac{dF(L)}{dL} = \frac{1}{4\bar{r}} \left[\frac{1}{\rho L} \right]^{1/2} \tag{26}$$

Once differentiation is complete, $\bar{r}$ can be replaced by its equivalent, as given in Eq. 14. The resulting value of the derivative is 0. 5, indicating that the condition expressed in Eq. 23 is met for the F(L) of Eq. 25 and, therefore, should be met for the F(L) of Eq. 15 with ρ replaced by O_o/A_o, where A_o is the area of zone o. This necessitates the addition of one more item to the list of zonal data needed, namely, the area of the zone.

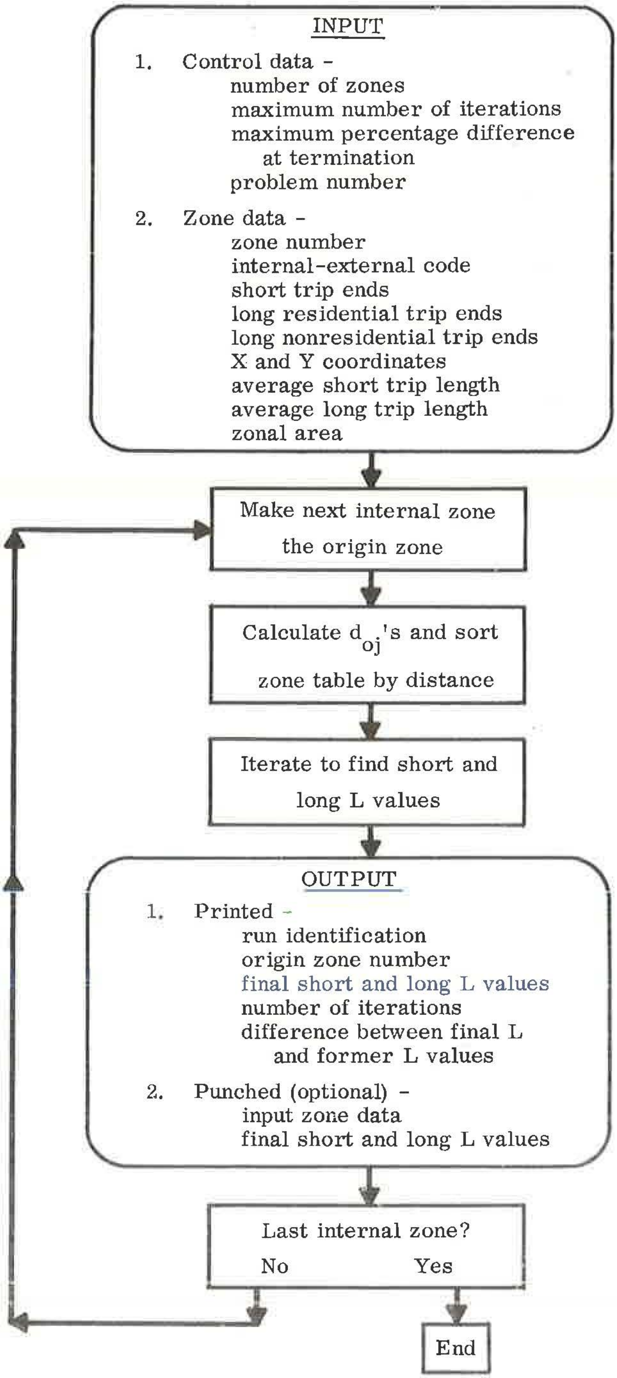


Figure 4. Generalized flow diagram of the L value calibration program.

TABLE 1

COMPARISON OF 1960 FOX RIVER VALLEY SURVEYED AND ASSIGNED TRIPS (Single L Value Model)[a]

Trip Type	Surveyed Trips	Assigned Trips	Ratio of Assigned to Surveyed
I-I[b]	42, 700	44, 200	1. 04
I-B[c]	45, 400	58, 800	1. 30
I-E[d]	49, 700	62, 800	1. 26
Total	137, 800	165, 800	1. 20

[a]Data derived from ref. 6.
[b]Trips with both origin and destination in the internal area.
[c]Trips with one end in the internal area and the other end in the buffer area.
[d]Trips with one end in the internal area and the other end in the external area.

TABLE 2

COMPARISON OF 1960 FOX RIVER VALLEY ACTUAL AND ASSIGNMENT (Multiple L Value Model)[a]

District of Origin	Actual[b] $\bar{r}_s$	Assigned $\bar{r}_s$	Ratio of Assigned to Actual
11-12	2. 18	2. 2	1. 06
21-22	2. 60	2. 7	1. 04
31-32	2. 85	2. 9	1. 02
41-42	2. 20	2. 5	1. 14
51-52	2. 82	2. 9	1. 03

[a]Data derived from ref. 6.
[b]$\bar{r}_s$ = average short trip lengths.

TABLE 3

COMPARISON OF 1960 FOX RIVER VALLEY SURVEYED AND ASSIGNED TRIPS (Multiple L Value Model)[a]

Trip Type[b]	Surveyed Trips	Assigned Trips	Ratio of Assigned to Surveyed
I-I	42, 700	45, 900	1. 07
I-B	45, 400	51, 400	1. 13
I-E	49, 700	54, 300	1. 09
Total	137, 800	151, 600	1. 10

[a]Data derived from ref. 6.
[b]See explanation accompanying Table 1.

Although the foregoing iterative calibration method can be expected to be accurate and uses only easily determined zonal data, it would be far from easy to use if all calculations had to be performed by hand. It was to satisfy this requirement that a computer program was written to accept as input the zonal data and control information needed, to calculate iteratively both short and long L values for any or all zones in the assignment area, and to punch out these L values on cards which can be used directly as part of the assignment input. A generalized flow diagram of the program appears in Figure 4. The fact that L values are found only for zones coded "internal" means that any number of selected zones can be calibrated, or that all zones which will actually be used to send trips in the assignment program can be calibrated.

The program has been written in FORTRAN II and running time, when calibrating all zones, is slightly less than that for an assignment using the same computer. The number of iterations necessary to achieve an accuracy of 0. 1 percent ranges from about 8 to 11.

RESULTS OBTAINED USING CALIBRATION PROCEDURES

Single L Value Method

The single L value calibration method was used in all assignments to the entire CATS area run before August 1965. These assignments have included not only the 1956 existing runs and the 1980 future runs, but also runs for a number of intervening years. In each case, a combination of trial and error methods and the use of the single L value calibration method have resulted, finally, in an acceptable assignment. The number of preliminary assignments has varied greatly, and in some cases has been reduced to one.

In two of the smaller areas within the Chicago metropolitan region, the Fox River Valley area and the Lake County area, the single L value calibration method and single L value assignments were tried a number of times, but never could be made to give acceptable results. An example of the problems involved is indicated in Table 1 which compares the final single L value run in the Fox River Valley with survey data. Although entirely internal trips have been quite accurately duplicated, trips between the internal area and the buffer and external areas are greatly overestimated. Results of this kind lead to the realization that multiple L values are necessary in the small area assignments.

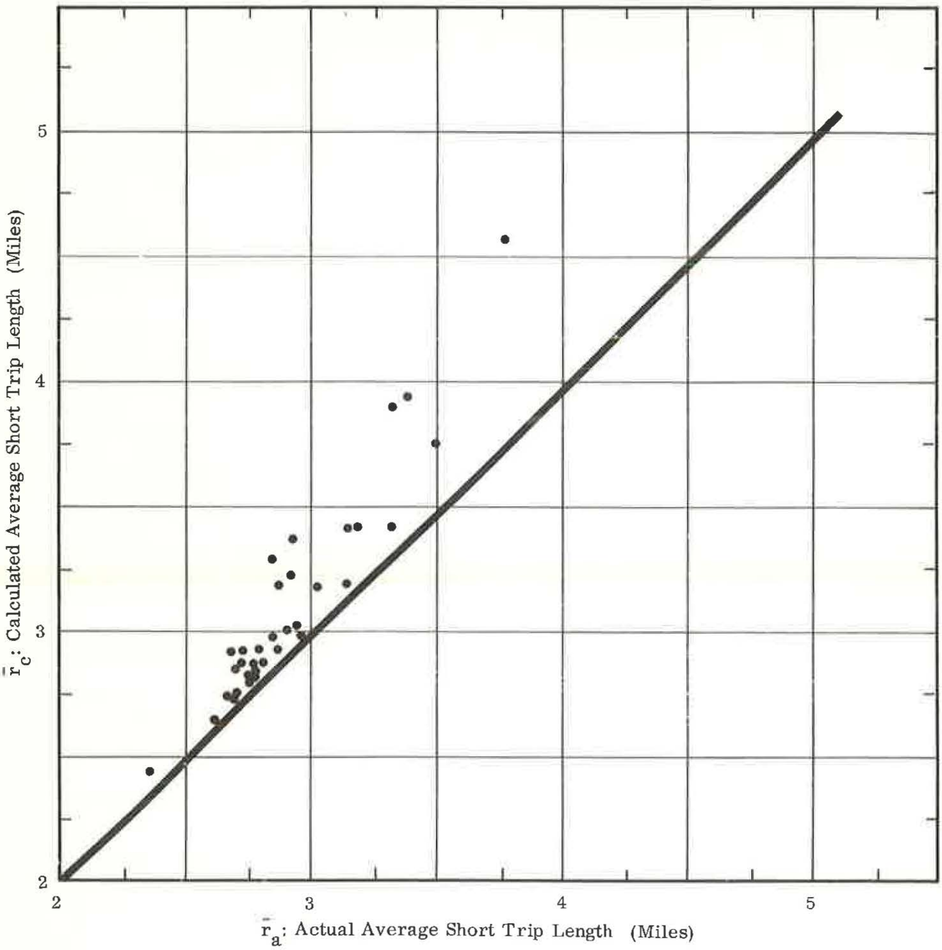


Figure 5. Comparison of calibration method input ($\bar{r}_a$) and select area assignment output ($\bar{r}_c$) average short trip lengths.

Multiple L Value Methods

Empirical Method. —The empirical method was used to calibrate multiple L values for present and future Fox River Valley and Lake County assignments. In both cases, the results indicate not only that the actual average trip lengths of the density classes are closely approximated by the assignments, but also that there is a general improvement in the quality of the assignments, as reflected in comparisons of trip survey and assignment origin and destination data. Table 2 gives actual and assignment average short trip lengths for final Fox River Valley assignments. Agreement is very good. Origin and destination comparisons in Table 3 indicate a general improvement, amounting to ten percent for all internal trips. Although trips with both origin and destination in the internal area are more poorly estimated when the multiple L value model is used, the total error is more uniformly distributed among the three groups of trips than it was when the single L value model was used.

Statistical Method. —No assignments have been calibrated by use of the statistical calibration method. The collection of the necessary test data was very time-consuming. And inasmuch as the results obtained through the empirical method for groups of zones in the Fox River Valley and Lake County assignments were considered sufficiently accurate, the more involved method was not attempted.

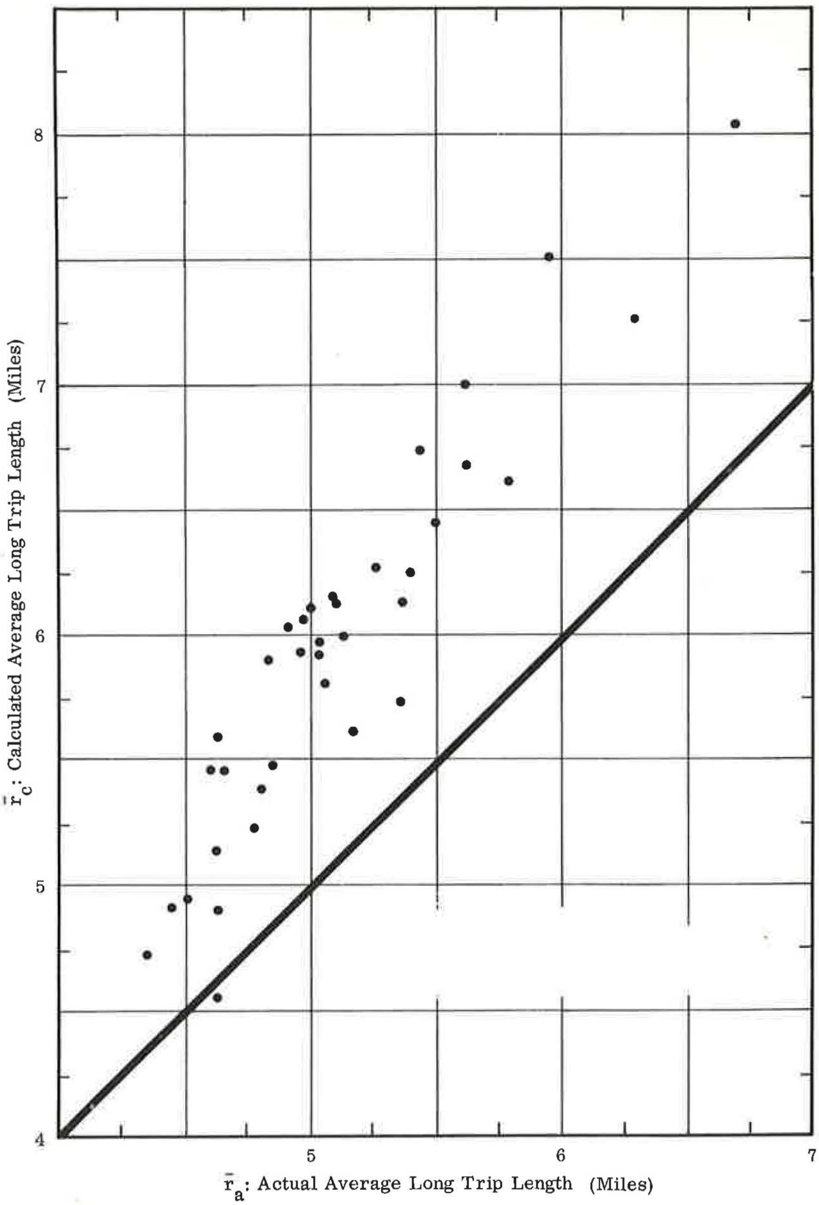


Figure 6. Comparison of iterative calibration method input ($\bar{r}_a$) and select area assignment output ($\bar{r}_c$) average long trip lengths.

Iterative Method. —The iterative method has not been used to calibrate any "production" assignment runs, but has been used for runs testing the ability of the Opportunity Model to distribute trips in a very small area. The area chosen is a 36-square mile section of Chicago lying between 1. 5 and 7. 5 miles north and west from the CBD. This area is identified as the "select area" (Fig. 1). The attempt to run a select area assignment was unique not only because of the small size of the area compared to the size of the metropolitan region, but also because the area cannot be considered to be even partially self-contained, as are the Fox River Valley and Lake County areas.

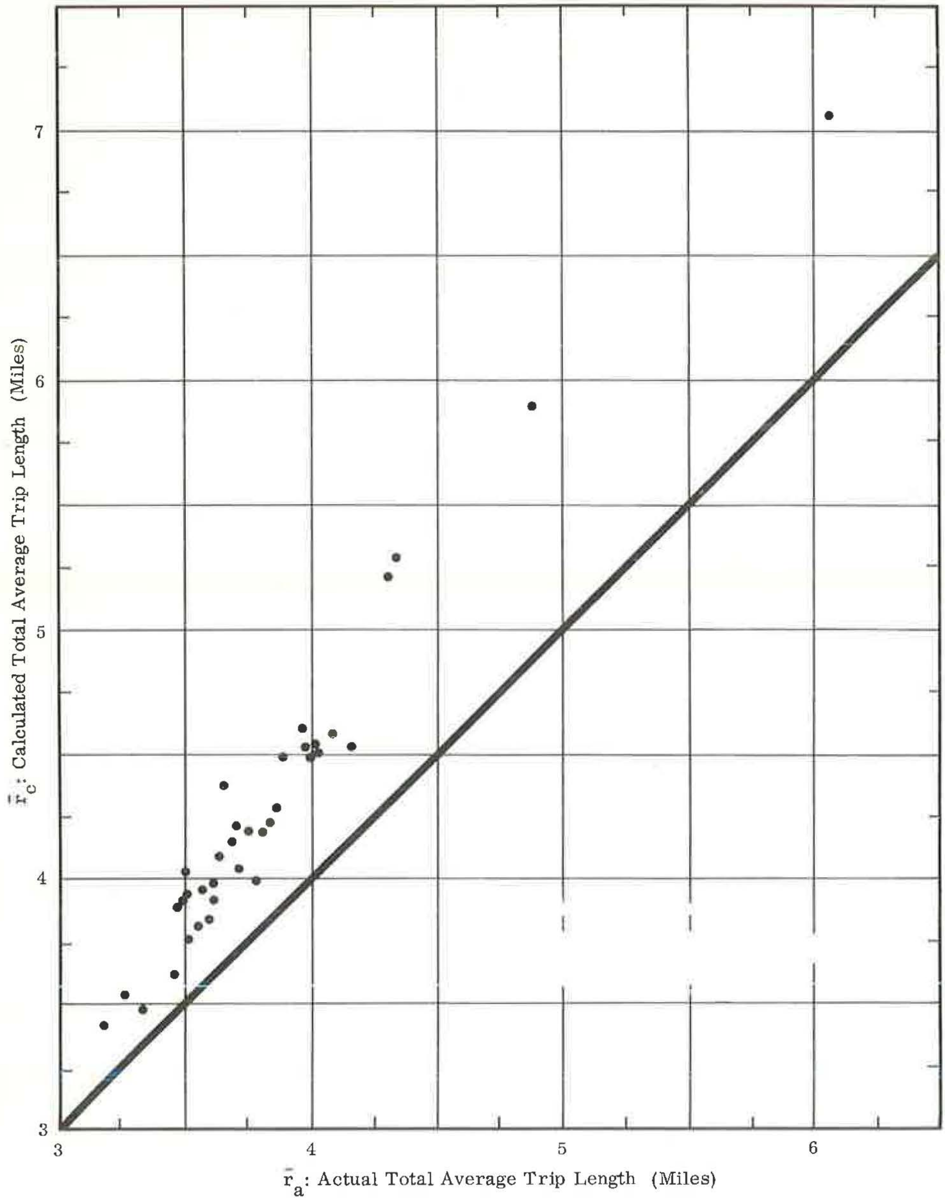


Figure 7. Comparison of iterative calibration method input ($\bar{r}_a$) and select area assignment output ($\bar{r}_c$) total average trip lengths.

The select area assignments involved the 1956 network and trip ends. Surveyed 1956 zonal average trip lengths were determined for both short and long trips. There were 36 one-square-mile zones for which trip end and average trip length information was available. Because it was desired to have as much detail as possible, the zones were divided to obtain 144 quarter-square-mile zones. The allocation of the three subpopulations of trip ends to the four smaller zones within each survey zone was based on the surveyed number of auto driver trip ends per quarter square mile, because these data were available and short and long trip ends were not available by quarter square mile. It was assumed that the average trip lengths for the square-mile zones would hold, also, for each of the four smaller zones.

TABLE 4

STATISTICAL MEASURES OF THE ACCURACY OF THE ITERATIVE CALIBRATION METHOD

Statistic[a]	Trip Population		
	Short	Long	All
$\bar{r}_c$ vs $\bar{r}_a$			
Mean of all $\bar{r}_a$ (mi)	3.00	5.12	3.83
$\bar{e}$ (mi)	0.211	0.828	0.458
MSE (sq mi)	0.1080	0.8063	0.2586
Due to e (sq mi)	0.0445	0.6853	0.2099
σ^2 (sq mi)	0.0635	0.1210	0.0487
RMSE (mi)	0.328	0.897	0.508
σ (mi)	0.252	0.348	0.221
Avg. range of variation of r_c's for quarter square miles within square mile zones (mi)	0.082		0.212

[a] $\bar{r}_c$ = output average trip length determined from select area assignments Nos. 9 and 10;

$\bar{r}_a$ = input average trip length, determined from CATS survey data;

$\bar{e}$ = average error in $\bar{r}_c$, equal to $\frac{\Sigma(\bar{r}_c - \bar{r}_a)}{n}$;

MSE = mean square error, equal to $\frac{\Sigma(\bar{r}_c - \bar{r}_a)^2}{n}$;

RMSE = root mean square error, equal to $\sqrt{MSE}$;

σ^2 = variance of errors, equal to MSE $- \bar{e}^2$; and

σ = standard deviation of errors, equal to $\sqrt{\sigma^2}$

The iterative calibration program was used to determine short and long L values for each of the 144 smaller zones. After running the assignment system using these calibrated L values, short, long, and total average trip lengths resulting from the assignment were calculated for each of the original 36 one-square-mile zones. Figures 5-7 show plots of actual average trip lengths ($\bar{r}_a$) and assignment-calculated average trip lengths ($\bar{r}_c$) are shown for short trips, long trips, and all trip averages.

A systematic error exists in this process of proceeding from $\bar{r}_a$ to L to $\bar{r}_c$. The output $\bar{r}_c$'s all are higher than they should be. It is believed that the major cause of this error is the distance ranking of destination zones used in the calibration program, whereas a time ranking is used in the assignment system. A distance ranking of destination zones minimizes average trip length subject to the L value. The time ranking must, therefore, result in at least as large and probably a larger average trip length than the distance ranking. The amount of divergence depends on the pattern of variations in speed on the various links of the network, as these variations are what cause the time and distance rankings to differ.

Various quantitative measures of the accuracy of the iterative calibration method are given in Table 4. The average errors found when $\bar{r}_c$ is compared with $\bar{r}_a$ for one-square-mile zones range from 0.211 miles to 0.828 miles, between seven and sixteen percent of the mean values. The dispersion of these errors is measured by their root mean square. The dispersions range from eleven to eighteen percent of the mean values. It is known now that systematic errors due to the distance ranking of destination zones exist, and it will be possible, in future applications, to adjust input averages downward by the amount of this systematic error. The standard deviation of the errors gives some indication of the accuracy of the method when this adjustment is used. Inasmuch

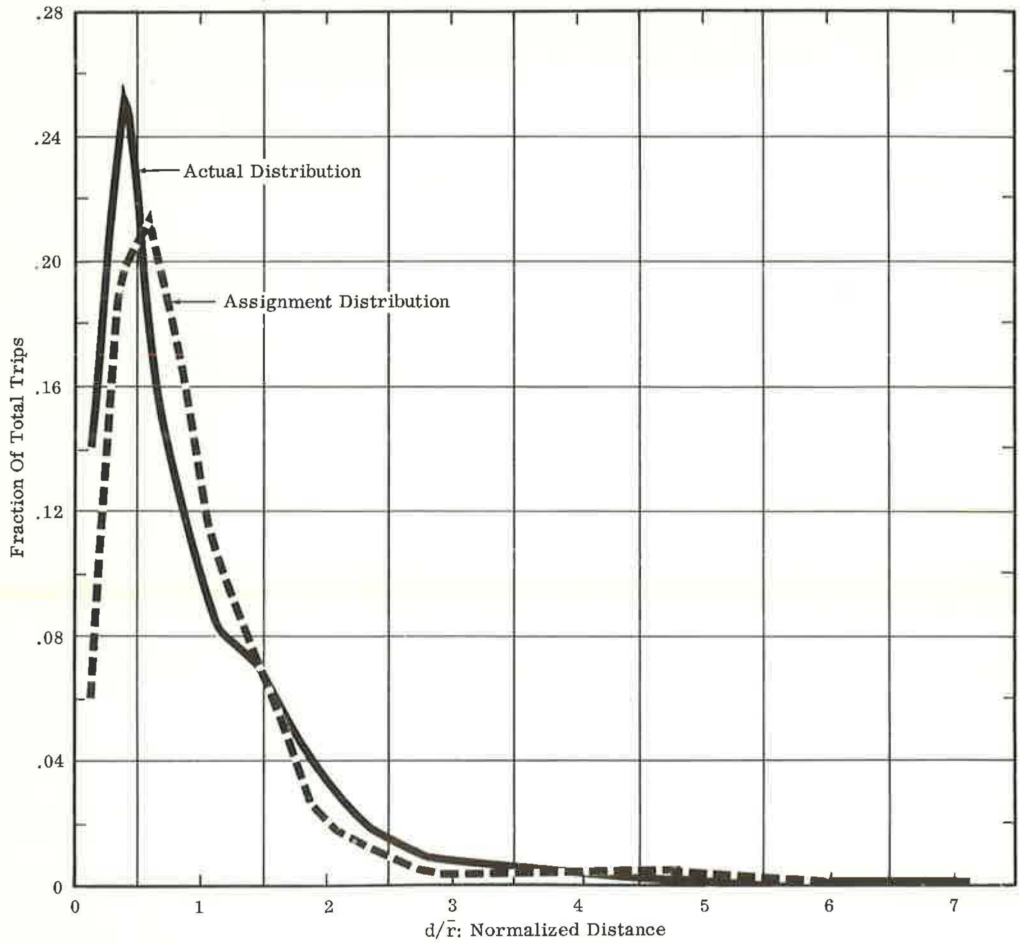


Figure 8. Comparison of actual and assignment output trip length distributions.

as these standard deviations all are less than ten percent, it appears that this calibration method will result in output average trip lengths which are within ten percent of the input averages two-thirds of the time. Furthermore, the average error will be very close to zero, so the total vehicle mileage of an assignment area will be very close to the observed amount.

To measure the ability of the calibration method to obtain the same average trip length in different zones, the range of variation of the output average trip lengths for each of the four zones within each square mile was determined. Table 4 gives the average range of variation for short trips and for all trips. These averages indicate that maximum zone-to-zone variations for equal input averages are only about three percent for short trips and six percent for all trips.

FUTURE CALIBRATION DIRECTIONS

Although the iterative calibration method does a good job of duplicating actual average trip lengths, further checks of the select area assignments indicate that the match of actual trip length distributions by assignment output distributions is poor (Fig. 8). The horizontal axis has been normalized in terms of the average trip length to indicate that the poor match is not the result of differences in the average trip lengths. Trips whose distances are from zero to about 0.5 of the average trip length are underestimated, trips from 0.5 to 1.5 of the average are overestimated, and trips longer than 1.5 times the average are underestimated.

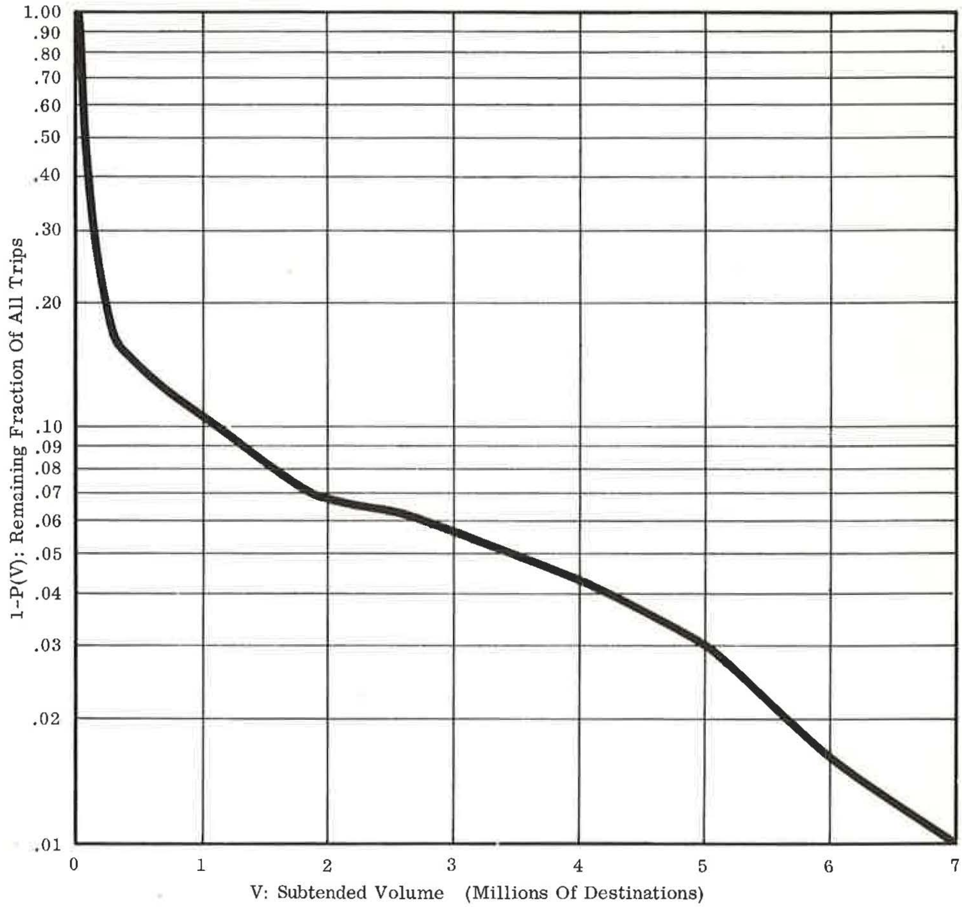


Figure 9. Cumulative distribution of all trips from CATS zone 487 according to number of opportunities.

These discrepancies can be interpreted in at least two ways and corrected in at least three ways. The first interpretation is that the model is all right, but that the trip split into long and short trips is faulty. This interpretation leads to two possible corrections: (a) keep the present trip split, but modify the observed average short and long trip lengths to obtain the correct total average trip length and the correct distribution; and (b) change the present trip split so that, using the short and long trip lengths corresponding to this split, the correct distribution is obtained.

A second possible interpretation of the trip distribution discrepancy is that the Opportunity Model's hypothesis of a constant probability of trip satisfaction (L value) is in error. Perhaps the probability of trip satisfaction is a function of V, the subtended volume. If this function could be found, trip splits would not be necessary, and the correct trip distribution would result when applying the modified model.

The three possible corrections mentioned are discussed next under the following headings: average trip length changes, trip split changes, and model changes.

Average Trip Length Changes

Investigation of the select area assignments indicates that if the average short trip length were set at about 0. 6 of its actual value and if the average long trip length were modified upward so that the average total trip length remained the same, the two curves shown in Figure 8 would nearly coincide. Therefore, it is possible to change average trip lengths arbitrarily so that trip distributions will be matched. It is felt, however, that this type of correction is too arbitrary to be valid and should not be used unless acceptable methods of correction are unavailable.

Trip Split Changes

It is felt that a more acceptable correction of the Opportunity Model would be to change the definitions used in splitting trips into long and short subpopulations. Investigation has shown that some groups of trips presently classified "long" have shorter averages than groups classified "short." Also, experience with many assignments has indicated to CATS researchers that more "short" trips are needed so that trip distributions will be matched. Investigation presently is continuing to determine which long trips should be added to the short trip population.

Model Changes

The Opportunity Model implies a linear semilog relationship between 1-P(V) and V. However, this relationship can be demonstrated only for one of the trip subpopulations at a time. Figure 1, for example, shows the relationship for long nonresidential trips only. When the relationship is graphed for total trips, a curve like that in Figure 9 is obtained. A straight line would be a poor fit to this curve, but perhaps a relationship of the form $L = aV^b$ would provide a good fit. If so, trip splits would be unnecessary. All trips originating from a zone could be distributed by use of the following equation:

$$T_{ij} = O_i \left[e^{-aV_j^{b+1}} - e^{-aV_{j+1}^{b+1}} \right] \qquad (27)$$

It would be necessary to change the second hypothesis of the Opportunity Model to allow for a variable L value instead of a constant. The second hypothesis could be changed to read:

> The probability of a destination being accepted, if it is considered, is a function of the number of destinations which already have been considered.

It is planned to investigate this approach to improving the Opportunity Model. The investigation will largely consist of curve-fitting, using data similar to those in Figure 9 and of determining methods of predicting the parameters needed to relate the variable L value to subtended volume.

SUMMARY AND CONCLUSIONS

The Opportunity Model has been analyzed by interpreting its hypotheses, its mathematical formulation, and its parameter, the L value. The L value also has been related to trip parameters. These analyses have served as the basis of a number of calibration methods which have been presented. The results obtained when these methods were applied to assignments at CATS have been given. Planned methods of improving both the calibration techniques and the Opportunity Model have been discussed.

The most promising calibration method developed so far is the iterative method, which provides a means of duplicating observed or predicted average trip lengths with a standard error of less than ten percent with one pass through a calibration program and one pass through the assignment system. Problems in matching observed trip length distributions indicate either that calibration methods must be concerned with more than matching averages, or that the Opportunity Model itself must be improved. CATS' future trip distribution research is expected to investigate both of these possibilities.

REFERENCES

1. Chicago Area Transportation Study. Final Rept., Vol. II, Chicago, 1960.
2. Hildebrand, F. B. Introduction to Numerical Analysis. McGraw-Hill, pp. 443, 1956.
3. Interim Report. Factors and Trends in Trip Lengths. Proj. 7-4, National Cooperative Highway Research Project, Feb. 1965.
4. Leonard, H. J., and Delano. A New Model for Distributing Short and Long Distance Travel. Presented at ASCE Transportation Engineering Conference, Minneapolis, May 1965.
5. Muranyi, T. C. Trip Distribution and Traffic Assignment. In Traffic Assignment Conference, pp. 21-22, Dec. 11, 1963.
6. Muranyi, T. C., and Miller, J. D. New Applications of the Opportunity Model to Traffic Planning and Design. Presented to Origin and Destination Committee, HRB Annual Meet., Jan. 1964.
7. Project Statement. Factors and Trends in Trip Lengths. Proj. 7-4, National Cooperative Highway Research Project, Feb. 1964.
8. Rathnau, A. S. Operations Manual—Traffic Assignment System. No. 34,320, CATS, Chicago, 1964.
9. Rathnau, A. S. Reference Manual—Traffic Assignment System. No. 34,330, CATS, Chicago, 1965.

Effect of Zone Size on Zonal Interchange Calculations Based on the Opportunity Model in a Homogeneous Region

W. STEARNS CASWELL, Subdivision of Transportation Planning and Programming, New York State Department of Public Works

•ONE of the prerequisites for a successful transportation study is a conceptually reasonable and operationally realistic model for the distribution of trips between zones. One such model is the opportunity model, developed by the Chicago Area Transportation Study and now being used by the Subdivision of Transportation Planning and Programming of the New York State Department of Public Works.

The opportunity model is theoretically based on the behavior of individual trips. In practice, however, it is necessary to deal with trips grouped together into zones, with all trips in a zone assumed to originate at the zone centroid. This procedure gives rise to errors in the estimation of trip interchange between zones. The magnitude of these errors is dependent upon zone size as well as trip density and propensity of people and activities to interact with each other.

It is the purpose of this paper (a) to develop formulas which exhibit the dependence of zonal interchange calculations upon the variables mentioned, with emphasis on the variable of zone size, and (b) to present tables showing the relative error as a function of zone size, trip density, and the trip interaction constant used in the opportunity model. These tables will suggest appropriate zone sizes to use in order to hold a constant level of bias for the entire study region.

The current computer assignment program utilizing the opportunity model treats successively zones with equal time path values from the origin zone. This succession is established in an essentially random fashion. This results in a bias, depending on the number of zones (volume of trips) having the same time path value, since the opportunity model would imply that these zones be treated simultaneously. In this paper the successive treatment of these zones is referred to as the "random selection" method, and the simultaneous treatment is called the method of "proportional split." The next section deals with the bias arising from the use of the method of random selection.

The dependence of the calculation of zonal trip interchange upon zone size is then discussed. The central assumption here is that the smaller the zones the more nearly correct is the calculation of zonal trip interchange. This assumption is warranted since the formulas involved in this calculation are based on the assumption of a continuous trip opportunity surface. The ideal procedure, of course, is to consider each trip origin or destination as a zone unto itself, but it is obviously impossible to do this in practice. In fact, since trips must take place along essentially rectangular route patterns, the minimum zone size is dependent on street spacing. It was decided for this study to use zones having quarter-mile sides as the smallest practical case. However, the discussion is entirely general so that an arbitrarily small zone size may be used if desired.

Tables have been prepared which exhibit the relative errors introduced by "random selection," by the use of various zone sizes, and by the combination of these two dis-

Paper sponsored by Committee on Origin and Destination.

torting influences on zonal interchange calculations. These tables are included in the final section.

COMPARISON OF TWO METHODS OF CALCULATING TRIP INTERCHANGE WHEN ZONES ARE GROUPED IN RINGS

The following discussion is concerned with the calculation of the interchange of trips between an origin zone and the individual zones of a ring of zones, all of which are at the same travel time from the origin zone. Two methods are contrasted here, those of proportional split and random selection. In both methods the formula used for the calculation of the interchange v_{ij} between zones i and j is

$$V_{ij} = V_i \left[e^{-tV} - e^{-t(V+V_j)} \right]$$

where

V_i = the volume of zone i,
V_j = the volume of zone j,
V = the volume of destinations available before reaching zone j (the subtended volume), and
t = the probability that a possible destination is acceptable after having been reached.

Throughout this discussion, it is assumed that all zones are square and have the same volume ds^2, where d is a given (constant) trip density and s is the length of each zone side.

Since we deal with square zones, a method of grouping zones in time rings must be determined. To do this some simplifications are in order. In particular, travel time is identified with distance between zones. There are at least two ways to measure distance on a rectangular grid—airline and right-angle. Since the latter method is much more amenable mathematically as well as more realistic for small zones, it has been chosen for use here.

				4				
			4	3	4			
		4	3	2	3	4		
	4	3	2	1	2	3	4	
4	3	2	1	0	1	2	3	4
	4	3	2	1	2	3	4	
		4	3	2	3	4		
			4	3	4			
				4				

Figure 1. Procedure for determining rings: zones are numbered in order of increasing right-angle distance from zone 0; set of zones with same number constitutes a ring.

In Figure 1 a study area has been partitioned into square zones, and an origin zone (numbered 0) has been chosen. The other zones in the area are numbered in order of increasing right angle distance from zone 0, using the right-angle distance between centers of adjacent zones as a unit. A set of zones with the same number constitutes a ring. The ring which consists of the zones numbered q will be called ring R_q, where q is called the ring number.

In the random selection method the p zones in ring R_q have been ordered in some fashion, say $q_1, q_2, q_3, \ldots, q_p$. These zones are then considered in the given order for the purpose of selection. In this case

$$V_{oq_1} = V_o \left[e^{-t[V+(i-1)ds^2]} - e^{-t(V+ids^2)} \right]$$

$$= V_o e^{-tV} \left[e^{-(i-1)tds^2} - e^{-itds^2} \right]$$

where V is the volume of destinations available before reaching ring R_q.

Let $x = e^{-tds^2}$. Then

$$V_{oq_i} = V_o e^{-tV} (x^{i-1} - x^i)$$

The total interchange between the origin zone and ring R_q is thus

$$\begin{aligned} V_{oR_q} &= \sum_{i=1}^{p} V_o e^{-tV}(x^{i-1} - x^i) \\ &= V_o e^{-tV} \sum_{i=1}^{p} x^{i-1} (1 - x) \\ &= V_o e^{-tV}(1 - x) \sum_{i=1}^{p} x^{i-1} \\ &= V_o e^{-tV}(1 - x) \left(\frac{1 - x^p}{1 - x}\right) \\ &= V_o e^{-tV}(1 - x^p) \\ &= V_o \left[e^{-tV} - e^{-t(V+pds^2)}\right] \end{aligned} \qquad (1)$$

On the other hand, in the proportional split method, the interchange $\bar{V}_{oq}$ between the origin zone and zone q_i is determined simply by taking 1/p of the total interchange V_{oR_q}; i.e.,

$$\bar{V}_{oq} = (1/p)V_o \left[e^{-tV} - e^{-t(V+pds^2)}\right] \qquad (2)$$

which is, of course, the mean of V_{oq_i} for $i = 1, 2, \ldots, p$. This method of determining the interchange can be defended conceptually. The probability that a trip originating in the origin zone will end in ring R_q is

$$P = e^{-tV} - e^{-t(V+pds^2)}$$

where $V_j = pds^2$ indicates that the opportunities in all of ring R_q are in competition with each other. The probability that a trip will end in zone q_i is 1/p, since all zones in R_q are regarded as having the same friction value with respect to the origin zone. Thus, the probability that a trip will end in ring R_q and zone q_i is (1/p)P. When this probability is multiplied by the volume of trips originating in the origin zone, we have $\bar{V}_{oq}$. As can be seen from Eqs. 1 and 2,

$$V_{oR_q} = \sum_{i=1}^{p} V_{oq_i} = \sum_{i=1}^{p} \bar{V}_{oq} = V_o \left[e^{-tV} - e^{-t(V+pds^2)}\right]$$

For the individual zones, however, $V_{oq_i} \neq \bar{V}_{oq}$. The deviation of V_{oq_i} from $\bar{V}_{oq}$ is given in terms of relative errors in Tables 1-8.

It will be important later to express V_{oR_q} as a function of the ring number q. This can be done as follows: Note that the number of zones in rings R_1, R_2, . . ., R_q increase in arithmetic progression; that is, there are four zones in ring R_1, eight in ring R_2, twelve in R_3, . . ., and 4_q in ring R_q (Fig. 1). Thus the subtended volume V for ring R_q is

$$\begin{aligned} V &= 4ds^2 + 8ds^2 + 12ds^2 + \ldots + 4(q-1)ds^2 \\ &= 4ds^2(1 + 2 + 3 + \ldots + q - 1) \\ &= 4ds^2 [(1/2)(q-1)q] \\ &= 2(q-1)q\, ds^2 \end{aligned}$$

Combining these results, we see that

$$V_{oR_q} = (ds^2)\, e^{-tds^2} \left[e^{-2q(q-1)tds^2} - e^{-2q(q+1)tds^2} \right] \tag{3}$$

DEPENDENCE OF ZONE TRIP INTERCHANGE ON ZONE SIZE

When applying the opportunity model (or any other "continuous" model), the calculation of zonal trip interchange by any program which considers all trip origins or destinations as concentrated at a zone centroid is dependent upon zone size. The purpose of the following discussion is to support this statement by exhibiting the dependence analytically.

Suppose that zones of a certain size, say one mile by one mile, are used in calculating zonal interchanges. In these calculations the formula dictated by the opportunity model is used:

$$V_{ij} = V_i \left[e^{-tV} - e^{-t(V+V_j)} \right]$$

Suppose further that the one-mile zones are subdivided into half-mile zones, that is, zones with one-quarter square mile for area. These smaller zones might be termed "subzones" of the larger one-mile zones. An assumption made here is that the smaller the zones the more nearly correct is the calculation of zonal interchange. If the interchange between the one-mile zones i and j is calculated first by using the usual formula directly and than by aggregating all the interchanges between the subzones which make up zones i and j respectively, the difference is the error which we propose to measure. Formulas will be derived and generalized so that this error may be computed for an arbitrary subdivision of zones. All assumptions concerning trip distribution, etc., are the same as given previously. All zones are assumed to be grouped in rings with respect to an origin zone as described earlier. For the immediate purpose of this section the method of proportional split discussed before will be used.

One way of setting up the problem is to start with a certain zone grid size and then to refine this grid by subdivision to an arbitrarily small zone size. It turns out that the problem is more tractable mathematically if the reverse procedure is followed; i.e., start with an arbitrarily small zone grid size and then increase the zone size by grouping zones into larger ones. Figure 2 shows an example where the small zones have been grouped into zones which are four times as large on a side. In this process of overlaying grids on smaller ones, the smallest or "starting" grid will be designated grid 1, the grid composed of zones which are twice as large on a side as those of grid 1 will be called grid 2, and so forth. Thus the grid k is made up of zones which are k times as large on a side as those of grid 1. Hereafter the zones of grid 1 will be referred to as basic zones.

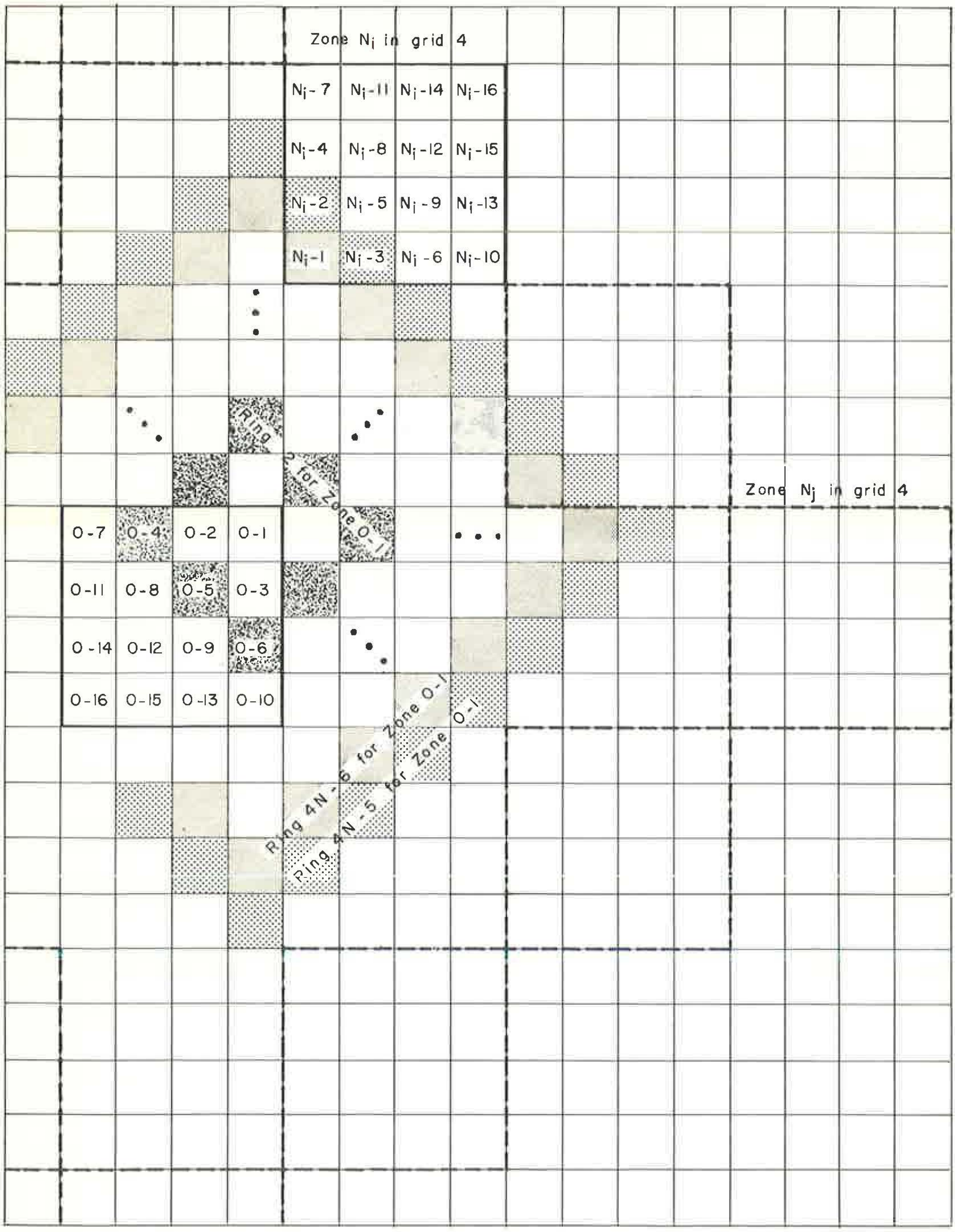


Figure 2. The aggregation of basic zones for k = 4.

Before we discuss the general case it might be well to consider an example. In Figure 2 we have pictured the case when k = 4. In order to refer easily to the basic zones in this case a two-character code is used to identify them. The grid 4 zones have been numbered consecutively within grid 4 rings, the grid 4 origin zone having the number 0, zones in ring 1 having the numbers 1, 2, 3, 4, zones in ring 2 having the numbers 5 through 12, etc. Following this scheme, the first character in the code identifies the grid 4 zone in which lies the basic zone being coded.

The second character in the code indexes the basic zone within its "parent" grid 4 zone. Since each grid 4 zone contains 16 basic zones, the integers 1, 2, . . ., 16 suffice for this second character.

To see why the particular order for these second characters was used, it is helpful to imagine the grid 1 rings radiating outward from the basic zones in the grid 4 origin zone. Each basic zone generates its own structure of concentric grid 1 rings. Picture these rings as they sweep across the grid 4 zones. The only criterion for the choice of order for the second character in a basic zone label is that the basic zones be numbered consecutively within a grid 1 ring in which they fall. It should be clear that in applying this criterion it is immaterial which basic origin zone has generated the grid 1 ring in question.

As an example of the coding procedure, basic zones n_i - 4, n_i - 5, n_i - 6 all lie in the ith grid 4 zone of grid 4 ring n (with respect to the origin zone in grid 4), and these basic zones have been given the second characters 4, 5, 6 because they all fall in the same grid 1 ring (with respect to the basic zones in the grid 4 origin zone).

For the purpose of generalization, suppose that the zones n_i and n_j are in the nth ring in grid 4. (This may take a stretch of the imagination since these zones are actually shown in ring 3. The reasons for doing things this way will become clear when the derivations are followed through.) It turns out that zones n_i and n_j, while they are in the same ring with respect to the origin zone, must be treated differently in calculating the interchange. The criterion here is that n_i is diagonally placed in the grid pattern from the origin zone while n_j is parallel.

Zones like n_i will be called side zones and zones like n_j corner zones. Symmetry assures that all side zones in a given ring may be treated alike as may the four corner zones. For the moment we will concentrate on side zones.

In grid 4 the interchange between the origin zone and zone n_i will be calculated by summing the interchanges between the sixteen basic zones which make up the origin zone and the sixteen subzones of zone n_i, a total of 256 grid 1 interchanges. For each of these grid 1 interchanges the formula (from Eqs. 2 and 3)

$$V_{ij} = e^{-tds^2} (ds^2/4m) \left[e^{-2m(m-1)tds^2} - e^{-2m(m+1)tds^2} \right] = f(m)$$

will be used, where m is the number of the grid 1 ring in which zone j lies with respect to zone i, and s is the length of the sides of the zones in grid 1. For a fixed s, t, and d, V_{ij} is a function of m, hence its designation by f(m). Note that when m = 0, the expression given for V_{ij} is undefined. The interchange f(0) is to be interpreted as the interchange of a grid 1 zone with itself, and is

$$f(0) = ds^2 \left(1 - e^{-tds^2}\right)$$

One of our purposes is to express m in terms of n, the ring number in grid 4. An examination of Figure 2 will indicate the method. For example, zone n_i - 1 is in the grid 1 ring number 4n - 6 with respect to zone 0 - 1, zones n_i - 2 and n_i - 3 are in the grid 1 ring number 4n - 5 with respect to zone 0 - 1, zones n_i - 4, n_i - 5, and n_i - 6 are in the grid 1 ring numbered 4n - 4 for zone 0 - 1, and so forth. As another example, zones n_i - 2 and n_i - 3 are in the grid 1 ring numbered 4n - 4 for zones 0 - 2 and also for zone 0 - 3.

To find the total interchange between the origin zone and zone n_i in terms of the smaller interchanges the following matrix will be helpful:

$$
\begin{array}{c}1\\2\\3\\4\\3\\2\\1\end{array}
\begin{bmatrix}
f(4n-6) & 2f(4n-5) & 3f(4n-4) & 4f(4n-3) & 3f(4n-2) & 2f(4n-1) & f(4n) \\
f(4n-5) & 2f(4n-4) & 3f(4n-3) & 4f(4n-2) & 3f(4n-1) & 2f(4n) & f(4n+1) \\
f(4n-4) & 2f(4n-3) & 3f(4n-2) & 4f(4n-1) & 3f(4n) & 2f(4n+1) & f(4n+2) \\
f(4n-3) & 2f(4n-2) & 3f(4n-1) & 4f(4n) & 3f(4n+1) & 2f(4n+2) & f(4n+3) \\
f(4n-2) & 2f(4n-1) & 3f(4n) & 4f(4n+1) & 3f(4n+2) & 2f(4n+3) & f(4n+4) \\
f(4n-1) & 2f(4n) & 3f(4n+1) & 4f(4n+2) & 3f(4n+3) & 2f(4n+4) & f(4n+5) \\
f(4n) & 2f(4n+1) & 3f(4n+2) & 4f(4n+3) & 3f(4n+4) & 2f(4n+5) & f(4n+6)
\end{bmatrix}
$$

To read this matrix, start with zone 0 - 1 in Figure 2 and the first row of the matrix, reading from left to right. (The numbers down the left-hand side of the matrix may be ignored for the moment.) This first row has for entries the values of the various interchanges which zone 0 - 1 contributes to the total interchange V_{on_i}. Now f(4n - 6) is the interchange between zones 0 - 1 and n_i - 1; f(4n - 5) is the interchange between zone 0 - 1 and either zone n_i - 2 or zone n_i - 3, hence the coefficient 2; f(4n - 4) is the interchange between zone 0 - 1 and each of zones n_i - 4, n_i - 5, and n_i -6, hence the coefficient 3. The remaining entries in the first row are determined similarly, moving up from lower left to upper right through zone n_i.

Next, the second row of the matrix has for entries the values of the several interchanges which either zone 0 - 2 or zone 0 - 3 contributes to V_{on_i}. The reason for the column of coefficients at the left of the matrix may be given now. The 2 placed before the second row is a factor by which each entry in the second row is to be multiplied, since the interchanges which either 0 - 2 or 0 - 3 contributes are the same. The other coefficients in the left-hand column are determined similarly. As a final example, the interchange between zone 0 - 4 and n_i - 11 is f(4n), which is also the interchange between zone 0 - 4 and either of zones n_i - 12 or n_i - 13, so that 3f(4n) is entered in the third row, fifth column of the matrix. The factor 3 at the left of the third row indicates that zones 0 - 5 and 0 - 6 contribute exactly the same interchange as does zone 0 - 4.

Finally, to obtain the total interchange V_{on_i} between the origin zone and zone n_i, the elements of each row are added together. These sums are multiplied by the corresponding factors appearing at the left of the rows, and then all of these weighted sums are added together. Collecting terms,

$$
\begin{aligned}
V_{on_i} = {} & (1\cdot 1)\ f(4n-6) \\
& + (1\cdot 2 + 2\cdot 1)\ f(4n-5) \\
& + (1\cdot 3 + 2\cdot 2 + 3\cdot 1)\ f(4n-4) \\
& + (1\cdot 4 + 2\cdot 3 + 3\cdot 2 + 4\cdot 1)\ f(4n-3) \\
& + (1\cdot 3 + 2\cdot 4 + 3\cdot 3 + 4\cdot 2 + 3\cdot 1)\ f(4n-2) \\
& + (1\cdot 2 + 2\cdot 3 + 3\cdot 4 + 4\cdot 3 + 3\cdot 2 + 2\cdot 1)\ f(4n-1) \\
& + (1\cdot 1 + 2\cdot 2 + 3\cdot 3 + 4\cdot 4 + 3\cdot 3 + 2\cdot 2 + 1\cdot 1)\ f(4n) \\
& + (2\cdot 1 + 3\cdot 2 + 4\cdot 3 + 3\cdot 4 + 2\cdot 3 + 1\cdot 2)\ f(4n+1) \\
& + (3\cdot 1 + 4\cdot 2 + 3\cdot 3 + 2\cdot 4 + 1\cdot 3)\ f(4n+2) \\
& + (4\cdot 1 + 3\cdot 2 + 2\cdot 3 + 1\cdot 4)\ f(4n+3) \\
& + (3\cdot 1 + 2\cdot 2 + 1\cdot 3)\ f(4n+4) \\
& + (2\cdot 1 + 1\cdot 2)\ f(4n+5) \\
& + (1\cdot 1)\ f(4n+6) \qquad (4)
\end{aligned}
$$

$$\begin{aligned} V_{on_i} = {} & f(4n-6) + 4f(4n-5) + 10f(4n-4) + 20f(4n-3) \\ & + 31f(4n-2) + 40f(4n-1) + 44f(4n) + 40f(4n-1) \\ & + 31f(4n+2) + 20f(4n+3) + 10f(4n+4) + 4f(4n+5) \\ & + f(4n+6) \end{aligned} \tag{5}$$

We wish to generalize these results for the kth grid. The reader will observe that for the kth grid the variable m runs from kn - (2k - 2) to kn + (2k - 2). The row and column coefficients in the matrix run from 1 to a maximum of k and back to 1. The required matrix is

$$\begin{array}{c} 1 \\ 2 \\ 3 \\ \cdot \\ \cdot \\ \cdot \\ k \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{array} \left[\begin{array}{lllll} f[kn-(2k-2)] & 2f[kn-(2k-3)] & 3f[kn-(2k-4)] & \ldots kf[kn-(k-1)] \ldots & f(kn) \\ f[kn-(2k-3)] & 2f[kn-(2k-4)] & 3f[kn-(2k-5)] & \ldots kf[kn-(k-2)] \ldots & f(kn+1) \\ f[kn-(2k-4)] & 2f[kn-(2k-5)] & 3f[kn-(2k-6)] & \ldots kf[kn-(k-3)] \ldots & f(kn+2) \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ f[kn-(k-1)] & 2f[kn-(k-2)] & 3f[kn-(k-3)] & \ldots kf(kn) & \ldots f[kn+(k-1)] \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ f(kn) & 2f(kn+1) & 3f(kn+2) & \ldots kf[kn+(k-1)] \ldots & f[kn+(2k-2)] \end{array}\right]$$

To obtain the coefficient in the expression for the total interchange V_{on_i} in the kth grid, we follow the example given for grid 4. Note that in Eqs. 4 and 5 the coefficient of f(4n - h) is the same as that of f(4n + h), $h > 0$; this is still true in grid k because of symmetry. Thus we need only to study the form of those coefficients up to and including the term involving f(kn). Now for grid 4, the first four coefficients are $(1\cdot 1)$, $(1\cdot 2 + 2\cdot 1)$, $(1\cdot 3 + 2\cdot 2 + 3\cdot 1)$, and $(1\cdot 4 + 2\cdot 3 + 3\cdot 2 + 4\cdot 1)$. For grid k, the first k coefficients are $(1\cdot 1)$, $(1\cdot 2 + 2\cdot 1)$, $(1\cdot 3 + 2\cdot 2 + 3\cdot 1)$, $(1\cdot 4 + 2\cdot 3 + 3\cdot 2 + 4\cdot 1)$, ..., $[1\cdot k + 2(k-1) + 3(k-2) + \ldots + k\cdot 1]$. The form of these coefficients is clearly

$$\sum_{i=1}^{p} i(p+1-i), \text{ where } p = 1, 2, \ldots, k$$

The remaining coefficients in the grid 4 example do not follow this pattern, but note that

$$\begin{aligned} & 1\cdot 3 + 2\cdot 4 + 3\cdot 3 + 4\cdot 2 + 3\cdot 1 \\ & \quad = 1\cdot 5 + 2\cdot 4 + 3\cdot 3 + 4\cdot 2 + 5\cdot 1 \\ & \quad - (1\cdot 2 + 2\cdot 1), \\ & 1\cdot 2 + 2\cdot 3 + 3\cdot 4 + 4\cdot 3 + 3\cdot 2 + 2\cdot 1 \\ & \quad = 1\cdot 6 + 2\cdot 5 + 3\cdot 4 + 4\cdot 3 + 5\cdot 2 + 6\cdot 1 \\ & \quad - (1\cdot 4 + 2\cdot 2 + 2\cdot 2 + 4\cdot 1), \end{aligned}$$

$$1\cdot 1 + 2\cdot 2 + 3\cdot 3 + 4\cdot 4 + 3\cdot 3 + 2\cdot 2 + 1\cdot 1$$
$$= 1\cdot 7 + 2\cdot 6 + 3\cdot 5 + 4\cdot 4 + 5\cdot 3 + 6\cdot 2 + 7\cdot 1$$
$$- (1\cdot 6 + 2\cdot 4 + 3\cdot 2 + 2\cdot 3 + 4\cdot 2 + 6\cdot 1)$$

The terms which are being subtracted on the right-hand side can be rewritten as

$$(1\cdot 2 + 2\cdot 1) = 2(1\cdot 2) = 4(1\cdot 1)$$
$$(1\cdot 4 + 2\cdot 2 + 2\cdot 2 + 4\cdot 1) = 2(1\cdot 4 + 2\cdot 2) = 4(1\cdot 2 + 2\cdot 1)$$
$$(1\cdot 6 + 2\cdot 4 + 3\cdot 2 + 2\cdot 3 + 4\cdot 2 + 6\cdot 1) = 2(1\cdot 6 + 2\cdot 4 + 3\cdot 2)$$
$$= 4(1\cdot 3 + 2\cdot 2 + 3\cdot 1)$$

Generalizing this subtrahend for grid k,

$$1\cdot(2k - 2) + 2(2k - 4) + \ldots + (k - 1)2 + 2(k - 1) + \ldots + (2k - 2)\cdot 1$$
$$= 2[1\cdot(2k - 2) + 2(2k - 4) + \ldots + 2(k - 1)]$$
$$= 4[1\cdot(k - 1) + 2(k - 2) + \ldots + (k - 1)\cdot 1]$$
$$= 4 \sum_{i=1}^{p-k} i[p + 1 - (k + i)]$$

when $p = k + 1, k + 2, \ldots, 2k - 1.$

The minuend, of course, remains in the form

$$\sum_{i=1}^{p} i(p + 1 - i), \quad p = k + 1, k + 2, \ldots, 2k - 1$$

Thus, the coefficients after k terms have the form

$$\sum_{i=1}^{p} i(p + 1 - i) - 4 \sum_{i=1}^{p-k} i[p + 1 - (k + i)]$$

Now,

$$\sum_{i=1}^{p} i(p + 1 - i) = \sum_{i=1}^{p} [(p + 1) i - i^2]$$
$$= (p + 1) \sum_{i=1}^{p} i - \sum_{i=1}^{p} i^2$$
$$= (p + 1) (1/2) p (p + 1) - (1/6) p (p + 1) (2p + 1)$$
$$= (1/6) p (p + 1) [3(p + 1) - (2p + 1)]$$
$$= (1/6) p (p + 1) (p + 2) = C(p + 2, 3)$$

where $C(x, y) = \frac{x!}{y!(x - y)!}$

Similarly,

$$4 \sum_{i=1}^{p-k} i[p + 1 - (k + i)] = 4C(p + 2 - k, 3)$$

If we follow the usual convention that C(x, y) = 0, when $x < y$, then all the coefficients may be written as

$$C(p + 2, 3) - 4C(p + 2 - k, 3), \text{ where } p = 1, 2, \ldots, 2k - 1$$

for when $p \leqq k$, then $p + 2 - k < 3$, so that $C(p + 2 - k, 3) = 0$.

Collecting all these results and combining terms with like coefficients, we have

$$V_{on_s} = \sum_{p=1}^{2k-2} [C(p + 2, 3) - 4C(p + 2 - k, 3)] \left(f\left\{kn - [2k - (p + 1)]\right\} + f\left\{kn + [2k - (p + 1)]\right\}\right) + [C(2k + 1, 3) - 4C(k + 1, 3)]f(kn) \qquad (6)$$

This formula gives the interchange, in grid k, between an origin zone and any one of the side zones within ring n in terms of the individual interchanges between the subzones of which the k x k zones are composed. A similar procedure leads to a formula for the interchange, in grid k, between an origin zone and any one of the four corner zones in ring n:

$$V_{on_c} = \sum_{p=1}^{k} [p^2k - 2C(p + 1, 3)] f[kn - (k - p)] + \sum_{p=2}^{k} 2C(p + 1, 3) f[kn + (2k - p)] + \sum_{p=1}^{k-1} \left\{2C(k + p + 1, 3) [p^2k + 6C(p + 1, 3)]\right\} f[kn + (k - p)] \qquad (7)$$

Equations 6 and 7 are subject to the restriction that $k \geq 2$, which does no harm since for $k = 1$ there is no "aggregate" interchange. For equation 6 it is necessary that $n \geq 2$, since the first ring has no side zones.

In all of the above, it has been convenient to assume $n \geq 1$. When we wish to consider the aggregate interchange of a grid k zone with itself, we might think of this as the case where $n = 0$. In the tables this case has been designated in this manner, but for uniformity of notation only; $n = 0$ has no real physical meaning. The necessary formula must be developed independently of those given above but can be found by the same technique:

$$V_{00} = k^2F(0) + 4 \sum_{j=3}^{k+1} C(j,3)f(2k-j+1)$$
$$+ 4 \sum_{j=3}^{k} \left\{ C(k+j,3) - [(j-1)^2k + 2C(j,3)] \right\} f(k-j+1)$$
$$+ 4[C(k+2,3) - k]f(k-1) \qquad (8)$$

USE OF ERROR TABLES

Sample tables of the relative errors which were referred to in earlier sections are given here. (The tables were prepared by Ralph J. Marshall using a program written by him in FORTRAN.) For each table the trip interaction constant T has been held. Values of trip density per square mile (D) have been listed down the left side of each table. The variable K is the zone size multiplier. The zone area in each column is $(K/4)^2$ square miles.

The entries in the tables have been grouped in triples. Reading from the top, the first error (A) in each triple is the relative error due to the use of the method of random selection, the third error (C) is that due to the aggregation of trip endings into zones, and the second error (B) is that due to a combination of the use of random selection with that of aggregation.

A table has been prepared for each value of $T = x \cdot 10^{-6}$, x = 1, 5, 10, 15, 20, 30, 40, and 50. In the case that knowledge is desired of errors for values of T, of trip density, or of zone size other than those listed, a linear interpolation should give results within the proper order of magnitude.

As an example of at least one use of these tables, suppose that one wishes to hold the bias arising from the influences discussed here to a certain level, say a relative error of 10 percent, for a segment of the region under study. In order to use one of the sample tables given here, suppose further that the trip interaction constant determined for use in this segment is approximately $T = 15 \times 10^{-6}$, and that the trip density averages 2,000. A study of Table 4 will indicate that a zone size in excess of one square mile should be avoided in this segment of the study area.

TABLE 1

U660 DATE JUL 02,1965
PAGE 1

ERRORS FOR T -.00000100

			VALUE FOR K										
			2	3	4	5	6	7	8	9	10	11	12
		A	0.33%	0.49%	0.65%	0.81%	0.97%	1.12%	1.28%	1.43%	1.58%	1.72%	1.88%
D	100	B	0.34%	0.50%	0.65%	0.82%	0.98%	1.12%	1.29%	1.44%	1.59%	1.74%	1.90%
		C	0.04%	0.05%	0.06%	0.08%	0.10%	0.12%	0.15%	0.18%	0.21%	0.24%	0.27%
		A	0.52%	0.77%	1.02%	1.26%	1.49%	1.74%	1.96%	2.18%	2.40%	2.65%	2.87%
	250	B	0.52%	0.77%	1.03%	1.27%	1.51%	1.76%	1.99%	2.21%	2.43%	2.69%	2.92%
		C	0.06%	0.08%	0.12%	0.15%	0.20%	0.25%	0.30%	0.35%	0.41%	0.47%	0.54%
		A	0.73%	1.08%	1.42%	1.74%	2.07%	2.36%	2.69%	2.95%	3.25%	3.49%	3.83%
V	500	B	0.73%	1.09%	1.43%	1.75%	2.10%	2.39%	2.74%	3.01%	3.32%	3.58%	3.94%
		C	0.09%	0.14%	0.19%	0.26%	0.33%	0.41%	0.50%	0.58%	0.68%	0.77%	0.89%
		A	1.02%	1.49%	1.96%	2.40%	2.87%	3.20%	3.65%	4.01%	4.50%	4.66%	5.00%
A	1,000	B	1.03%	1.51%	1.99%	2.43%	2.92%	3.27%	3.74%	4.13%	4.65%	4.84%	5.21%
		C	0.15%	0.23%	0.32%	0.43%	0.56%	0.68%	0.83%	0.98%	1.17%	1.29%	1.47%
		A	1.23%	1.81%	2.38%	2.87%	3.32%	3.90%	4.36%	4.68%	5.16%	5.57%	5.88%
L	1,500	B	1.25%	1.84%	2.42%	2.93%	3.40%	4.01%	4.50%	4.86%	5.39%	5.84%	6.21%
		C	0.21%	0.31%	0.44%	0.58%	0.73%	0.93%	1.13%	1.31%	1.54%	1.77%	1.99%
		A	1.42%	2.07%	2.69%	3.25%	3.83%	4.43%	4.87%	5.48%	6.00%	6.41%	6.70%
U	2,000	B	1.44%	2.11%	2.74%	3.33%	3.94%	4.58%	5.06%	5.73%	6.31%	6.79%	7.16%
		C	0.25%	0.38%	0.54%	0.71%	0.92%	1.16%	1.38%	1.68%	1.97%	2.25%	2.52%
		A	2.15%	3.10%	3.97%	4.78%	5.53%	6.47%	7.26%	7.85%	8.23%	9.53%	9.58%
E	5,000	B	2.21%	3.19%	4.11%	4.98%	5.82%	6.87%	7.79%	8.54%	9.06%	10.62%	10.81%
		C	0.49%	0.74%	1.04%	1.40%	1.80%	2.32%	2.85%	3.35%	3.81%	4.68%	5.01%
		A	2.92%	4.14%	5.42%	6.57%	7.18%	8.10%	8.73%	10.53%	10.75%	12.71%	14.96%
	10,000	B	3.03%	4.32%	5.70%	7.00%	7.78%	8.93%	9.79%	11.98%	12.41%	14.80%	17.53%
		C	0.81%	1.21%	1.77%	2.42%	3.00%	3.76%	4.45%	5.71%	6.20%	7.58%	9.13%
		A	3.97%	5.53%	7.25%	8.23%	9.58%	10.56%	13.38%	16.82%	17.19%	21.04%	25.51%
	20,000	B	4.19%	5.88%	7.85%	9.10%	10.84%	12.20%	15.62%	19.76%	20.19%	24.63%	29.72%
		C	1.34%	2.00%	2.99%	3.90%	5.07%	6.10%	8.05%	10.38%	10.58%	12.81%	15.24%
		A	5.99%	7.59%	10.42%	13.09%	15.44%	21.32%	28.73%	28.81%	35.31%	41.87%	48.08%
	50,000	B	6.55%	8.46%	11.92%	15.28%	18.14%	24.92%	33.30%	32.98%	40.38%	48.05%	55.66%
		C	2.64%	3.74%	5.79%	7.90%	9.52%	12.91%	16.83%	16.05%	19.58%	23.57%	28.04%
		A	7.77%	11.49%	16.60%	21.79%	25.58%	34.63%	43.86%	52.14%	58.47%	62.21%	43.37%
	100,000	B	8.84%	13.32%	19.52%	25.41%	29.29%	39.51%	50.35%	60.88%	70.45%	78.72%	58.03%
		C	4.21%	6.75%	10.26%	13.08%	14.26%	19.02%	24.72%	31.43%	39.29%	48.23%	38.56%

COEFFICIENT OF VARIATION DUE TO
A- RANDOM SELECTION
B- RANDOM SELECTION AND AGGREGATION
C- AGGREGATION

TABLE 2 U660 DATE JUL 02,1965 PAGE 2

ERRORS FOR T -.00000500

			VALUE FOR K										
			2	3	4	5	6	7	8	9	10	11	12
		A	0.73%	1.08%	1.42%	1.74%	2.07%	2.36%	2.69%	2.95%	3.25%	3.49%	3.83%
D	100	B	0.74%	1.09%	1.43%	1.75%	2.10%	2.39%	2.74%	3.01%	3.32%	3.58%	3.94%
		C	0.12%	0.15%	0.20%	0.26%	0.33%	0.41%	0.50%	0.58%	0.68%	0.77%	0.89%
		A	1.14%	1.65%	2.16%	2.64%	3.10%	3.53%	3.97%	4.55%	4.78%	5.19%	5.53%
	250	B	1.15%	1.67%	2.19%	2.69%	3.17%	3.61%	4.09%	4.70%	4.96%	5.42%	5.80%
		C	0.19%	0.27%	0.38%	0.50%	0.65%	0.80%	0.97%	1.18%	1.34%	1.55%	1.75%
		A	1.58%	2.32%	2.92%	3.58%	4.14%	4.71%	5.42%	6.05%	6.57%	6.95%	7.18%
V	500	B	1.61%	2.36%	2.99%	3.68%	4.28%	4.89%	5.67%	6.38%	6.97%	7.44%	7.76%
		C	0.30%	0.45%	0.62%	0.84%	1.07%	1.33%	1.66%	2.00%	2.34%	2.66%	2.94%
		A	2.15%	3.10%	3.97%	4.78%	5.53%	6.47%	7.26%	7.85%	8.23%	9.53%	9.58%
A	1,000	B	2.21%	3.19%	4.11%	4.98%	5.82%	6.87%	7.79%	8.54%	9.06%	10.62%	10.81%
		C	0.50%	0.74%	1.04%	1.40%	1.80%	2.32%	2.84%	3.35%	3.81%	4.68%	5.01%
		A	2.63%	3.81%	4.63%	5.71%	6.38%	7.29%	7.97%	9.56%	9.90%	11.58%	11.49%
L	1,500	B	2.71%	3.94%	4.83%	6.02%	6.82%	7.90%	8.77%	10.66%	11.20%	13.26%	13.31%
		C	0.67%	1.01%	1.38%	1.91%	2.41%	3.04%	3.66%	4.71%	5.24%	6.46%	6.71%
		A	2.92%	4.14%	5.42%	6.57%	7.18%	8.10%	8.73%	10.53%	10.75%	12.71%	14.96%
U	2,000	B	3.03%	4.32%	5.70%	7.00%	7.78%	8.93%	9.79%	11.98%	12.41%	14.80%	17.53%
		C	0.81%	1.21%	1.77%	2.42%	3.00%	3.76%	4.45%	5.71%	6.20%	7.57%	9.13%
		A	4.50%	6.06%	7.78%	9.78%	11.50%	12.85%	16.61%	17.41%	21.80%	26.96%	25.61%
E	5,000	B	4.78%	6.51%	8.51%	10.97%	13.18%	14.98%	19.51%	20.44%	25.50%	31.35%	29.42%
		C	1.62%	2.36%	3.47%	4.96%	6.45%	7.69%	10.24%	10.71%	13.22%	15.99%	14.48%
		A	5.99%	7.59%	10.42%	13.09%	15.44%	21.32%	28.73%	28.81%	35.31%	41.87%	48.08%
	10,000	B	6.55%	8.46%	11.92%	15.28%	18.14%	24.92%	33.30%	32.98%	40.38%	48.05%	55.66%
		C	2.64%	3.74%	5.79%	7.90%	9.52%	12.91%	16.83%	16.04%	19.58%	23.57%	28.04%
		A	7.77%	11.49%	16.60%	21.79%	25.58%	34.63%	43.86%	52.14%	58.47%	62.21%	43.37%
	20,000	B	8.84%	13.32%	19.51%	25.41%	29.29%	39.51%	50.34%	60.88%	70.45%	78.72%	58.03%
		C	4.21%	6.75%	10.26%	13.08%	14.26%	19.02%	24.71%	31.43%	39.29%	48.23%	38.56%
		A	10.71%	19.41%	28.39%	42.98%	55.50%	62.36%	42.29%	37.78%	31.45%	24.64%	18.28%
	50,000	B	12.87%	22.65%	32.22%	49.00%	65.37%	78.92%	59.83%	64.07%	67.38%	69.87%	71.47%
		C	7.13%	11.67%	15.24%	23.53%	34.54%	48.37%	42.32%	51.75%	59.59%	65.39%	69.09%
		A	17.10%	31.80%	51.66%	62.57%	40.31%	32.07%	22.53%	14.25%	8.22%	4.37%	2.16%
	100,000	B	20.20%	35.69%	59.47%	79.23%	61.51%	66.55%	69.95%	71.50%	71.10%	69.05%	65.96%
		C	10.76%	16.20%	29.46%	48.60%	46.46%	58.32%	66.22%	70.07%	70.62%	68.92%	65.92%

COEFFICIENT OF VARIATION DUE TO
A- RANDOM SELECTION
B- RANDOM SELECTION AND AGGREGATION
C- AGGREGATION

TABLE 3

U660 DATE JUL 02,1965
PAGE 3

ERRORS FOR T -.00001000

			VALUE FOR K 2	3	4	5	6	7	8	9	10	11	12
		A	1.02%	1.49%	1.96%	2.40%	2.87%	3.20%	3.65%	4.01%	4.50%	4.66%	5.00%
D	100	B	1.04%	1.51%	1.99%	2.44%	2.92%	3.27%	3.74%	4.13%	4.65%	4.84%	5.21%
		C	0.19%	0.24%	0.33%	0.43%	0.56%	0.68%	0.83%	0.98%	1.16%	1.29%	1.47%
		A	1.58%	2.32%	2.92%	3.58%	4.14%	4.71%	5.42%	6.05%	6.57%	6.95%	7.18%
	250	B	1.61%	2.36%	2.99%	3.68%	4.28%	4.89%	5.67%	6.38%	6.97%	7.44%	7.76%
		C	0.31%	0.45%	0.62%	0.84%	1.07%	1.33%	1.66%	2.00%	2.34%	2.66%	2.94%
		A	2.15%	3.10%	3.97%	4.78%	5.53%	6.47%	7.26%	7.85%	8.23%	9.53%	9.58%
V	500	B	2.21%	3.19%	4.11%	4.98%	5.82%	6.87%	7.79%	8.54%	9.06%	10.62%	10.81%
		C	0.50%	0.74%	1.04%	1.40%	1.80%	2.32%	2.84%	3.35%	3.80%	4.68%	5.01%
		A	2.92%	4.14%	5.42%	6.57%	7.18%	8.10%	8.73%	10.53%	10.75%	12.71%	14.96%
A	1,000	B	3.03%	4.32%	5.70%	7.00%	7.78%	8.93%	9.79%	11.98%	12.41%	14.79%	17.53%
		C	0.81%	1.21%	1.77%	2.42%	3.00%	3.76%	4.45%	5.71%	6.20%	7.57%	9.13%
		A	3.62%	5.07%	6.37%	7.41%	8.73%	9.74%	10.38%	12.76%	15.57%	15.57%	18.63%
L	1,500	B	3.79%	5.34%	6.80%	8.06%	9.68%	11.02%	11.97%	14.86%	18.26%	18.30%	21.86%
		C	1.12%	1.67%	2.38%	3.17%	4.19%	5.16%	5.96%	7.61%	9.55%	9.61%	11.43%
		A	3.97%	5.53%	7.25%	8.23%	9.58%	10.56%	13.38%	16.82%	17.19%	21.04%	25.51%
U	2,000	B	4.19%	5.88%	7.85%	9.10%	10.84%	12.20%	15.62%	19.76%	20.19%	24.63%	29.72%
		C	1.34%	2.00%	2.99%	3.90%	5.07%	6.10%	8.05%	10.38%	10.58%	12.80%	15.24%
		A	5.99%	7.59%	10.42%	13.09%	15.44%	21.32%	28.73%	28.81%	35.31%	41.87%	48.08%
E	5,000	B	6.55%	8.46%	11.92%	15.28%	18.14%	24.92%	33.30%	32.98%	40.38%	48.05%	55.65%
		C	2.64%	3.74%	5.79%	7.89%	9.52%	12.91%	16.83%	16.04%	19.58%	23.57%	28.04%
		A	7.77%	11.49%	16.60%	21.79%	25.58%	34.63%	43.86%	52.14%	58.47%	62.21%	43.37%
	10,000	B	8.84%	13.32%	19.51%	25.41%	29.29%	39.51%	50.34%	60.88%	70.45%	78.72%	58.03%
		C	4.21%	6.74%	10.26%	13.07%	14.26%	19.02%	24.71%	31.43%	39.29%	48.23%	38.55%
		A	10.41%	15.41%	28.71%	35.27%	48.05%	57.94%	62.81%	42.16%	38.11%	32.55%	26.45%
	20,000	B	12.35%	18.13%	33.06%	40.07%	55.33%	69.33%	80.83%	60.17%	63.97%	67.00%	69.40%
		C	6.65%	9.55%	16.38%	19.02%	27.43%	38.08%	50.88%	42.93%	51.37%	58.57%	64.16%
		A	17.10%	31.80%	51.66%	62.57%	40.31%	32.07%	22.53%	14.25%	8.22%	4.37%	2.16%
	50,000	B	20.20%	35.69%	59.47%	79.23%	61.51%	66.55%	69.95%	71.50%	71.10%	69.05%	65.96%
		C	10.76%	16.20%	29.46%	48.60%	46.46%	58.32%	66.22%	70.07%	70.62%	68.92%	65.92%
		A	28.15%	55.41%	41.93%	31.28%	18.21%	8.70%	3.52%	1.23%	0.38%	0.10%	0.02%
	100,000	B	31.19%	63.85%	58.33%	65.78%	69.91%	70.25%	67.28%	62.56%	57.48%	52.74%	48.58%
		C	13.43%	31.72%	40.55%	57.87%	67.49%	69.71%	67.18%	62.55%	57.48%	52.74%	48.58%

COEFFICIENT OF VARIATION DUE TO
A- RANDOM SELECTION
B- RANDOM SELECTION AND AGGREGATION
C- AGGREGATION

TABLE 4

U660 DATE JUL 02,1965
PAGE 4

ERRORS FOR T -.00001500

			VALUE FOR K										
			2	3	4	5	6	7	8	9	10	11	12
		A	1.23%	1.81%	2.38%	2.87%	3.32%	3.90%	4.36%	4.68%	5.16%	5.57%	5.88%
D	100	B	1.25%	1.84%	2.42%	2.93%	3.40%	4.01%	4.50%	4.86%	5.39%	5.84%	6.21%
		C	0.24%	0.32%	0.44%	0.58%	0.73%	0.93%	1.12%	1.30%	1.54%	1.77%	1.99%
		A	1.90%	2.72%	3.62%	4.28%	5.07%	5.62%	6.37%	6.97%	7.41%	7.64%	8.73%
	250	B	1.95%	2.79%	3.72%	4.43%	5.29%	5.91%	6.76%	7.47%	8.02%	8.37%	9.65%
		C	0.41%	0.60%	0.86%	1.14%	1.49%	1.82%	2.25%	2.68%	3.08%	3.43%	4.12%
		A	2.63%	3.81%	4.63%	5.71%	6.38%	7.29%	7.97%	9.56%	9.90%	11.58%	11.49%
V	500	B	2.71%	3.94%	4.83%	6.02%	6.82%	7.90%	8.77%	10.66%	11.20%	13.26%	13.31%
		C	0.67%	1.01%	1.38%	1.91%	2.40%	3.04%	3.66%	4.71%	5.24%	6.45%	6.71%
		A	3.62%	5.07%	6.37%	7.41%	8.73%	9.74%	10.38%	12.76%	15.57%	15.57%	18.63%
A	1,000	B	3.79%	5.34%	6.80%	8.06%	9.68%	11.02%	11.97%	14.86%	18.26%	18.30%	21.86%
		C	1.12%	1.67%	2.38%	3.17%	4.19%	5.16%	5.96%	7.61%	9.55%	9.61%	11.43%
		A	4.14%	6.05%	7.18%	9.01%	10.53%	11.70%	14.96%	15.64%	19.45%	23.96%	29.14%
L	1,500	B	4.39%	6.46%	7.82%	10.03%	12.00%	13.57%	17.53%	18.37%	22.80%	27.95%	33.80%
		C	1.45%	2.26%	3.11%	4.42%	5.75%	6.89%	9.14%	9.64%	11.90%	14.40%	17.13%
		A	4.63%	6.38%	7.57%	9.89%	11.49%	15.26%	16.48%	21.14%	26.71%	25.82%	30.71%
U	2,000	B	4.96%	6.91%	8.84%	11.24%	13.33%	17.89%	19.36%	24.73%	31.06%	29.65%	35.15%
		C	1.77%	2.65%	3.81%	5.33%	6.76%	9.34%	10.15%	12.84%	15.84%	14.57%	17.11%
		A	6.90%	9.09%	12.61%	16.08%	23.73%	26.13%	33.97%	42.01%	49.51%	55.77%	60.27%
E	5,000	B	7.71%	10.39%	14.73%	18.88%	27.63%	29.93%	38.80%	48.16%	57.44%	66.12%	73.91%
		C	3.44%	5.03%	7.62%	9.89%	14.16%	14.60%	18.73%	23.56%	29.12%	35.51%	42.77%
		A	9.46%	14.05%	20.83%	26.64%	37.87%	48.79%	57.38%	62.26%	43.03%	40.30%	36.07%
	10,000	B	11.03%	16.54%	24.29%	30.41%	43.16%	56.36%	68.51%	78.79%	58.76%	62.38%	65.34%
		C	5.68%	8.72%	12.49%	14.65%	20.70%	28.23%	37.42%	48.27%	40.01%	47.61%	54.49%
		A	12.58%	23.70%	33.91%	49.47%	60.25%	43.11%	39.04%	32.39%	24.94%	17.98%	12.23%
	20,000	B	15.11%	27.39%	38.35%	56.95%	73.43%	58.05%	63.06%	66.85%	69.67%	71.42%	71.98%
		C	8.38%	13.74%	17.92%	28.23%	41.98%	38.87%	49.52%	58.48%	65.05%	69.12%	70.94%
		A	26.63%	45.64%	62.07%	39.39%	28.68%	17.51%	9:24%	4.30%	1.79%	0.67%	0.23%
	50,000	B	30.34%	51.33%	77.01%	61.72%	67.39%	70.56%	70.84%	68.54%	64.68%	60.26%	55.92%
		C	14.55%	23.50%	45.59%	47.52%	60.98%	68.35%	70.23%	68.40%	64.65%	60.26%	55.92%
		A	41.25%	42.86%	32.50%	16.75%	6.45%	1.96%	0.48%	0.10%	0.02%	0.00%	0.00%
	100,000	B	44.99%	53.96%	64.13%	69.13%	68.57%	63.84%	57.73%	51.95%	47.01%	42.87%	39.40%
		C	17.97%	32.77%	55.28%	67.07%	68.27%	63.81%	57.73%	51.95%	47.01%	42.87%	39.40%

COEFFICIENT OF VARIATION DUE TO
A- RANDOM SELECTION
B- RANDOM SELECTION AND AGGREGATION
C- AGGREGATION

TABLE 5

U660 DATE JUL 02,1965
PAGE 5

-ERRORS FOR T -.00002000

			VALUE FOR K										
			2	3	4	5	6	7	8	9	10	11	12
		A	1.42%	2.07%	2.69%	3.25%	3.83%	4.43%	4.87%	5.48%	6.00%	6.41%	6.70%
D	100	B	1.45%	2.11%	2.74%	3.33%	3.94%	4.58%	5.06%	5.73%	6.31%	6.79%	7.16%
		C	0.30%	0.39%	0.54%	0.71%	0.92%	1.16%	1.38%	1.67%	1.97%	2.25%	2.51%
		A	2.15%	3.10%	3.97%	4.78%	5.53%	6.47%	7.26%	7.85%	8.23%	9.53%	9.58%
	250	B	2.21%	3.19%	4.11%	4.98%	5.82%	6.87%	7.79%	8.54%	9.06%	10.62%	10.81%
		C	0.50%	0.74%	1.04%	1.40%	1.80%	2.31%	2.84%	3.35%	3.80%	4.68%	5.01%
		A	2.92%	4.14%	5.42%	6.57%	7.18%	8.10%	8.73%	10.53%	10.75%	12.71%	14.96%
V	500	B	3.03%	4.32%	5.70%	7.00%	7.78%	8.93%	9.79%	11.98%	12.41%	14.79%	17.53%
		C	0.81%	1.21%	1.77%	2.42%	3.00%	3.75%	4.45%	5.70%	6.20%	7.57%	9.13%
		A	3.97%	5.53%	7.25%	8.23%	9.58%	10.56%	13.38%	16.82%	17.19%	21.04%	25.51%
A	1,000	B	4.19%	5.88%	7.85%	9.10%	10.83%	12.20%	15.61%	19.76%	20.18%	24.63%	29.72%
		C	1.34%	2.00%	2.99%	3.90%	5.07%	6.10%	8.05%	10.37%	10.58%	12.80%	15.24%
		A	4.63%	6.38%	7.97%	9.89%	11.49%	15.26%	16.48%	21.14%	26.71%	25.82%	30.71%
L	1,500	B	4.96%	6.91%	8.84%	11.24%	13.33%	17.89%	19.36%	24.73%	31.05%	29.65%	35.15%
		C	1.77%	2.65%	3.81%	5.33%	6.76%	9.34%	10.15%	12.84%	15.84%	14.57%	17.11%
		A	5.42%	7.18%	8.72%	10.74%	14.96%	16.83%	22.37%	29.14%	28.46%	34.25%	40.14%
U	2,000	B	5.87%	7.89%	9.86%	12.44%	17.53%	19.76%	26.13%	33.77%	32.60%	39.18%	46.01%
		C	2.24%	3.27%	4.60%	6.27%	9.15%	10.35%	13.50%	17.07%	15.89%	19.01%	22.50%
		A	7.77%	11.49%	16.60%	21.79%	25.58%	34.63%	43.86%	52.14%	58.47%	62.21%	43.37%
E	5,000	B	8.84%	13.32%	19.51%	25.41%	29.29%	39.51%	50.34%	60.88%	70.45%	78.72%	58.02%
		C	4.21%	6.74%	10.26%	13.07%	14.26%	19.01%	24.71%	31.42%	39.29%	48.23%	38.55%
		A	10.41%	15.41%	28.71%	35.27%	48.05%	57.94%	62.81%	42.16%	38.11%	32.55%	26.45%
	10,000	B	12.35%	18.13%	33.05%	40.07%	55.33%	69.73%	80.83%	60.17%	63.96%	67.00%	69.39%
		C	6.65%	9.55%	16.38%	19.01%	27.43%	38.08%	50.87%	42.93%	51.37%	58.56%	64.15%
		A	13.67%	25.48%	43.79%	58.44%	43.20%	38.53%	30.61%	22.08%	14.66%	9.04%	5.21%
	20,000	B	16.39%	28.83%	49.71%	69.82%	57.43%	63.20%	67.40%	70.28%	71.67%	71.48%	69.91%
		C	9.04%	13.50%	23.51%	38.20%	37.84%	50.11%	60.05%	66.72%	70.16%	70.91%	69.72%
		A	28.15%	55.41%	41.93%	31.28%	18.21%	8.70%	3.52%	1.23%	0.38%	0.10%	0.02%
	50,000	B	31.18%	63.85%	58.33%	65.78%	69.90%	70.25%	67.27%	62.56%	57.48%	52.74%	48.58%
		C	13.43%	31.72%	40.55%	57.87%	67.49%	69.70%	67.18%	62.55%	57.47%	52.74%	48.58%
		A	51.47%	39.73%	22.34%	8.18%	2.15%	0.43%	0.07%	0.01%	0.00%	0.00%	0.00%
	100,000	B	56.69%	58.48%	66.77%	68.20%	63.41%	56.48%	50.03%	44.69%	40.35%	36.77%	33.78%
		C	23.76%	42.90%	62.93%	67.71%	63.37%	56.48%	50.03%	44.69%	40.35%	36.77%	33.78%

COEFFICIENT OF VARIATION DUE TO
A- RANDOM SELECTION
B- RANDOM SELECTION AND AGGREGATION
C- AGGREGATION

TABLE 6

ERRORS FOR T -.00003000

			VALUE FOR K									
			2	3	4	5	6	7	8	9	10	11	12
		A	1.71%	2.48%	3.27%	3.96%	4.53%	5.08%	5.81%	6.43%	6.91%	7.23%	8.25%
D	100	B	1.75%	2.54%	3.35%	4.08%	4.69%	5.31%	6.11%	6.82%	7.39%	7.81%	8.99%
		C	0.39%	0.52%	0.73%	0.98%	1.24%	1.53%	1.90%	2.27%	2.63%	2.97%	3.58%
		A	2.63%	3.81%	4.63%	5.71%	6.38%	7.29%	7.97%	9.56%	9.90%	11.58%	11.49%
	250	B	2.72%	3.94%	4.83%	6.02%	6.82%	7.90%	8.77%	10.66%	11.20%	13.26%	13.31%
		C	0.68%	1.01%	1.38%	1.91%	2.40%	3.04%	3.66%	4.71%	5.24%	6.45%	6.71%
		A	3.62%	5.07%	6.37%	7.41%	8.73%	9.74%	10.38%	12.76%	15.57%	15.57%	18.63%
V	500	B	3.79%	5.34%	6.80%	8.05%	9.68%	11.02%	11.97%	14.85%	18.26%	18.30%	21.86%
		C	1.12%	1.67%	2.38%	3.17%	4.19%	5.16%	5.96%	7.61%	9.54%	9.61%	11.43%
		A	4.63%	6.38%	7.97%	9.89%	11.49%	15.26%	16.48%	21.14%	26.71%	25.82%	30.71%
A	1,000	B	4.96%	6.90%	8.84%	11.24%	13.32%	17.89%	19.36%	24.73%	31.05%	29.65%	35.15%
		C	1.77%	2.64%	3.81%	5.33%	6.75%	9.34%	10.15%	12.83%	15.84%	14.56%	17.10%
		A	5.53%	7.85%	9.57%	11.90%	13.91%	19.04%	25.51%	25.93%	31.95%	38.19%	46.32%
L	1,500	B	6.02%	8.67%	10.89%	13.85%	16.35%	22.30%	29.68%	29.74%	36.53%	43.71%	51.00%
		C	2.37%	3.69%	5.18%	7.07%	8.59%	11.62%	15.17%	14.57%	17.72%	21.27%	25.24%
		A	6.36%	8.72%	10.37%	15.56%	18.62%	26.11%	27.32%	34.38%	41.57%	48.37%	56.26%
U	2,000	B	7.02%	9.81%	12.03%	18.27%	21.82%	30.33%	31.28%	39.29%	47.67%	56.00%	63.91%
		C	2.97%	4.49%	6.09%	9.56%	11.37%	15.44%	15.24%	19.01%	23.33%	28.23%	33.77%
		A	9.46%	14.05%	20.83%	26.64%	37.87%	48.79%	57.38%	62.26%	43.03%	40.30%	36.07%
E	5,000	B	11.03%	16.53%	24.29%	30.40%	43.16%	56.36%	68.51%	78.78%	58.75%	62.38%	65.34%
		C	5.68%	8.72%	12.49%	14.65%	20.69%	28.22%	37.42%	48.27%	40.01%	47.61%	54.48%
		A	12.58%	23.70%	33.91%	49.47%	60.25%	43.11%	39.04%	32.39%	24.94%	17.98%	12.23%
	10,000	B	15.11%	27.39%	38.35%	56.95%	73.43%	58.05%	63.06%	66.85%	69.67%	71.42%	71.98%
		C	8.38%	13.74%	17.92%	28.22%	41.97%	38.87%	49.51%	58.48%	65.05%	69.12%	70.93%
		A	20.76%	37.75%	57.34%	42.79%	35.92%	25.63%	16.01%	8.93%	4.50%	2.08%	0.88%
	20,000	B	24.16%	42.27%	67.57%	57.86%	64.39%	68.79%	71.10%	71.07%	68.99%	65.61%	61.69%
		C	12.36%	19.01%	35.75%	38.95%	53.44%	63.84%	69.28%	70.50%	68.84%	65.58%	61.68%
		A	41.25%	42.86%	32.50%	16.75%	6.45%	1.96%	0.48%	0.10%	0.02%	0.00%	0.00%
	50,000	B	44.99%	53.95%	64.13%	69.12%	68.57%	63.84%	57.73%	51.95%	47.00%	42.87%	39.40%
		C	17.96%	32.77%	55.28%	67.06%	68.27%	63.81%	57.73%	51.95%	47.00%	42.87%	39.40%
		A	41.91%	28.29%	9.17%	1.78%	0.23%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%
	100,000	B	48.73%	62.61%	66.46%	60.93%	52.69%	45.58%	40.04%	35.68%	32.19%	29.32%	26.92%
		C	24.86%	55.86%	65.82%	60.91%	52.69%	45.58%	40.04%	35.68%	32.19%	29.32%	26.92%

COEFFICIENT OF VARIATION DUE TO
A- RANDOM SELECTION
B- RANDOM SELECTION AND AGGREGATION
C- AGGREGATION

ERRORS FOR T -.00003999

VALUE FOR K

			2	3	4	5	6	7	8	9	10	11	12
		A	1.96%	2.87%	3.65%	4.50%	5.00%	5.90%	6.69%	7.32%	7.78%	8.03%	9.17%
D	100	B	2.02%	2.94%	3.76%	4.66%	5.22%	6.22%	7.11%	7.87%	8.45%	8.82%	10.18%
		C	0.47%	0.65%	0.89%	1.22%	1.51%	1.95%	2.41%	2.87%	3.30%	3.67%	4.42%
		A	2.92%	4.14%	5.42%	6.57%	7.18%	8.10%	8.73%	10.53%	10.75%	12.71%	14.96%
	250	B	3.04%	4.32%	5.70%	7.00%	7.78%	8.93%	9.79%	11.98%	12.41%	14.79%	17.52%
		C	0.82%	1.21%	1.76%	2.42%	3.00%	3.75%	4.44%	5.70%	6.20%	7.57%	9.13%
		A	3.97%	5.53%	7.25%	8.23%	9.58%	10.56%	13.38%	16.82%	17.19%	21.04%	25.51%
V	500	B	4.19%	5.88%	7.84%	9.10%	10.83%	12.20%	15.61%	19.76%	20.18%	24.63%	29.71%
		C	1.35%	2.00%	2.99%	3.90%	5.07%	6.09%	8.05%	10.37%	10.57%	12.80%	15.23%
		A	5.42%	7.18%	8.72%	10.74%	14.96%	16.83%	22.37%	29.14%	28.46%	34.25%	40.14%
A	1,000	B	5.87%	7.89%	9.86%	12.44%	17.53%	19.76%	26.13%	33.77%	32.59%	39.17%	46.01%
		C	2.24%	3.27%	4.59%	6.27%	9.14%	10.35%	13.50%	17.07%	15.88%	19.01%	22.49%
		A	6.36%	8.72%	10.37%	15.56%	18.62%	26.11%	27.32%	34.38%	41.57%	48.37%	54.26%
L	1,500	B	7.02%	9.81%	12.03%	18.27%	21.82%	30.33%	31.28%	39.29%	47.67%	56.00%	63.91%
		C	2.97%	4.49%	6.08%	9.56%	11.37%	15.44%	15.24%	19.01%	23.33%	28.22%	33.77%
		A	7.25%	9.57%	13.37%	17.18%	25.50%	27.88%	36.09%	44.31%	51.73%	57.62%	61.49%
U	2,000	B	8.13%	10.97%	15.65%	20.15%	29.62%	31.88%	41.22%	50.92%	60.36%	69.03%	76.68%
		C	3.67%	5.36%	8.13%	10.52%	15.07%	15.46%	19.91%	25.09%	31.09%	38.01%	45.81%
		A	10.41%	15.41%	28.71%	35.27%	48.05%	57.94%	62.81%	42.16%	38.11%	32.55%	26.45%
E	5,000	B	12.35%	18.12%	33.05%	40.06%	55.32%	69.33%	80.82%	60.16%	63.96%	67.00%	69.39%
		C	6.64%	9.54%	16.37%	19.01%	27.42%	38.07%	50.86%	42.92%	51.37%	58.56%	64.15%
		A	13.67%	25.48%	43.79%	58.44%	43.20%	38.53%	30.61%	22.08%	14.66%	9.04%	5.21%
	10,000	B	16.39%	28.83%	49.70%	69.82%	57.42%	63.20%	67.39%	70.28%	71.67%	71.48%	69.91%
		C	9.04%	13.50%	23.51%	38.19%	37.84%	50.10%	60.04%	66.72%	70.15%	70.90%	69.71%
		A	28.65%	47.94%	62.78%	37.89%	26.35%	15.30%	7.65%	3.36%	1.31%	0.46%	0.15%
	20,000	B	32.45%	54.10%	79.56%	62.71%	68.11%	70.78%	70.35%	67.42%	63.16%	58.58%	54.23%
		C	15.23%	25.06%	48.87%	49.97%	62.81%	69.10%	69.93%	67.33%	63.15%	58.58%	54.23%
		A	51.47%	39.73%	22.34%	8.18%	2.15%	0.43%	0.07%	0.01%	0.00%	0.00%	0.00%
	50,000	B	56.69%	58.47%	66.77%	68.20%	63.40%	56.47%	50.02%	44.68%	40.34%	36.76%	33.77%
		C	23.75%	42.90%	62.92%	67.71%	63.37%	56.47%	50.02%	44.68%	40.34%	36.76%	33.77%
		A	40.76%	17.99%	3.50%	0.38%	0.02%	0.00%	0.00%	0.00%	0.00%	0. %	0. %
	100,000	B	52.59%	63.68%	62.00%	53.06%	44.79%	38.55%	33.82%	30.12%	27.16%	10.99%	10.00%
		C	33.24%	61.08%	61.90%	53.06%	44.79%	38.55%	33.82%	30.12%	27.16%	10.99%	10.00%

COEFFICIENT OF VARIATION DUE TO
A- RANDOM SELECTION
B- RANDOM SELECTION AND AGGREGATION
C- AGGREGATION

TABLE 8 U660 DATE JUL 02,1965
PAGE 8

ERRORS FCR T -.00004999

			VALUE FOR K										
			2	3	4	5	6	7	8	9	10	11	12
		A	2.15%	3.10%	3.97%	4.78%	5.53%	6.47%	7.26%	7.85%	8.23%	9.53%	9.58%
D	100	B	2.22%	3.19%	4.11%	4.98%	5.82%	6.87%	7.79%	8.54%	9.06%	10.62%	1C.81%
		C	0.55%	0.75%	1.04%	1.40%	1.80%	2.31%	2.84%	3.35%	3.8C%	4.67%	5.00%
		A	3.25%	4.67%	6.00%	7.13%	7.59%	9.63%	1C.43%	10.86%	13.09%	15.70%	15.45%
	250	B	3.39%	4.89%	6.35%	7.67%	8.34%	10.75%	11.85%	12.55%	15.26%	18.41%	18.15%
		C	C.97%	1.45%	2.10%	2.84%	3.45%	4.78%	5.63%	6.28%	7.84%	9.62%	9.52%
		A	4.50%	6.06%	7.78%	9.78%	11.5C%	12.85%	16.61%	17.41%	21.8C%	26.96%	25.61%
V	5C0	B	4.78%	6.50%	8.51%	10.96%	13.18%	14.97%	19.50%	20.44%	25.49%	31.34%	29.41%
		C	1.62%	2.36%	3.46%	4.95%	6.44%	7.69%	1C.23%	1C.70%	13.21%	15.98%	14.46%
		A	5.99%	7.59%	10.42%	13.09%	15.44%	21.32%	28.73%	28.81%	35.31%	41.87%	48.08%
A	1,000	B	6.55%	8.46%	11.92%	15.28%	18.13%	24.92%	33.29%	32.97%	4C.37%	48.04%	55.65%
		C	2.64%	3.73%	5.78%	7.89%	9.51%	12.90%	16.82%	16.03%	19.57%	23.56%	28.C2%
		A	6.90%	9.09%	12.61%	16.C8%	23.73%	26.13%	33.97%	42.01%	49.51%	55.77%	6C.27%
L	1,500	B	7.71%	10.39%	14.73%	18.88%	27.63%	29.93%	38.79%	48.16%	57.43%	66.11%	73.90%
		C	3.44%	5.03%	7.61%	9.89%	14.15%	14.59%	18.73%	23.55%	29.11%	35.50%	42.76%
		A	7.77%	11.49%	16.60%	21.79%	25.58%	34.63%	43.86%	52.14%	58.47%	62.21%	43.37%
U	2,000	B	8.83%	13.32%	19.51%	25.41%	29.29%	39.51%	50.34%	6C.88%	7C.44%	78.71%	58.02%
		C	4.20%	6.74%	10.25%	13.07%	14.25%	19.01%	24.70%	31.42%	39.28%	48.22%	38.54%
		A	10.71%	19.41%	28.39%	42.98%	55.5C%	62.36%	42.29%	37.78%	31.45%	24.64%	18.28%
E	5,000	B	12.86%	22.65%	32.22%	48.99%	65.36%	78.91%	59.82%	64.06%	67.37%	69.86%	71.46%
		C	7.12%	11.66%	15.23%	23.52%	34.53%	48.35%	42.31%	51.74%	59.58%	65.38%	69.08%
		A	17.10%	31.80%	51.66%	62.57%	4C.31%	32.C7%	22.53%	14.25%	8.22%	4.37%	2.16%
	10,000	B	20.20%	35.69%	59.47%	79.22%	61.51%	66.54%	69.94%	71.49%	71.09%	69.05%	65.95%
		C	10.75%	16.19%	29.45%	48.58%	46.45%	58.31%	66.21%	70.06%	70.61%	68.91%	65.92%
		A	28.15%	55.41%	41.93%	31.28%	18.21%	8.70%	3.52%	1.23%	0.38%	0.10%	0.02%
	20,000	B	31.18%	63.84%	58.32%	65.78%	69.90%	70.24%	67.27%	62.55%	57.47%	52.74%	48.57%
		C	13.42%	31.71%	40.54%	57.87%	67.48%	69.70%	67.18%	62.54%	57.47%	52.74%	48.57%
		A	58.29%	34.24%	14.53%	3.85%	C.70%	0.09%	0.01%	0.00%	0.00%	0.00%	C.C0%
	50,000	B	65.80%	61.06%	67.38%	64.97%	57.74%	50.39%	44.35%	39.55%	35.68%	32.51%	29.85%
		C	30.53%	50.55%	65.79%	64.86%	57.74%	50.39%	44.35%	39.55%	35.68%	32.51%	29.85%
		A	36.76%	10.86%	1.30%	0.08%	C.C0%	0.00%	C.00%	0.C0%	0. %	0. %	0. %
	100,000	B	54.24%	62.67%	56.43%	46.62%	39.07%	33.58%	29.45%	26.22%	10.45%	9.43%	8.58%
		C	39.89%	61.72%	56.42%	46.62%	39.07%	33.58%	29.45%	26.22%	10.45%	9.43%	8.58%

COEFFICIENT OF VARIATION DUE TO
A- RANDOM SELECTION
B- RANDOM SELECTION AND AGGREGATION
C- AGGREGATION

Modal Split Model in the Penn-Jersey Transportation Study Area

ANTHONY R. TOMAZINIS, Transportation Planning Consultant, Delaware Valley Regional Planning Commission

•THE analysis and projection of mass transit trips of a metropolitan region have been in the center of interest of several major urban transportation studies in the recent past. This interest includes a desire for more accurate, detailed and comprehensive projection of transit system utilization at given future time intervals within the urban regions and, in a few cases, a desire to incorporate the effect of the transportation systems (highway and transit) themselves on the magnitude and the particular characteristics of the travel demand in a region.

The selection of the mode of travel by each individual has frequently been recognized as an event of substantial complexity, involving considerations of such diverse nature as personal preferences, availability of alternative means of travel, and sensitivity and meaningfulness of the means of measuring transportation systems. Various methods of incorporating these elements of the problem have been proposed and put to use by several study groups in the past few years. Relationships, frequently called mathematical models, were developed, and with various degrees of accuracy and confidence were put to use in simulating and projecting the transit travel pattern in various urban areas.

A concerned effort toward an accurate and reliable analysis and projection of the transit trips within the Philadelphia metropolitan area has also been part of the work program of the Penn-Jersey Transportation Study staff since early in 1961. Since then, several attempts were carried out in a continuous and cumulative effort to derive a model which will meet technical as well as policy objectives of the Study. The various individual projects which were undertaken can be considered as falling within three major phases of work: (a) the test of the 1947 O-D data and the multivariable model initially tested, (b) the simplified model of the 1960 O-D data and the initial 1975 projections, and (c) the complete modal split model for 1975 and 1985 projections.

THE 1947 TEST MODEL

One of the very first attempts to reproduce transit trip rates in the region was the one based on data of sample districts of the 1947 O-D survey. Data limitations made it necessary to limit the test to 15 districts of the Study area. The test was intended to investigate in a preliminary manner the relevance of several of the variables which appear initially significant in the modal split problem. Ten such variables were finally selected for the test, among a much larger number of conceptually suitable variables which were initially defined. The selected variables were formed on a district basis and were named as follows:

1. Car ownership rate (cars per 100 persons);
2. Density of residential development (D. U. per gross residential acre);
3. Transit system accessibility by cost codes;
4. Transit system accessibility by time codes;
5. Transit system serving capacity (vehicle departures within 24 hours × total passenger capacity of each vehicle);
6. Percent of persons between 5 and 19 years old to total district population;
7. Employed resident labor force;

Paper sponsored by Committee on Origin and Destination.

8. Percent of resident labor force to total district population;
9. Reported median number of years of school completed by residents; and
10. Reported median household income.

The dependent variable was the percent of total trips in the district which were made by mass transit (including railroad internal trips, subway and surface trips).

The geographic distribution of the test districts was extended to include the Philadelphia and Camden CBDs and several other districts of the Pennsylvania and New Jersey sides of the region, located at various distances from the Philadelphia CBD. The test was primarily a single multiple regression analysis including several runs and a step-by-step incorporation of the variables.

The results of this test are shown in Tables 1, 2, and 3. Table 1 gives the matrix of the correlation coefficients of each pair of variables. This table reveals several high intercorrelations such as the ones between car ownership rates and residential density, between the median household income and the median years of school completed, or between several other pairs of variables. Also one can point out the very low relationship demonstrated between the two types of system accessibility, first derived on the basis of time travel and then on the basis of travel cost, in the transit and highway systems of 1947.

The correlations of the ten variables with the percent of total transit trips of each district produced a coefficient of correlation of R = 0. 995 and an S_{yx} = 2. 32 percent. The step-by-step incorporation of the variables revealed, however, that for several rather apparent reasons a number of the variables contributed in a minor manner to the overall relationship. Examination of the significance of each variable with the help of the traditional tools, such as the simple correlation coefficients, the partial correlation coefficients, and the beta coefficients, indicated that if a different order of successive incorporation of the variables were adopted (than the one suggested by their listing), we could reach high levels of correlation with fewer than ten variables. Tests of this nature produced the results given in Table 2. From this table it became clear that an R of 0. 99 and an S_{yx} of 2. 80 percent could be reached using only six variables. Even four variables appear to be capable of producing an R of 0. 98, if properly selected. The actual level of simulation achieved by each of these sets of variables is given in Table 3, on a district-by-district basis.

Among the additional conclusions which this preliminary analysis indicated was that the car ownership rate appeared to be the most significant variable in the 1947 set of circumstances and also that the level of transit service and the income appeared to be of equal and of high significance. Next in line of significance appeared to be the percent of labor force in the district, its transit system time accessibility, and the percent of people between 5 and 19 years of age.

THE 1960 ORIGINAL MODAL SPLIT RELATIONSHIPS

The second phase of the investigation on the modal split problem was carried out on the basis of the data of the 1960 O-D survey. This phase included several differences from the previous one; it also attempted to capitalize on the conclusions of the analysis of the 1947 data and other previous works, and the whole effort became essentially part of the trip generation procedure of the Study.

The trip generation analysis in PJTS emphasized trips in five groups of trip purpose. Home origin trips were divided into trips from home to work and trips from home to all non-work purposes. Non-home origin trips were divided into three groups—trips from work to home, trips from non-work origins to home and trips from non-home origins to non-home destinations. In addition to the above five individual trip purposes, attempts were made to develop relationships for total home origin trips and total non-home origin trips. In terms of trip generating types of land use, the trip generation analysis emphasized the derivation of relationships for trips from residential land use (home and non-home origin all under L. U. Code 0) from Offices (L. U. Code 1), from a combination of Retail and Services and Passenger Transportation land uses (L. U. Codes 2 + 5), from another combination of Manufacturing, Wholesale and Goods Transportation land uses (L. U. Codes 3 + 4 + 6), and from the combination of Public Buildings and Community Facilities (L. U. Codes 7 + 8).

TABLE 1

CORRELATION MATRIX
(Simple correlation coefficients between the indicated pairs of variables)

Variable	Y	1	2	3	4	5	6	7	8	9	10
Y	1. 000	-0. 787	0. 625	0. 663	0. 503	0. 675	-0. 311	0. 515	0. 551	-0. 491	-0. 663
1	-0. 787	1. 000	-0. 912	-0. 429	-0. 373	-0. 481	0. 035	-0. 532	-0. 061	0. 823	0. 872
2	0. 625	-0. 912	1. 000	0. 185	0. 253	0. 295	0. 100	0. 636	-0. 090	-0. 852	-0. 783
3	0. 663	-0. 429	0. 185	1. 000	0. 119	0. 920	-0. 807	-0. 068	0. 602	0. 074	-0. 557
4	0. 503	-0. 373	0. 253	0. 119	1. 000	0. 055	0. 030	0. 383	0. 230	-0. 404	-0. 103
5	0. 675	-0. 481	0. 295	0. 920	0. 055	1. 000	-0. 758	0. 008	0. 443	0. 015	-0. 626
6	-0. 311	0. 035	0. 100	-0. 807	0. 030	-0. 758	1. 000	0. 207	-0. 456	-0. 418	0. 163
7	0. 515	-0. 532	0. 636	-0. 068	0. 383	0. 008	0. 207	1. 000	0. 027	-0. 597	-0. 320
8	0. 551	-0. 061	-0. 090	0. 602	0. 230	0. 443	-0. 456	0. 027	1. 000	0. 186	-0. 192
9	-0. 491	0. 823	-0. 852	0. 074	-0. 404	0. 015	-0. 418	-0. 597	0. 186	1. 000	0. 614
10	-0. 663	0. 872	-0. 783	-0. 557	-0. 103	-0. 626	0. 163	-0. 320	-0. 192	0. 614	1. 000

TABLE 2

STATISTICAL RESULTS OF VARIABLE TESTING: 1947 MODEL

Case A				Case B				Case C			
According to Betas				According to Partial Coefficients				According to Simple Coefficients			
Variable	Successive R	Successive S	Betas of the Final Equation	Variable	Successive R	Successive S	Betas of the Final Equation	Variable	Successive R	Successive S	Betas of the Final Equation
1	0. 78	13. 9	-1. 169	8	0. 55	18. 8	0. 555	1	0. 78	13. 9	-1. 306
5	0. 85	11. 7	0. 718	5	0. 73	15. 5	0. 718	5	0. 85	11. 7	0. 681
10	0. 89	10. 4	0. 723	10	0. 81	13. 4	0. 722	10	0. 89	10. 4	0. 658
8	0. 98	4. 6	0. 555	1	0. 98	4. 6	-1. 169	3	0. 90	10. 1	-0. 485
3	0. 99	3. 3	-0. 290	6	0. 99	3. 3	0. 174	2	0. 90	10. 1	-0. 108
6	0. 99	2. 8	0. 174	3	0. 99	2. 8	-0. 290	8	0. 99	3. 2	0. 579

Predictive Equations

Cases A and B: $Y_1 = -149.221 - 4.031X_1 + 0.051X_5 + 0.023X_{10} + 4.012X_8 - 40.580X_3 + 0.758X_6$

Case C: $Y_1 = -115.131 - 4.505X_1 + 0.048X_5 + 0.021X_{10} + 68.090X_3 - 0.191X_2 + 4.190X_8$

TABLE 3

SIMULATION RESULTS: 1947 MODEL

Test District	Actual Mass Transit Trips of Each District	Actual Percent of Total Trips Made by Mass Transit	Calculated Mass Transit Trips (Percent of Total Trips): Original Run — Percent (10 variables)	Original Run — Residual (10 variables)	According to Betas or Partial R's — Percent (6 variables)	According to Simple R's — Percent (6 variables)
000	316. 171	86. 00	86. 03	+0. 03	86. 80	86. 45
012	23. 402	63. 00	64. 96	+1. 96	65. 98	66. 24
017	31. 254	73. 00	73. 21	+0. 21	73. 97	71. 86
021	31. 441	45. 00	46. 49	+1. 49	45. 10	43. 55
060	9. 803	44. 00	40. 44	-3. 55	41. 50	42. 86
039	18. 852	41. 00	40. 53	-0. 47	37. 85	37. 43
041	82. 774	70. 00	69. 71	-0. 29	66. 00	66. 20
054	50. 157	68. 00	65. 27	-2. 73	63. 55	64. 31
063	41. 572	65. 00	63. 54	-1. 46	63. 62	61. 10
202	20. 916	41. 00	40. 07	-0. 93	42. 89	41. 31
421	7. 429	26. 00	31. 44	+5. 44	31. 38	32. 26
451	2. 961	18. 00	16. 31	-1. 69	17. 86	19. 26
092	24. 411	69. 00	70. 58	+1. 58	72. 05	72. 46
084	75. 960	80. 00	80. 23	+0. 23	80. 60	83. 70
471	5. 363	14. 00	14. 17	+0. 17	13. 81	13. 97

TABLE 4

ORIGINAL TRANSIT TRIP EQUATIONS: 1960 DATA

Y	Equation	R	S. E.	B_1	B_2
Y_1	$52.45 + 0.002X_6 - 0.8X_1 + 0.04X_{55} - 1.2X_2$	0.81	6%		
Y_{383}	$0.25 - 0.74 \log X_2 - 0.03 \log X_6$	0.84	10%	-0.82	0.02
Y_{383}	$0.09 - 0.54 \log X_2 + 0.08 \log X_{55}$	0.82	10%	-0.62	0.22
Y_{385}	$0.12 - 0.45 \log X_2 + 0.13 \log X_{55}$	0.87	9%	-0.52	0.38
Y_{385}	$0.73 - 0.92 \log X_2 + 0.35 \log X_6$	0.88	9%	-1.04	0.22
Y_{388}	$148.44 - 92.38X_2 + 0.08X_{55}$	0.63	40.65 (trips)	0.60	0.03
Y_{400}	$16.86 + 0.02X_{45}$	0.75	10%		
Y_{401}	$0.22 + 0.0006X_{46}$	0.69	13%		
Y_{403}	$0.20 + 0.29 \log X_{57}$	0.78	13%		
Y_{335}	$0.07 + 0.17 \log X_8$	0.76	11%		
Y_{405}	$0.02 + 0.17 \log X_8$	0.75	11%		
Y_{407}	$0.22 + 0.0002X_{45}$	0.62	16%		
Y_{409}	$0.06 + 0.0006X_{46}$	0.74	11%		
Y_{411}	$0.17 + 0.0009X_{48}$	0.75	17%		
Y_{341}	$0.04 + 0.0005X_{78}$	0.68	5%		
Y_{342}	$0.05 + 0.0009X_8$	0.84	5%		
Y_{343}	$0.05 + 0.0002X_{45}$	0.75	8%		
Y_{344}	$0.01 + 0.0006X_{58}$	0.82	6%		
Y_{345}	$0.07 + 0.11 \log X_{57}$	0.66	6%		
Y_{429}	$0.06 + 0.0004X_{58}$	0.64	6%		
Y_{266}	$-0.14 + 0.22 \log X_{57}$	0.76	9 (trips)		
Y_{285}	$0.24 + 0.002X_{58}$	0.70	27 (trips)		

Y_1 = Percent of total mass transit trips from each district (excl. RR trips).
Y_{383} = Total mass transit trips from home per total person trips from home.
Y_{385} = Home to work mass transit trips per home to work person trips.
Y_{388} = Home to non-work mass transit trips per 100 persons.
Y_{400} = Work to home M.T.T. from Office L.U. per jobs from Office L.U.
Y_{401} = W-H M.T.T. from Retail L.U. per W-H person trips from Retail L.U.
Y_{403} = W-H M.T.T. from Manufacturing L.U. per W-H person trips from Manufacturing L.U.
Y_{335} = W-H M.T.T. from L.U. 3, 4 and 6 per W-H person trips from L.U. 3, 4 and 6.
Y_{405} = Non-work to home M.T.T. per non-work to home person trips.
Y_{407} = NW-H M.T.T. from Office L.U. per NW-H person trips from Office L.U.
Y_{409} = NW-H M.T.T. from Retail L.U. per NW-H person trips from Retail L.U.
Y_{411} = NW-H M.T.T. from Manufacturing L.U. per NW-H person trips from Manufacturing L.U.
Y_{341} = NH-NH mass transit trips per NH-NH person trips.
Y_{342} = NH-NH M.T.T. from L.U. 1-X per NH-NH person trips from L.U. 1-X.
Y_{343} = NH-NH M.T.T. from L.U. 1 per NH-NH person trips from L.U. 1.
Y_{344} = NH-NH M.T.T. from L.U. 2 and 5 per NH-NH person trips from L.U. 2 and 5.
Y_{345} = NH-NH M.T.T. from L.U. 3, 4 and 6 per NH-NH person trips from L.U. 3, 4 and 6.
Y_{429} = NH-NH M.T.T. from L.U. 2 and 5 per jobs from L.U. 2 and 5.
Y_{266} = Total M.T.T. from L.U. 3, 4 and 6 per jobs from L.U. 3, 4 and 6.
Y_{285} = Total M.T.T. from L.U. 2 and 5 per jobs from L.U. 2 and 5.
X_1 = Total cars per total population.
X_{45} = Office jobs per Office acres.
X_{46} = Retail jobs per Retail acres.
X_{57} = Jobs from L.U. 3, 4 and 6 per acres of L.U. 3, 4 and 6.
X_8 = Jobs per non-residential acres.
X_{48} = Manufacturing jobs per Manufacturing acres.
X_{78} = Total auto driver trips per non-residential acres.
X_{58} = Jobs from L.U. 2 and 5 per acres of L.U. 2 and 5.

Within the framework of the trip generation analysis, attempts were made to establish relationships for trip generation rates of each trip purpose from each major type of land use and for each mode of travel. Trips were classified as "total person trips," "auto trips," "auto driver trips," and "total transit trips." The mass transit trip generation rates were formed as trips per household, per person, and trips per job, or as a percent of total person trips made by mass transit in each trip type. This procedure, if successful, would permit the derivation or projection of transit trips in two ways—once directly on the basis of the mass transit trip relationships and again as the residual of the subtraction of auto trips from total person trips in each major trip type.

The trip generation analysis in PJTS has completed both these projection procedures before the actual selection of procedure was made. The effort to develop acceptable and reliable individual relationships directly for each type of transit trip (a total of 35 types of transit trips) did not produce finally a complete set of consistently acceptable and reliable equations.[1] While several types of transit trips produced substantially acceptable relationships, many other types of transit trips resulted in relationships largely unreliable. Making use of conclusions previously reached and the array of available data, the variables used in these attempts were car ownership rate (X_2), median household income (X_6), net residential density (X_{55}), and the various net job densities for each type of job. Table 4 gives the better equations of this attempt. The remaining equations resulted in R's below 0. 60, using either logarithmic or arithmetic forms of the variables.

As a result of the unsatisfactory consistency and reliability of these transit trip equations, the first round of the 1975 transit trip projection was completed almost exclusively by subtracting auto trips from total person trips. This was done for each of the five trip purposes (H-W, H-NW, W-H, NW-H, NH-NH) and for each of the six types of land use aggregations (L. U. 0, 1, 2+5, 3+4+6, 7+8, 9) which were the land use types finally agreed upon to be projected in the 1975 land use plans. For home origin trips the projection of person trips and auto trips was made on the basis of predictive equations which have had generally better predictive capability than the transit trip equations. For non-home origin trips, the projections of person trips were made by use of a combination of mean rates and statistical relationships. Auto non-home origin trips were projected primarily with predictive equations which utilized job density and, in several cases, the proportion of jobs, by type, in each district vs the jobs in the region.

The projection procedure resulted in several imperfections. First, the necessity of using means for the projection of several types of person trips and auto trips from non-home origins precluded the direct introduction in these cases of the influence of rider or area variables such as residential or job density in each district. Second, a detailed investigation of the results of the predictive equations of the home origin trips, on a district basis, revealed that although these equations produced highly satisfactory results on a regional basis (e.g., a simulation error of 19 percent of actual auto trips in 1960), they did overstate significantly auto trips in the central part of the City of Philadelphia where the vast majority of transit trips of the region took place. Further, the same equations were found to generally understate auto trips in the New Jersey districts and overstate auto trips in the Pennsylvania districts. Figures 1 and 2 show this overstatement of auto trips (and consequently understatement of transit trips). Finally, nowhere in the projection process is the effect of the system (and of its potential changes) directly or indirectly incorporated.

The projections were improved by incorporating in an elementary manner, at least, the effects of the 1975 alternative systems. This incorporation was made by an ad hoc, generally upward adjustment of the total number of transit trips. For each district, use was made of a relationship between highway and transit travel time to the Phila-

---

[1]The 35 types of transit trips were three trip purposes from home (H-W, H-NW, Total) and four trip purposes (W-H, NW-H, NH-NH, Total) from each of the following land uses: Residential (0), Offices (1), Retail and Services (2), Passenger Transportation (5), from the combination of (2) and (5), from the combination of Wholesale with Stocks (3), Manufacturing (4), and Goods Transportation (6), from the combination of Public Buildings (7) and Community Services (8), and from Recreation Land Use (9).

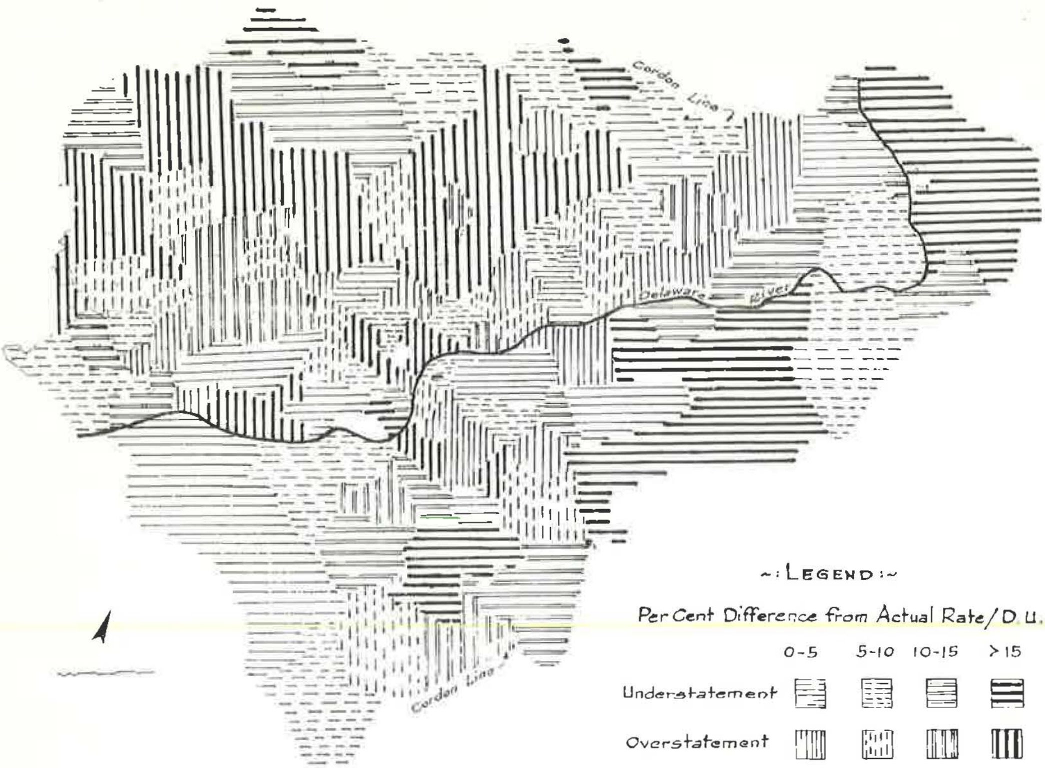


Figure 1. Residuals of auto home to work trip generation equations.

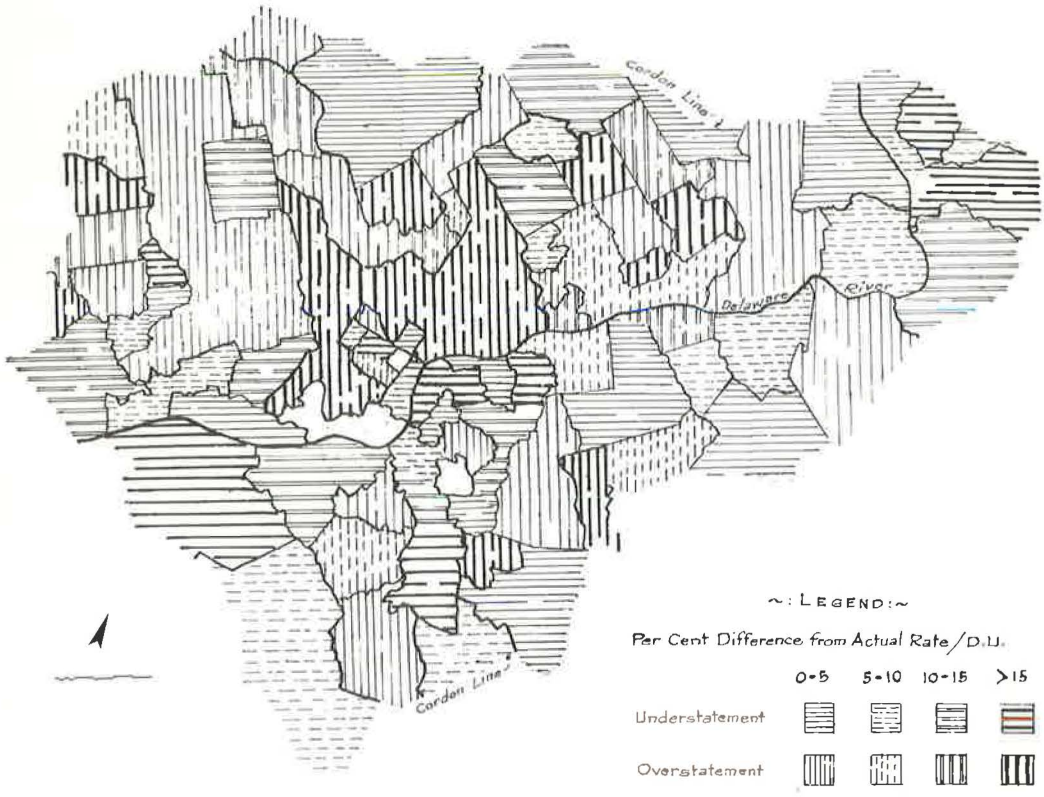


Figure 2. Residuals of auto home to non-work trip generation equations.

delphia CBD and the percent of total transit trips in each district. Changes of this relationship between 1960 and 1975 were used as the basis of an adjustment of total transit trips projected for 1975 in each district. This adjustment served an intermediate purpose and helped to indicate, at least partially, the effect that might be produced by different rates of transit/highway capital investment in the 1975 plans. The results of this adjustment were considered sufficient for the occasion but, at the same time, they helped to emphasize the need for continuation of the modal split analysis and for the completion of a systematic and comprehensive method with which the effects of the car ownership rates, the density of development, and the transportation system of the region are directly and simultaneously incorporated in the analysis and projection process.

THE 1975 TRANSPORTATION PLANS

The need for a direct and as complete as possible incorporation of the effects of the transportation system on the process of determining the selection of the mode of travel becomes better understood when one relates this objective with the extent of future changes in the transportation system of the region. Within the context of PJ circumstances, it may be pointed out that the proposals for the 1975 transportation system for the PJ region include two alternative highway plans and three alternative transit plans. These five plans produce six combinations of systems (2 × 3) including both highway and transit facilities. The minimum highway plan anticipates $1, 269 million capital investment while the minimum transit plan anticipates $163 million capital investment in transit facilities. Correspondingly, the maximum highway and maximum transit plans anticipate $1, 632 million and $718 million capital expenditures each. By adding the highway system cost and the transit system cost of each combination of plans, the total estimate is found varying from a minimum of $1, 432 million to a maximum of $2, 350 million in capital expenditures. Each plan anticipates a technology for both highway and transit basically similar to the present-day modern and operational technology of these systems. In terms of amount of facilities, the highway plans anticipate 226. 4 or 330. 4 miles of new expressways and the transit plans anticipate 7. 6, 12. 2, or 33. 1 miles of subway in three or nine line-extensions plus the conversion of three railroad lines (62. 9 miles) into electrified suburban-commuter rapid lines. The detailed specifications of the 1975 plans are shown in Figures 3, 4, 5, 6, and 7.

SYSTEM VARIATIONS AND TRANSIT TRIPS

A method incorporating the effects of alternative transportation systems into the planning process becomes meaningful exercise only when it can, indeed, trace and effectively take into account any or all of the particular effects of the different systems. One of these effects, and perhaps one of the most significant ones for a transportation study which such a method would be asked to trace and measure, is the effect of alternative transportation systems on the magnitude and the characteristics of the travel demand in the region.

Clearly there should be at least two basic technical concerns in establishing alternative transportation systems. One is the amount of travel demand served by a system and the sufficiency and efficiency of the system in doing so. The second is the additional effect that each system will have in determining the basic characteristics of travel demand. We usually recognize the system effects in the distribution of trips when we use one of the synthetic models (e.g., a gravity or a probability model) which greatly influence the particular interchanges of travel according to the transportation system of the region. We also recognize partially the system's effect in the assignment of trips by assigning trips on these facilities which form the "minimum path" or the "best alternative" path. Frequently, however, it has been proved difficult to incorporate the effects of a transit and highway system in estimating the number of transit trips which the combination of transportation systems facilitate and induce in a region. In the case of PJTS, the acceptance for testing of three alternative transit systems, varying by $600 million, increases in meaning decidedly if the effect of each system could be associated with the number of transit trips induced and served by each system.

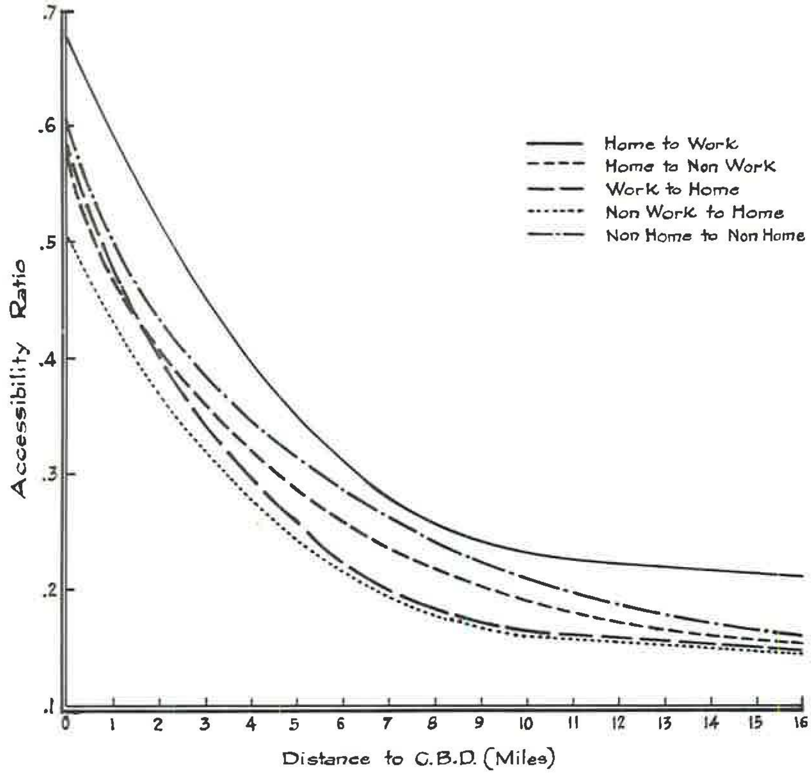


Figure 8. Transit system accessibility, 1960.

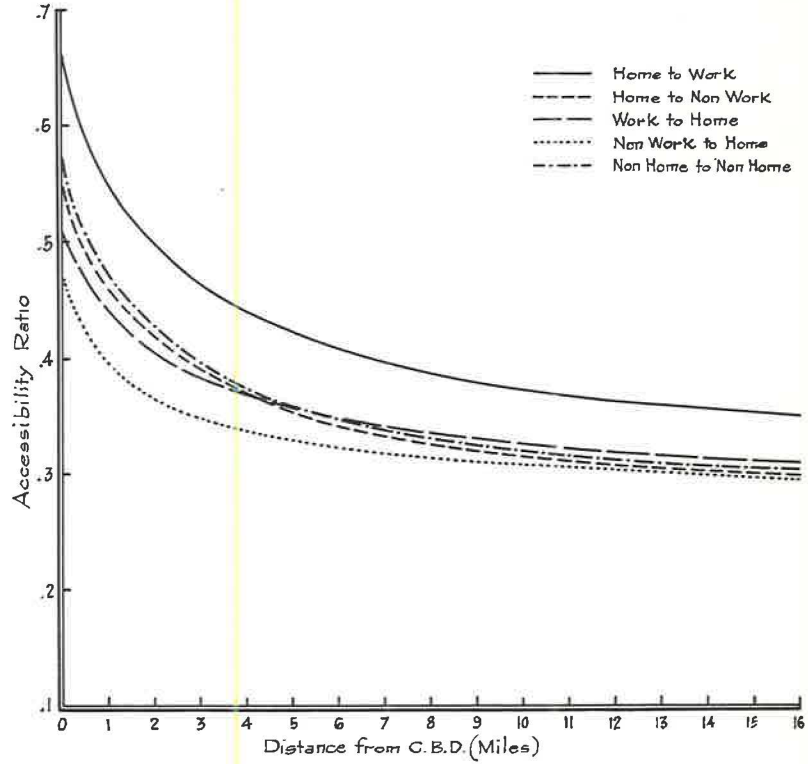


Figure 9. Transit system accessibility, 1975 (System 82—maximum transit, minimum highway).

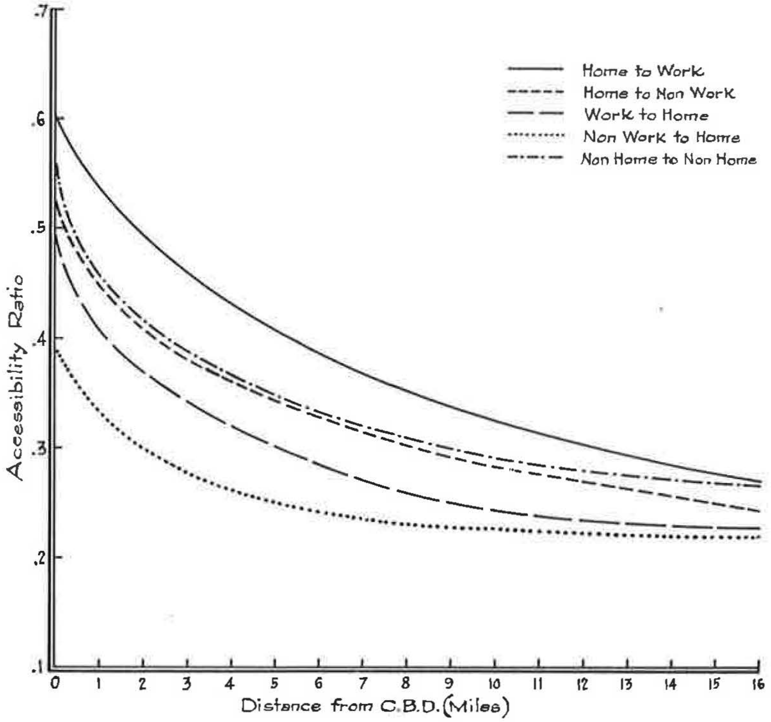


Figure 10. Transit system accessibility, 1975 (System 83-A—mimimum transit, maximum highway).

of person trip destinations reached by transit out of the total person trip destinations reached by highway. This weighting of the mean increases the significance of the close-by person trip destinations by taking them into account repeatedly in each successive cumulative estimation of trip destination by cost interval. This mean ratio is in essence the highway and transit system accessibility for each district. As long as transit travel is in any respect slower or costlier or more restrictive than highway travel this ratio is always below 1. 00, usually varying to about 0. 10. The highest values are found in the districts with the best transit service available, usually the center of the region and the other transportation centers.

Accessibility ratios can be formed for each set of a highway system, a transit system, and a land use distribution. Figures 8 to 10 indicate the values of this variable with the 1960 set of inputs and with the 1975 land use for each alternative pair of projections and two combinations of systems, the minimum highway-maximum transit system (82) and maximum highway-minimum transit system (83-A).

Since 1961, when the present concept of "transit system accessibility" was first publicly proposed, several other professional attempts were made to develop a system accessibility relevant to modal split (6, 7, 8). Most frequently discussed is the use of the gravity model denominator as an index of accessibility. Comparing the PJ concept of system accessibility with the gravity model denominator, one can see the similarities and the differences. For instance, both utilize the land use pattern. The difference is primarily in the simultaneous use of the highway and transit system, in the weight which is placed in the nearby trip destinations in the PJ model and in the manner in which the land use pattern is taken into account.

The PJ modal split analysis focused separately on each of the five trip purposes which were universally analyzed and projected within the Study. In each case the relevant person trip destinations were taken into account, as follows:

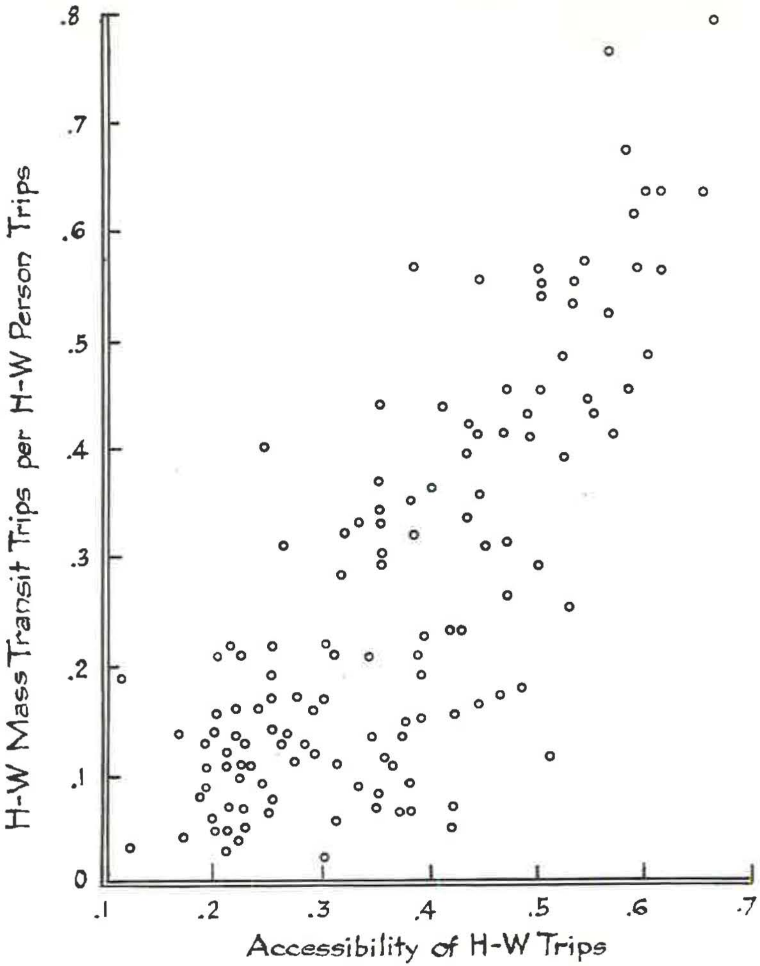


Figure 11. Home to work mass transit trip rates vs accessibility ratios.

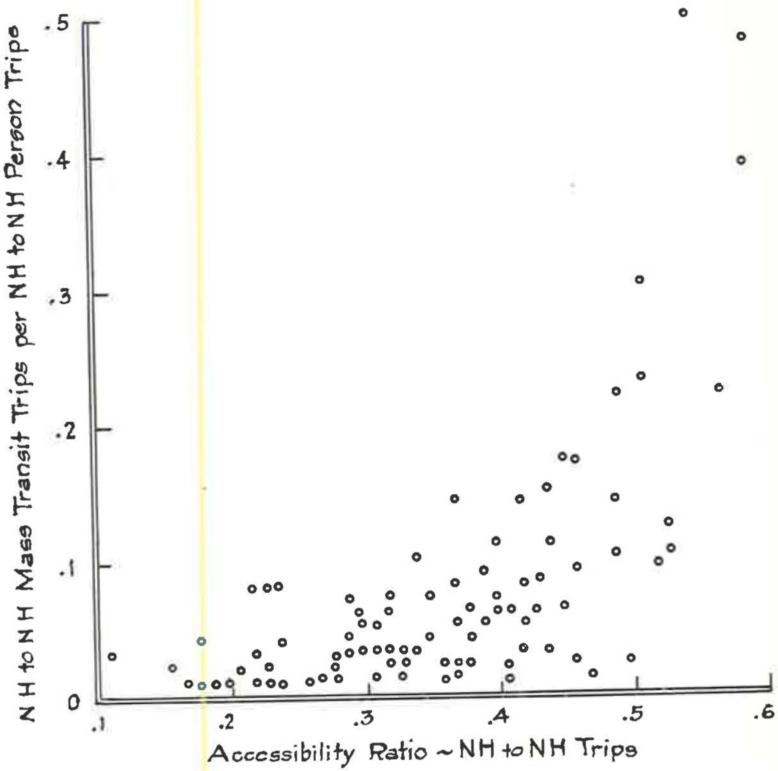


Figure 12. Non-home to non-home mass transit trip rates vs accessibility ratios.

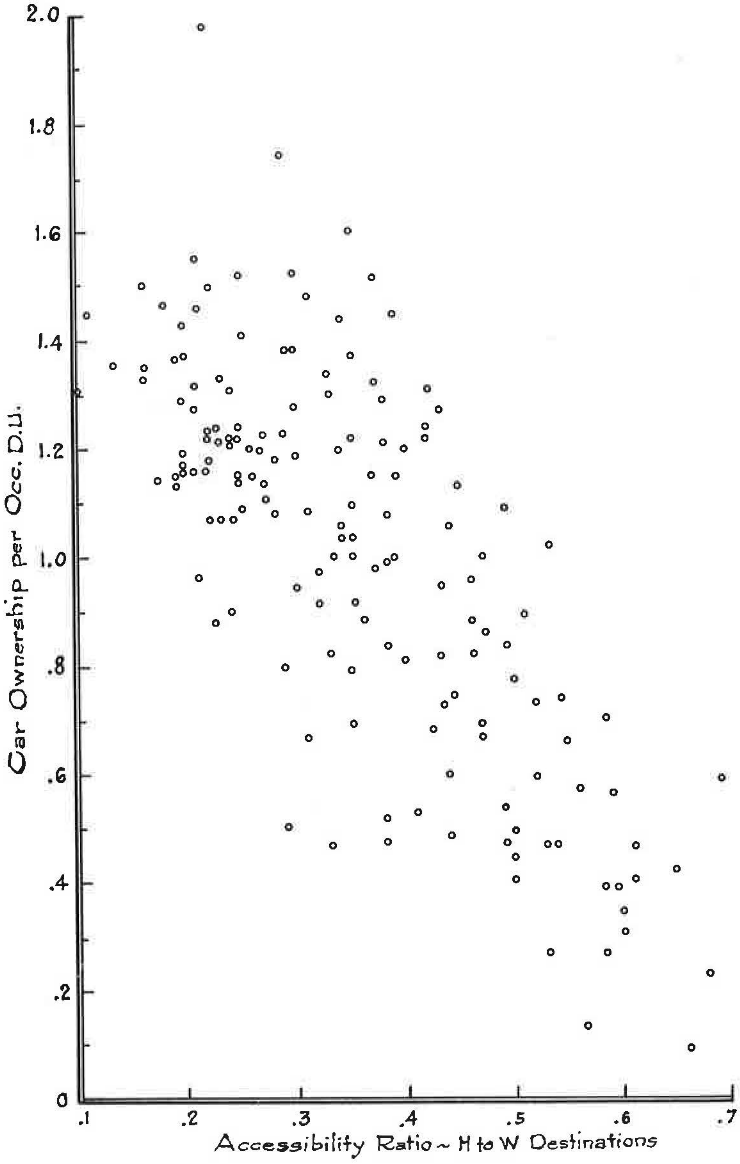


Figure 13. Car ownership rates vs H-W accessibility ratios.

Home to work trips = work to home destinations
Home to non-work trips = non-work to home destinations
Work to home trips = home to work destinations
Non-work to home trips = home to non-work destinations
Non-home to non-home trips = non-home to non-home destinations

Transit system accessibility for each purpose of trips was formed and initially checked in the form of scatter diagrams. Figures 11 and 12 provide a good sample of the relationship between transit system accessibility and percent of transit trip generated in each district. (The percent of transit trips in each district is determined on the basis of all trips generated in the district for a given travel purpose, regardless of the land use of origin.)

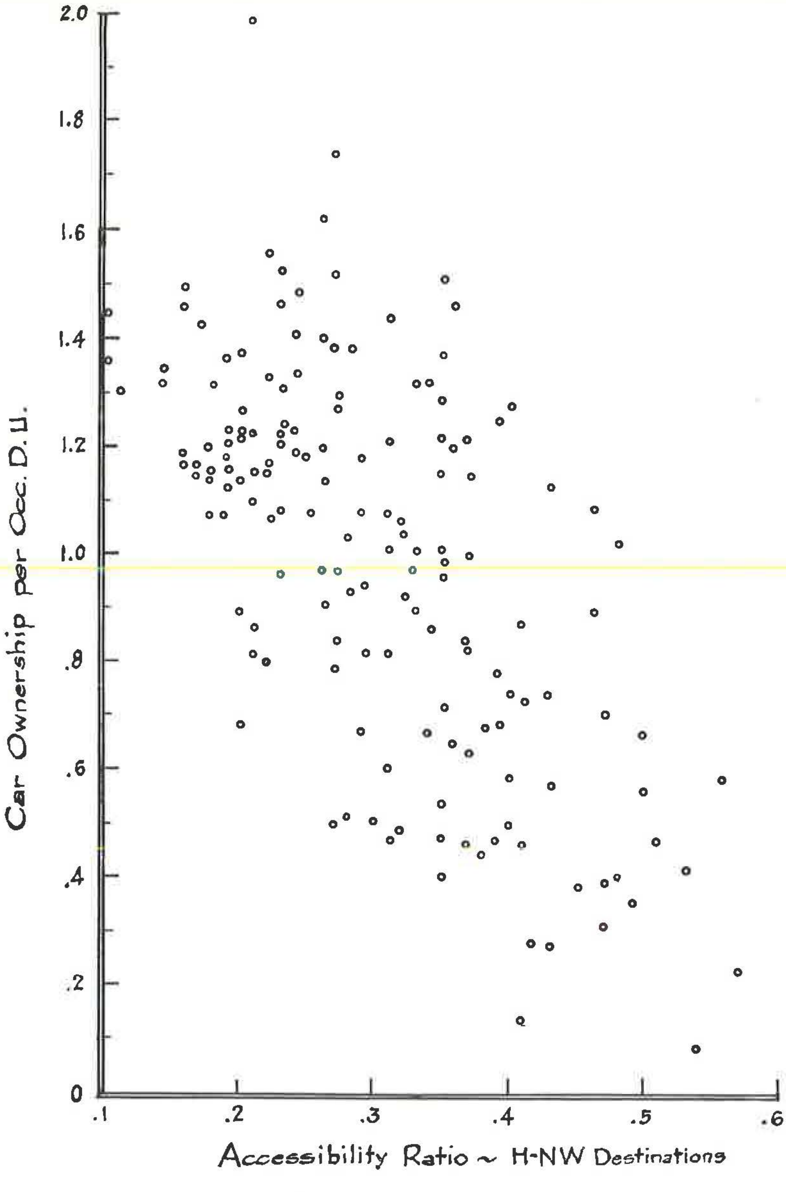


Figure 14. Car ownership rates vs H-NW accessibility ratios.

Another interesting test was an investigation of the relationship between transit system accessibility and mean car ownership rate in the district. Figures 13 and 14 show the results together with a simplified income stratification of each district. The available evidence clearly supports the contention that transit accessibility (or availability) is associated with and to a certain degree influences the rate of car ownership in an area. It was found that the greater the transit accessibility to work and to non-work trip destinations, the smaller the car ownership rate in the district. Correlation analysis between the two variables verified this relationship. The correlation coefficient between car ownership and system accessibility of work trip and non-work trip destinations was correspondingly -0. 65 and -0. 56 which, although not high enough to stand alone, clearly indicates an existing relationship. (Car ownership rates were

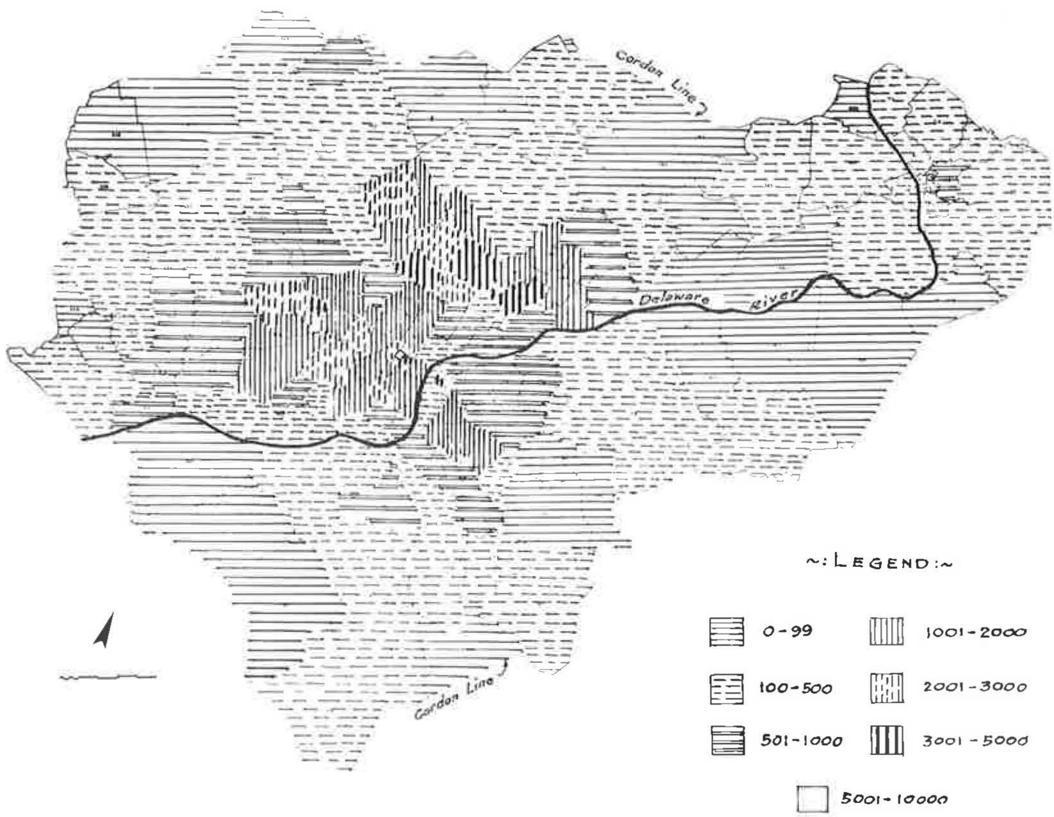


Figure 15. Average daily transit service frequency.

also found to be closely associated with residential density, size of household and, most of all, with median household income. To the extent that the present investigation did not include all these pertinent and possible relationships, the present findings should be considered as conditional and only partially indicative of a direct, causative relationship between transit system accessibility and car ownership rates. Clearly, residential density is also related to transit system accessibility due to the present-day distribution of densities and transit systems in our metropolitan areas.)

A second system variable was also explored to a great extent and was finally incorporated as part of the predictive modal split model. This variable expresses the frequency of transit service available to a district. It is measured in terms of the number of transit vehicle departures occurring in a district within 24 hours, in all transit lines serving the district. Originally a subway train departure was multiplied by the mean number of subway cars in a train. However, this highly overstated the statistical significance of the subway and resulted in substantial overstatement of rates at the simulation tests. Accordingly, a departure in the final form of the model signifies just a departure of a means of transit without indicating whether a bus, a subway train, or a commuter train is involved. Of course, the previously discussed variable of system accessibility already places greater weight on subways and commuter railroads because of the generally higher speeds that these facilities provide. No additional consideration was given (beyond what was warranted statistically) to these systems also because of their various conflicting characteristics. For instance, climbing stairs, limiting the number of stops, concern for public safety, lighting considerations, and fear of missing a particular departure might be considered as balancing out most of the security of finding a seat in a five-car train. What might remain as additional significance can be considered as simply an adjustment of supply to demand without any discernable additional effect of the supply of facilities on the demand for service. Figure 15 shows the distribution of vehicle departures by district in the PJ region.

THE PJ MODEL

The preparatory work based on the 1947 data and on the first two rounds of analysis of the 1960 data was helpful in discovering several consistent, general, and significant relationships between transit trips originating in a district and the various characteristics of the riders and the district. The development of the conceptual framework within which the transportation system serving a district can be considered and the formulation and initial testing of two specific system variables were helpful in making available the basic components with which a more complete modal split model for the PJ region could be formed.

The basic approach in this case can be considered as focusing on two items: (a) that mass transit trips could properly be divided into the five distinctly different trip purposes but need not be divided into trips by type of land use at the origin; and (b) that, where possible, the predictive model should include components which express the rider, the density of development and the transportation system.

The reasoning behind the division of transit trips by purpose is basically a realization that trips made for various purposes are made under different conditions and for various considerations. For instance, work trips have a dominant character which suggests that these trips might have the first choice in the family means of travel. Trips from home to non-work purposes are usually made without strict time schedules and frequently demonstrate preferences for easily obtainable travel means. Trips with no home connections are usually made in the commercial centers of the region and are the shortest trips on record. Therefore, proximity of service might be crucial in selecting the system of travel. Trips from non-work purposes to home are usually made without strict time schedule but by people already in action for some time; therefore, they are made by people who might appreciate proximity of service above other features (such as speed of travel). Finally, trips from the three major central business districts of the region were considered to be of sufficiently particular nature to warrant special but systematic treatment.

In view of these considerations, it was decided that the model should be formed in such a manner that it would be able to depict the individual influences on travel mode selection. Thus the following types of trips and relationships were selected in simulating and projecting the transit trips in the PJ region.

1. Home to Work Trips: to be related with car ownership rates (indicating car availability), with residential density (indicating proximity of transit service), with system accessibility (indicating travel speed and cost difference), and with frequency of transit service.
2. Home to Non-Work Trips: to be related with car ownership rates, residential density, system accessibility, and frequency of transit service.
3. Work to Home Trips: (a) from the Philadelphia, Camden, and Trenton Central Business Districts; (b) from the rest of the region. In both cases trips to be related with total job density (indicating proximity of transit service), system accessibility, and frequency of transit service.
4. Non-Work to Home Trips: (a) from the Philadelphia, Camden, and Trenton Central Business Districts; (b) from the rest of the region. In both cases trips to be related with total job density, system accessibility, and frequency of transit service.
5. Non-Home to Non-Home Trips: These trips are primarily from non-residential areas and especially from the central business districts; to be related with total job density, system accessibility and frequency of transit service.

Before any selection of mathematical relationships could be made, the dependent and independent variables were related in an arithmetic, logarithmic and non-parametric manner. The final relationships were selected on the basis of their overall as well

TABLE 5

MODAL SPLIT MODEL FOR THE DVRPC REGION
(Generation of Transit Trip Origins in Each District Using Transit and Highway System Variables)

Trip Purpose	Equation	Area	R	S	F	1960 Simulation Error	Std. Dev. Y	Mean $\overline{Y}$
H-W	Y_{385} = Percent of person trips made by transit = $0.1626 - 0.4107 \log X_2 + 0.0709 \log X_{55} + 0.0702 \log X_{43} + 0.0001 X_{54}$	Cordon Area	0.902	0.08	149.20	-0.82	0.18	0.26
H-NW	$\log Y_{387} = -2.1544 - 0.7255 \log X_2 + 0.3314 \log X_{55} + 0.2142 \log X_{56} + 0.3058 \log X_{54}$	Cordon Area	0.839	0.20	76.23	-7.19	0.35	0.08
W-H	$\log Y_{397} = -1.6234 + 0.1553 \log X_8 + 2.0913 X_{19}$	CBD Phila., Camden & Trenton	0.914	0.13	28.01	-0.81	0.21	0.43
	$\log Y_{397} = -2.0513 + 0.2285 \log X_8 + 0.5700 X_{19} + 0.3104 \log X_{54}$	Cordon Area	0.864	0.16	109.47	+4.17	0.25	0.19
	$\log Y_{405} = -2.0556 + 0.6244 \log X_8 + 0.6143 X_{24}$	CBD Phila., Camden & Trenton	0.909	0.17	26.18	+8.11	0.25	0.32
NW-H	$\log Y_{45} = -2.3800 + 0.4332 \log X_8 + 0.0371 X_{24} + 0.3198 \log X_{54}$	Cordon Area	0.807	0.21	68.25	+3.50	0.35	0.07
NH-NH	$Y_{342} = -0.0257 + 0.0007 X_8 + 0.1971 X_{67}$	Cordon Area	0.894	0.04	200.17	+4.82	0.09	0.07

X_2 = Total cars per total Occ. D.U.
X_8 = Jobs per non-residential acre.
X_{19} = Accessibility ratio for W-H trips.
X_{24} = Accessibility ratio for NW-H trips.
X_{43} = Accessibility ratio for H-W trips.
X_{54} = Average transit vehicles per day.
X_{55} = Perm. Occ. D.U. per residential acre.
X_{56} = Accessibility ratio for H-NW trips.
X_{67} = Accessibility ratio for NY-NW trips.

TABLE 6

STATISTICAL RESULTS OF VARIABLE TESTING
(Forming the Model Presented in Table 5)

Trip Purpose	Area	Number of Variables	"t" Values				Beta Values			
			Car Ownership	Density	Accessi-bility	Frequency of Service	Car Ownership	Density	Accessi-bility	Frequency of Service
H-W	Cordon Area	Four	6.82	2.59	1.15	6.00	-0.48	0.21	0.07	0.25
H-NW	Cordon Area	Four Re-run	2.94	3.77	1.06	6.15	-0.23	0.35	0.06	0.36
W-H	CBD Phila., Camden & Trenton	Three	—	1.15	2.77	0.11[a]	—	0.25	0.68	0.02[a]
		Two	—	1.26	3.46	—	—	0.25	0.70	—
	Cordon Area	Three	—	5.80	2.22	5.57	—	0.41	0.16	0.39
NW-H	CBD Phila., Camden & Trenton	Three	—	3.23	0.40	0.36[a]	—	0.83	0.06	0.07[a]
		Two	—	3.45	0.53	—	—	0.80	0.12	—
	Cordon Area	Three Re-run	—	8.10	1.15	3.51	—	0.62	0.07	0.24
NH-NH	Cordon Area	Three	—	7.40	4.05	0.07[a]	—	0.69	0.29	-0.01[a]
		Two	—	7.00	4.19	—	—	0.73	0.24	—

[a]Excluded from the final predictive equation.

as particular simulative and projective capability. Table 5 presents the model in its complete form, together with the statistical yardsticks which indicate its level of reliability. Table 6 gives the values of "t" test of the coefficients and the beta weights of the variables.

Several significant observations can be made on the basis of the findings of the model as given in Tables 5 and 6. First, it can be seen that the car ownership variable varies in significance between home to work and home to non-work trips. In the first instance car ownership appears to be the most significant variable, twice as important as density and frequency of service, and seven times as important as transit system accessibility. In the second instance the car ownership variable appears less important than residential density and even less important than frequency of transit service. Transit system accessibility is again the least important variable for this type of trip.

Second, it can be observed that the residential density or the density of total jobs have a varying degree of significance for the various travel purposes. For instance, one can notice that job density is most closely associated with NW-H trips from the three CBDs of the region, and next with NH-NH trips from the entire region. It appears that proximity to transit lines (which is the essential connotation of the density of development variable) is the primary factor in transit trip production in these two cases. Job density is next found to be closely associated with NW-H trips from the rest of the region and with W-H trips from the non-CBD part of the region. Surprisingly, the model reveals that job density (proximity of transit lines) is least significant for the work to home trips from the three CBDs for which trips transit system accessibility (or the implied speed of transit lines) appears the first and foremost factor. With regard to residential density, it appears that the variable is more important for home to non-work trips than for home to work trips, indicating once more that proximity of service or convenience of using the transit facility is much more important for non-work trips than work trips.

Third, the significance of system accessibility is shown to be different for each type of trip, in absolute and relative measures. For H-W and H-NW trips the absolute contribution of the system accessibility is rather small. Clearly the speed of the transit system has a very small association with the number of H-W or H-NW trips made by transit. For W-H from the three CBD trips, the situation is reversed. It appears that people going home from work and from a CBD area place special emphasis on a fast transit system. This is shown clearly by the beta values and the "t" test of the coefficients. For W-H trips from non-CBD origins, however, the significance of the system accessibility variable is much smaller than either the job density or the frequency of transit service variable. For NW-H trips from the three CBDs of the region or for the same trips originating from the rest of the region, the system accessibility variable is found also to be little associated with the rate of transit trips. Finally, for NH-NH trips, the system accessibility (or speed and cost of transit system) emerges also as a relatively important factor as evidenced by both the pertinent statistical yardsticks.

Fourth, it can be seen that the significance of the frequency of transit service variable is most important for W-H trips from the non-CBD part of the region. For this type of trip the frequency of transit service appears as significant as the job density variable. Next in significance is the contribution of frequency of transit service for the H-NW trips from the region for which frequency of service is at least as important as the residential density variable and of considerably greater importance than the other two variables. Third in line comes the contribution of the frequency of transit service for NW-H and then for H-W trips from the entire region.

It is of interest to notice in the examination of the contribution of each variable that the significance of each variable shifts from case to case and varies both absolutely and relatively. This type of situation suggests that aggregation of trip types can obviously cause significant disparities in simulation and unsatisfactory projections. One might speculate that inappropriate aggregations of trip types in the past could have been, on occasion, the root of controversies in this subject and could have contributed to the unsatisfactory performance of several modal split models developed in the past.

Table 6 also gives the results of the "t" test of the coefficients of each variable for each trip type. For H-W and H-NW trips three variables are above the 1 percent level of significance while the accessibility is significant only at the 25 percent level. For W-H trips from the cordon area, all three variables are found to be above the 1 percent level of significance. For W-H trips from the three CBDs, the frequency of service variable is found to be statistically below any acceptable level of significance and therefore is eliminated. For NW-H trips from the three CBDs the same variable is also found to be statistically below acceptable levels of significance and therefore is also eliminated. For this type of trip the accessibility variable is found to be at a very low level of significance but it has been retained as a variable because it is three times as significant as the frequency of service variable and because of the desire to avoid simple correlations with only one variable. For the NW-H trips from the cordon area, the two variables are above the 1 percent level of significance while the accessibility variable reaches only the 25 percent level of significance. Finally, for the NH-NH trips the frequency of service variable is found to be statistically insignificant (and therefore is eliminated) while the other two variables are above the 1 percent level of significance.

In a more direct form, the model indicates clearly that for H-W trips the car ownership rate, the frequency of transit service, and the density variables play the primary role in that order. For H-NW trips the emphasis is shifted to density and frequency of transit service as primary variables, followed by the car ownership rate variable. For W-H and NW-H trips from the entire region, of primary importance is the job density and frequency of transit service (especially for W-H trips). For NH-NH trips, of primary importance by far is job density followed by system accessibility. For this type of trip as well as for W-H and NW-H trips starting from the three CBDs of the region, the frequency of transit service plays an extremely minute role, if any. Perhaps it should be repeated here that for W-H trips from the three CBDs, the most important variable is by far the system accessibility (indicating, perhaps, the desire of workers to get home as fast as possible) while for NW-H trips from the three CBDs of the region, the most important variable proves to be the job density by a large margin (indicating, perhaps, the desire of close-by, convenient transit lines for this type of trip).

In conclusion, according to PJ findings, transit H-W trips depend primarily on car availability and then on frequency of service; transit H-NW trips depend primarily on proximity of transit line and then on frequency of transit service; transit W-H trips depend primarily on proximity of service and then on frequency of service when these trips originate from non-CBD parts of the region but they depend primarily on travel speed and cost when the same trips originate in the three CBDs of the region; transit NW-H trips depend primarily on proximity of service and then on frequency of service when these trips originate in non-CBD parts of the region but they depend almost exclusively on proximity of service when they originate in the three CBDs of the region; finally, transit NW-NH trips depend primarily on proximity of service and then on travel speed and cost.

SIMULATION OF 1960 TRAVEL PATTERN

The tests of the quality of the model took several forms. The first and most generalized measures of accuracy in the simulation process were the generally used statistical yardsticks of correlation analysis. As seen in Table 5, the correlation coefficients of the equations varied from 0.807 for non-work to home trips to 0.914 for work to home trips from the three CBDs of the region. The standard error of estimate of the equations varied from 4.3 percent to 21 percent of the mean values of NH-NH and NW-H trips, respectively. The F statistic is also significant in all cases and the values of betas and of the "t" test verify the contribution and the significance of each variable taken into account. Obviously, statistical stability of the coefficients of the equations may also be expected for projection purposes.

Additional tests of accuracy of simulation were carried out by purpose of travel, sector of trip origin and total trips in the region. In terms of trips by purpose (Table 7),

TABLE 7

COMPARISON OF SIMULATION RESULTS OF 1960 O-D TRANSIT TRIPS
(By Purpose of Travel)

Category	H-W		H-NW		W-H		NW-H		NH-NH		Total	
	Trips	Percent Diff.	Trips	Percent Diff.	Trips	Percent Diff.	Trips	Percent Diff.	Trips	Percent Diff.	Trips	Percent Diff.
Actual 1960	336,278	—	265,489	—	323,252	—	261,582	—	86,821	—	1,273,422	—
Original simulation	277,709	-17.42 (-7.5)	246,462	-7.17 (-1.0)	315,372	-2.44 (-1.1)	205,730	-21.35 (-2.9)	99,086	+14.13 (+0.8)	1,144,359	-10.14 (-1.9)
Model simulation with accessibility	326,199	-3.00	234,078	-11.83	359,947	+11.35	216,314	-17.31	100,763	+16.06	1,237,301	-2.84
Model simulation with accessibility plus frequency of service[a]	333,532	-0.82	246,391	-7.19	336,726	+4.17	270,733	+3.50	100,763	+16.06	1,288,145	+1.16

[a]The numbers reported in this line represent the simulation of all the transit trips in the region, by individual purpose. The specific simulation errors of the equation corresponding to only the districts included in the equations vary slightly from the above and are correspondingly -0.82, -7.19, +4.17, +3.50 and +16.06.

Note: Percent estimates within parentheses represent the simulation error of the trips expressed on the basis of auto trips in each trip purpose.

the results indicate that for total home to work trips the simulation produces less than 1 percent overall error. Non-work to home trips from the region present the greatest difficulty in simulation with an overall simulation error of the districts forming the equation of -11. 25 percent. However, when the simulation is expanded to cover the entire region, the simulation error is reduced to only +3. 50 percent. The reverse is observed with regard to the NH-NH trips, for which the districts of the equation produce only 4. 82 percent simulation error, but when the equation is expanded to cover the whole region, the simulation error rises to 16 percent of actual total trips.

In order to ascertain the degree of simulation of the various parts of the region, the total area was divided into 12 sectors (Fig. 16). Six of these sectors include the Philadelphia area, three sectors include the Pennsylvania suburbs and three sectors include the New Jersey areas. The results are shown in Table 8. One can notice that individual sectors frequently have simulation errors well above the overall simulation error of the five trip purposes combined. The difficulty in achieving higher accuracy by each individual sector is manifest in all modal split models attempted in this as well as in other studies. If uniformity, consistency, and theoretical foundation is to be retained in a modal split model, individual differences by areas are bound to be greater than the overall simulation discrepancies in any trip purpose. In our case the differences frequently counterbalance each other by purpose of trips and by sector in the same vicinity. In most cases they are also well within acceptable design limits in terms of volume of trips or percent error, or both. The remaining few rather large differences are found in sectors with small volumes in 1960 and low sample accuracy. Finally, in terms of the overall simulation error for all trip purposes and the entire region, the discrepancy is found to be less than 2 percent of the actual number of trips.

Tables 7 and 8 also give the results of the two previous simulation efforts of mass transit trips. The original simulation presents the results of the transit trips as

Figure 16. Twelve sectors for modal split tests.

TABLE 8

COMPARISON OF SIMULATION RESULTS OF 1960 O-D TRANSIT TRIPS
(On Basis of 12 Sectors)

Sector	1960 Actual Trips	Originally Simulated		Model Simulation With Accessibility		Model Simulation With Accessibility and Frequency	
		Trips	Percent Diff.	Trips	Percent Diff.	Trips	Percent Diff.
1	229,689	207,342	-9.73	227,280	-1.05	232,412	+1.19
2	221,076	185,688	-16.01	193,062	-12.67	212,629	-3.82
3	98,820	91,843	-7.06	90,139	-8.78	100,689	+1.89
4	180,847	129,430	-28.43	160,296	-11.36	179,867	-0.54
5	165,159	114,993	-32.02	126,193	-25.40	148,210	-12.38
6	111,466	88,602	-20.51	92,486	-17.03	98,030	-12.05
7	9,059	21,429	+136.55	14,523	+60.32	10,384	+14.63
8	38,425	48,883	+27.22	49,374	+28.49	44,393	+15.53
9	104,010	102,142	-1.80	123,785	+19.01	120,953	+16.29
10	24,834	53,661	+116.08	60,305	+142.83	41,516	+67.17
11	71,510	82,223	+14.98	81,933	+14.58	81,185	+13.53
12	14,527	18,123	+24.75	17,925	+23.39	17,877	+23.06
Total	1,273,422	1,144,359	-10.14	1,237,301	-2.84	1,288,145	+1.16

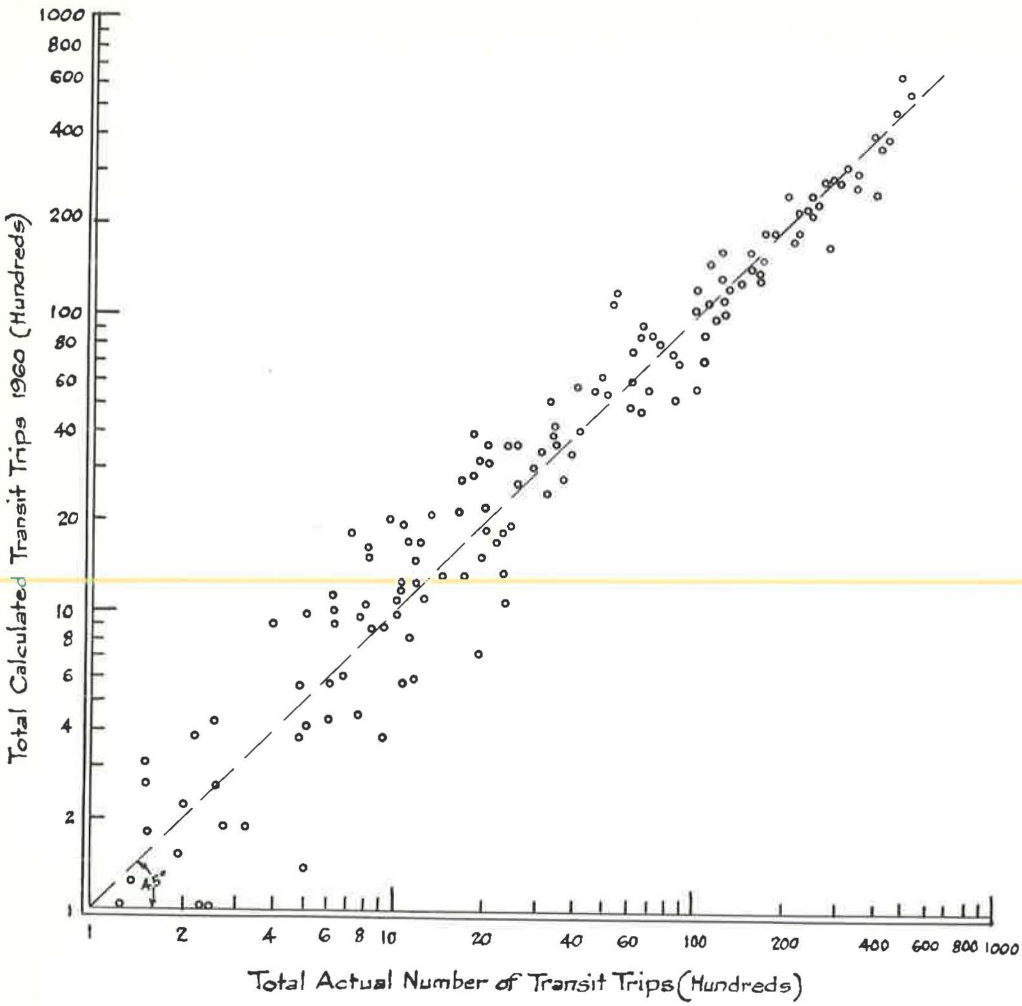
produced by subtracting auto trips from total person trips. The second effort presents the results of the modal split model without the use of the frequency of service variable. The comparison points out the improvements of the completed model in terms of total trips, trips by purpose, and trips in each of the twelve test sectors. The final results by sector, for all trip purposes combined, indicate that the model reaches an acceptable level of simulation even though it is not capable of producing results always below the 5 percent level of accuracy which is usually the acceptable error in simulation of other parts of the travel demand analysis. Trips by each individual district are frequently, of course, found to carry much greater simulation error. However, even in these cases, when trips from all trip purposes are taken together, the total simulation error decreases in most cases to very reasonable levels. Figure 17 shows the satisfactory degree of the simulation discrepancies on the district level.

ALTERNATIVE TRANSIT TRIP PROJECTIONS FOR 1975

The satisfactory simulation of the 1960 travel pattern on the basis of a set of general principles and consistent mathematical relationships made possible the projection of the 1975 transit travel pattern in the PJTS area on the same basis.

The projection process usually starts by projecting the independent variables which enter in the projection model. In our case the independent variables were five: car ownership rate for each district in the region, to be used with H-W and H-NW trips; residential density in each district, to be used with H-W and H-NW transit trips; system accessibility, to be individually projected for each district and for each transit trip purpose and combination of highway and transit system; transit service frequency, to be projected for each district and for each of the three transit alternative plans and to be uniformly used with all five transit trip types; and total job density for each district, to be used with W-H, NW-H and NH-NH trips.

Car ownership rates for each district in the region were projected on the basis of the relationships developed for 1960 and using a projection of average family income, residential net density and size of households. Although at a later stage of analysis a relationship was found between the rate of cars per household and transit system accessibility of a district, no such relationship was put to use in the projection of car



Figrue 17. Total calculated vs total actual transit trips on district-to-district basis.

ownership rates for the 1975 plans. Also, due to the rather short period of time between 1960 and 1975, the problem of car ownership saturation rates did not emerge to any significant extent.

Residential net density in each district was projected on the basis of trend extension and consideration of the growth patterns evidenced in the various parts of the region. In general, a slight reduction of density in the densely developed residential parts of the region and a moderate increase of residential net density in the largely undeveloped parts of the region were frequently projected.

The system accessibility for each of the five travel purposes was estimated using 1975 person trip destinations (by trip purpose) in each district and a pair of the proposed highway and transit systems. The estimations were completed for three combinations of system—the minimum highway and maximum transit systems (system 82), and the maximum highway and the minimum transit systems (system 83-A). Figures 18 to 22 present the overall trends of system accessibility for each trip purpose and according to a classification of the districts on the basis of their distance from the Philadelphia CBD. It is of special interest to notice that for all five trip purposes the system accessibility results essentially in substantial gains of accessibility for those

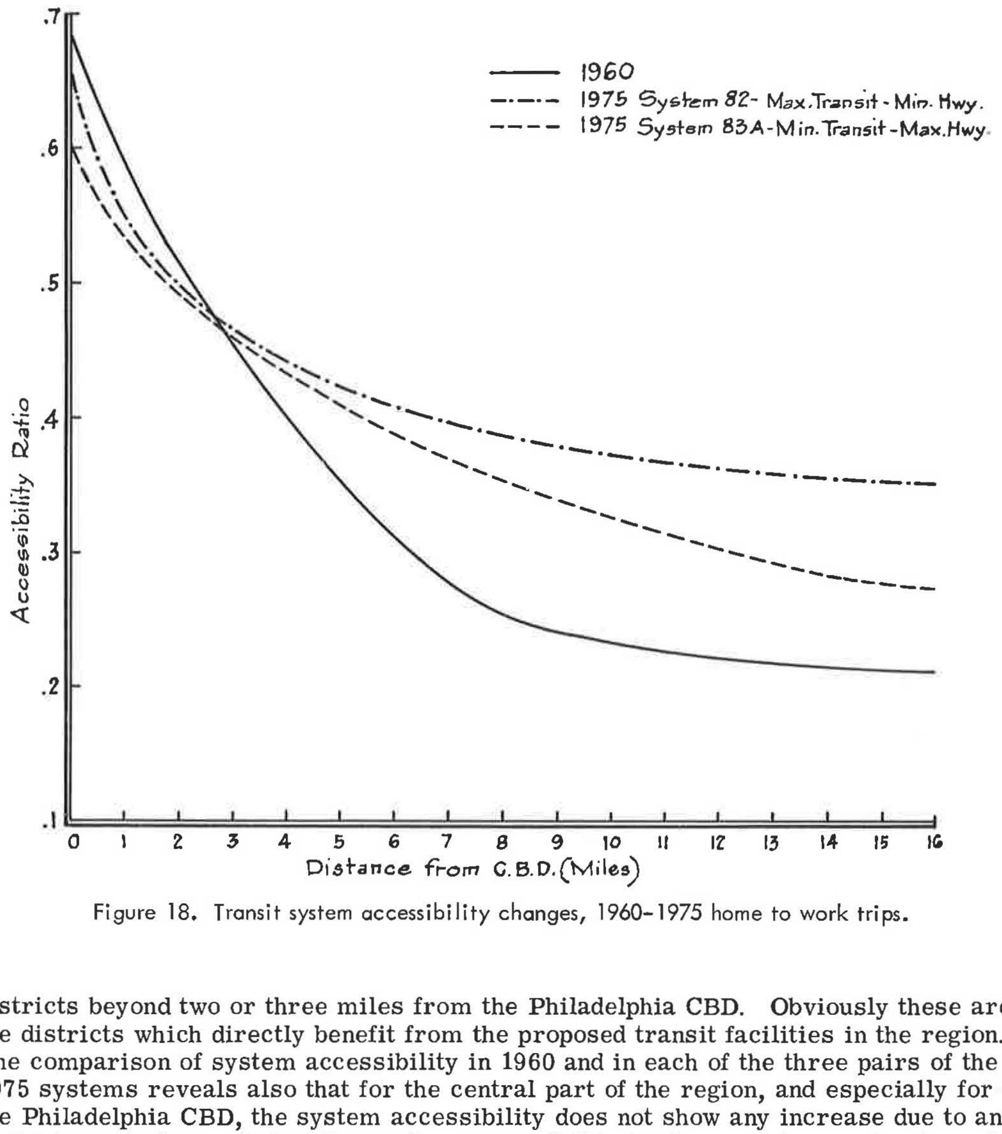


Figure 18. Transit system accessibility changes, 1960-1975 home to work trips.

districts beyond two or three miles from the Philadelphia CBD. Obviously these are the districts which directly benefit from the proposed transit facilities in the region. The comparison of system accessibility in 1960 and in each of the three pairs of the 1975 systems reveals also that for the central part of the region, and especially for the Philadelphia CBD, the system accessibility does not show any increase due to any of the transit plans. In fact the system accessibility is shown as decreasing from 3 to 10 percent for the various travel purposes.

The projection of the frequency of transit service followed a uniform approach by accepting a policy determination that in 1975 the maximum transit plan will include a 25 percent increase in the frequency of service, uniformly experienced in all major lines of the region. For the minimum plan the present-day frequency of transit service was accepted in each line. New lines will have a frequency of service comparable to similar lines of today. For special districts, in which particular developments were expected, the frequency of service was projected accordingly.

Total job density for 1975 was accepted to be essentially similar to that prevailing in 1960. Special districts, where particular policy considerations or developmental plans were in existence or expected, were adjusted accordingly. For the rest of the region the projected effective job density was accepted as primarily determined by the dominating density of present development in each district.

On the basis of these projected variables, the 1975 alternative transit trip projections for the entire region were carried out for two pairs of systems, the system with

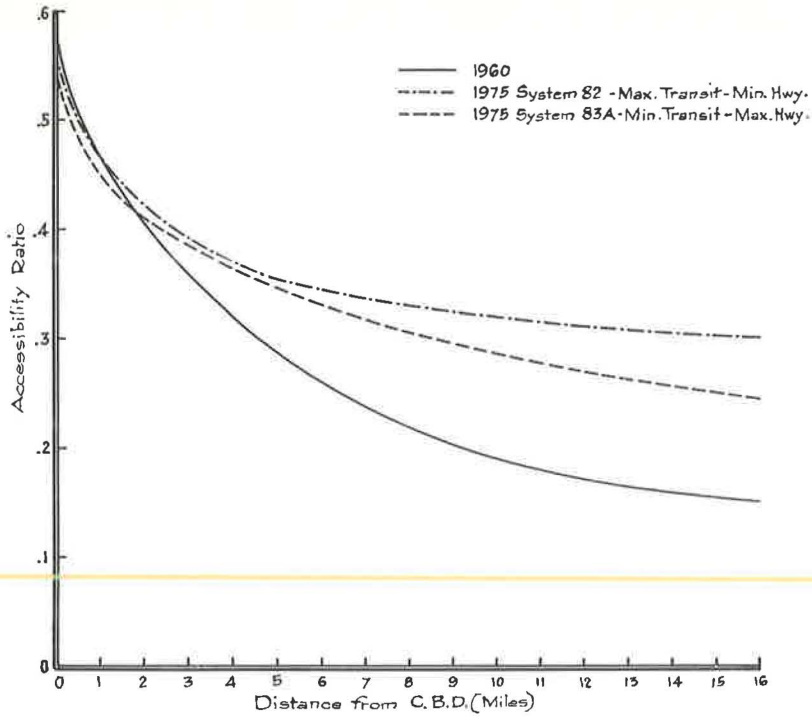


Figure 19. Transit system accessibility changes, 1960–1975 home to non-work trips.

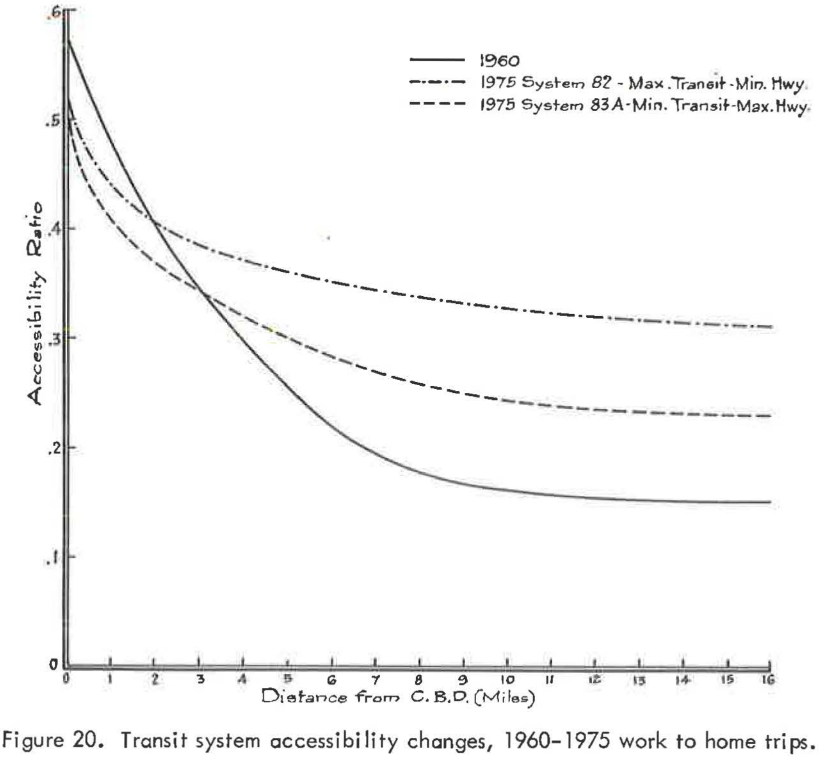


Figure 20. Transit system accessibility changes, 1960–1975 work to home trips.

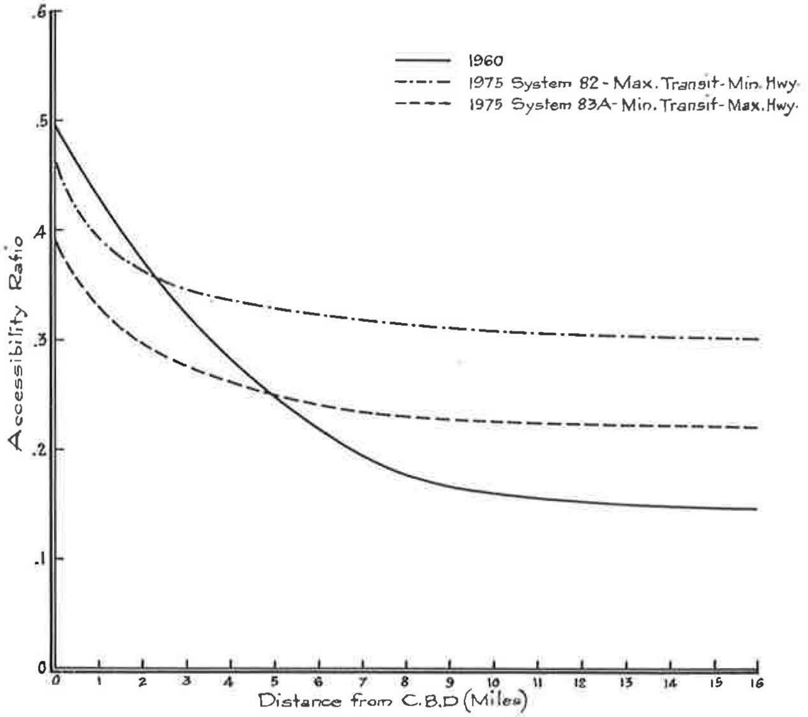


Figure 21. Transit system accessibility changes, 1960-1975 non-work to home trips.

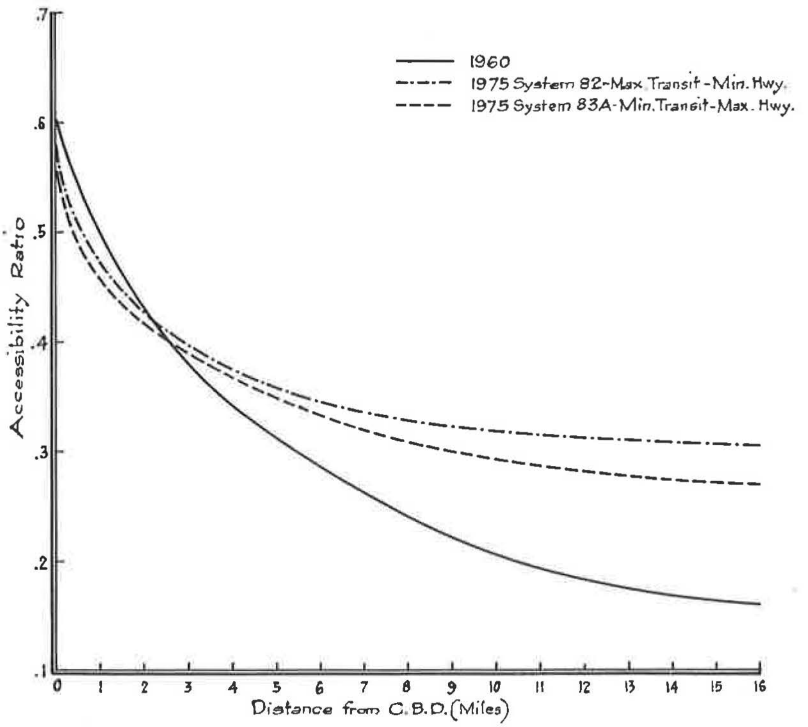


Figure 22. Transit system accessibility changes, 1960-1975 non-home to non-home trips.

TABLE 9

ALTERNATIVE TRANSIT TRIP PROJECTIONS, 1975
(By Trip Purpose)

System	H-W		N-NW		W-H		NW-H		NH-NH[a]		Total	
	Trips	Percent Change From 1960	Trips	Percent Change From 1960	Trips	Percent Change From 1960	Trips	Percent Change From 1960	Trips	Percent Change From 1960	Trips	Percent Change From 1960
Actual 1960	336, 278	—	265, 489	—	323, 252	—	261, 582	—	86, 821	—	1, 273, 422	—
Original 1975 projection	360, 925	+7. 33	197, 811	-25. 49	360, 737	+11. 60	252, 962	-3. 30	121, 928	+40. 44	1, 294, 363	+1. 64
Model projection (system 83-A)	339, 018	+0. 81	322, 852	+21. 65	342, 378	+5. 92	330, 750	+26. 44	162, 486	+87. 15	1, 497, 484	+17. 60
Model projection (system 82)	357, 349	+6. 27	328, 447	+23. 71	363, 016	+12. 30	339, 225	+29. 68	175, 996	+102. 71	1, 564, 033	+22. 82

[a]Excluding L.U. O for NH-NH.

TABLE 10

ALTERNATIVE TOTAL TRANSIT TRIPS PROJECTIONS, 1975
(By Sector)

Sector	1960[a] Trips	Original 1975 Projection		Model Projection			
				System 82		System 83-A	
		Trips	Percent Growth 1960-1975	Trips	Percent Growth 1960-1975	Trips	Percent Growth 1960-1975
1	229,689	272,974	+18.85	249,450	+8.60	232,915	+1.40
2	221,076	164,805	-25.45	254,791	+15.25	245,602	+11.09
3	98,820	75,936	-23.16	117,033	+18.43	115,492	+16.87
4	180,847	140,181	-22.49	214,656	+18.69	209,098	+15.62
5	169,159	120,315	-28.87	183,296	+8.36	177,686	+5.04
6	111,466	104,611	-6.15	128,436	+15.22	122,220	+9.65
7	9,059	20,713	+128.65	21,889	+141.63	21,495	+137.28
8	38,425	73,012	+90.01	59,354	+54.47	57,329	+49.20
9	104,010	135,305	+30.09	149,147	+43.40	143,790	+38.25
10	24,834	61,933	+149.39	40,732	+64.02	40,069	+61.35
11	71,510	97,623	+36.52	116,414	+62.79	107,518	+50.35
12	14,527	26,955	+85.55	28,835	+98.49	24,270	+67.07
Total	1,273,422	1,294,363	+1.64	1,564,033	+22.82	1,497,484	+17.60

[a]Excluding L.U. 0 for NH-NH.

TABLE 11

PROJECTION OF TOTAL TRANSIT TRIPS, 5 PURPOSES
(By Area of Dominant Transit Line Influence)

Area Served by Subsystem	1960 Total Transit Trips	Total Transit Trips, 1975		
		Original Projection	Model Projection for System 82	Model Projection for System 83-A
Philadelphia CBD	229,689	272,974	249,450	232,915
Camden CBD	19,632	17,462	24,717	22,691
Trenton CBD	6,867	13,670	8,204	8,132
Area of Market Street Subway[a]	374,499	393,369	431,517	406,948
Area of Frankford Extension[a]	415,135	434,210	466,652	440,343
Area of Jenkintown Extension[a]	458,841	430,098	497,126	472,561
Area of Northeast Max. Plan	466,723	443,445	528,420	505,434
Area of Northeast Min. Plan[a]	519,631	484,801	586,528	560,419
Area of Broad Street Subway[a]	444,096	420,234	498,717	477,528
Area of Kirkwood Line[a]	51,038	62,479	74,893	70,556
Area of Woodbury Line[a]	45,741	51,741	63,295	56,135
Area of Moorestown Line[a]	34,638	39,629	49,468	43,768

[a]Several districts are taken into account for more than one line.

minimum highway and maximum transit investment and the system with maximum highway and minimum transit investment (systems 83-A and 82).

The results of these projections are given in Tables 9, 10, and 11. In all three tables a comparison is made among the 1960 actual trip estimates, the projections carried out with the original and simplified modal split procedure, and the two 1975 trip projections using systems 82 and 83-A.

Table 9 gives the results on a region basis for each of the five travel purposes, and for total transit trips. In all cases it becomes clear that significantly different transit trip projections are derived. The contribution of each system on transit trip production becomes evident for each travel purpose. Differences in this contribution are also apparent. Clearly, H-NW and NW-H trips appear to be the most sensitive to transit system improvements. Also NH-NH trips appear to be the type of trips with the greatest percent increase due to transit system improvements and to job increases in the region.

Table 10 gives the results on a sector-to-sector basis for all five trip purposes. Examination of this table reveals that the transit trip growth is expected to vary greatly by sector. According to the 1960 relationships and the projections of the independent variables, the greatest need and use for the mass transit system exist (or will be) in the suburban areas of the PJTS region. Whereas the City of Philadelphia demonstrates as a whole a stability or a relative inelasticity in the transit trip demand, the suburban area indicates a capacity to double or triple the present-day transit trips, if appropriate transit improvements are established there. Table 10 also shows that the difference of transit trips by sector between the two alternative projections has a relative consistency in terms of the area and the magnitude of trip changes. Clearly, a change in the transit system does not produce radical results in individual sectors—only incremental changes, positive or negative, are registered in each sector, although particular small districts within the sector may register a far greater rate of change than their sector as a whole.

Table 11 presents the results on a line-by-line basis. Each proposed line of system 82 or 83-A has been assumed to affect a number of individual and neighboring districts. The changes in the number of trips in each of these districts were taken into account in establishing the potential changes which each line might serve and/or induce. Note again that the significant changes occur mostly in the lines which serve suburban areas of the region. Of interest also is the substantial increase of transit trips appearing within the CBD of the City of Camden, which is expected to be the convergence point of three new transit lines from Philadelphia to the New Jersey areas.

All three tables also present a comparison of trip estimates reached on the basis of the original method. The comparison indicates clearly that the total understatement of trips in the region as well as the biased estimation of trips in the central area of the region and of its suburbs was clearly carried on to the projection phase of the work. The produced total understatements and consistence biases are especially evident for the City of Philadelphia. The differences between this set of projections and the ones achieved with the completed model of modal split are apparent by purpose of trips, by sector of trip origin and, of course, at the estimation of total trips. In all cases the previous imperfections have been eliminated.

An interesting aspect of this modal split model and its application for projection purposes is the fact that the basic form of the model indicates clearly that the relationships are addressed exclusively to the forces which affect the transit trip generation of each trip purpose. Nothing compels each set of relationships to produce estimates of one type of transit trip origins in the region equal to estimates of any other type of transit trip origins in the region. Accordingly, the use of this model for projection purposes would clearly produce estimates responsive to the various changes of the pertinent variables but also estimates which might or might not be in any type of equilibrium. For instance, home origin transit trip projections would reflect trip estimates according to the effects which changes in car ownership, residential density and job accessibility would have on home origin transit trips. In contrast to this, projections of W-H and NW-H trip estimates would reflect the effect which changes in job density and home accessibility would have on non-home origin transit trips. These effects might very well not be coincidental or, in other cases, conflicting; e.g., increases of car ownership would clearly tend to decrease home origin transit trips while increases of job density would tend to increase the non-home transit trip origins. The result in arithmetic terms would be that the projected estimates of H-W and H-NW transit trips would be lower than the projected estimates of W-H and NW-H transit trips. In reality the total number of transit trips from and to home will be the result of the combined influence of conditions at home as well as at the non-home trip origin. In essence, therefore, the average influence of the forces affecting the transit trip at its origin and destination should be taken into account when projections of this type of trip are undertaken. This approach is followed in our 1975 estimates in producing the final transit trip projections for the region.

CONCLUSIONS

The completion of this phase of the modal split model in the PJTS area makes possible the derivation of several conclusions of direct relevance to this study and perhaps to other similar efforts. Among them the following six should be stressed.

1. The selection of five transit trip types to be analyzed, simulated, and projected has produced satisfactory and reasonably reliable results. In essence this classification of trips reveals that other types of transit trips, formed on the basis of land use types at the origin or destination of trips, might not be necessary and that a classification of trips on the basis of the purpose of travel is both feasible and revealing of the different factors which influence transit trip generation. In our case the classification of trips into home origin and non-home origin, and into work and non-work trips proved to be of direct significance in achieving high correlations, low simulation errors, and reasonable projections.

2. The modal split model incorporates simultaneously one variable descriptive of the rider, another variable descriptive of the area of trip origination, and two variables directly descriptive of the systems serving the area and the region. The simultaneous incorporation of the effect of these variables is considered of significance and contributes to the reliability and accuracy of transit trip projections in the region. Of special significance is the incorporation of system variables to those non-system ones. The simultaneous incorporation of variables diminishes the significance of the question of whether one or another factor affects transit trip generation in advance of the other variables and at the same time increases the level of confidence of the resulting equations by permitting a greater number of observations to form the statistical basis of the formation of the predictive equations.

3. The system variables selected for this modal split model are essentially three. The first one is the implied density (or proximity) of transit lines in a measure of residential or job density in a district. On the basis of present-day experience, this implied association by density of development and density of transit lines is clearly justifiable. Transit lines are indeed more numerous and more frequent in areas with higher developmental density than in areas with low developmental density. The second system variable is the system accessibility, as previously defined. This variable combines three different features of the highway and transit system in addition to two area features. It combines the physical existence of a facility, as well as the travel speed and the associated costs (tolls, fares, etc.) for each highway and transit system. The area features it combines are the amount of relevant destinations included in each district within the metropolitan complex. The third system variable is the frequency of transit service which is descriptive of the availability of transit service in a district. This variable permits the recognition of the difference between a line with a few vehicles serving a district and another line with frequent and extensive service serving another district.

The need or desirability of incorporating additional system variables in a modal split model has frequently been pointed out in the literature. Although such need or desirability is generally recognized, the actual incorporation of variables expressing comfort or convenience has not been possible in the present modal split model. It appears that it would require, first, an objective, quantitative and meaningful definition of the concepts of comfort and convenience before their effective incorporation can be achieved. Additionally, data from "before" and "after" actual experiments will be required before detailed conclusions can be formed with regard to the extent of influence such variables may have on travel mode selection.

4. Closely related with the previous observations is the finding that each of the five trip purposes indicates a particular dependence on only one or two of the five variables incorporated in the model. Thus, if one desires to associate each of the five trip purposes with the one or two most important variables in explaining the district-to-

district transit trip variation, one can cite the following major associations: H-W transit trips with car ownership and frequency of transit service; H-NW transit trips with frequency of transit service and proximity of transit service (density variable); W-H and NW-H transit trips from the region at large with proximity of transit service (density variable) and frequency of transit service; NH-NH transit trips with proximity of transit service (density variable) and transit system accessibility; W-H transit trips from the CBDs of the region with transit system accessibility and proximity of transit service (density variable); and, finally, NW-H transit trips from the CBDs of the region with proximity of transit service (density variable). Implicit in these findings appears to be an understanding that any major change in each of the variables will have its most direct effect on the type of transit trip with which the variable is most closely associated. Of course, changes of transit trips at any future date will also have to be associated with changes in the variables which are directly associated with the complementary type of each transit trip purpose.

5. The verified relationships between transit trips and car ownership, density of development, and system variables open new possibilities for policy consideration in planning future transportation systems in an urban area. The formulation and testing of alternative transportation systems thus takes on additional meaning. Variations in the extent of facilities in the transit or highway system, variation in highway speeds, highway tolls, and transit fares, or variations in the frequency of transit service in each line can be directly reflected in the projection of transit trips produced in a region, just as planned or expected changes in the future car ownership rate and in the residential or job density in the various districts of a region can be directly incorporated in the projection process of transit trips in the region.

6. Another significant observation is in regard to the sensitivity and range of applicability of this model. Examination of the equations and of the means of deriving accessibility measures indicates that individual changes of each variable have limited impact on the final results. The projection of two transit trip estimations on the basis of a common land use plan but with two distinctly different transportation systems presents a good indication of the sensitivity of the model to various sets of transportation facilities. The two system variables, which respond directly to system variations, have produced most of the differences observed between original projections and final projections, as well as all of the differences between systems 82 and 83-A. Clearly, a transportation system should include substantial changes before any real effect on transit rates would become significant. The model itself also has a restrictive characteristic on the effect of any of the variables. It is apparent that although the model is directly responsive to variations in the transportation system and density of development, these variations should be of significant magnitude and extent before any appreciable impact will register in the number of transit trips produced in the particular sector(s) and in the region in general. Thus, individual facilities can seldom be expected to produce changes of the required magnitude and therefore such facilities could seldom be expected to register any region-wide effect, beyond the effect on a few districts through which they might pass and thus serve directly.

Finally, it should be pointed out that the flexibility acquired with the modal split process should not foster unreasonable application of the model. Although the relationship between each variable and the transit trip rate is already conditioned by the parameters of the predictive equations, there is still need for cautious application of the model within reasonable limits of change of each variable. It would be, for instance, an unreasonable application of the model if a multiple increase of each variable were assumed for the sake of experimentation and a corresponding increase of transit trips were expected. Incremental increases or decreases of present rates should be considered as the proper objective of the modal split model. Radical changes in the travel patterns and preferences in urban regions, as well as in the transportation systems, are not usually subject to the predictive capacity of one or of a handful of variables derived and correlated within the present-day set of circumstance. No predictive model based on manifest present-day behavior and on existing trends and circumstances should be expected to project radical changes or to anticipate any fundamental reversal of preferences and situations.

REFERENCES

1. Campbell, E. M. Transportation Quality: Research Challenge and Promise. Traffic Quarterly, pp. 395-407, July 1963.
2. Claffey, P. J. User Criteria for Transit Planning. Jour. Urban Planning and Develop. Div., Proc. ASCE, p. 4029, Sept. 1964.
3. Moses, L. N. Economics of Consumer Choice in Urban Transportation. The Dynamics of Urban Transportation National Symposium, Oct. 1962.
4. A Model for Estimating Travel Mode Usage in Washington, D. C. Traffic Research Corp., 1962.
5. Hill, D. M., and Von Cube, H. G. Development of a Model for Forecasting Travel Mode Choice in Urban Areas. Highway Research Record 38, p. 78, 1963.
6. Schofer, J. L., and Voorhees, A. M. Modal Split Model. Paper presented at HRB 43rd Annual Meeting to Committee on Origin and Destination.
7. Schmidt, J. W. Modal Split Rationale.
8. Basmaclyan, H., and Schmidt, J. W. Development and Application of a Modal Split Model for the Puget Sound Region. Staff Rept. No. 12, Puget Sound Regional Transportation Study, May 1964.

Effect of Trip Direction on Interzonal Trip Volumes: Test of a Basic Assumption of Trip Distribution Models

DAVID E. BOYCE, Assistant Professor of City and Regional Planning, University of Pennsylvania

An assumption common to existing urban trip distribution models concerning the location of zones is tested by an analysis of variance of interzonal trip data. The assumption is that interzonal volumes are independent of zonal location but are a function of zonal separation. A test for the existence of a trip direction effect with respect to the urban center is developed and applied.

•COMPARED with many problem areas in urban transportation planning, trip distribution models are generally regarded as a relatively well-developed and operational methodology. In an absolute sense, however, these models must be regarded as quite crude, yielding only a first approximation of the urban travel pattern. For example, Heanue and Pyers (5) in a comparative evaluation of these models, found standard errors of 10 to 90 percent of mean interzonal volumes computed on the basis of a highly aggregated spider network assignment, suggesting a need for further research.

The study reported here investigates a basic assumption common to all trip distribution models through an empirical analysis of interzonal trip data. The assumption of interest concerns the use of zonal separation to specify the relative location of origin and destination zones, thereby ignoring the effect of the direction of trips with respect to the urban center on interzonal volumes. An analysis of variance model involving four factors—trip direction, trip length, size of destination zone, and zone of origin—is developed to test this assumption and to investigate general characteristics of urban travel patterns. Separate analyses are conducted for arterial and mass transit trips using data compiled by the Chicago Area Transportation Study.

The following discussion of the assumptions of trip distribution models emphasizing the description of zonal locations leads to a statement of the objectives of the research. Following a brief discussion of the statistical model and tests of hypotheses, the results of the analysis are described. A discussion of the implications of the study for trip distribution models and urban theory concludes the paper. A more detailed discussion of the statistical methodology used in the analysis may be found in Boyce (1).

ASSUMPTIONS OF TRIP DISTRIBUTION MODELS

Currently available trip distribution models range from empirical relationships involving many parameters estimated from observed origin-destination patterns to simple probability models derived from postulates concerning the nature of urban travel. The former, which includes a number of variations on the gravity model concept, distributes trips as a function of the number of trip ends in possible destination zones and an inverse power function or an exponential function involving travel time between zones. This trip distribution gravity model is a descendent of the gravity model concept in the social sciences (6). It is the most widely used of the operational trip distribution models, in part because of the availability of computer software developed by the U.S. Bureau of Public Roads (14) and others.

Paper sponsored by Committee on Origin and Destination.

A second class of models, the opportunity models, distributes trips as a function of the size of destination zone and the number of trip ends or opportunities for completing a trip closer to the origin zone than the destination zone is to the origin zone. The intervening opportunities model, developed by Schneider (10, 11), is derived from simple postulates regarding urban travel, and has been successfully applied in a number of transportation studies (3). A related model, similar in concept but quite differently formulated, is the competing opportunity model of Tomazinis (13). Both models are similar in concept to the earlier work of Stouffer (12). Harris (4) has extended the opportunity model formulation, and derived a modified form of the gravity model partly from the assumptions of Schneider's model.

The following assumptions are implied in both types of trip distribution models:

1. Travel time is an adequate measure of accessibility or ease of travel between zones and can be determined with sufficient accuracy for the time period of interest.
2. The choice of length of trip is entirely influenced by the proximity of the attractive forces that cause trips to occur and is not influenced by particular social and economic characteristics of a zone; rather, these characteristics are reflected in the number of trip ends per zone.
3. The average travel pattern is independent of zonal locations (e.g., whether the two zones are radial or circumferential to each other) and is rather a function of the separation between zones.

The use of travel time or a function of travel time as a measure of accessibility in trip distribution models (assumption 1) is a well-established procedure in nearly all urban transportation studies. Studies of the relationship between interzonal travel time and straight line distance between zones indicate that travel time may be some simple transformation of airline distance. However, the use of travel time permits the inclusion of terminal times and takes account of geographical features, such as rivers and other types of barriers, more adequately than distance measures.

The economic and social differences among zones (assumption 2) appear to affect the number of trip ends per zone more than the zone-to-zone movement of trips. These economic and social differences are explicitly accounted for in trip generation methodology (8). There is little evidence that such differences are important in trip distribution models, although some versions of the gravity model include factors for making special adjustments for such conditions.

The question of the effect of location of zones on interzonal trip volumes (assumption 3), or equivalently, the question of the direction of trips with respect to the study area, is the central focus of this study. Both classes of models, gravity and opportunity, regard destination zones as being arranged along a straight line in order of their travel times from the origin. From this assumption it follows that from a given origin, trips in all directions are equally likely for destination zones of equal size and travel time from the origin. Therefore, the estimates of interzonal volumes are unaffected by the location of these destinations in relation to the center of the urban area and the overall pattern of urban activities. Since both distance and direction are required to determine a destination location from a given origin, these models give no consideration to direction of trips. Therefore, these models cannot distinguish between radial, circumferential, or other direction-determined types of trips. For this reason assumption 3 is necessary to the conceptual formulation of these trip distribution models.

The principal objective of this study, then, is to test the assumption that interzonal trip volume is independent of the direction of trips, i.e., independent of the location of zones. A more general purpose is to devise a general statistical procedure for examining questions concerning variations in interzonal trip volumes. An analysis of variance model involving four classification factors is derived for this purpose.. Two series of experimental designs specifying combinations of these factors are applied to a sample of interzonal trip data. Hypotheses on the effect of each factor and combinations of factors on variations in interzonal volumes are tested using this methodology. Following a brief description of the interzonal trip and network data used in the study, the statistical methodology and test results are described.

Data for the Chicago area in 1956 were made available for this study through the cooperation of the Chicago Area Transportation Study (CATS). The CATS interzonal trip file is based on a home interview survey of 1/30 of the households residing in the 1,236 square mile study area in 1956 (2). The travel pattern of the sample households was aggregated into the 582 analysis zones, and expanded to provide an estimate of the total interzonal travel pattern. Therefore, the data consist of 24-hr person and vehicle trip volumes between zones by mode of trip. Interzonal travel was recorded in the survey for about 40,000 pairs of zones or about 24 percent of the possible zonal interchanges.

In addition to the interzonal trip file, descriptions of the transportation networks in 1956 are necessary to determine interzonal travel time. Coded networks for 1956 with the travel time over each link were furnished by CATS. The arterial network, consisting of all freeways and principal arterials, is composed of 2,396 nodes and 8,807 one-way links. The mass transit network is a combination of suburban railroad, subway-elevated, and principal bus routes in the study area and consists of 1,749 nodes and 6,212 one-way links. Minimum travel time paths through these networks were computed using a modification of a program by Martin (7) based on the Road Research Laboratory algorithm.

A METHODOLOGY FOR TESTING HYPOTHESES ON VARIATION IN INTERZONAL TRIP VOLUMES

A statistical model based on an experimental design was formulated for the purpose of testing hypotheses concerning the effects of specified factors on the variation of interzonal trip volume. The statistical model is the mixed model in the analysis of variance, in particular the T^2 procedure and the method of multiple comparisons described by Scheffé (9).

The analysis of variance is a statistical technique for analyzing measurements, depending on several kinds of factors operating simultaneously to determine which factors are important and to estimate their magnitude. The choice of analysis of variance for this study is appropriate for two reasons. First, no relationship between interzonal volume and trip direction is specified for testing. The use of linear regression analysis, for example, would require that a linear function between the interzonal volume and trip direction be assumed. Second, trip direction is not treated as a specific direction or azimuth from the center of the area but, instead, as a sector defined by two radials with a given included angle. The question of the variation of trip volumes due to a factor of this qualitative type is properly answered by an analysis of variance.

Before the statistical model is introduced, the four factors and their levels are defined. Then the model is stated in terms of these factors, and meaning of the model and tests of hypotheses are discussed.

Direction of Trips—Factor A

The direction of interzonal trips from origin to destination with respect to the center is the principal factor of interest in this study. Trip direction sectors are defined by I radials ($i = 1, \ldots, I$) extending from the origin with equal included angles. The I sectors formed in this manner exhaust all possible trip directions from a given origin zone.

An "absolute deviation" definition of trip direction sectors is necessary in this analysis because of the position of Lake Michigan in the otherwise symmetrical CATS study area (Fig. 1). The sectors are numbered in a consistent manner with sector one bisected by a radial through the center. Sectors two through (I-1) are divided equally to the left and right of sector one. Sector I is opposite sector one. For any given origin zone, destination zones may be classified by these trip direction sectors. Zones in sector one, for example, are described as "toward the center," and zones in sector I are "away from the center."

In the analysis of variance, four sectors are specified in most designs (Fig. 1a). For two designs, six sectors are employed to determine the sensitivity of the analysis to the number of sectors specified.

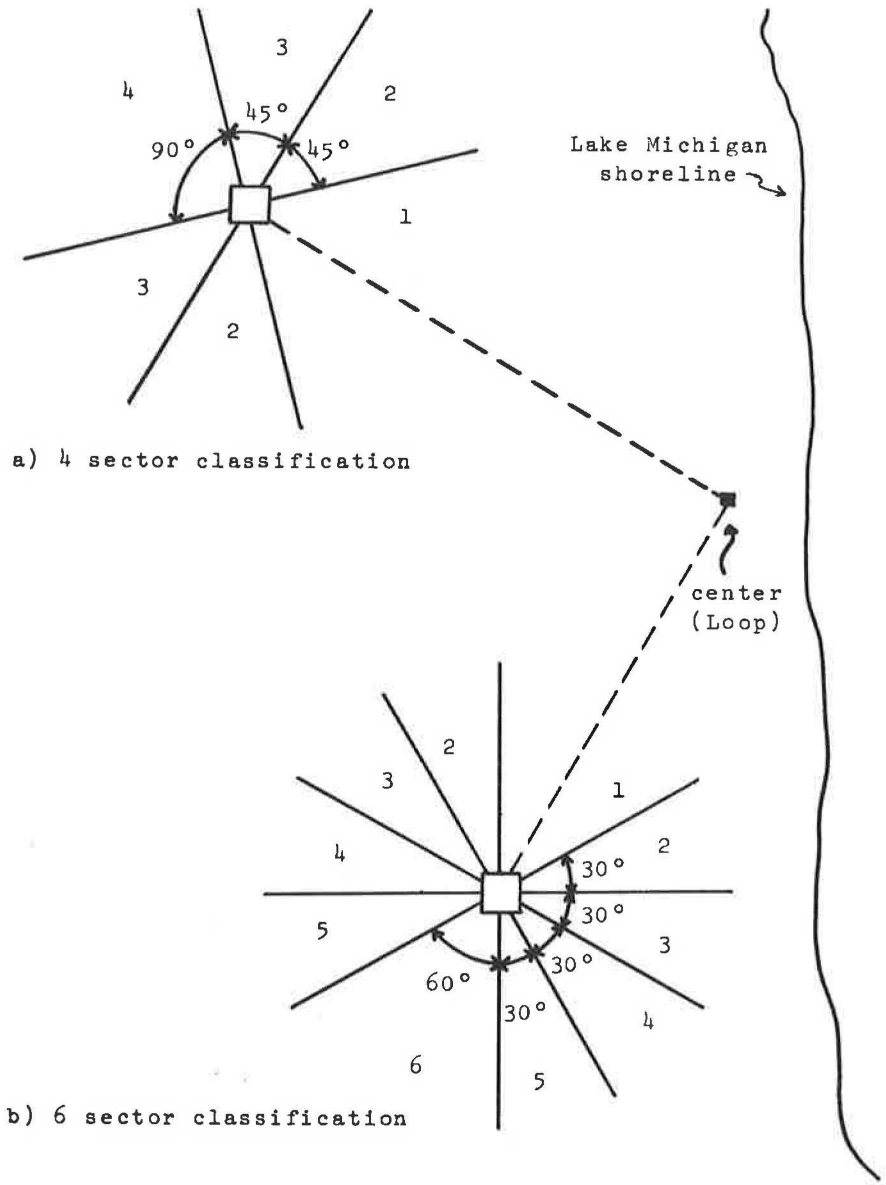


Figure 1. Definition of absolute deviation sectors.

Length of Trips—Factor B

Trip length is a second classification factor included in the analysis for its well-known relation to interzonal volume. A trip length level is an interval of elapsed travel time from origin to destination over the arterial or transit network, depending on the mode being analyzed. For most of the designs, three intervals of equal length are defined so that in each interval the number of destination zones "linked" to the origin by an observed interzonal volume is as uniform as possible. For arterial trips, intervals of 2 to 15, 15 to 28, and greater than 28 minutes are defined. For mass transit trips, intervals of 4 to 12, 12 to 20, and greater than 20 minutes are defined. No interzonal travel times of less than 2 minutes for the arterial mode and 4 minutes for the transit mode are recorded.

Size of Destination Zone—Factor C

The third classification factor, size of destination zone, is specified in certain designs to determine its effect on interzonal volume. Size of zone is measured by the total number of destinations; two levels are sufficient to determine the existence of an effect. The size classes are established so that there are an equal number of linked destination zones in each class. For arterial trips, 20,800 trips is the median size; for transit trips, 8,400 trips is the median.

Origin Zone by Ring—Factor D

Factor D is composed of a sample of zones from the population of 582 zones in the study area. The seven concentric rings of zones about the center of the study area defined by CATS were used to organize the zones into roughly homogeneous subpopulations. A 15 percent sample of zones was drawn from each ring. Each analysis is performed on the sample of zones from a given ring, each zone being a level of this random factor.

Examination of the trip volumes for the sample zones from the CATS rings revealed that certain rings were unsatisfactory for analysis. The density of trips from rings 6 and 7 is too sparse to permit the analysis. Rings 0 and 1 were also discarded because of the small number of zones involved and their proximity to the center. Rings 2 and 3 were combined into a composite ring to obtain a sample of sufficient size for the analysis. The data used in the analysis, then, consist of 11 zones from rings 2-3, 13 zones from ring 4, and 15 zones from ring 5. Separate analyses are performed for arterial and mass transit person trips for the sample origins from each ring.

Interzonal Person Trips—Criterion Variable

Based on the definitions of the four factors, the criterion variable in the analysis of variance is now defined. Consider interzonal trips originating from the m^{th} sample origin zone in a given ring. The destination zone of each interzonal volume may be classified according to its direction from the origin, its travel time from the origin, and its size. Therefore, let Y_{ijkm} equal the number of trips from the m^{th} origin zone in a given ring to a destination in direction i, at trip length j, and of size class k. The value of this random variable, Y_{ijkmn}, is the n^{th} observation for the treatment or cell, ijkm. A series of N such observations drawn at random from the population of interzonal volumes originating at zone m constitutes N replications of this experiment (Fig. 2).

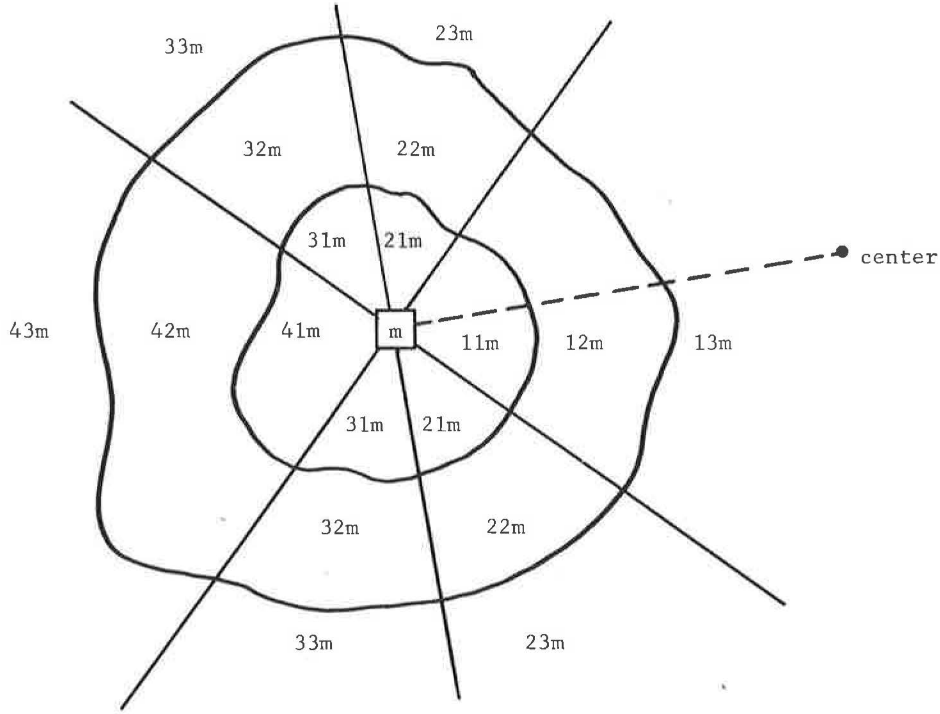
Two problems are associated with the use of this criterion variable in the analysis of variance. First, the definition calls for one-way volumes, whereas the CATS data are nondirectional volumes between zones. Because of the high symmetry of interzonal trips, the CATS data are taken as an acceptable substitute for the one-way volumes. This procedure is based on a study by CATS (2) of the difference in one-way interzonal volumes for pairs of zones selected at random, demonstrating that the distribution of differences approximates that expected from sampling error alone.

The second question concerns the frequency distribution of the criterion variable, assumed to be normally distributed in the analysis of variance model. Studies of the observed volumes for the sample origin zones show a distribution highly skewed to the right. A natural logarithmic transformation of the data eliminates most of the skewness and maintains the kurtosis within acceptable limits. This transformation proved acceptable and is implied in all of the following analyses.

Discussion of Statistical Methodology

The mixed model for the case of two fixed factors (A and B) and one random factor (D) is summarized as follows. The model equation is

$$Y_{ijmn} = \mu + \alpha_i^A + \alpha_j^B + a_m^D + \alpha_{ij}^{AB} + a_{im}^{AD} + a_{jm}^{BD} + a_{ijm}^{ABD}$$



The classification of destination zones in each area is specified by the notation ijm in that area.

Figure 2. Classification of destination zones.

where

μ = grand mean;

α_i^A = main effect due to i^{th} trip direction sector;

α_j^B = main effect due to j^{th} trip length interval;

a_m^D = main effect due to m^{th} origin zone;

α_{ij}^{AB} = interaction of i^{th} sector and j^{th} interval;

a_{im}^{AD} = interaction of i^{th} sector and m^{th} origin zone;

a_{jm}^{BD} = interaction of j^{th} interval and m^{th} origin zone;

a_{ijm}^{ABD} = interaction of i^{th} sector, j^{th} interval, and m^{th} origin; and

e_{ijmn} = error term.

Of principal concern are the main effects of factor A, trip direction, and the interactions with factors B and C, trip length and size of destination zone. The hypothesis concerning trip direction is stated as

$$H_A: \alpha_i^A = 0, i = 1, \ldots, I$$

The meaning of testing this hypothesis of no main effects is to evaluate the magnitude of a series of comparisons among mean directional interzonal volumes. The rejection of H_A indicates that significant differences (greater than could occur by chance at the α level of significance) do exist between some of these mean volumes. Scheffé's T^2 procedure, available for the foregoing model, provides an exact test of this hypothesis. Similar tests apply for H_B and H_C for models involving these factors.

The hypotheses for the interactions with trip direction are

$$H_{AB}: \alpha_{ij}^{AB} = 0, i = 1, \ldots, I$$

$$H_{AC}: \alpha_{ij}^{AC} = 0, i = 1, \ldots, I$$

Approximate F-tests based on a conventional analysis of variance table are used for these hypotheses. The absence of interactions is interpreted to mean that the main effects are additive for the various combinations of levels considered. Other hypotheses of interest are tested in a similar manner.

By plotting the profiles of the factor response surface, a graphic interpretation may be given to the tests of main effects and interactions. A horizontal profile is indicative of the absence of main effects, whereas a sloping or irregular profile suggests that a significant effect exists. Interaction between two factors can be detected by comparing the slope of profile curves. In order for the curves to be parallel, the response surface must be composed of the sum of the main effects. Nonparallel profiles, then, indicate the presence of an interaction between the two factors.

FINDINGS AND INTERPRETATION OF ANALYSIS OF VARIANCE

Summarized here are the findings and interpretation of the analysis of variance of two series of experimental designs. First, the exact T^2-tests for the five designs given in Table 1 are representative of the principal research findings. These designs apply a

TABLE 1

T^2-TESTS FOR PRINCIPAL DESIGNS

Ring, Mode	No. of Sectors	No. of Intervals	Trip Direction[a]			Trip Length[b]		
			d.f.[c]	F[d]	$F_{.05}$[e]	d.f.	F	$F_{.05}$
2-3, A	4	3	3, 8	1.67	4.07	2, 9	43.95	4.26
4, A	4	3	3, 10	1.99	3.71	2, 11	168.84	3.98
5, A	4	3	3, 12	3.11	3.49	2, 13	55.09	3.80
2-3, T	4	3	3, 8	8.90	4.07	2, 9	44.74	4.26
4, T	4	3	3, 8	9.52	4.07	2, 9	49.50	4.26

[a] H_A: $\alpha_i^A = 0, i = 1, \ldots, 4.$

[b] H_B: $\alpha_j^B = 0, j = 1, \ldots, 3.$

[c] d.f. = degrees of freedom.

[d] $F = [(M-h+1)/(M-1)(h-1)]T^2$, h = I, J.

[e] $F_{.05}$ = value of the F distribution at the 5 percent level of significance with (h - 1(, (M - h + 1) degrees of freedom, h = I, J.

TABLE 2

APPROXIMATE F-TESTS FOR PRINCIPAL DESIGNS

Design: Ring, Mode	Source of Variation	d.f.[a]	SS	MS	F[b]	$F_{.05}$[c]
2-3, A	A: Trip direction	3	2.29	0.76	1.09	2.62
	B: Trip length	2	206.04	103.02	147.17	3.02
	D: Origin zone	10	24.49	2.45	3.50	1.85
	AB interaction	6	3.30	0.55	0.79	2.12
	Residual	494-1[d]	346.01	0.70		
	Total	515-1	582.13			
4, A	A: Trip direction	3	5.39	1.80	2.40	2.62
	B: Trip length	2	365.91	182.96	243.95	3.02
	D: Origin zone	12	19.25	1.60	2.13	1.78
	AB interaction	6	4.23	0.71	0.95	2.12
	Residual	586	437.03	0.75		
	Total	609	831.81			
2-3, T	A: Trip direction	3	15.18	5.06	9.20	2.62
	B: Trip length	2	114.11	57.05	103.73	3.02
	D: Origin zone	10	26.52	2.65	4.82	1.85
	AB interaction	6	14.50	2.42	4.40	2.12
	Residual	455-1[d]	249.26	0.55		
	Total	476-1	419.57			
4, T	A: Trip direction	3	20.50	6.83	11.58	2.62
	B: Trip length	2	91.47	45.74	77.53	3.02
	D: Origin zone	10	22.14	2.21	3.75	1.85
	AB interaction	6	7.85	1.31	2.22	2.12
	Residual	406-4[d]	237.16	0.59		
	Total	427-4	379.12			

[a] d.f. = degrees of freedom.
[b] $F = [(M-h+1)/(M-1)(h-1)]T^2$, h = I, J.
[c] $F_{.05}$ = value of the F distribution at the 5 percent level of significance with (h - 1), (M - h + 1) degrees of freedom, h = I, J.
[d] Loss of d.f. due to estimation of missing values.

four trip direction sector by three trip length interval classification to the five sets of sample origin zones—three arterial and two transit—suitable for analysis.

The T^2-tests are used to test hypotheses of no main effects for the trip direction and trip length factors (H_A and H_B). The hypotheses for trip length are clearly rejected for both arterial and transit trips, verifying the well-known decreasing relationship between interzonal volume and trip length. The hypotheses on trip direction, H_A, cannot be rejected for arterial trips at the 5 percent level of significance. For transit trips, however, this hypothesis is rejected, indicating that significant variation in mean direction interzonal volume does exist for this mode. Scheffé's method of multiple comparisons is used to identify levels responsible for rejection of the hypotheses. For H_A, higher mean volumes in the direction of the center are responsible for the rejection of the transit designs. For H_B all combinations of levels are significant.

The approximate F-tests for the first four designs are given in Table 2. These statistics are used to test the interaction hypothesis, H_{AB}. In each case, all the interactions except AB have mean squares less than, or approximately equal to, the error mean square. These terms are, therefore, pooled to form a residual mean square. The approximate F-test is then the ratio of MS_{AB} to $MS_{residual}$. The table indicates that H_{AB} cannot be rejected for arterial trips, but is significant for transit trips. Therefore, for transit trips, there exists a variation between factors A and B beyond that accounted for by the addition of the two main effects. F-tests for H_A and H_B from Table 2 are consistent with the exact T^2-tests previously discussed. In addition, H_D, the hypothesis of no main effects for zone of origin, is rejected by the F-tests, as would be expected due to the variation in composition by zone.

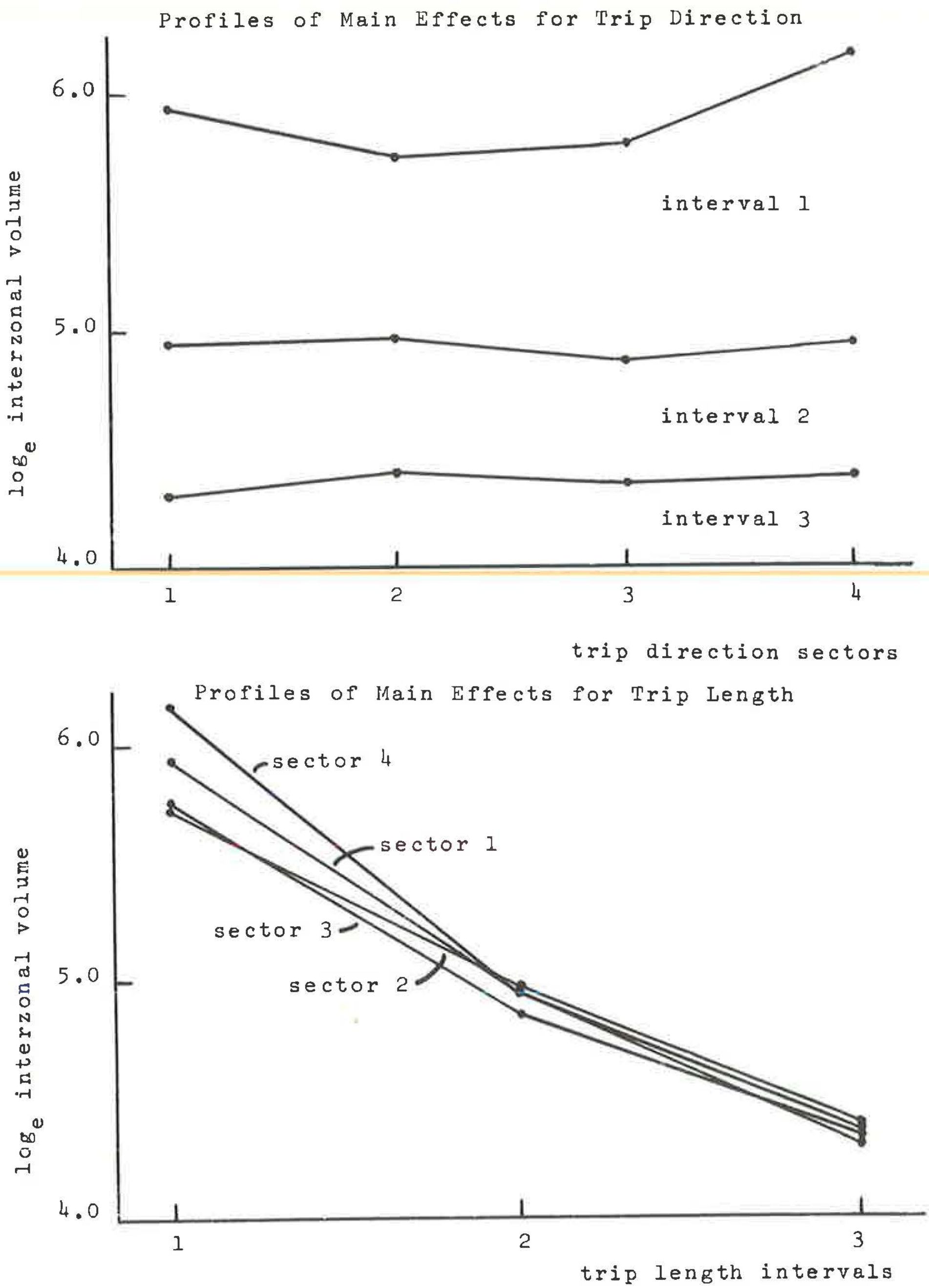


Figure 3. Ring 2-3 arterial trips.

Figures 3 and 4 show the profile curves from these designs. Figure 3 for ring 2-3 arterial trips shows the flat profiles for the trip direction factor and the sloping trip length profiles. The profiles are quite parallel, indicative of the absence of an interaction between the two factors. Figure 4 for ring 2-3 transit trips suggests the reasons for the rejection of H_A and H_{AB} for the transit case. Sector 1 has a much higher mean volume than other sectors. The nonparallel tendency of the profiles leading to a rejection of H_{AB} is also pronounced in these curves.

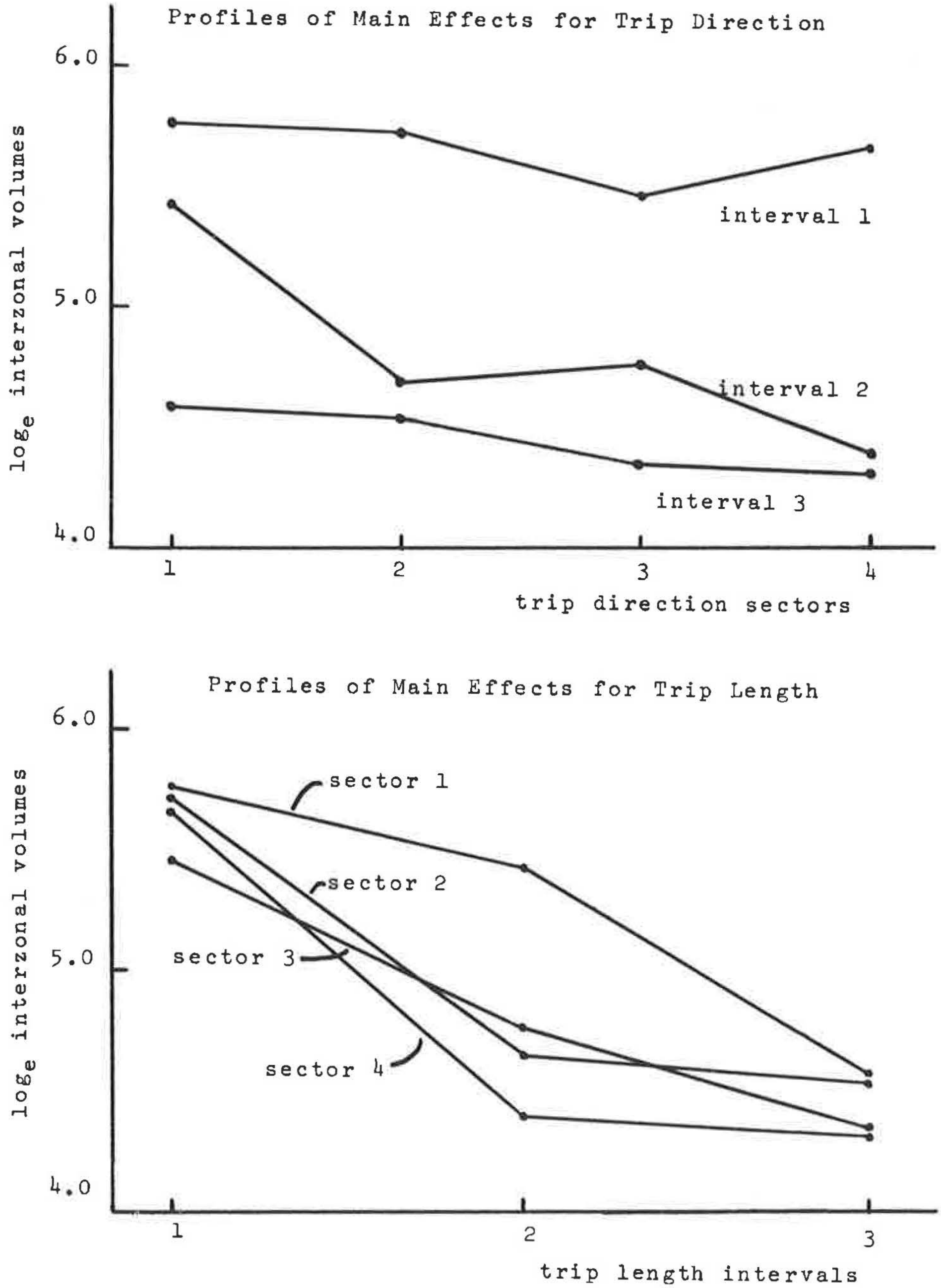


Figure 4. Ring 2-3 transit trips.

Summarizing these principal findings, a trip direction effect is significant for transit trips, higher in the direction of the center, but not for arterial trips. A significant interaction also occurs for transit trips, indicating that the effect of trip length on mean volumes differs by direction of trips. Finally, there is a small effect associated with the zone of origin. Several additional ring 4 designs consisting of a six sector definition for trip direction and from two to nine intervals for trip length were analyzed to test the sensitivity of these principal findings to the definition of levels. No significant deviations from these findings were identified in the analysis of these designs.

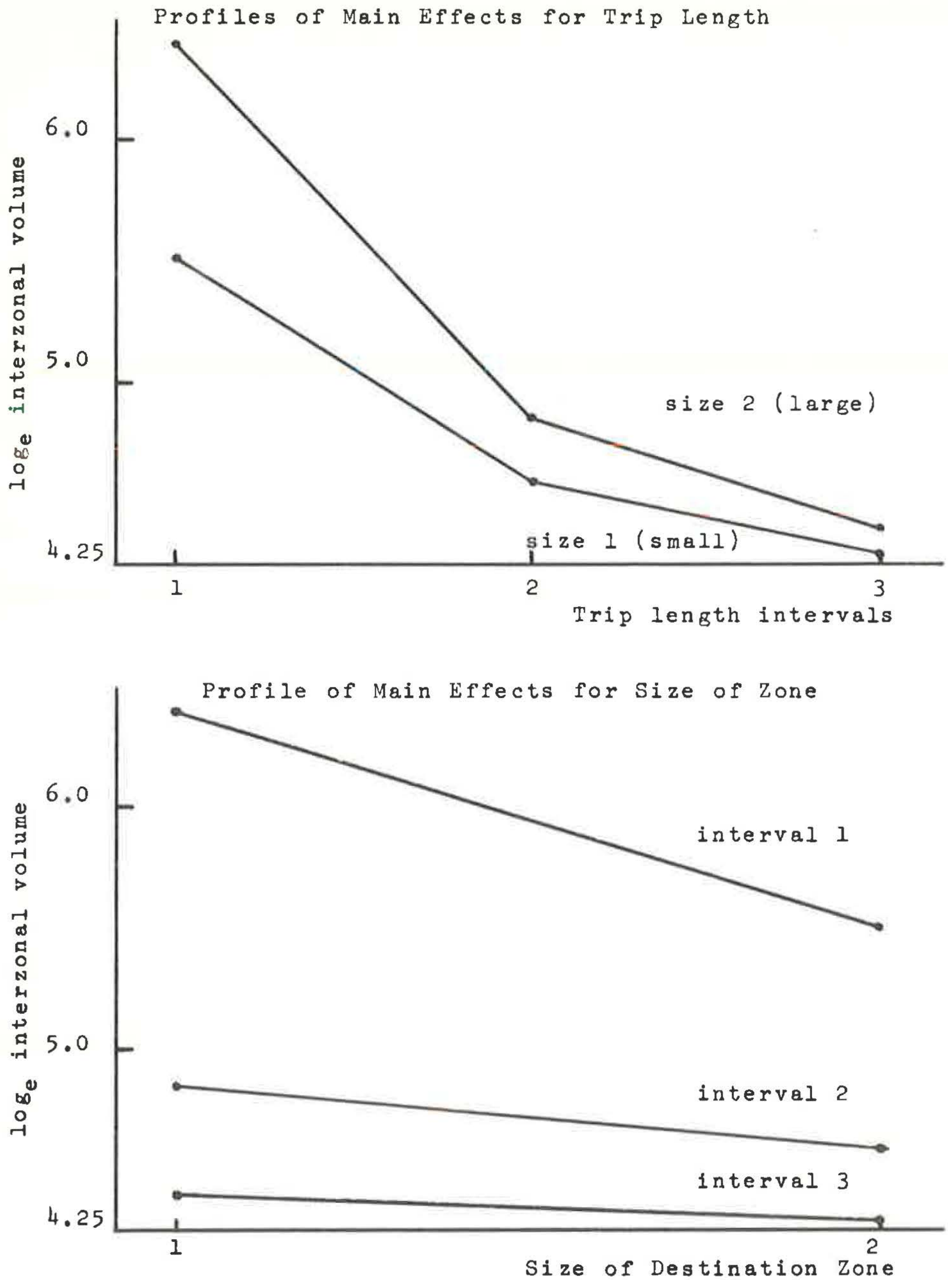


Figure 5. Ring 4 arterial trips.

The analyses of a second series of designs are now considered, to determine the effect of size of destination zone on the foregoing results reported. T^2-tests for a four sector by two size of zone design for the ring 4 arterial trips result in the rejection of H_C, the hypothesis of no main effect for size of zone. However, H_A and H_{AC} are not rejected. Due to data limitations, this design could not be implemented for ring 4 transit trips.

T^2-tests and F-tests were computed for a three interval by two size of zone design for both ring 4 arterial and transit trips. In each design H_B and H_C are rejected.

Also, H_{BC}, the interaction hypothesis, is rejected for both designs. The profile curves for ring 4 arterial trips suggest the reason for the rejection of H_{BC} (Fig. 5). The size of zone mainly affects interzonal volumes to nearby destinations. For zones in trip length intervals two and three, the size of zone profiles are nearly flat. A similar result is found for ring 4 transit trips but only in the third trip length interval.

IMPLICATIONS FOR TRIP DISTRIBUTION MODELS AND ANALYSIS OF TRAVEL PATTERNS

The primary purpose of this study is to evaluate the assumption that interzonal volumes are independent of the location of zones and are, instead, a function of the separation of zones. An evaluation of the assumption is provided by testing the hypothesis that mean interzonal volumes classified by the direction of trip are constant. The implications of the findings described in the foregoing for this assumption are clear. For arterial trips, the assumption is valid; for transit trips the assumption is not substantiated.

This observed bias in trip direction toward the center for transit trips coincides with the radial location of the rail transit network in the study area in 1956. Although a radial freeway system was a major component of the arterial network in 1956, a similar bias in trip direction for arterial trips is not evident. Two reasons are suggested for this apparent inconsistency. First, the difference in interzonal travel time between rail transit and bus transit is probably much greater than between freeway and arterial facilities. Therefore, there is a more pronounced bias in the transit network toward the center than for the arterial network. Second, there are inherent disadvantages in trip distribution models that consider arterial and transit trips separately. A tripmaker's selection of a destination zone is probably not based on a choice among a rail transit ride,. a bus ride, or a combination of the two as suggested by trip distribution models, but rather a choice between some combination of transit and auto trips. One can speculate that if rail transit service were available in all directions (a very unlikely event due to considerations of economic feasibility), the bias in trip direction toward the center would be reduced. At the same time, such a rail transit network might result in fewer arterial trips in noncentral directions resulting in a trip direction effect for arterial trips.

This problem of choice between modes has been studied separately from the trip distribution problem in urban transportation planning. The questions raised by the foregoing discussion document the need for future research to develop models in which the choice of mode and selection of destination zone are determined within a single formulation. At that time, the finding of this study regarding a trip direction bias in mass transit trips should be considered.

A secondary objective of this research was to develop a statistical methodology for the analysis of urban travel patterns. The mixed model in the analysis of variance is an appropriate technique for this problem and is applied successfully in this study. The application of the multivariate T^2 technique in this study reveals the potential contribution of multivariate statistical analysis to problems of spatial behavior and structure.

In addition to the results on trip direction discussed, the mixed model enables the testing of hypotheses on other factors related to urban travel patterns. As expected, the variance associated with trip length is significant throughout. The size of the destination zone is also a significant factor. The significant interaction between size of zone and trip length indicates the influence of size decreases as trip length increases.

The uniformity of the total trip pattern by trip direction is a second finding demonstrated by the study results. For the Chicago area with its large and important central area, the almost negligible impact of the center on the travel pattern suggests the need for a revised approach in some studies of urban location and structure. Many theoretical studies of urban location make the simplifying assumption that employment and trade occur only at the center. Although such assumptions may greatly facilitate the development of theory, it appears that in this case the assumption may be so far removed from reality that the resulting theory yields few operational constructs or testable hypotheses. A relatively simple but probably much more useful assumption is that employment and trade are uniformly distributed over the urban area, and the center is merely the point at which the initial development occurred. From this assumption, distribution of trip length can be obtained, and the transportation cost necessary for a theoretical development computed. In view of the results of this study, it appears that such a reorientation of theory could lead to the development of a more adequate urban theory than is presently available.

ACKNOWLEDGMENT

The use of the facilities of the Computation Center, Battelle Memorial Institute, and the cooperation of the Chicago Area Transportation Study are gratefully acknowledged.

REFERENCES

1. Boyce, D.E. The Effect of Direction and Length of Person Trips on Urban Travel Patterns. Journal of Regional Science, Vol. 6, No. 1.
2. Final Report, Chicago Area Transportation Study, Vol. I, pp. 36-38, Chicago, 1959.
3. Final Report, Chicago Area Transportation Study, Vol. II, Chicago, 1960.
4. Harris, B. A Note on the Probability of Interaction at a Distance. Journal of Regional Science, Vol. 5, No. 2, 1964.
5. Heanue, K. E., and Pyers, C. E. A Comparative Evaluation of Trip Distribution Procedures. Highway Research Record 114, pp. 20-37, 1966.
6. Isard, W. Methods of Regional Analysis. MIT Technology Press and John Wiley, Cambridge, Mass., Chap. 11, 1960.
7. Martin, B. V. Minimum Path Algorithms for Transportation Planning. MIT, Cambridge, Mass., 1963.
8. Oi, W. Y., and Shuldiner, P. W. An Analysis of Urban Travel Demands. Northwestern Univ. Press, Evanston, Ill., 1962.
9. Scheffé, H. The Analysis of Variance. John Wiley, New York, 1959.
10. Schneider, M. Gravity Models and Trip Distribution Theory. Papers, Regional Science Assoc., Vol. 5, 1959.
11. Schneider, M. Panel Discussion on Inter-Area Travel Formulas. HRB Bull. 253, pp. 128-137, 1960.
12 Stouffer, S. A. Intervening Opportunities: A Theory Relating Mobility and Distance. American Sociological Review, Vol. 5, 1948.
13. Tomazinis, A. R. A New Method of Trip Distribution in an Urban Area. HRB Bull. 347, pp. 77-99, 1962.
14. Calibrating and Testing a Gravity Model with a Small Computer. U.S. Bureau of Public Roads, Washington, D.C., 1962.

Tri-State Transportation Commission's Freight Study Program

ROBERT T. WOOD, Tri-State Transportation Commission, New York

•THE Tri-State Transportation Commission's work program is unique among all urban studies because it measures and describes both local and intercity freight traffic by all modes. The completed picture will be fuzzy in spots, but it will be comprehensive.

Objectives of this first, descriptive, phase of Tri-State's freight program are limited. The first objective is to develop good estimates of the region's total traffic flow by mode. This phase has been completed, and the results are shown in Tables 1 and 2 and Figure 1. The differing roles of each mode are evident, as well as the way areas differ in their rates of freight generation per capita. The second objective is to measure and describe each mode's traffic, and the characteristics of the freight-carrying portion of the system.

PURPOSE

The purpose is to develop a base for future projection. By determining the freight requirements of the region's present activities and by using forecasts of those activities, future freight demands can be forecast. Through trend analysis and other techniques, it is possible to estimate each mode's future share of traffic by broad commodity category.

DIFFERENT APPROACHES FOR EACH MODE

The amount and quality of available freight traffic data vary considerably, depending on which mode is being discussed.

Rail and waterborne freight traffic are measured through use of secondary sources. We have obtained special tabulations of the ICC's 1 percent Rail Carload Waybill Sample, covering traffic into, out of and within the Tri-State Region. After some travail in matching the data with data collected by the Port of New York Authority, this ICC source has been validated, subject to some qualifications. The principal source for waterborne freight traffic has been the publications of the Corps of Army Engineers. In this case also, certain discrepancies were found with locally generated data. The major source, "Waterborne Commerce of the United States, Part 1," is not an origin and destination survey, but its detailed commodity analysis of the traffic, plus a breakdown into domestic, foreign, local and other categories, makes it possible to draw firm conclusions on areas of origin and destination.

Air cargo and pipeline sources require more personal contact and interpretation. Air mail, express, and freight tons emplaned are published by airport, but there is no published origin and destination or commodity analysis. Individual carriers have made studies of their own traffic, however. Pipeline volumes are not published except on a company-wide basis, and then only for regulated pipelines. Individual lines must be contacted for this information.

Truck freight data are practically nonexistent. Surveys had to be run by Tri-State to develop meaningful information. These surveys are the central subject of this paper.

Paper sponsored by Committee on Origin and Destination.

Table 1

FREIGHT TRAFFIC OF THE TRI-STATE REGION

1962 - 1963

by mode of carriage

	WATER	TRUCK	RAIL	PIPELINE	AIR	TOTAL
Millions of Tons Carried						
into the region	82 *40*	68	43	9	*	202
percent of total	40.5%	33.7%	21.3%	4.5%	*	100.0%
out of the region	13 *7*	68	10	6	*	97
percent of total	13.4%	70.1%	10.3%	6.2%	*	100.0%
within the region	69	195	4	*	*	268
percent of total	25.7%	72.8%	1.5%	*	*	100.0%
Miles of Average Haul						
into the region	2707 *3553*	115	669	90	1428 *2826*	1,286
out of the region	3677 *6000*	115	671	266	1495 *2754*	658
within the region	43	14	40	18	N.A.	22
Billions of Ton-miles						
into the region	222.0 *142.1*	7.8	29.0	0.8	0.2 *0.1*	259.8
percent of total	85.4%	3.0%	11.2%	0.3%	0.1%	100.0%
out of the region	47.8 *42.0*	7.8	6.3	1.6	0.3 *0.2*	63.8
percent of total	74.9%	12.2%	9.9%	2.5%	0.5%	100.0%
within the region	3.0	2.6	0.2	*	*	5.8
percent of total	51.8%	44.8%	3.4%	*	*	100.0%
Millions of Dollars of Freight Revenue or Cost						
into the region	713 *440*	583	483	3	62 *34*	1,844
percent of total	38.6%	31.6%	26.2%	0.2%	3.4%	100.0%
out of the region	285 *239*	583	156	8	83 *50*	1,115
percent of total	25.6%	52.3%	14.0%	0.7%	7.4%	100.0%
within the region	55	1,800	10	*	*	1,865
percent of total	2.9%	96.6%	0.5%	*	*	100.0%
Revenue or Cost per Ton-mile (Cents)						
into the region	.32 *.31*	7.40	1.66	0.39	26.00 *26.31*	0.71
out of the region	.60 *.57*	7.40	2.46	0.39	25.78 *26.32*	1.75
within the region	1.80	69.00	6.11	N.A.	N.A.	32.00

ROMAN: Total Traffic (including foreign)
Italics: Foreign Traffic
*: Less than 50 million ton-miles, $500,000, 500,000 tons or 0.05%
N.A.: Not applicable
SOURCE: Tri-State Transportation Commission

Freight tons, average length of haul, ton-miles, revenue (to the common carrier) or cost (to the private carrier) and ton-mile revenue or cost are shown in this table. Tons listed as "into" and tons shown "out of" each represent tons handled once in the region. Tons "within" the region represent two handlings, one at origin and one at destination. Figures are based in part on preliminary findings.

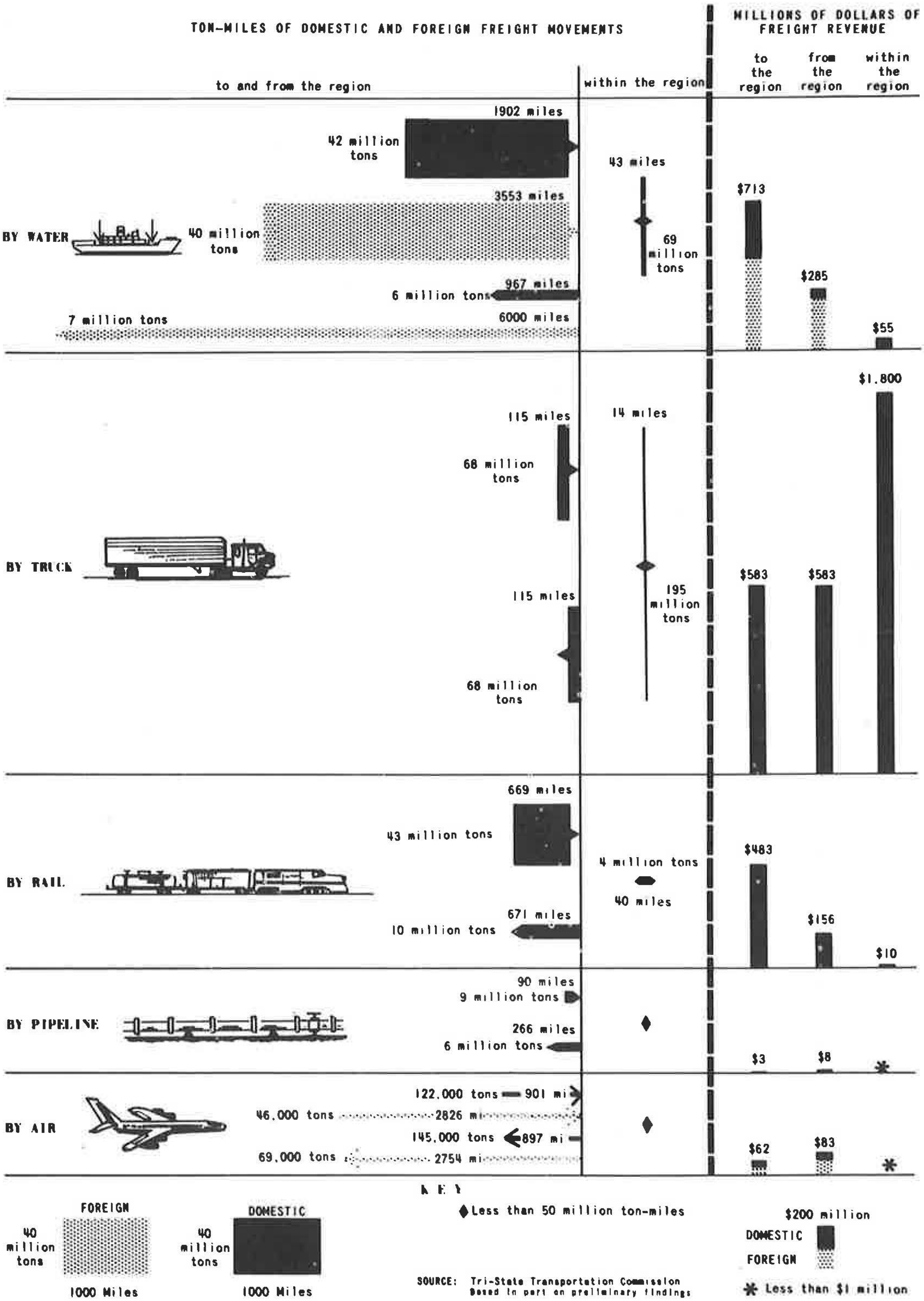


Figure 1. Freight traffic of the Tri-State Region, 1962-1963. Left side shows tons carried as height of bar and average length of haul as horizontal width of bar for 5 different modes of freight carriage. The area of the bar, therefore, measures ton-miles. Arrows show direction of movement into, out of and within the region. Through traffic is not included. On the right, three bars show amount of money spent for each mode's freight service—inbound, outbound and within the region. Note height of money bars compared to size of ton-mile bars.

TABLE 2
FREIGHT GENERATION RATES FOR SELECTED AREAS
(Domestic Intercity Tons)

Category	United States (1962)		Tri-State Region (1962)		Northeast Corridor[a] (1960)	
	Received	Shipped	Received	Shipped	Received	Shipped
Total tons (millions)	4,812	4,812	214	142	625	411
Population (millions)	185.8		18.5		38.1	
Tons per person	25.9	25.9	11.6	7.7	16.4	10.8

[a]Source: Intercity Freight Transportation Requirements of the Washington-Boston Corridor in 1980.

TRUCK SURVEYS: TWO SEPARATE APPROACHES

There are two obvious conclusions about truck freight and measurement of it: (a) the major truck distribution patterns have been identified; and (b) methods for measuring the freight which moves through these channels are not fully satisfactory.

Measuring internal truck freight was deliberately limited to determining the total freight distributed throughout the area by locally registered trucks, and the cargo-carrying characteristics and utilization of the fleet. Experience gained in executing this survey was immeasurably helpful in developing the Goods Movement External Truck Survey (GOMETS) performed in 1964.

The Internal Truck Survey—Three Percent Sample, 13,000 Trucks

Since plans for a Truck-Taxi Survey in the classical form were well under way in early 1963, the obvious move was to put questions on freight traffic into that survey. We investigated and pretested at two levels. Large and small companies were called on for a judgment sample, first with general questions, then with a questionnaire. Since these people—traffic managers, dispatchers, contract truckers, petroleum distributors, etc.—all seemed to know what they carried and how much of it, we selected a random sample of 21 registrations and carried out pretest interviews.

Since many operators make from 30 to 120 stops a day to pick up or deliver freight, and since this project concerned only the total freight distributed, only questions about the total load carried from the base and the total returned to the base were asked. This greatly reduced the burden on the respondent. This system works fine for route deliverymen, but there are a variety of other situations described later.

In the face of this wide variety of truck freight operations, the interviewer was instructed to follow the manual concerning the recording of freight information, but to add footnotes as necessary to explain the action when a load was delivered and another picked up at the same stop. In spite of this, there was a serious control problem. It was solved by assigning one coder to do all the freight coding, and then supervising him closely. Considerable telephone follow-up work was performed by quality control, field supervisors, and the coder himself. The control problem turned out to be manageable for three reasons:

1. About ⅓ of registered trucks do not move on any given day;
2. Up to half of those that move do not carry any freight; and
3. About ¾ of those which carry freight carry only one commodity.

Coding Specifications and Procedures

This is the system set up for organizing freight data obtained in the field.

1. For delivery routes—show total load carried from the base and the total returned to the base.
2. For a serviceman (TV repairs, oil burner repairs, etc.)—show his tools and equipment on the first trip from the base. If additional material is delivered or installed at a stop (such as a hot water heater)—show this item on the trip where it was delivered.
3. On department store deliveries, or on a for-hire truck line's pickup and delivery routes where a different item is handled at each stop, show each item on the trip on which it was delivered, or on the trip at whose origin it was picked up.
4. Unspecified situations—show the article only once, either at the pickup or delivery point.
5. Where more than one commodity is handled at one stop, use a multiple commodity card on which additional commodities are shown.

Commodity Coding

The commodity code used in this project was the Standard Transportation Commodity Code developed by the Bureau of the Budget and used in the 1963 Census of Transportation. Due to the small number of observations expected, detail was carried only to the three-digit level. Appended to this commodity code was a density code developed from the New England Motor Rate Bureau Coordinated Classification No. 11, by which commodities are grouped in classes of 1-5, 5-10, 10-20, and over 20 pounds per cubic foot.

It was found that the STC Code, which was developed for intercity transportation, was not adequate for some common situations that arise in local movements. Codes were therefore added to take care of service equipment (e.g., TV repair), to show the types of people carried by truck (migrant labor, longshoremen, etc.) and to take care of unusual situations, such as the delivery of autos or road equipment under dealer plates.

Evaluation—Internal Truck Freight Survey

Our 20-20 hindsight shows several areas in which the questionnaire and the procedures were lacking.

1. The questionnaire did not adequately provide for recording the freight generated in the distribution patterns described above. As a minimum step, a simple field code ("p" or "d") opposite the description of the item, to show whether it was picked up or delivered, would have saved a great deal of editing time.
2. In most cases we depended on the driver's information. Validation of interviewers' work against company records might have been worthwhile where large companies were concerned.
3. A more extensive program of pretesting would have uncovered more of these various truck distribution patterns, and would have led to clearer questions on freight and clearer coding instructions.

THE GOMETS SURVEY

The one large remaining gap in the developing freight picture was the freight carried by truck across the cordon line. Trucks had been interviewed in the External Survey in 1963, but due to interviewing time restrictions no questions on freight had been asked and only the origin and the destination of the trip that was intercepted at the cordon line were obtained.

Pretesting was carried out at two levels: (a) by interviewing drivers at truck stops, which are gas stations specializing in serving over-the-road trucks; and (b) by setting up and operating a roadside interviewing station under actual conditions.

It was decided to ask the freight-related questions in as simple a form as possible, in essence:

Do you have freight aboard?

If you do:
Where did you get it?
Where are you taking it?

If you don't have freight aboard,
Where was your last previous stop?
Where are you going next?

It became evident during the pretest that interviewers would have to get information on the home base of the truck, the identifying number of the truck and/or trailer, and the dispatcher, for each sample. This was done with little complaint from the drivers. Most control problems arose from these patterns:

1. Multi-stop (deliverymen);
2. Multi-commodity freight, terminal to terminal (Less-Than-Truckload or LTL); and
3. Multi-commodity and multi-stop combined.

In many cases the driver would have a sealed pouch containing the freight bills for his load. He could not or would not open this pouch. Such loads were coded as "miscellaneous freight," and a special subsample of 10 percent received intensive follow-up to develop weight, commodities, and number of shipments.

Coding Specifications and Procedures

This GOMETS questionnaire was much more successful in determining the origins and destinations of the freight than the Truck Survey questionnaire. The principal remaining problem was to identify double crossings of the cordon. The coding rule was to code all of the truck's stops until it crosses the cordon again or gets rid of its load. Where more than one shipment was handled at a stop, special multi-commodity forms and cards were used. The follow-up work required gave rise to a special problem of controlling the flow of schedules between coding and quality control.

Commodity Code

The commodity code used was the same as that used for the internal Truck Survey.

Special Points About the GOMETS Survey

This survey was designed to obtain actual weight of the load, from documents or from the driver's knowledge of his own business. Most other roadside freight surveys have either sampled and weighed loaded trucks and unloaded trucks by vehicle type, or have obtained weight of the load in terms of percent loaded—¼, ½, ¾, or full. In either case the characteristics of particular types of traffic have not been made clear.

This is the first roadside survey to make use of the Standard Transportation Commodity Code which describes the traffic carried by truck much better than the old ICC classification that it replaced.

The sample was drawn to achieve round-the-clock representation from 10 p.m. Sunday night through 10 p.m. Friday night for the cordon line as a whole.

The classification data of the 1963 External Survey was used as a frame. All of the top eleven routes, with ¾ of the tractor-trailer traffic, were sampled. Nineteen of the 71 remaining routes were sampled at a lower rate. The probability of selection of each route was made proportionate to truck traffic. All traffic was classified during each 8-hr shift. Trucks were sampled from the traffic stream at a rate designed to keep the interview site full. This was done because on-the-spot inspection by our sampling group showed no bias from this procedure, and sampling on any other basis was impractical. All four-tire trucks were declared out-of-scope.

Evaluation—GOMETS Survey

The approach developed for this survey is basically sound. Good origin and destination data were obtained for freight crossing the line. The traffic characteristics of the freight crossing the cordon were successfully measured. While information on most loads is obtainable in the field, it is absolutely necessary to obtain vehicle information and to provide a follow-up procedure, much of which can be handled by phone.

The Analysis of Land-Use Linkages

FRANK E. HORTON, Assistant Professor of Geography, University of Iowa, and
PAUL W. SHULDINER, U.S. Department of Commerce and Northwestern University

•EXAMINATION of urban travel activity patterns suggests that person trip sets may be considered to be analogous to a closed circuit movement; that is, the tripmaker generally leaves his home base, makes one or more stops, and returns to the home base. The complexity of the system in which these movements are negotiated makes it extremely difficult to deal explicitly with all relevant or potentially relevant variables. However, this very complexity suggests the use of relatively simple probabilistic models which allow the analyst to vary the components of the system without recourse to an ultra-sophisticated theoretical framework and without being handicapped by concern with a multitude of parameter changes. The Markov chain model is one such simple tool.

Let us assume that each group of land uses within the spatial structure of urban land is a member of a finite collection of states which a tripmaker may choose as a trip end. Assume further that the movement of a tripmaker is part of a process such that if he is in a given state, i, there is some probability, p_{ij}, that he will move to another given state, j, in any given time period, t. Under these assumptions, there is a simple probabilistic model which may be used to describe and analyze such a situation. This simple, time-dependent probability model is known as a finite Markov chain.

A SIMPLE MARKOV CHAIN MODEL

As an example[1] of the use of simple Markov chain models in the analysis of travel characteristics, consider the closed system of three land-use parcels shown in Figure 1, and a tripmaker who may be on any of the three parcels. The rule of the model is that in each time period the tripmaker must leave the particular parcel on which he is located and must either travel to one of the other parcels or return to the original parcel. Using the length of the trip as a criterion (although it would be possible to use any meaningful criterion) and the parcel arrangement shown in Figure 1, we would assume that a tripmaker located on parcel 1 would be most likely to return there, may go to parcel 2, and would be least likely to go to parcel 3. If we assign probabilities to each move, their sum will be unity because the tripmaker must move in some way. The probabilities of ending a trip on each land-use parcel from each of the three possible starting positions may be expressed in a 3 by 3 matrix in which the rows designate land-use parcels of destination, and each row sums to one. A matrix of this type is known as a transition matrix; each of the land-use parcels represents a Markov "state," and each of the elements of the matrix is known as transition probabilities. A transition matrix for the system outlined above could take the following form:

	1	2	3	Σ
1	.5	.4	.1	1.0
2	.3	.5	.2	1.0
3	.1	.3	.6	1.0

Paper sponsored by Committee on Origin and Destination.

[1]This example was adapted from (1), pp. 33-34b.

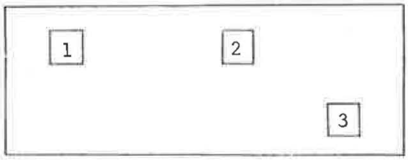


Figure 1. Closed land-use system.

Given the transition probabilities, it must also be known in what state our tripmaker will be at the beginning of the process. This is done by establishing a row or vector in which each element of the row represents the probability of the tripmaker beginning on a given land-use parcel. For example, if we know the tripmaker always begins on parcel 1 in time (0), the initial vector would take the form (1, 0, 0). If, on the other hand, there is an equal probability that the tripmaker will begin on any of the parcels, the initial vector will appear as (1/3, 1/3, 1/3). This row of values is known as a probability vector. The sum of the elements of a probability vector is always equal to one.

Let us assume that there is an equal probability that the tripmaker will start at any one of the three land-use parcels. Multiplying the initial probability vector by the transition matrix will yield a probability vector whose elements define the probability of the tripmaker being in any of the three states or land-use parcels. Thus:

$$(1/3, 1/3, 1/3) \begin{bmatrix} .5 & .4 & .1 \\ .3 & .5 & .2 \\ .1 & .3 & .6 \end{bmatrix} = (.30, .40, .30)$$

The probability vector (.30, .40, .30) becomes an intermediate probability vector establishing the probabilities that the tripmaker is on any of the land-use parcels at the end of time period (1). Furthermore, since we assume the transition probabilities remain constant through time, the probability that the tripmaker will be on land-use X at the end of the second time period may be established by multiplying the new probability vector by the matrix of transition probabilities. The relationship may be expressed by the following equation:

$$\underset{1 \times n}{R^{(t+1)}} = \underset{1 \times n}{R^{(t)}} \times \underset{n \times n}{P}$$

where

- R = the probability row vector with n_j as any element in the vector,
- j = (1 . . . n),
- P = the transition matrix with a_{ij} as any element in the matrix,
- i = (1 . . . n), and
- t = the time period designation, $t = (1 \ldots m)$, and t is not an exponent.

Given these assumptions, there are many questions about a particular process that the Markov chain model can answer. For example, the model can establish the probability that a tripmaker will be on each land-use parcel in any given time period, t. A Markovian approach can also establish whether or not there exists an equilibrium or balance such that the probability of being on each land-use parcel remains constant after a particular time period, t. Stated in another way, this means that it is possible to establish a finite number of time periods in which the tripmaker will move from any Markov "state" to all other Markov "states." This situation exists for every "regular" transition matrix. A "regular" transition matrix is such that there exists some exponent, g, such that P^g has no zero elements. The transition matrix used in our example is a "regular" transition matrix since the matrix P to the first power contains no zero elements. With the knowledge that there exists an equilibrium situation, it is possible to establish the following set of equations:

$$r_1 = .5r_1 + .3r_2 + .1r_3 \qquad (1)$$

$$r_2 = .4r_1 + .5r_2 + .3r_3 \quad (2)$$

$$r_3 = .1r_1 + .2r_2 + .6r_3 \quad (3)$$

Because Eqs. 1, 2 and 3 are not independent equations, another equation must be introduced in order to solve for r_1, r_2, and r_3:

$$1 = r_1 + r_2 + r_3 \quad (4)$$

Any two of Eqs. 1, 2, or 3 and Eq. 4 will constitute a solvable set of equations. The equilibrium vector resulting from our example is (14/46, 19/46, 13/46).

Another interesting problem which may be solved within a Markovian framework is the derivation of "mean first passage time." The mean first passage time is the number of time periods it takes the tripmaker to return to the "state" or land-use parcel from which he started. Thus, it is possible to derive the average number of stops on multipurpose trips.

All of the problems discussed above are capable of being solved through the use of computer programs currently available (2, 3). A more detailed discussion of Markov chain analysis is given by Kemeny and Snell (4).

APPLICATION OF THE MODEL TO THE ANALYSIS OF TRAVEL BEHAVIOR

The finite Markov chain model described above may be adapted to the analysis of travel behavior through the derivation of the following tables, or matrices.

The F Matrix

A matrix of trip origins and destinations is developed from standard O-D data. In this instance, let us assume that this matrix consists of the number of trips, f_{ij}, which start at a given land use, i, and end at a given land use, j. An example of such a matrix is shown below.

Land Use at Origin	Land Use at Destination					
	1	2	......	j	......	n
1	f_{11}	f_{12}	......	f_{1j}	......	f_{1n}
2	f_{21}	f_{22}	......	f_{2j}	......	f_{2n}
i	f_{i1}	f_{i2}	......	f_{ij}	......	f_{in}
n	f_{n1}	f_{n2}	......	f_{nj}	......	f_{nn}

The P Matrix

The proportion of trips, P_{ij}, which go from any origin, i, to each destination, j, can be readily computed by dividing the number of trips from i to j by the total number of trips originating at land-use i. Thus:

$$P_{ij} = \frac{f_{ij}}{\sum_{j=1}^{n} f_{ij}}$$

In the context of the Markov chain model, P_{ij} is defined as a maximum likelihood estimator of the probability of a tripmaker from state i (in this example, land-use i) moving to state j (here, land-use j). The P matrix would look just like the F matrix, except that the relative frequency (i.e., the probability) of trips moving from i to j would be shown, rather than the absolute frequency of such movements. The sum of the P_{ij}'s for any originating land use must equal 1.0; that is,

$$\sum_{j=1}^{n} P_{ij} = 1.0$$

The A Matrix

If all elements of the P matrix are greater than zero, i.e., if

$$P_{ij} > 0 \text{ for all } i, j$$

a limiting matrix, A, can be derived through repeated multiplication of P by itself. Symbolically,

$$A = P^n$$

As a result of repeated multiplication and as a direct consequence of the structure of the P matrix, each row of A is identical. One interpretation of the element, a_j, of any row is that a_j is equal to the expected percentage of tripmakers which will be found at the jth land use at some random time during the day.

The S Matrix

The regular Markov chain can be restructured so that one or more of the diagonal elements, P_{ij}, of the P matrix is made equal to one. Such a Markov chain is called an absorbing chain, since once an entry is made in the k, j cell for which $P_{ij} = 1$, no exit is permitted. By operating upon this modified P matrix another matrix, S, can be developed; the elements, u_{ij}, may be interpreted as the mean number of times a tripmaker will be in a nonabsorbing land use, u_j, given that it starts at a nonabsorbing land use, u_i.

The V Matrix

The A matrix discussed previously can be operated upon to yield a matrix, V, in which each of the elements, v_{ij}, represents the expected variation in the expected number of stops, v_{ij}, at each land use.

The V Vectors

Further manipulation of the S matrix will yield the expected average number of stops and variances of trips with a particular land use designated as the first stop. Further operations on S will also lead to the expected mean number of stops per trip set, and the variance for the system as a whole. All of the operations outlined can be carried out by using two programs developed by Marble (3).

TABLE 1

LAND USE AT ORIGIN RELATED TO LAND USE AT DESTINATION—WACO, 1964

From \ To	1. HOME	2. AGFOFI	3. MFGOUR	4. MFGNDU	5. TRCOOI	6. COMRET	7. COMSVC	8. WHOSAL	9. PUBQPU	10. PUBOPE	Sum
1. HOME	0	50	359	486	307	3,729	1,547	250	5,511	283	12,432
2. AGFOFI	56	4	0	2	0	15	5	0	9	0	91
3. MFGDUR	338	0	9	10	3	55	20	4	18	2	459
4. MFGNDU	436	0	10	12	2	82	33	7	25	2	609
5. TRCOOI	286	0	6	4	30	71	25	2	34	4	462
6. COMRET	4,174	11	44	53	52	1,148	285	51	333	52	6,203
7. COMSVC	1,356	7	12	20	21	454	210	20	187	18	2,305
8. WHOSAL	251	2	5	5	3	60	13	17	20	2	378
9. PUBQPU	4,681	17	21	56	36	645	258	29	794	40	6,577
10. PUBOPE	311	1	1	0	3	42	13	4	41	34	450

Note: Definitions of abbreviations used in Tables 1 through 6:

1. HOME = Home
2. AGFOFI = Agriculture, Forestry, Fishing
3. MFGDUR = Manufacturing, Durable
4. MFGNDU = Manufacturing, Nondurable
5. TRCOOI = Transportation, Communication, Other Industrial
6. COMRET = Commercial, Retail
7. COMSVC = Commercial, Service
8. WHOSAL = Wholesale
9. PUBQPU = Public and Quasi-Public
10. PUBOPE = Public Open Space

TABLE 2

10 × 10 LAND-USE MATRIX OF TRANSITION PROBABILITIES—WACO, 1964

From \ To	1. HOME	2. AGFOFI	3. MFGDUR	4. MFGNDU	5. TRCOOI	6. COMRET	7. COMSVC	8. WHOSAL	9. PUBQPU	10. PUBOPE
1. HOME	.00	.00[a]	.03	.04	.02	.30	.12	.02	.44	.02
2. AGFOFI	.62	.04	.00	.02	.00	.16	.05	.00	.10	.00
3. MFGDUR	.74	.00	.02	.02	.01	.12	.04	.01	.04	.00
4. MFGNDU	.72	.00	.02	.02	.00[a]	.13	.05	.01	.04	.00[a]
5. TRCOOI	.62	.00	.01	.01	.06	.15	.05	.00[a]	.07	.01
6. COMRET	.67	.00[a]	.01	.01	.01	.19	.05	.01	.05	.01
7. COMSVC	.59	.00[a]	.01	.01	.01	.20	.09	.01	.08	.01
8. WHOSAL	.66	.01	.01	.01	.01	.16	.03	.04	.05	.01
9. PUBQPU	.71	.00[a]	.00[a]	.01	.01	.10	.04	.00[a]	.12	.01
10. PUBOPE	.69	.00[a]	.00[a]	.00	.01	.09	.03	.01	.09	.08

[a]Less than .005.

Note: See Table 1 for definition of abbreviations.

TABLE 3

ONE ROW OF THE LIMITING MATRIX A (PERCENTAGE OF THE GROUP THAT WILL BE FOUND IN A PARTICULAR STATE AT SOME RANDOM TIME DURING THE DAY) GENERATED FROM MATRIX P OF THE 10 × 10 LAND-USE MATRIX

1 HOME	2 AGFOFI	3 MFGDUR	4 MFGNDU	5 TRCOOI
40.6	0.3	1.5	2.1	1.5
6 COMRET	**7 COMSVC**	**8 WHOSAL**	**9 PUBQPU**	**10 PUBOPE**
20.8	7.8	1.3	22.8	1.4

Note: See Table 1 for definition of abbreviations.

TABLE 4

PREDICTED NUMBER OF STOPS ON WACO PERSON TRIPS

Starting State	Mean	Variance
2. AGFOFI	1.6	.8
3. MFGDUR	1.4	.6
4. MFGNDU	1.4	.7
5. TRCOOI	1.6	.8
6. COMRET	1.5	.7
7. COMSVC	1.6	.9
8. WHOSAL	1.5	.7
9. PUBQPU	1.4	.7
10. PUBOPE	1.5	.7
SYSTEM	1.5	.7

Note: See Table 1 for definition of abbreviations.

ANALYSIS AND SOME PRELIMINARY RESULTS

The models outlined were applied to land-use and trip data for the Waco, Texas, area. Four F matrices were developed: (a) a 10 × 10 trip purpose matrix; (b) a 10 × 10 major land-use matrix; (c) a 21 × 21 commercial land-use matrix; and (d) a 28 × 28 major and commercial land-use matrix. The results generated by the application of the two Markov models to these data are given in Tables 1 through 18. The tables corresponding to the matrices described in the previous section are as follows: F matrices—Tables 1, 7, and 13; P matrices—Tables 2, 8, and 14; 1 row of the A matrices—Tables 3, 9, and 15; the vectors, V'—Tables 4, 10, and 18; S matrices—Tables 5, 11, and 16; and V matrices—Tables 6, 12, and 17. Results using the 28 × 28 land-use matrix are not illustrated in the tables.

TABLE 5

PREDICTED NUMBER OF STOPS BY LAND USE AT FIRST STOP TRIP—WACO, 1964

First Land Use \ Stops At	2. AGFOFI	3. MFGDUR	4. MFGNDU	5. TRCOOI	6. COMRET	7. COMSVC	8. WHOSAL	9. PUBQPU	10. PUBOPE
2. AGFOFI	1.05	.00[a]	.03	.00[a]	.26	.08	.00[a]	.14	.00[a]
3. MFGDUR	.00[a]	1.02	.03	.01	.18	.06	.01	.07	.01
4. MFGNDU	.00[a]	.02	1.02	.01	.20	.08	.02	.07	.01
5. TRCOOI	.00[a]	.02	.01	1.07	.25	.08	.01	.12	.01
6. COMRET	.00[a]	.01	.01	.01	1.27	.07	.01	.09	.01
7. COMSVC	.00[a]	.01	.01	.01	.30	1.12	.01	.13	.01
8. WHOSAL	.01	.02	.02	.01	.24	.06	1.05	.09	.01
9. PUBQPU	.00[a]	.01	.01	.01	.16	.06	.01	1.16	.01
10. PUBOPE	.00[a]	.00[a]	.00[a]	.01	.16	.05	.01	.13	1.08

[a]Less than .005.

Note: See Table 1 for definition of abbreviations.

TABLE 6

VARIANCE IN NUMBER OF STOPS BY LAND USE AT FIRST STOP ON TRIP—WACO, 1964

First Land Use \ Stops At	2. AGFOFI	3. MFGDUR	4. MFGNDU	5. TRCOOI	6. COMRET	7. COMSVC	8. WHOSAL	9. PUBQPU	10. PUBOPE
2. AGFOFI	.05	.01	.04	.01	.26	.11	.10	.17	.01
3. MFGDUR	.00[a]	.02	.03	.01	.18	.08	.02	.09	.01
4. MFGNDU	.00[a]	.02	.07	.01	.20	.09	.02	.09	.01
5. TRCOOI	.00[a]	.02	.02	.07	.25	.11	.01	.14	.02
6. COMRET	.00[a]	.01	.02	.02	.29	.09	.02	.11	.02
7. COMSVC	.01	.01	.02	.02	.29	.11	.02	.16	.02
8. WHOSAL	.01	.02	.02	.02	.23	.08	.05	.11	.01
9. PUBQPU	.00[a]	.01	.02	.01	.17	.08	.01	.15	.01
10. PUBOPE	.00[a]	.01	.01	.01	.17	.07	.02	.14	.09

[a]Less than .005.

Note: See Table 1 for definition of abbreviations.

TABLE 7

TRIP PURPOSE AT ORIGIN RELATED TO TRIP PURPOSE AT DESTINATION—WACO, 1964

From \ To	1 HOME	2 WORK	3 PERBUS	4 MEDDEN	5 SCHOOL	6 SOCREC	7 CHMODE	8 EATMEA	9 SHOP	10 SERPAS	Sum
1 HOME	00	3230	1231	181	2031	2444	23	294	1800	1953	13,187
2 WORK	2920	701	199	24	11	106	1	474	205	205	4,891
3 PERBUS	1113	145	430	12	16	153	2	70	299	89	2,330
4 MEDDEN	157	13	18	8	2	25	1	8	40	10	282
5 SCHOOL	1574	32	25	10	22	122	3	54	57	41	1,940
6 SOCREC	2617	54	108	18	49	652	1	86	216	156	3,957
7 CHMODE	18	2	0	2	5	2	0	1	1	2	33
8 EATMEA	340	440	50	2	44	85	1	6	72	53	1,093
9 SHOP	2216	60	173	12	4	211	0	71	686	124	3,557
10 SERPAS	1680	368	116	19	59	168	1	63	200	641	3,315

Note: Definitions of abbreviations used in Tables 7 through 12:

1 HOME = Home
2 WORK = Work
3 PERBUS = Personal Business
4 MEDDEN = Medical Dental
5 SCHOOL = School
6 SOCREC = Social Recreation
7 CHMODE = Change Mode
8 EATMEA = Eat Meal
9 SHOP = Shop
10 SERPAS = Serve Passenger

TABLE 8

10 × 10 PURPOSE MATRIX OF TRANSITION PROBABILITIES—WACO, 1964

From \ To	1 HOME	2 WORK	3 PERBUS	4 MEDDEN	5 SCHOOL	6 SOCREC	7 CHMODE	8 EATMEA	9 SHOP	10 SERPAS
1 HOME	.00	.24	.09	.01	.15	.19	.00[a]	.02	.14	.15
2 WORK	.60	.14	.04	.00[a]	.00[a]	.02	.00[a]	.10	.04	.05
3 PERBUS	.48	.06	.18	.01	.01	.07	.00[a]	.03	.13	.04
4 MEDDEN	.56	.05	.06	.03	.01	.09	.00[a]	.03	.14	.04
5 SCHOOL	.81	.02	.01	.01	.01	.06	.00[a]	.03	.03	.02
6 SOCREC	.66	.01	.03	.00[a]	.01	.16	.00[a]	.02	.05	.04
7 CHMODE	.56	.06	.00	.06	.16	.06	.00	.03	.00	.06
8 EATMEA	.31	.40	.05	.00[a]	.04	.08	.00[a]	.01	.07	.05
9 SHOP	.62	.02	.05	.00[a]	.00[a]	.06	.00	.02	.19	.03
10 SERPAS	.51	.11	.03	.01	.02	.05	.00[a]	.02	.06	.19

[a]Less than .005.

Note: See Table 7 for definition of abbreviations.

TABLE 9

ONE ROW OF THE LIMITING MATRIX A (PERCENTAGE OF THE GROUP THAT WILL BE FOUND IN A PARTICULAR STATE AT SOME RANDOM TIME DURING THE DAY) GENERATED FROM MATRIX P OF THE 10 × 10 PURPOSE MATRIX—WACO, 1964

1	2	3	4	5
HOME	WORK	PERBUS	MEDDEN	SCHOOL
11.3	14.4	6.7	.8	6.3
6	7	8	9	10
SOCREC	CHMODE	EATMEA	SHOP	SERPAS
11.3	.1	3.3	10.2	9.5

TABLE 10

PREDICTED NUMBER OF STOPS ON WACO PERSON TRIPS

Starting State	Mean	Variance
2 WORK	1.8	1.4
3 PERBUS	1.9	1.5
4 MEDDEN	1.8	1.3
5 SCHOOL	1.3	.7
6 SOCREC	1.6	1.1
7 CHMODE	1.7	1.2
8 EATMEA	2.2	1.6
9 SHOP	1.7	1.2
10 SERPAS	1.9	1.5
SYSTEM	1.7	1.3

Note: See Table 7 for definition of abbreviations.

TABLE 11

EXPECTED NUMBER OF STOPS BY FIRST PURPOSE AND TYPE OF STOP—WACO, 1964

First Purpose \ Stops At	2 WORK	3 PERBUS	4 MEDDEN	5 SCHOOL	6 SOCREC	7 CHMODE	8 EATMEA	9 SHOP	10 SERPAS
2 WORK	1.25	.08	.01	.01	.07	.00[a]	.13	.10	.10
3 PERBUS	.14	1.26	.01	.02	.13	.00[a]	.06	.23	.09
4 MEDDEN	.11	.11	1.03	.01	.15	.00[a]	.05	.23	.08
5 SCHOOL	.05	.03	.01	1.02	.09	.00[a]	.04	.06	.04
6 SOCREC	.05	.06	.01	.02	1.22	.00[a]	.04	.10	.07
7 CHMODE	.12	.03	.07	.16	.12	1.00	.06	.05	.11
8 EATMEA	.53	.11	.011	.05	.14	.00[a]	1.07	.15	.12
9 SHOP	.06	.09	.01	.01	.11	.00[a]	.04	1.27	.07
10 SERPAS	.20	.08	.01	.03	.11	.00[a]	.05	.13	1.27

[a]Less than .005.

Note: See Table 7 for definition of abbreviations.

TABLE 12

VARIANCE IN NUMBER OF STOPS BY FIRST PURPOSE AND TYPE OF STOP—WACO, 1964

First Purpose \ Stops At	2 WORK	3 PERBUS	4 MEDDEN	5 SCHOOL	6 SOCREC	7 CHMODE	8 EATMEA	9 SHOP	10 SERPAS
2 WORK	.02	.12	.01	.03	.12	.00[a]	.17	.16	.13
3 PERBUS	.22	.27	.02	.03	.19	.00[a]	.11	.28	.14
4 MEDDEN	.17	.14	.03	.03	.18	.00[a]	.09	.24	.18
5 SCHOOL	.08	.04	.01	.01	.10	.00[a]	.52	.08	.05
6 SOCREC	.10	.08	.01	.03	.23	.00[a]	.06	.14	.09
7 CHMODE	.17	.07	.07	.17	.15	.00[a]	.09	.11	.12
8 EATMEA	.50	.16	.02	.07	.21	.00[a]	.02	.24	.18
9 SHOP	.12	.11	.01	.02	.14	.00[a]	.07	.29	.10
10 SERPAS	.26	.12	.02	.04	.16	.00[a]	.10	.19	.29

[a]Less than .005.

Note: See Table 7 for definition of abbreviations.

TABLE 13

LAND USE AT ORIGIN RELATED TO LAND USE AT DESTINATION, SHOPPING TRIPS—WACO, 1964

From \ To	1 HOME	2 FOODRU	3 EATDRI	4 GENMER	5 APPACC	6 FUHFHA	7 MOTVAC	8 GASSTA	9 LUBDHA	10 LIQBER	11 MISRET	12 FININS	13 PERSER	14 BUSSER	15 AUTOGA	16 MISREP	17 AMUREC	18 MEDDEN	19 OTHPRO	20 OFFBLD	21 MISSER
1 HOME	0	1,379	641	485	95	54	276	216	106	19	459	331	442	42	88	37	174	239	31	46	28
2 FOODRU	1,716	159	25	28	7	2	4	17	3	2	26	18	28	0	6	1	5	8	0	2	0
3 EATDRI	706	39	52	30	3	6	13	4	7	1	16	25	12	7	7	1	11	8	1	2	0
4 GENMER	549	65	27	66	21	7	5	1	4	1	39	12	9	2	2	0	1	4	0	0	1
5 APPACC	83	18	6	26	14	1	1	1	0	0	8	1	3	0	0	0	0	3	0	0	0
6 FUHFHA	52	5	9	7	0	3	2	3	3	0	3	1	1	0	0	1	1	0	0	0	0
7 MOTVAC	253	24	22	7	1	4	21	8	2	0	2	11	7	1	1	1	0	4	0	2	1
8 GASSTA	226	23	8	4	2	1	6	4	2	0	6	9	2	0	1	3	5	0	0	0	2
9 LUBDHA	101	6	7	3	1	1	3	1	10	0	3	-3	4	0	0	1	0	2	0	2	1
10 LIQBER	23	1	2	0	0	1	1	0	0	2	0	1	0	0	0	0	0	0	0	0	0
11 MISRET	465	40	22	36	8	2	5	6	5	0	45	8	19	0	1	2	0	5	0	0	2
12 FINIS	294	22	21	26	4	1	18	7	3	1	19	38	15	0	4	0	1	10	0	0	0
13 PERSER	442	50	20	21	1	1	6	12	2	0	14	11	25	0	0	0	1	6	2	5	2
14 BUSSER	36	4	5	3	0	0	0	0	0	0	1	4	1	0	1	0	1	0	0	0	0
15 AUTOGA	79	7	6	4	0	1	2	0	4	2	2	6	1	1	5	2	4	3	0	0	0
16 MISREP	44	1	2	0	1	0	0	0	0	0	2	0	0	2	0	2	0	0	0	0	0
17 AMUREC	154	6	42	0	0	0	1	2	0	1	4	0	5	0	1	0	5	0	0	1	0
18 MEDDEN	201	28	16	15	2	0	2	1	0	0	14	9	8	0	2	0	1	11	1	0	1
19 OTHPRO	32	1	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0
20 OFFBLD	45	6	1	3	0	0	0	0	0	0	0	1	2	0	0	0	1	1	0	2	0
21 MISSER	29	2	2	3	0	0	1	1	0	0	1	0	0	0	0	0	0	1	0	0	2

Note: Definitions of abbreviations used in Tables 13 through 18:

1 HOME = Home
2 FOODRU = Food and Drugs
3 EATDRI = Eating and Drinking Establishments
4 GENMER = General Merchandise
5 APPACC = Apparel and Accessories Stores
6 FUHFHA = Furniture, House Furnishings and Household Appliance Stores
7 MOTVAC = Motor Vehicle and Accessories: Boats and Motors, and Farm Equipment
8 GASSTA = Gasoline Service Stations
9 LUBDHA = Lumber, Building Material and Hardware Stores
10 LIQBER = Retail Liquor and Beer Establishments
11 MISRET = Miscellaneous Retail Stores
12 FININS = Finance, Insurance and Real Estate
13 PERSER = Personal Services
14 BUSSER = Business Services
15 AUTOGA = Automobile Repair and Services, Garages, and Marine Servicing
16 MISREP = Miscellaneous Repair Services
17 AMUREC = Indoor Amusement and Recreation
18 MEDDEN = Medical and Health Services
19 OTHPRO = Other Professional Services
20 OFFBLD = Office Buildings Not Elsewhere Classified
21 MISSER = Miscellaneous Services

TABLE 14

COMMERCIAL LAND-USE MATRIX OF TRANSITION PROBABILITIES—WACO, 1964

From \ To	1 HOME	2 FOODRU	3 EATDRI	4 GENMER	5 APPACC	6 FUHFHA	7 MOTVAC	8 GASSTA	9 LUBDHA	10 LIQBER	11 MISRET	12 FININS	13 PERSER	14 BUSSER	15 AUTOGA	16 MISREP	17 AMUREC	18 MEDDEN	19 OTHPRO	20 OFFBLD	21 MISSER
1 HOME	.00	.27	.12	.09	.02	.01	.05	.04	.02	.00[a]	.09	.06	.09	.01	.02	.01	.03	.05	.01	.01	.01
2 FOODRU	.83	.08	.01	.01	.00[a]	.00[a]	.00[a]	.01	.00[a]	.00[a]	.01	.01	.01	.00	.00[a]	.00[a]	.00[a]	.00[a]	.00	.00[a]	.00
3 EATDRI	.62	.06	.08	.05	.00[a]	.01	.02	.01	.01	.00[a]	.02	.04	.02	.01	.01	.00[a]	.02	.01	.00[a]	.00[a]	.00
4 GENMER	.67	.08	.03	.08	.03	.01	.01	.00[a]	.00[a]	.00[a]	.05	.01	.01	.00[a]	.00[a]	.00	.00[a]	.00[a]	.00	.00	.00[a]
5 APPACC	.50	.11	.04	.16	.08	.01	.01	.01	.00	.00	.05	.01	.02	.00	.00	.00	.00	.00	.00	.00	.00
6 FUHFHA	.57	.05	.10	.08	.00	.03	.02	.03	.03	.00	.03	.01	.01	.00	.00	.01	.01	.00	.00	.00	.00
7 MOTVAC	.68	.06	.06	.02	.00[a]	.01	.06	.02	.01	.00	.01	.03	.02	.00[a]	.00[a]	.00[a]	.00	.01	.00	.01	.00[a]
8 GASSTA	.74	.06	.03	.01	.01	.00[a]	.02	.01	.01	.00	.02	.03	.01	.00	.00[a]	.01	.02	.00	.00	.00	.01
9 LUBDHA	.68	.04	.05	.02	.01	.01	.02	.01	.07	.00	.02	.02	.03	.00	.00	.01	.00	.01	.00	.01	.01
10 LIQBER	.74	.03	.06	.00	.00	.03	.03	.00	.00	.06	.00	.03	.00	.00	.00	.00	.00	.00	.00	.00	.00
11 MISRET	.69	.06	.03	.05	.01	.00[a]	.01	.01	.01	.00	.07	.01	.03	.00	.00[a]	.00[a]	.00	.01	.00	.00	.00[a]
12 FINIS	.61	.05	.04	.05	.01	.00[a]	.03	.01	.01	.00[a]	.04	.08	.03	.00	.09	.00	.00[a]	.02	.00	.00	.00
13 PERSER	.71	.08	.03	.03	.00[a]	.00[a]	.01	.02	.00[a]	.00	.02	.02	.04	.00	.00	.00	.00[a]	.01	.00[a]	.01	.00[a]
14 BUSSER	.64	.07	.09	.05	.00	.00	.00	.00	.00	.00	.02	.07	.02	.00	.02	.00	.02	.00	.00	.00	.00
15 AUTOGA	.61	.05	.05	.03	.00	.01	.02	.00	.03	.02	.02	.05	.01	.01	.04	.02	.03	.02	.00	.00	.00
16 MISREP	.81	.02	.04	.00	.02	.00	.00	.00	.00	.00	.04	.00	.00	.03	.00	.03	.00	.00	.00	.00	.00
17 AMUREC	.69	.03	.19	.00	.00	.00	.00[a]	.01	.00	.00[a]	.02	.00	.02	.00	.00[a]	.00	.02	.00	.00	.00[a]	.00
18 MEDDEN	.64	.09	.05	.05	.01	.00	.01	.00[a]	.00	.00	.04	.03	.03	.00	.01	.00	.00[a]	.03	.00[a]	.00	.00[a]
19 OTHPRO	.84	.03	.05	.00	.00	.00	.03	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.05	.00	.00
20 OFFBLD	.73	.10	.02	.05	.00	.00	.00	.00	.00	.00	.00	.02	.03	.00	.00	.00	.02	.02	.00	.03	.00
21 MISSER	.69	.05	.05	.07	.00	.00	.02	.02	.00	.00	.02	.01	.00	.00	.00	.00	.00	.02	.00	.00	.05

[a]Less than .005.

Note: See Table 13 for definition of abbreviations.

TABLE 15

ONE ROW OF THE LIMITING MATRIX A (PERCENTAGE OF THE GROUP THAT WILL BE FOUND IN A PARTICULAR STATE AT SOME RANDOM TIME DURING THE DAY) GENERATED FROM MATRIX P OF THE 21 × 21 LAND USE MATRIX

1	2	3	4	5	6	7
HOME	FOODRU	EATDRI	GENMER	APPACC	FUHFHA	MOTVAC
41.7	14.9	7.6	6.1	1.3	.7	3.0
8	9	10	11	12	13	14
GASSTA	LUBDHA	LIQBER	MISTRET	FININS	PERSER	BUSSER
2.2	1.2	.2	5.3	4.0	4.8	.5
15	16	17	18	19	20	21
AUTOGA	MISREP	AMUREC	MEDDEN	OTHPRO	OFFBLD	MISSER
1.0	.4	1.7	2.4	.3	.5	.3

Note: See Table 13 for definition of abbreviations.

10 × 10 Land-Use Matrix

A glance at the matrix of transition probabilities (Table 2) indicates that the commercial retail land uses are highly linked with all other land uses. Quite unexpectedly the linkages between public-quasi-public land uses and the other groups are also quite high. This is most likely a measure of the magnitude of employment at the air base in Waco, and also a function of the number of school trips in the area. Within the Waco area 74 out of every 100 trips from the home may be expected to end at a commercial retail or public-quasi-public land use. If one adds the probability of going to commercial service land use, approximately 85 percent of the trips emanating from home will end at three of the ten land-use classes. As one would expect, the majority of the trips emanating from non-home bases terminate at the home base.

Table 3 gives the expected percentage of the tripmakers in any one state. The large number of trip ends concentrated in the four land uses of home, commercial retail, commercial service, and public-quasi-public are reflected in the expected percentage of trip makers found in these states. Although the figures in Table 3 indicate the expected percentage of tripmakers in any one state at some random time during the day, it would be more reasonable to interpret these figures reflecting percentages of tripmakers in any one state during the working hours of the day. We can expect that of the universe of the tripmakers in Waco, fully 92 percent will be found on the four land uses named. Table 4 gives the expected mean number of stops on multiple-leg trips starting from the nine non-home land uses, and the expected variances in number of stops before terminating at the home. This particular breakdown of the land uses yields values of total expected stops which show little variation between land-use classes. Once again, the distribution of expected trip ends (Table 5) is highly skewed in favor of commercial retail, commercial service, and public-quasi-public land uses. For example, trips which have as a first stop wholesale land use, have an expected 0.05 stops at wholesale land uses after the initial stop; 0.24 stops are expected to terminate at commercial retail land uses given the same initial stop.

10 × 10 Purpose Matrix

Table 8 gives the limiting transition probabilities for the 10 × 10 purpose matrix. Of trips beginning from the home base, there is an expected probability of 0.24 that people are leaving for work. The majority of trips have home as a purpose. However, the purposes of eat meal and personal business have probabilities of only 0.31 and 0.48, respectively, that the next purpose will be to go home. These low figures indicate the large number of eat meal trips which ultimately end up back at the work place. In personal business it illustrates that a large number of stops are made on personal trips (in Table 10 personal business has one of the largest number of expected stops). Other highly linked purposes are as follows:

1. Work and work, eat meal;
2. Personal business and personal business, social-recreation;
3. Medical dental and social-recreation;

TABLE 16

PREDICTED NUMBER OF STOPS BY LAND USE AT FIRST STOP ON SHOPPING TRIP—WACO, 1964

First Purpose \ Stops At	2 FOODRU	3 EATDRI	4 GENMER	5 APPACC	6 FUHFHA	7 MOTVAC	8 GASSTA	9 LUBDHA	10 LIQBER	11 MISRET	12 FININS	13 PERSER	14 BUSSER	15 AUTOGA	16 MISREP	17 AMUREC	18 MEDDEN	19 OTHPRO	20 OFFBLD	21 MISSER
2 FOODRU	1.09	.02	.02	.01	.00[a]	.00[a]	.01	.00[a]	.00[a]	.02	.01	.01	.00[a]	.00[a]	.00[a]	.00[a]	.01	.00[a]	.00[a]	.00[a]
3 EATDRI	.10	1.11	.07	.01	.01	.03	.01	.02	.00[a]	.04	.05	.03	.01	.01	.00[a]	.02	.02	.00[a]	.00[a]	.00[a]
4 GENMER	.11	.05	1.11	.03	.01	.01	.00[a]	.01	.00[a]	.06	.02	.02	.00[a]	.00[a]	.00[a]	.00[a]	.01	.00[a]	.00[a]	.00[a]
5 APPACC	.17	.06	.20	1.10	.01	.01	.01	.00[a]	.00[a]	.08	.02	.03	.00[a]	.00[a]	.00[a]	.00[a]	.02	.00[a]	.00[a]	.00[a]
6 FUHFHA	.10	.13	.10	.01	1.04	.03	.04	.04	.00[a]	.05	.03	.02	.00[a]	.00[a]	.01	.02	.00[a]	.00[a]	.00[a]	.00[a]
7 MOTVAC	.10	.08	.04	.01	.01	1.07	.03	.01	.00	.02	.04	.03	.00[a]	.00[a]	.00[a]	.00[a]	.02	.00[a]	.01	.00[a]
8 GASSTA	.10	.04	.03	.01	.00[a]	.03	1.02	.01	.00[a]	.03	.04	.01	.00[a]	.00[a]	.01	.02	.00[a]	.00[a]	.00[a]	.01
9 LUBDHA	.07	.07	.04	.01	.01	.03	.01	1.07	.00[a]	.03	.03	.04	.00[a]	.00[a]	.01	.00[a]	.02	.00[a]	.02	.00
10 LIQBER	.05	.09	.01	.00[a]	.04	.04	.00[a]	.00[a]	1.07	.01	.04	.01	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]
11 MISRET	.09	.05	.07	.02	.01	.01	.01	.01	.00[a]	1.08	.02	.04	.00[a]	.00[a]	.00[a]	.00[a]	.01	.00[a]	.00[a]	.00[a]
12 FININS	.08	.07	.08	.01	.00[a]	.05	.02	.01	.00[a]	.06	1.10	.04	.00[a]	.01	.00[a]	.00[a]	.03	.00[a]	.00[a]	.00[a]
13 PERSER	.11	.05	.05	.00[a]	.00[a]	.01	.02	.01	.00[a]	.03	.03	1.05	.00[a]	.00[a]	.00[a]	.00[a]	.01	.00[a]	.01	.00[a]
14 BUSSER	.10	.11	.08	.00[a]	.00[a]	.01	.00[a]	.00[a]	.00[a]	.03	.09	.03	1.00	.02	.00[a]	.02	.01	.00[a]	.00[a]	.00[a]
15 AUTOGA	.09	.08	.05	.00[a]	.01	.02	.00[a]	.04	.02	.03	.06	.02	.01	1.04	.02	.04	.03	.00[a]	.00[a]	.00[a]
16 MISREP	.04	.05	.01	.02	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.05	.01	.00[a]	.04	.00[a]	1.04	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]
17 AMUREC	.06	.22	.02	.00[a]	.00[a]	.01	.01	.00[a]	.01	.03	.01	.03	.00[a]	.01	.00[a]	1.03	.00[a]	.00[a]	.01	.00[a]
18 MEDDEN	.13	.07	.07	.01	.00[a]	.01	.01	.00[a]	.00[a]	.06	.04	.04	.00[a]	.01	.00[a]	.01	1.04	.00[a]	.00[a]	.00[a]
19 OTHPRO	.04	.06	.01	.00[a]	.00[a]	.03	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	1.06	.00[a]	.00[a]
20 OFFBLD	.12	.03	.06	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.01	.02	.04	.00[a]	.00[a]	.00[a]	.02	.02	.00[a]	1.03	.00[a]
21 MISSER	.08	.07	.09	.00[a]	.00[a]	.03	.03	.00[a]	.00[a]	.04	.01	.00[a]	.00[a]	.00[a]	.00[a]	.03	.03	.00[a]	.00[a]	1.05

[a]Less than .005.

Note: See Table 13 for definition of abbreviations.

TABLE 17

VARIANCE IN NUMBER OF STOPS BY LAND USES AT FIRST STOP ON SHOPPING TRIP—WACO, 1964

First Purpose \ Stops At	1 HOME	2 FOODRU	3 EATDRI	4 GENMER	5 APPACC	6 FUHFHA	7 MOTVAC	8 GASSTA	9 LUBDHA	10 LIQBER	12 FININS	13 PERSER	14 BUSSER	15 AUTOGA	16 MISREP	17 AMUREC	18 MEDDEN	19 OTHPRO	20 OFFBLD	21 MISSER
2 FOODRU	.09	.03	.03	.01	.00[a]	.01	.01	.00[a]	.00[a]	.02	.02	.02	.00[a]	.00[a]	.00[a]	.00[a]	.01	.00[a]	.00[a]	.00[a]
3 EATDRI	.13	.09	.09	.01	.02	.04	.02	.02	.00[a]	.06	.06	.04	.01	.02	.00[a]	.02	.02	.00[a]	.01	.00[a]
4 GENMER	.14	.07	.09	.04	.01	.02	.01	.01	.00[a]	.08	.03	.03	.00[a]	.01	.00[a]	.01	.01	.00[a]	.00[a]	.00[a]
5 APPACC	.21	.09	.22	.10	.01	.02	.02	.00[a]	.00[a]	.10	.03	.05	.00[a]	.01	.00[a]	.01	.03	.00[a]	.00[a]	.00[a]
6 FUHFHA	.14	.15	.12	.01	.03	.04	.05	.04	.00[a]	.07	.04	.04	.01	.00[a]	.01	.02	.01	.00[a]	.00[a]	.00[a]
7 MOTVAC	.12	.09	.05	.01	.02	.06	.03	.01	.00[a]	.03	.05	.04	.01	.01	.00[a]	.01	.02	.00[a]	.00[a]	.00[a]
8 GASSTA	.11	.06	.04	.01	.00[a]	.03	.01	.01	.00[a]	.04	.04	.02	.00[a]	.01	.01	.02	.01	.00[a]	.00[a]	.01
9 LUBDHA	.10	.08	.06	.01	.01	.03	.02	.08	.00[a]	.05	.04	.05	.00[a]	.00[a]	.01	.01	.02	.00[a]	.02	.00[a]
10 LIQBER	.08	.10	.03	.00[a]	.04	.05	.01	.01	.07	.02	.05	.01	.00[a]	.00[a]	.00[a]	.01	.01	.00[a]	.00[a]	.00[a]
11 MISRET	.12	.06	.09	.02	.01	.02	.02	.01	.00[a]	.07	.03	.05	.00[a]	.01	.00[a]	.00[a]	.02	.00[a]	.00[a]	.00[a]
12 FINIS	.13	.09	.10	.02	.01	.06	.03	.01	.00[a]	.07	.09	.06	.00[a]	.01	.00[a]	.01	.03	.00[a]	.00[a]	.00[a]
13 PERSER	.13	.06	.06	.01	.01	.02	.03	.01	.00[a]	.04	.03	.04	.00[a]	.00[a]	.00[a]	.01	.02	.00[a]	.01	.00[a]
14 BUSSER	.13	.13	.09	.01	.01	.02	.01	.01	.00[a]	.05	.09	.04	.00[a]	.02	.00[a]	.03	.01	.00[a]	.00[a]	.00[a]
15 AUTOGA	.12	.10	.07	.01	.01	.03	.01	.04	.02	.05	.07	.03	.01	.04	.02	.04	.04	.00[a]	.00[a]	.00[a]
16 MISREP	.05	.06	.03	.02	.00[a]	.01	.00[a]	.00[a]	.00[a]	.05	.02	.01	.04	.00[a]	.04	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]
17 AMUREC	.09	.21	.04	.01	.01	.02	.02	.00[a]	.00[a]	.04	.03	.04	.01	.01	.01	.02	.01	.00[a]	.01	.00[a]
18 MEDDEN	.15	.09	.09	.01	.01	.01	.01	.01	.00[a]	.07	.05	.05	.00[a]	.01	.00[a]	.01	.04	.00[a]	.00[a]	.00[a]
19 OTHPRO	.05	.07	.01	.00[a]	.00[a]	.03	.00[a]	.00[a]	.00[a]	.00[a]	.01	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.00[a]	.06	.00[a]	.00[a]
20 OFFBLD	.14	.04	.07	.01	.00[a]	.01	.01	.00[a]	.00[a]	.02	.03	.05	.00[a]	.00[a]	.00[a]	.02	.02	.00[a]	.03	.00[a]
21 MISSVC	.11	.08	.10	.01	.01	.04	.03	.01	.00[a]	.06	.02	.02	.00[a]	.00[a]	.00[a]	.01	.03	.00[a]	.00[a]	.05

[a]Less than .005.

Note: See Table 13 for definition of abbreviations.

TABLE 18

PREDICTED NUMBER OF STOPS ON WACO PERSON TRIPS–COMMERCIAL LAND USE

Starting State	Mean	Variance
2 FOODRU	1.2	.3
3 EATDRI	1.5	.7
4 GENMER	1.5	.7
5 APPACC	1.7	.9
6 FUHFHA	1.6	.8
7 MOTVAC	1.5	.7
8 GASSTA	1.4	.5
9 LUBDHA	1.5	.7
10 LIQBER	1.4	.6
11 MISRET	1.4	.6
12 FININS	1.6	.8
13 PERSER	1.4	.6
14 BUSSER	1.5	.7
15 AUTOGA	1.6	.8
16 MISREP	1.3	.4
17 AMUREC	1.5	.7
18 MEDDEN	1.5	.7
19 OTHPRO	1.2	.3
20 OFFBLD	1.4	.5
21 MISSER	1.4	.6
SYSTEM	1.4	.6

Note: See Table 13 for definition of abbreviations.

TABLE 19

DISTRIBUTION OF TRIP PURPOSES BY DESTINATION FOR SEVERAL URBAN AREAS[a]

Area	Home	Work	Shopping	School	Social-Rec.
San Francisco	37.7	27.3	9.2	2.8	12.2
Sacramento	31.4	33.6	10.5	2.6	10.1
Cedar Rapids	38.4	22.4	9.8	1.1	6.2
Chicago	43.3	20.3	7.6	1.9	12.7
Waco[b]	37.3	14.4	10.2	6.3	11.3

[a]Information from Marble (2), Table 3, p. 153.
[b]As predicted by the Markov Chain model.

4. Social-recreation and social-recreation;
5. Change mode and school;
6. Eat meal and work, social-recreation;
7. Shop and shop; and
8. Serve passenger and serve passenger, work.

Although the commercial retail and commercial service land uses had high linkages with all other land uses, the purpose shop is not highly linked with all purposes.

Table 9 indicates that the majority of tripmakers would be traveling for four purposes: (a) to go home, (b) to work, (c) for social-recreation, and (d) to shop. Data from other studies are compared with the results of the Waco analysis in Table 19. Waco differs considerably from the other urban places in the percentage of tripmakers in the purpose states of work (much lower) and school (much higher). The presence of Baylor University in Waco could well account for the high percentage of tripmakers going to school. The rest of the percentages seem to be fairly well in line.

The purpose of shopping dominates the expected number of stops in the transition states, given some non-home purpose first stop. Serve passenger and personal business also have large expected stop values. The variation in trip lengths (measured in number of stops) is quite large when using the purpose classification. Trip lengths vary from 2.2 stops for the eat meal purpose to 1.3 stops for the school purpose.

21 × 21 Commercial Land Use

The 21 × 21 commercial land-use matrix was derived from shopping trips. This particular matrix was used to decrease the level of aggregation and derive information as to the internal pattern of trip distributions and lengths within two major land-use categories.

Table 14 gives the resulting transition probabilities. It appears that food and drug, eat and drink, general merchandise, miscellaneous retail stores, finance-insurance-real estate, and personal service have the highest level of interaction with all other land uses. Further, the higher the order of the good (or service) the greater the linkages between the same land use; i.e., a trip for a low order or convenience good, such as food or drugs, has a relatively small probability that a person will stop at another land use or activity which dispenses food or drugs. A person who is interested in obtaining a higher order good such as furniture or apparel generates a higher probability of a stop on the same trip at a similar land use (Table 15).

Table 16 indicates several interesting changes in the linkage structure given a first stop at a particular land use. For example, activities which distribute food and drugs showed a high level of interaction with most of the other land-use categories. On the other hand, when food and drug activities become the first stop on the trip, one may expect very little interaction with other types of land uses. In all other cases, food and drug activities again exhibit high levels of interaction. The categories indicated above as having high probabilities of linkages maintain that level of interaction when the model is restructured into an absorbing chain.

The mean and variation in the expected length of trips which stop at a given land use are indicated in Table 18. The longest trips are associated with apparel and accessories (1.7) and the shortest with food and drug (1.2). It is clear that the number of stops is a function of distance to some extent. Higher order goods, usually found at nucleated planned or unplanned shopping centers (i.e., furniture, apparel, banking facilities, etc.) and at farther distances from the home than lower order goods, have an expectation of longer trips (again measured in number of stops).

VALUE OF THE MODELS IN CONTINUING RESEARCH

The Markov models used in the analysis thus far have demonstrated their value in the analysis of origin and destination data. Such models have considerable potential for adding meaningful insights into travel behavior on multipurpose trips. These models can also be useful in generating valuable information for use in the analysis of nonresidential trip generation. For example, linkage parameters can be used to "explain" spatial variation in the trip attraction of a given nonresidential land use.

The Markov models also add a new dimension to the analysis of household travel behavior. Given sets of socioeconomically homogeneous households, trip length (measured in number of stops), linkages, and purpose distribution information may be added to the current body of knowledge associated with household travel behavior. Time-series data on the same sets of family units provide a base for the examination of the stability of the aforementioned relationships through time.

Finally, there is the spatial connotation that probabilistic models can be given. By using areal units, such as census tracts or traffic analysis zones, in addition to land use or trip purpose to define a given state of the Markov process, information as to the spatial distribution of trip ends may be derived. Although trip distribution is beyond the scope of the research which is currently being conducted, future research efforts might well include the development of probabilistic models which can integrate generation and distribution.

ACKNOWLEDGMENTS

The support and assistance of the U.S. Bureau of Public Roads, and the efforts of the Texas Highway Department, the Waco Area Transportation Study, and the Texas Transportation Institute of the University of Texas in obtaining and processing the data used in this analysis are gratefully acknowledged.

REFERENCES

1. Brown, Lawrence A. The Diffusion of Innovation: A Markov Chain-Type Approach. Discussion Paper No. 3, Department of Geography, Northwestern Univ., 1964.
2. Marble, Duane F. A Simple Markovian Model of Trip Structures in a Metropolitan Region. Proc. Western Section, Regional Science Association, pp. 150-156, 1964.
3. Marble, Duane F. Two Computer Programs for the Analysis of Simple Markov Chains. RSRI Discussion Paper Series: No. 6, Philadelphia, 1964.
4. Kemeny, J. G., and Snell, J. L. Finite Markov Chains. D. Van Nostrand & Co., pp. 24-42, 1960.

Direct Estimation of Traffic Volume at a Point

MORTON SCHNEIDER, Director of Research, Tri-State Transportation Commission

•THE IMMEDIATE impetus for the work reported here is the growing practical need for an easy and reliable way to estimate average traffic volume, a need that is making itself felt in many quarters and which might possibly be described as urgent in an area as large, complex, and irregular as the Tri-State Region.

There are also reasons other than the immediately practical for this undertaking. Conventional traffic assignment (calculating zonal interchanges and stringing them through a network), in addition to being arduous and complicated, is an essentially unfinished process. A better instrument should be found to offer more convenient traffic estimates and also to provide more confidence in technique and, hopefully, a theoretical basis for larger problems. Although this paper does not present any fundamentally different view of travel, it does at least state a new tactic.

The following discussion is concerned with the problem of estimating the number of vehicles passing one point on a street during a fairly long period of time, such as one day. It seems likely that the mathematics could be rephrased to yield turning movements as well as simple volumes; with some ingenuity perhaps the calculations could be extended to a system-wide set of estimates, and possibly the concepts could be enlarged to cover other modes of movement. However, none of these problems has been very well thought out yet, and they are not considered here. Also, discussion of the applied aspects of the ideas developed here are mostly reserved for future reports.

DERIVATION OF TRAFFIC VOLUME EXPRESSION

The argument leans heavily on certain directional symmetries imputed to the traffic system: all streets are two-way (for every section of street going in one direction there is another section a negligible distance away allowing movement in the opposite direction at the same speed), the total long-term (i. e., daily) flow past a point in one direction is the same as that in the other direction, and one or two other considerations. This should not prevent handling one-way streets in practice through considerations of reasonableness. Also, it is assumed throughout that every moving vehicle is coming from an origin and going to a destination, and that the terms "origin" and "destination" are concepts suitable for all uses to which they are put.

Imagine a street running north and south and a point of interest on the street at which traffic volume is to be calculated. Put a single hypothetical destination at a point of interest which is placed in such a way that it is completely accessible to both northbound and southbound vehicles and does not in the least interfere with traffic. The introduction of this destination changes nothing in the traffic situation except that sometime in the course of a day an additional vehicle has to occupy it; therefore it is necessary to subtract one from the traffic volume when it is finally computed.

There is a certain probability, P_n, that any northbound vehicle at the point of interest will stop at this hypothetical destination and another probability, P_s, that any southbound vehicle will stop. These probabilities multiplied by their respective traffic volumes, Q_n and Q_s, give the expected number of vehicles that will accept the destination, and that expected number must, of course, be one:

$$P_n Q_n + P_s Q_s = 1 \qquad (1)$$

Paper sponsored by Committee on Origin and Destination.

or if $Q_n = Q_s$

$$P_n + P_s = \frac{1}{Q} \qquad (2)$$

This uncomplicated proposition expresses a definite relationship between the probability of a vehicle leaving the traffic stream and the number of vehicles in the stream, and might serve as a point of departure for lines of reasoning quite different from the one followed next.

If P_n and P_s can be evaluated, then Q is determined and the problem is solved. Every northbound vehicle approaching the point of interest has declared, by being where it is, its intention of finding a destination in some fairly well-defined geographical region lying generally north of the point of interest. Assume that these vehicles are distributed among destinations within this north domain according to some function of position relative to the point of interest, so that

$$dP = CF\ dV \qquad (3)$$

where dP is the probability that a vehicle approaching the point of interest will go to one of dV destinations clustered around a point at which the function F has a definite value. C is a constant determined by the condition that the vehicle must find its destination in the north domain:

$$\int_n dP = C \int_n FdV = 1, \text{ so } C = 1/\int_n FdV \qquad (4)$$

The symbol $\int_n$ denotes integration over the entire surface of the north domain; in general, $\int_D FdV$ may as well be called a domain integral and be replaced by the symbol I_D. Thus the probability of having a destination in some particular region R within the north domain is, integrating Eq. 3, I_R/I_n.

At this point only two things need be stipulated about the function F: it should have a finite value at the point of interest and it should be of such a form that any domain integral will be finite, no matter how large the domain (assuming that the density of destinations is never infinite). These conditions are not restrictive; any sensible function would fulfill them.

The probability of a northbound vehicle taking the destination at the point of interest becomes

$$P_n = \frac{I_o}{I_n} \qquad (5)$$

where I_o contains only that destination at the point of interest. As everything can be framed in exactly the same way from the southbound point of view,

$$P_s = \frac{I_o}{I_s} \qquad (6)$$

But F can always be scaled to make I_o equal one, so that substituting Eqs. 5 and 6 into Eq. 2 gives

$$\frac{1}{Q} = \frac{1}{I_n} + \frac{1}{I_s} \qquad (7)$$

or

$$Q = \frac{I_s I_n}{I_s + I_n} \qquad (8)$$

However, Eq. 8 does not really amount to much as it stands. If F is assumed to be a descending function of simple distance-like parameters such as travel time, cost, etc., then Eq. 8 shows no sensitivity, or rather a perverse sensitivity, to competing facilities. For example, if the street of interest is a run-of-the-mill arterial and a parallel expressway a quarter of a mile away is opened up, the domain integrals in Eq. 8 will probably grow larger, leading to the result that an expressway competing with an arterial causes the volume on the arterial to increase. Plainly the network configuration must somehow enter into the distribution. But although it is hard indeed to think of network configuration as an explicit parameter, there is a tolerably easy revision of the distribution concept that amounts to the same thing.

Consider again the northbound stream of traffic at the point of interest, this time in the presence of a nearby expressway. Presumably, the stream is full of vehicles that have recently left the expressway. But these vehicles are not free to find a destination anywhere in the north domain. The fact that they have left the expressway implies that they are going to some subregion which excludes all places more easily accessible by remaining on the expressway. In general, any stream of traffic may be regarded as being composed of free vehicles able to go anywhere and fixed vehicles restricted by some event in their past history to a lesser destination domain. These lesser domains are referred to as n' and s', that is, north prime and south prime domains.

There is no reason why the fixed and free vehicles should not be subject to totally different distributions. However, it is highly plausible to suppose that the fixed distribution is the same function as the free but that it falls to zero everywhere outside the prime domain. This meets the essential condition that the two distributions tend to be the same, because the prime domain tends to be coextensive with the main domain, and it is helpful in other respects as well. Making this supposition, the probability of a vehicle being free and going to some group of destinations dV is $A_n (F/I_n)\, dV$, and the probability of a vehicle being fixed and going to the same group is $(1-A_n)(F'/I_{n'})\, dV$, where A_n is the fraction of free vehicles in the northbound stream and F' = F everywhere within the north prime domain but F' = 0 everywhere outside the prime. So the total probability of any vehicle going to this destination group, an amplified version of Eq. 3, is

$$dP = \left[A_n \frac{F}{I_n} + (1 - A_n) \frac{F'}{I_{n'}} \right] dV \tag{9}$$

and Eq. 5 becomes

$$P_n = I_o \left[A_n/I_n + (1 - A_n)/I_{n'} \right] \tag{10}$$

(remember that the point of interest itself is always inside the prime domain); also Eq. 6 expands into

$$P_s = I_o \left[A_s/I_s + (1 - A_s)/I_{s'} \right] \tag{11}$$

Once more, consider the stream of traffic northbound at the point of interest. All of these vehicles have originated somewhere in the south domain and have made their decision to terminate in the north domain. But generally there are regions in the south domain from which it is easier to get into the north domain by a route other than one leading past the point of interest, so that vehicles in the stream coming from these regions must be headed for some special part of the north domain, not merely to the north domain at large, and are by definition fixed. It can be argued that the special part they are going to is simply that region in the north which communicates with the south most easily via the point of interest, because if they were going somewhere else it would usually be possible to construct a better route than the one through the point of interest. On the other hand, vehicles in the stream which have originated in that part of the south where easiest passage to the north is through the point of interest exhibit no overt special intentions, other than ending up in the north domain, and are free.

So the fixed and free vehicles have been defined in terms of where they originate or, an equivalent way of looking at it, in terms of the route possibilities they have declined.

Furthermore, it appears that the destination area for northbound fixed vehicles—the north prime domain—is the same as the origin area for southbound free vehicles and the destination area for southbound fixed vehicles—the south prime domain—is the same as the origin area for northbound free vehicles. This is an important simplification without which troublesome complications set in. It should be regarded as an approximation to real behavior.

With regard to the point of interest, assume that the number of northbound vehicles originating in the south prime domain, the number of free vehicles, is equal to the number of southbound vehicles ending there. But the southbound vehicles going to the south prime domain have two components: the southbound fixed vehicles, all of which must go to the prime domain, and those of the southbound free group that happen to find their destinations in the prime. Transcribing this paragraph into notation gives

$$Q_n A_n = Q_s(1 - A_s) + Q_s \frac{I_s'}{I_s} A_s \tag{12}$$

and the Q's, of course, drop out if $Q_n = Q_s$. An exactly analogous equation can be written for the southbound free vehicles:

$$A_s = (1 - A_n) + \frac{I_n'}{I_n} A_n \tag{13}$$

Between them, these two equations determine A_n and A_s, yielding

$$\left.\begin{aligned} A_n &= \frac{r_s}{1 - (1 - r_s)(1 - r_n)} \\ \text{and} \quad A_s &= \frac{r_n}{1 - (1 - r_s)(1 - rn)} \end{aligned}\right\} \tag{14}$$

Thus

$$r_s = \frac{I_s'}{I_s}, \quad r_n = \frac{I_n'}{I_n}$$

Putting Eq. 14 into Eq. 10 and Eq. 11 and manipulating a little leads to modified forms of Eq. 5 and Eq. 6:

$$\left.\begin{aligned} P_n &= \frac{I_o}{I_n}\left[\frac{1}{1 - (1 - r_s)(1 - r_n)}\right] \\ P_s &= \frac{I_o}{I_s}\left[\frac{1}{1 - (1 - r_s)(1 - r_n)}\right] \end{aligned}\right\} \tag{15}$$

and now proceeding as in Eq. 7 produces, finally, the augmented counterpart of Eq. 8:

$$Q = \frac{I_s I_n}{I_s + I_n}\left[1 - (1 - r_s)(1 - r_n)\right] \tag{16}$$

Q is the one-way traffic past the point of interest, totaled over a long enough period of time (probably one day) for the symmetry postulates to hold.

BEHAVIOR OF TRAFFIC VOLUME EXPRESSION

So far only very weak delimitations have been imposed on the distribution function; it could be almost anything. Even so, there is a fair amount of visible character in Eq. 16.

The effective quality or competitive position of a street operates through the bracketed part of Eq. 16. The r's are the ratios of prime domain integrals to their respective main domain integrals; as the prime domains become a larger part of the main domains the r's and the entire bracketed expression grow larger. The bracketed expression achieves its maximum value, one, when the prime domains are so large as to include the entire main domains, a situation that would occur if the street of interest were, for instance, the only bridge across a long river. In this case the traffic itself, Q, would be greatest for any given I_s and I_n.

An expressway, because of its high speed, tends to have extensive prime domains, and therefore a large volume. Its extensiveness depends on its speed advantage and how far it is from other expressways. The prime domains of an ordinary arterial would usually be smaller, taking the form of strips running the length of the street and enclosing it, whereas those of a local street would be very small, pinched off after short distances. If the prime domains contained no destinations at all, the bracketed expression, and the volume, would be zero. Or, in stricter agreement with the theory, if the prime domains are so small that they include only the hypothetical destination at the point of interest, Eq. 16 reduces to a very close approximation of Q = 1 (a little less than one actually, expressing the slight possibility of the destination being its own origin).

The overall strength of the traffic field is measured by the left-hand part of Eq. 16, the factor in I_s and I_n. This strength increases as one or both domain integrals increase. Also, it goes to zero, taking the traffic with it, as either of the domain integrals goes to zero—a necessary property because a zero integral implies that there is no place a vehicle can go by passing the point of interest, as, for example, in the case of a street dead-ending at the ocean. For any given sum, $I_s + I_n$, the strength is maximized when $I_s = I_n$. Assuming the r's to remain constant, and without attempting a precise phrasing, this is to say that a given collection of destinations generates the most traffic at the point of interest when distributed evenly on both sides.

The domain integrals, of course, increase as destination density in the domains increases. Also, it may reasonably be suspected that they increase or decrease as destinations move nearer to or farther from the point of interest (although little is known about the distribution function). This leads to a final general inference from Eq. 16: that traffic at the point of interest tends to increase when surrounding destination masses increase or when these masses move closer, and tends to decrease when destination masses decrease or when they move farther away.

THE DISTRIBUTION FUNCTION

The function of F must be given precise definition in order to do any specific calculating from Eq. 16.

One convenient, acceptable function is

$$F = e^{-kt} \tag{17}$$

where t is travel time from the point of interest and k is a kind of natural constant. Or, more generally,

$$F = e^{-(k_1 t + k_2 u)} \tag{18}$$

where u is the cost incurred from the point of interest. Probably the simplest assumption that can be made about the distribution of vehicles among destinations is that all destinations are equally likely, subject to the constraint that average travel time must be finite even in an infinitely extensive universe of destinations.

Imagine the north domain to be divided into many cells, each containing the same number of destinations and having, therefore, the same a priori attractiveness for vehicles, and let the Q northbound vehicles at the point of interest distribute themselves among these cells so that the first cell receives q_1 vehicles, the second q_2, and so on. The Q vehicles can now be redistributed in such a way that the occupancy numbers q_1, q_2, etc., remain the same but not every vehicle is in the same cell as before.

The number of possible different arrangements of this kind for a particular set of occupancy numbers is

$$\frac{Q!}{q_1!\, q_2! \ldots q_n!} \tag{19}$$

The question can be asked: what set of occupancy numbers can be obtained in the most ways? This would be the set most likely to turn up at random because it can occur in more different ways than any other pattern. The set of occupancy numbers that can be obtained in the most ways is, of course, that set which maximizes Eq. 19, and it is the set in which all q's are equal.

Now the constraint that average (or total) travel time must be finite can be written

$$q_1 t_1 + q_2 t_2 + \ldots + q_n t_n = T \tag{20}$$

where t_i is the travel time to the ith cell and T is some finite constant. And the question in the preceding paragraph can be rephrased: what set of occupancy numbers consistent with Eq. 20 can be obtained in the most ways? This is a somewhat more sophisticated question, but it can be answered in essentially the same way: by determining the q's that maximize Eq. 19, although taking Eq. 20 into account. The procedure is to take the partial derivative of Eq. 19 with respect to each q_i, add to it a term proportional to the corresponding derivative of Eq. 20 (using Lagrange's multipliers), and set the sum equal to zero. This involves both manipulation and approximation, and the result is

$$q_i = e^{-\lambda t_i} \tag{21}$$

which is the same as Eq. 17 once the notation is adjusted to conform to previous usage. If Eq. 17 is used as the distribution function, I_0 in Eq. 5 and those following it naturally equals one without any further meddling.

Cost can be introduced in a completely analogous fashion by arguing that just as travel time must be limited, so must travel cost. This produces another constraint,

$$q_1 u_1 + q_2 u_2 + \ldots + q_n u_n = U \tag{22}$$

Equation 19 can now be maximized subject to both Eq. 20 and Eq. 22 and the result is equivalent to Eq. 18.

WORKING METHODS

With F defined, it becomes possible in principle to evaluate the integrals of whioh the traffic estimate is composed. The possibility in principle, however, scarcely helps when it comes time to go ahead and do it in practice, yet preserve the measure of convenience that is the most important aspect. If the conditions of destination density and network geometry were regular, the integrals could be evaluated by direct mathematical operations. In the real, unaccomodating world the integrals can still be calculated numerically, but any straightforward numerical technique would seem to be laborious.

There is a question of how much detail and precision the whole process deserves. The input information—network speeds and destination densities—is not really well defined or accurately obtainable, and the theoretical structure itself does not have the ring of final truth. Moreover, conventional assignment often yields wildly inaccurate results on the level of specific street estimates, yet is generally regarded as an acceptable methodology. In short, at present no great accuracy seems to be either possible or expected.

Further, a previous paper (1) derived traffic estimates, under drastically simplifying assumptions, from extremely rudimentary information: a single average trip-end density and single average spacings of local streets, arterials, and expressways. Although this was not a practicable procedure for a variety of reasons, it did produce estimates with a rough, order-of-magnitude realism. It is suggested that a small number of pieces of information above this bare one-point level would produce a great improvement without being too difficult.

The pieces of information might take the form of readings at points scattered throughout the region, constituting a kind of sample of the region. These readings would consist of the best route travel time from the point of interest to the point of reading and the average destination density around the point of reading. If the location of each reading point is known, the read values could be interconnected by an arbitrary interpolation and the necessary integrations performed; some description of the borders of the north, south, and prime domains would also be required. Thus the precision of the method would be directly related to the amount of work put into it, i. e., to the number of reading points. Inasmuch as the contribution of any area to traffic at the point of interest diminishes with distance, the readings can grow farther and farther apart as they move away from the point of interest. To be mathematically convenient, the reading points should lie in a regular pattern, and this pattern should be fixed so the person taking the readings is not free to make a biased choice of points.

Based on these considerations, the working method presently used is this. A template or transparent overlay is drawn showing radial lines emanating from a point; rings intersect the radials, with the spacing of successive rings becoming larger as they lie farther from the center. This template is overlaid on a map containing the street system, with speeds indicated and destination densities blocked in, with the center of the template right on the point of interest. For each intersection of ring and radial, the reader estimates best route travel time from the center, notes the ambient density, and enters these values on a form. In a separate operation, he writes down polar coordinate points (from the same template) which, when connected by straight lines, will reasonably delineate the prime domains. For simplicity, the north and south domains are considered to be demarcated by a straight east-west line (the directional terms are schematic, of course); this is a convenience of the moment, not an essential simplification, and will very likely be revised. The rule for drawing a prime domain is that the prime should include all points from which it is easier to cross the main domain line by passing the point of interest than by any other way, and should exclude all other points.

The forms containing these readings are key-punched and the cards fed into an IBM 1401 computer, which performs all the complicated calculations, ending in an estimate of traffic at the point of interest. Linear interpolations are made among the point values, allowing integral terms to be computed.

These methods seem to fall within a tolerable range of labor. The readings do not seem too hard to execute, and the computing is quick and easy. A total reading time of an hour or two and computing time (1401) of 10 or 15 min seems within shooting distance. A lot depends on where the point of diminishing returns lies in the number of readings. Also, none of this should be regarded as fixed; some better working scheme might very well emerge to replace it.

CONCLUSION

The mathematical and computational forms developed here appear at this time to represent traffic behavior. A dozen or so real cases have been calculated. Although the purpose of these calculations has been to regularize the technique rather than to subject it to strict tests, and although good input information is not yet available, it seems fair to say that the results, for so early a state of evolution, are quite promising. The calculations evidently admit a full range of traffic volumes, from local street to expressway, and the practical labor is within reason. Exact results ought properly to be considered meaningless right now, and when bad they are so considered. However, the comparison with listed traffic flows is fairly good for a first trial.

A number of practical problems are turning up. In some cases, delineation of prime domains is ambiguous—different people will draw them with appreciable differences—

perhaps calling for more carefully devised rules. The working method described in the foregoing introduces various statistical problems, all of which can probably be solved, concerning the number of reading points to be used and their pattern, whether or not the prime domains should be sampled differently from the main domains, and how to avoid statistical wastefulness, i. e., taking readings where there is no real gain in information or, conversely, throwing information away because of the arrangement of the template. There are also customary minor difficulities of thinning out errors in reading and transcribing.

ADDENDUM

Since this paper was written some developments have taken place which ought to be at least briefly mentioned here.

A large-scale computer program has been written, and is now being tested, to calculate traffic volumes throughout a very large network automatically. The program is designed to produce volumes on any specified links, on all links within a specified area or group of areas, on links of a selected class (such as expressways), or on all links in the entire system. Turning movements can also be requested pretty much at will, and the program will do its best to compute them. Other options, capacity restraint among them, are imaginable and may be added.

The map-and-template method has been pursued beyond the previous discussion, but not too far beyond. Although it was convenient for making experimental calculations and remains useful for situations in which no coded network is available, it is not at all competitive with a fully computerized system once the inputs for such a system have been prepared.

Several quirks in the behavior of Eq. 16 have turned up and the exact form given here probably will not long survive. It appears, though, that most of these aberrations can be removed by a plausible modification, entailing no loss of generality and leaving the basic reasoning intact.

In summation, there seems now to be a good chance of developing a durable methodology which will improve traffic-estimating technique through the sheer force of its flexibility. The user can focus his attention, effort, and budget on that aspect of traffic estimating that concerns him, from minor detail to generalized planning. A complete set of system-wide link volumes is seldom of much interest. Most often only a relative handful of estimates, expressing local finer resolution or the results of planning changes, is really wanted, and this handful can be obtained in a few seconds of computer time, without having to run a specially scaled, full assignment. At the same time, satisfactory and systematic estimates can still be made by any one of many possible variations on the map-and-template method when tying into a big computer system is not warranted. In short, the technique looks well-tempered: the precision (not to be confused with accuracy) and extent of the calculations—or, more generally, their expense—can be made consistent with the uses to which they are put.

Also, the direct estimation point of view seems adaptable to a larger scheme. Recently, a fragment of a theory has been worked out which includes travel within a more general framework, relating activity at a site to something that might be called accessibility of the site. On the face of it, this partial theory has nothing much to do with the work described here, except that the direct estimation mathematics can be reformulated within its context. When that is done, however, the term giving rise to the most stubborn (and disturbingly fundamental) of the quirks mentioned is precisely canceled out. Moreover, it turns out that the domain integrals introduced here have a distinct kinship with quantities appearing in the newer theory, and that it might be possible to lead into a land activity model using programs and materials developed for direct estimation.

Finally, since an explicit calculation has been stated which gathers the access of a piece of road to the geography in which it is embedded into a traffic flow on the road, the natural speculation arises: might not the process be reversed? Might not the pattern of traffic flows imply the geographic spread of activities, and traffic counts be used to measure activity on a piece of land?

ACKNOWLEDGMENTS

The idea presented here has evolved into a fairly major enterprise to which many members of the Tri-State Transportation Commission staff have contributed. Geoffrey J. H. Brown did much of the computational analysis and all of the demanding computer work central to the problem. Albert Woehrle supervised most of the map-and-template readings and input preparation in general. W. L. Mertz gave encouragement and support. Above all, this work is indebted to J. Douglas Carroll, Jr.

REFERENCE

1. Schneider, M. A Direct Approach to Traffic Assignment. Highway Research Record 6, pp. 71-75, 1963.

Covariance Analysis of Manufacturing Trip Generation

MICHAEL KOLIFRATH, Northwestern University, and
PAUL SHULDINER, U.S. Department of Commerce and Northwestern University

The variability of generation rates for trips to various subclasses of manufacturing land is tested to determine whether a significant increase in the precision with which such trips can be estimated is achieved by treating each subclass as a separate category in contrast to dealing with manufacturing land as a single class.

Separate regression equations are developed relating trips to manufacturing land at the two-digit Standard Industrial Classification (SIC) level as a function of (a) employment (estimated by first work trips), (b) floor space, and (c) truck trips for 21 CATS traffic districts in Chicago. On the basis of these equations, covariance analysis is used to test the utility of subdividing manufacturing land into subclasses for estimating trip attractions.

It is concluded for the set of data used in the analysis that: (a) when employment is used as the basis for estimating trip attractions to manufacturing land, no significant improvement in precision can be derived by using a finer (two-digit SIC) classification system; and (b) when floor area or truck trips are used as the independent variable, significant improvement in the precision of estimates can be obtained by going to the two-digit level.

•THE OBJECTIVE of this research is to determine whether there are significant differences in the trip generation rates of manufacturing land-use activities at the two-digit SIC level of manufacturing land use. Two alternatives are considered: (a) there are significant differences in trip generation rates between the two-digit manufacturing land-use types comprising the major manufacturing activity category; and (b) there are no significant differences in trip generation rates between two-digit manufacturing land-use groups. If the second hypothesis holds, no accuracy is lost by combining the twenty, two-digit manufacturing land-use types into a major one-digit manufacturing land-use class as a basis for estimating trip generation rates.

In this report a trip is defined as a one-way journey by a person traveling as a driver or passenger in an automobile, or as a passenger in a taxi, truck or mass transportation vehicle. The trip may have any land-use type as an origin, but must have manufacturing land as a destination. No walking trips, scooter or cycle trips are included. Truck driver trips are analyzed separately.

PROCEDURE

Regression Analysis

Regression analysis is used to determine best fit linear equations relating total trips to the independent variables. Standard errors of estimate, standard errors of

Paper sponsored by Committee on Origin and Destination.

the regression coefficients, coefficients of determination, and F-values are also provided by the analysis. The following regression equations are derived:

$$y = a_1 + b_1x_1 \quad (1)$$
$$y = a_2 + b_2x_2 \quad (2)$$
$$y = a_3 + b_3x_3 \quad (3)$$
$$y = a_4 + b_4x_1 + b_5x_2 \quad (4)$$
$$y = a_5 + b_6x_1 + b_7x_3 \quad (5)$$
$$y = a_6 + b_8x_2 + b_9x_3 \quad (6)$$
$$y = a_7 + b_{10}x_1 + b_{11}x_2 + b_{12}x_3 \quad (7)$$
$$x_3 = a_8 + b_{13}x_4 \quad (8)$$

where

y = total trips (manufacturing land-use destination),
x_1 = manufacturing floor area (sq ft),
x_2 = estimated manufacturing employment,
x_3 = truck driver trips (manufacturing land-use destination), and
$x_4 = x_2/x_1$.

The a_i and b_i are least squares estimates of the population parameters A_i and B_i.

These eight equations are derived for each of the twenty, two-digit manufacturing land-use groups within the manufacturing durable category containing two-digit groups 20 to 29, the manufacturing nondurable category containing two-digit groups 30 to 39, and for the general manufacturing category composed of all twenty of the two-digit sub-groups.

Essentially, this information provides best fit estimating equations of the assumed linear relationship between trips generated by manufacturing land and the independent variables.

The main objective of this study may be stated as follows: Can a single regression line be used to estimate generated trips by manufacturing land for all twenty, two-digit groups of manufacturing land-use activity?

To answer this question two tests must be performed. (A test for equality of variance throughout the range of the independent variable, homoscedasticity, is performed before testing the slopes of the regression equations. The effects of violation of this precondition on conclusions drawn from the covariance analysis are discussed later.)

1. Letting B_i represent the true slope of the regression function, are the twenty sample slopes (b_1, b_2 . . . b_{20}) estimates of the same true slope B_i?
2. Assuming $B_i = b_1 = b_2 = \ldots = b_{20}$, are group trip generation means equal after adjusting for differences in the covariates (independent variables)?

The null hypothesis associated with analysis of covariance may be stated as follows: There are no significant differences in the mean manufacturing trip generation rates between the twenty, two-digit manufacturing land groups. A basic assumption inherent in the covariance analysis technique is the equality of slope coefficients. Thus, the test for equality of slopes is performed first, and if the equality of slopes hypothesis cannot be rejected, covariance analysis is employed. Non-rejection of the null hypothesis associated with covariance analysis is equivalent to stating that one single regression line can be employed to estimate generated person trips by manufacturing land for all twenty of the two-digit manufacturing land-use groups.

DATA

The data employed in this analysis were provided by the Chicago Area Transportation Study (CATS). The trip data were obtained from tabulations derived from the 1956 O-D study. Floor area measurements were obtained from tabulations of the land-use survey of the Chicago region. Four specific groups of data were gathered for the 21 CATS traffic districts contained in Chicago: (a) manufacturing floor area (sq ft); (b) first work trips to manufacturing land uses; (c) total person trips to manufacturing land

TABLE 1

TOTAL PERSON TRIPS VS FLOOR AREA (10,000 sq ft)[a]

Land-Use Code	Std. Error of Est.	a	b	Std. Error of b	Var. Coeff. (%)	F-Value	R^2
20	206	94	0.48	0.14	29	12.5	0.415
21	528	125	1.66	0.18	11	90.2	0.826
22	225	178	1.07	0.20	19	27.8	0.596
23	778	285	1.86	0.28	15	45.0	0.714
24	627	528	1.34	0.08	7	288.8	0.941
25	1014	979	1.27	0.18	14	52.5	0.745
26	5005	2733	3.13	1.09	35	8.2	0.313
27	—	—	—	—	—	—	—
28	474	-49	3.44	0.39	11	78.3	0.822
29	646	587	1.48	0.12	8	141.1	0.881
30	1751	414	1.34	0.17	13	60.1	0.770
31	51	24	1.28	0.11	9	128.0	0.908
32	425	120	0.72	0.19	26	14.6	0.493
33	301	95	4.11	0.13	3	1064.0	0.985
34	418	99	2.21	0.16	7	200.1	0.918
35	1215	-775	4.66	0.68	15	47.1	0.723
36	418	321	1.50	0.14	9	118.7	0.868
37	173	136	0.41	0.12	29	11.1	0.396
38	262	104	2.47	0.69	27	12.8	0.445
39	299	42	2.50	0.16	6	231.8	0.932
20-29	2199	642	1.54	0.14	9	118.5	0.372
30-39	1047	274	1.54	0.08	5	416.1	0.693
20-39	1750	455	1.56	0.08	5	347.6	0.474

[a]Data derived from Chicago Area Transportation Study.

uses; and (d) truck trips with manufacturing land as the destination. These groups of data were summarized from tables derived from the initial coded data by two-digit manufacturing land-use activity and by CATS traffic districts. The numbers of observations were limited by the availability of floor area measurements. The floor area survey covered Chicago; an average of 20 to 21 district observations per land-use group was available.

CATS land-use coding system for manufacturing activities is quite similar to the SIC system, and only varies in the numerical system employed and not in the activity composition of each group.

First work trips to manufacturing land were used to estimate manufacturing employment. Although manufacturing employment estimates can be secured from the Illinois Employment Service, Bureau of Labor Statistics, these data are summarized by postal zones which are not comparable with CATS traffic districts. CATS is presently developing a method by which employment can be apportioned to appropriate traffic zones assuming equal employment intensity within two-digit land-use activities, but this conversion deck was not available at the time this study was conducted.

Because of disclosure limitations pertaining to individual firms, no data are available for zones which have only one firm in a specific land-use category. The Bureau of Labor Statistics records indicate a number of zones which fall into this category. Also, because of seasonal variation of employment, vacation periods and absenteeism, work trips may actually provide better estimates of employment for the specific survey day than those achieved by converting from postal zones. In any event, the accuracy of the results of this study depends heavily on the accuracy with which employment is estimated by first work trips, and some check would have been desirable.

RESULTS AND INTERPRETATION

Multiple Regression Analysis

Tables 1-8 were derived by means of linear regression analysis. They include the standard error of estimate, the regression constant, the regression coefficient,

TABLE 2

TOTAL PERSON TRIPS VS ESTIMATED EMPLOYMENT[a]

Land-Use Code	Std. Error of Est.	a	b	Std. Error of b	Var. Coeff. (%)	F-Value	R^2
20	55	13	1.202	0.059	4.9	413	0.958
21	58	-5	1.138	0.012	1.1	8959	0.998
22	72	53	1.084	0.052	4.7	442	0.959
23	116	-6	1.136	0.021	1.8	2819	0.994
24	143	27	1.146	0.015	1.3	5874	0.997
25	710	216	0.978	0.087	8.9	126	0.875
26	376	138	1.102	0.016	1.5	4632	0.996
27	145	46	1.079	0.052	4.8	435	0.960
28	519	25	0.934	0.118	12.6	62	0.786
29	155	1	1.168	0.022	1.9	2750	0.993
30	408	20	1.097	0.034	3.1	1023	0.983
31	51	14	0.969	0.094	8.7	133	0.911
32	99	-3	1.255	0.055	4.4	530	0.972
33	69	45	1.074	0.007	0.6	20566	0.999
34	66	-39	1.170	0.013	1.1	8618	0.998
35	136	12	1.198	0.017	1.4	5191	0.996
36	173	-107	1.278	0.046	3.9	778	0.977
37	52	12	1.403	0.081	5.8	300	0.946
38	43	5	1.164	0.035	3.0	1099	0.986
39	159	73	1.105	0.037	3.3	871	0.981
20-29	322	32	1.108	0.009	0.8	14682	0.987
30-39	201	33	1.118	0.009	0.8	16098	0.989
20-39	270	33	1.109	0.006	0.5	30375	0.988

[a]Data derived from Chicago Area Transportation Study.

TABLE 3

TOTAL PERSON TRIPS VS TRUCK TRIPS[a]

Land-Use Code	Std. Error of Est.	a	b	Std. Error of b	Var. Coeff. (%)	F-Value	R^2
20	192	98	2.6	0.6	23	17	0.49
21	373	71	11.0	0.8	7	197	0.91
22	198	208	2.5	0.4	16	42	0.69
23	747	253	10.3	1.5	14	50	0.74
24	1500	829	2.7	0.5	18	36	0.66
25	1593	993	4.8	1.5	31	10	0.37
26	3704	1499	10.6	2.0	19	30	0.62
27	510	400	4.7	1.1	23	18	0.50
28	209	187	13.0	0.6	5	471	0.97
29	667	173	9.6	0.8	8	131	0.87
30	1751	237	6.0	0.8	13	60	0.77
31	159	39	1.9	1.4	74	2	0.13
32	319	135	6.3	1.0	16	38	0.71
33	732	57	19.3	1.5	8	167	0.91
34	263	170	5.7	0.3	5	534	0.97
35	2161	1380	0.9	0.6	67	3	0.12
36	473	307	4.7	0.5	11	88	0.83
37	215	147	1.4	1.3	93	2	0.07
38	316	140	6.3	3.1	49	4	0.21
39	586	86	15.5	2.2	14	48	0.74
20-29	2073	684	4.88	0.39	8	159	0.44
30-39	1456	543	3.24	0.29	9	126	0.41
20-39	1838	633	4.16	0.25	6	279	0.42

[a]Data derived from Chicago Area Transportation Study.

TABLE 4

TOTAL PERSON TRIPS VS FLOOR AREA AND EMPLOYMENT[a]

Land-Use Code	Std. Error of Est.	a	Floor Area (b_1)	Employ. (b_2)	Multiple R^2
20	56	17	-0.003	1.24	0.959
21	59	-5	-0.001	1.14	0.998
22	73	53	0.005	1.05	0.959
23	119	-8	0.003	1.12	0.994
24	143	5	0.009	1.21	0.997
25	688	303	0.034	0.77	0.889
26	353	63	0.220	1.08	0.996
27	148	53	-0.001	1.08	0.961
28	333	-101	0.209	0.50	0.822
29	134	51	0.019	1.04	0.993
30	443	47	-0.021	1.24	0.986
31	52	18	0.053	0.57	0.914
32	102	-3	0.000	1.25	0.972
33	61	48	0.044	0.96	0.999
34	59	-35	0.017	1.09	0.998
35	136	-23	0.012	1.17	0.997
36	126	-47	0.038	1.00	0.989
37	53	13	0.002	1.37	0.947
38	36	-1	0.031	1.10	0.990
39	161	67	0.023	1.01	0.981
20-29	318	15	0.006	1.09	0.987
30-39	197	35	-0.007	1.16	0.989
20-39	270	29	0.003	1.10	0.988

[a]Data derived from Chicago Area Transportation Study.

TABLE 5

TOTAL PERSON TRIPS VS FLOOR AREA AND TRUCK TRIPS[a]

Land-Use Code	Std. Error of Est.	a	Floor Area (b_1)	Truck Trips (b_2)	Multiple R^2
20	170	37	0.03	1.90	0.62
21	349	57	0.05	8.14	0.93
22	193	180	0.04	1.78	0.72
23	634	101	0.10	6.10	0.82
24	641	553	0.14	-0.17	0.94
25	885	535	0.11	2.33	0.82
26	3778	1278	0.06	9.58	0.63
27	516	431	-0.01	4.83	0.52
28	206	144	0.08	11.52	0.97
29	578	351	0.08	4.67	0.91
30	1033	-200	0.08	3.57	0.92
31	—	—	—	—	—
32	318	106	0.02	5.18	0.74
33	273	68	0.03	3.52	0.99
34	225	128	0.07	4.13	0.98
35	1206	-883	0.52	-0.44	0.74
36	409	293	0.10	1.63	0.88
37	173	109	0.04	1.05	0.43
38	266	76	0.22	2.69	0.48
39	300	31	0.22	2.14	0.94
20-29	2012	513	0.068	3.44	0.48
30-39	974	217	0.129	1.26	0.74
20-29	1633	362	0.105	2.22	0.54

[a]Data derived from Chicago Area Transportation Study.

TABLE 6

TOTAL PERSON TRIPS VS EST. EMPLOYMENT AND TRUCK TRIPS[a]

Land-Use Code	Std. Error of Est.	a	Employ. (b_1)	Truck Trips (b_2)	Multiple R^2
20	47	19	1.37	-0.66	0.970
21	59	-5	1.15	-0.10	0.998
22	69	60	0.98	0.36	0.964
23	117	-11	1.11	0.32	0.994
24	143	21	1.13	0.07	0.997
25	712	132	0.92	0.77	0.881
26	266	58	1.04	0.94	0.998
27	141	45	1.16	-0.64	0.964
28	210	161	0.10	11.94	0.967
29	141	-2	1.06	1.04	0.995
30	483	-88	1.05	0.37	0.983
31	52	17	0.99	-0.23	0.913
32	100	-4	1.32	-0.46	0.973
33	70	47	1.09	-0.28	0.999
34	66	-27	1.09	0.39	0.998
35	124	1	1.19	0.07	0.997
36	114	-67	1.02	1.18	0.991
37	53	10	1.39	0.14	0.947
38	38	-5	1.13	0.92	0.989
39	159	81	1.17	-1.18	0.982
20-29	312	15	1.081	0.278	0.987
30-39	196	29	1.096	0.160	0.989
20-39	263	24	1.086	0.221	0.988

[a]Data derived from Chicago Area Transportation Study.

TABLE 7

TOTAL PERSON TRIPS VS FLOOR AREA AND EMPLOYMENT AND TRUCK TRIPS[a]

Land-Use Code	Std. Error of Est.	a	Floor Area (b_1)	Employ. (b_2)	Truck Trips (b_3)	Multiple R^2
20	46	25	0.006	1.45	-0.72	0.973
21	61	-5	0.001	1.15	-0.09	0.998
22	71	61	0.004	0.99	0.40	0.964
23	121	-12	0.002	1.10	0.31	0.994
24	133	-28	0.017	1.24	0.15	0.998
25	675	207	0.039	0.66	1.01	0.899
26	269	38	0.006	1.04	0.87	0.998
27	141	45	—	1.16	-0.65	0.964
28	203	104	0.057	0.12	9.88	0.971
29	133	37	0.014	1.01	0.58	0.996
30	454	69	0.025	1.29	-0.22	0.986
31	54	19	0.046	0.64	-0.15	0.915
32	104	-6	0.002	1.32	-0.52	0.974
33	63	49	0.044	0.97	-0.15	0.999
34	58	-24	0.017	1.02	0.34	0.999
35	128	8	0.003	1.19	0.08	0.997
36	110	-52	0.017	0.98	0.83	0.992
37	54	11	0.002	1.36	0.15	0.948
38	34	-7	0.025	1.09	0.70	0.992
39	162	76	0.024	1.07	-1.20	0.982
20-29	313	13	0.0014	1.080	0.255	0.987
30-39	191	31	0.0086	1.138	0.172	0.990
20-39	263	25	0.0008	1.088	0.229	0.988

[a]Data derived by Chicago Transportation Area Study.

TABLE 8

TRUCK TRIPS VS EMPLOYMENT INTENSITY (Emp./10,000 sq ft)[a]

Land-Use Code	Employ. Intensity	Std. Dev.	a	b	Std. Error of Est.	R^2
20	11	11	68	-0.72	72	0.013
21	37	40	—	—	—	—
22	70	195	68	0.06	117	0.010
23	27	18	142	-0.57	121	0.008
24	19	11	1325	-26.59	728	0.152
25	56	136	297	-0.41	249	0.050
26	103	145	442	-0.32	445	0.011
27	77	199	101	-0.11	108	0.044
28	87	145	33	-0.02	85	0.001
29	39	41	148	-0.22	183	0.003
30	17	7	497	-2.52	538	0.001
31	136	320	16	-0.01	31	0.007
32	21	25	61	-0.43	80	0.019
33	66	39	129	-0.98	116	0.102
34	34	42	156	-0.68	249	0.014
35	24	11	253	4.41	871	0.003
36	1386	6047	181	-0.01	222	0.011
37	56	93	13	0.26	33	0.367
38	47	54	17	0.02	25	0.003
39	48	99	35	-0.04	63	0.004
20-29	52	120	219	-0.203	378	0.0042
30-39	192	1984	141	-0.003	373	0.0002
20-39	119	1376	176	-0.005	376	0.0003

[a]Data derived from Chicago Area Transportation Study.

standard error, and the coefficient of determination (R^2) for each two-digit land-use type and the summary groups. As discussed previously, this analysis was performed to provide data concerning trip generation rates and to establish a basis for the covariance analysis.

Regression Intercepts

Table 1 indicates the significant parameters when floor area is employed as the estimator variable of total person trips. Origin intercept was not imposed on the regression equations in order to permit examination of the intercept constants. However, it is reasonable to assume that there is a minimum size (floor area) at which a manufacturing plant can operate efficiently. Thus the relationship between floor area and generated trips is not too meaningful between the origin and minimum size plant, and extrapolation in this area could lead to serious errors. (For example, two land uses indicated negative coefficients. If these data were to be strictly employed without regard to minimum size plant requirements, negative trips would be estimated when no floor area existed. Since the data are summarized by district, minimum plant size is lost and no conclusions in this respect should be drawn from the regression constant. For example, land-use 28, printing and publishing firms, would need at least 142,000 sq ft to produce zero trips, but the additional floor area required for a plant of sufficient operating size is not known.)

Employment, Floor Area and Employment Intensity Interrelationships

Regression coefficients, variation coefficients, and multiple coefficients of determination are more meaningful statistics for this analysis. The interrelationship of employment, floor area, and employment intensity are discussed by means of an example using these measures.

The R^2 values of Table 1 vary from 0.313 (for group 26, electrical machinery) to 0.985 (for group 33, clothing). What would this indicate in terms of trip generation characteristics? As manufacturing employment is by far the best predictor of trips generated by manufacturing land uses (see high R^2 values in Table 2) one would expect little variation of employment intensity to occur within the electrical machinery manufacturing activities in relation to the variation of employment intensity within group 33. The following data derived from Table 8 support this hypothesis:

Land-Use Group	Employment Intensity	Std. Dev.	Coeff. Var.
26	103.45	145	1.4
33	66.20	39	0.6

A similar analysis can be performed for the two major land-use categories (group 20-29 and 30-39):

Land-Use Group	Estimating Equation
20-29	Total trips = 642 + 1.54 (floor area)
30-39	Total trips = 274 + 1.54 (floor area)

In this particular case, the regression coefficients are the same, but the equations are separated by a constant (642-274) = 368 total trips. These same groups in Table 2 indicate that similar regression equations have been derived; thus the variation between the regression constants cannot be explained by employment. From this examination it could be concluded that the employment intensity is less for the first group than for the second. Again, Table 8 supports this argument:

Land Use Group	Employment Intensity (per 10,000 sq ft)
20-29	52
30-39	192

Other similar comparisons and relationships could be shown to occur, but the most significant result is the amount of total trip variation explained by estimated employment. Table 2 indicates that all but two values of R^2 are above 0.900 and a majority are above 0.960. When employment is used in multiple regression with either floor area or truck trips or both, little, if any, increase in explained trip variation occurs. Because of the possible multicolinearity among the prediction variables, the difficulty in projecting all three independent variables, and the lack of any significant additional explanation of the trip variation when multiple regression is employed, the tests for group classification are performed only for the simple regression models.

Truck Activity and Employment Intensity

In this analysis employment intensity was defined as the number of employees per 100,000 sq ft of floor area. Table 8 gives the significant parameters of the regression. Because of the very low values of the coefficient of determination, the hypothesis that the regression coefficients were equal to zero was tested ($b_i = 0$). In all cases the F-values derived were very low and the hypothesis could not be rejected with reasonable certainty. (F-value required for rejection at 90 percent level of confidence was 61.) It follows that there is no significant correlation within groups of employment intensity to truck trips although predominately negative regression coefficients are intuitively satisfying.

The between groups employment intensity to truck trips relationship does provide some useful information. A general statement could be made indicating the inverse relationship of truck trips to employment intensity between groups. Table 9 gives the ratios of mean employment intensity to intercept constant (in this case the mean truck trips). There seems to be a definite inverse relationship but it does not appear to be a linear function.

Summary of Regression Analysis Findings

1. In all groups except group 28, simple regression equations using employment as the independent variable provide the highest values of the coefficient of determination.
2. The addition of floor area and truck trips to employment in multiple regression analysis provides little, if any, added explanation of the total trip variation.
3. There seems to be no significant relationship between truck trips and employment intensity.

As previously discussed, tests for group classification are performed only for the three simple regression models. The accuracy of projected person trips depends chiefly on: (a) the accuracy with which the independent variable can be projected, and (b) the accuracy with which the regression equation estimates the total person trips. It is assumed that the accuracy with which the independent variable can be projected is greater at a one-digit level than at a two-digit level of classification. Essentially, then, the accuracy with which total trips can be made depends not only on the accuracy of the estimating equation but on the classification system employed as the basis of projection. For this reason all three simple regression models are analyzed with the covariance technique of group classification.

Covariance Analysis

The results for testing hypothesis I (equal slopes) are given in Table 10. In case one and case three the hypothesis is rejected and the conclusion drawn that there are significant differences in the regression coefficients between two-digit land-use groupings. It can then be said that because of variable slope coefficients between land-use groups,

TABLE 9

RATIO OF MEAN EMPLOYMENT INTENSITY TO AVERAGE TRUCK TRIPS
(Ranked according to employment intensity)

Land-Use Group	Employ./10,000 Sq Ft Floor Area	Avg. No. Truck Trips	Ratio
20	11	68	6
30	17	497	30
24	19	1325	60
32	21	60	3
35	24	252	10
23	27	142	6
34	34	156	5
29	39	147	3
38	46	17	0.3
39	48	34	0.8
37	56	13	0.2
25	56	297	4
33	66	128	2
22	70	70	1
27	76	100	0.7
28	87	33	0.3
31	136	15	0.1
36	1386	180	0.1

TABLE 10

F-VALUES FOR TESTING EQUALITY OF SLOPES

Regression Independent Variable	F (cal.) $v_1 = 19,\ v_2 = 348$	F_{95} (for rejection)
Case 1. Floor area	6.95	1.67
Case 2. Employment	1.22	1.67
Case 3. Truck trips	10.80	1.67

a single regression line cannot be used for all observations when floor area or truck trips are employed as the estimating variables. If one of these independent variables is used to predict trip generation of manufacturing land, the two-digit subclassification must be used as the basis for projection.

However, the hypothesis of equal slopes between groups cannot be rejected for case two, i. e., one may assume there is no difference in the regression coefficients between two-digit groups when employment is used as the independent variable.

As the assumption of equal slope coefficients between two-digit groups is satisfied for case two, covariance analysis can be employed to determine if the mean trip rates are equal after adjustment of employment variation between two-digit groups. Table 11 gives the results of this analysis. The conclusion drawn from the analysis is that there are no significant differences in means between two-digit groups after covariate adjustment.

Thus, for the data analyzed, the regression line relating employment and total person trips for the major one-digit manufacturing category may be used for all observations, and no addition in accuracy can be gained by subdividing the manufacturing category into two-digit groups. Table 12 provides a comparison of group mean generation rates before and after covariate adjustment.

Inasmuch as truck trips have been dealt with separately, it remains to develop a measure of truck activity. The regression analysis shows that the hypothesis $b_i = 0$ could not be rejected in any two-digit group. An analysis of variance produced an F-value of 5. 4 and the hypothesis of equal group means could not be rejected:

$$F \text{ calculated} = 5.44 < F\ (99\%) = 6.63$$

Because the ratio of between group truck trip variation to within group truck trip variation is not significant, it may be concluded that using the overall average of truck trips would be no less accurate than employing the average number of truck trips for each of the two-digit classes of land use and, in addition, would be more suitable for purposes of projection.

TABLE 11

RESULTS OF COVARIANCE ANALYSIS
(Independent Variable = Employment)

Source	DF	YY	Sum Squares (due)	Sum Squares (approx)	Mean Square
Treatment (between)	19	87699032			
Error (within)	368	9656656	9478100	25576894	70072
Treatment error	387	70213086	69382549	24577184	
Difference for testing adjusted treatment means				1000288	52645

$F(19,365) = 0.551 < F_{.01;19,365} = 6.70$

TABLE 12

COMPARISON OF MEAN GENERATION RATES BEFORE AND AFTER COVARIATE ADJUSTMENT

Land Use	Mean Trips	Adjusted Mean	Std. Error of Mean
20	253	1, 359	60
21	250	1, 354	58
22	187	1, 374	58
23	1, 555	1, 375	59
24	3, 023	1, 331	66
25	2, 104	1, 266	59
26	5, 850	1, 491	69
27	828	1, 358	59
28	407	1, 264	61
29	1, 503	1, 416	58
30	2, 941	1, 282	62
31	72	1, 342	69
32	459	1, 384	65
33	1, 087	1, 360	63
34	9, 229	1, 340	59
35	1, 516	1, 430	60
36	1, 135	1, 379	59
37	185	1, 383	61
38	251	1, 357	63
39	592	1, 414	61

GENERAL CONCLUSIONS

The results of the preceding analysis lead to the following conclusions for the data used in the study:

1. If floor area (or truck trips) are employed as the independent variables in order to predict generated trips by manufacturing land-use activities, it is not sufficient to project these measurements for the major one-digit classification of manufacturing. Because of unequal slopes between two-digit groups, projection must be based on the two-digit classification of manufacturing land-use activities.
2. The use of employment as the predictive variable provides the best estimates of total person trips generated by manufacturing land-use activities. The results of the analysis indicate that no substantial accuracy can be gained by using multiple regression estimating equations. The difficulties of projection and possible colinearity between variables would indicate that a simple linear regression equation using employment as the independent variable is an adequate estimate of total trips.
3. Using employment as the independent variable allows the use of the generalized manufacturing category as the basis of projection. There are no significant differences of slope coefficients or means between two-digit land-use groups. The regression line may, therefore, be used at the one-digit classification level without loss of accuracy.

The validity of these results depends to some extent on the degree to which the assumption of homoscedasticity is violated. The X^2 value derived from Bartletts' test was approximately 1. 5 times the required X^2 value for rejection at 95 percent confidence. At 0. 995 the hypothesis of equal variances could not be rejected. Violations of this assumption require that the computed F statistic be slightly larger than indicated by the F limit for rejection of hypothesis at a specified level of confidence. In the case where employment is used as the independent variable, the hypothesis of equality of group slope coefficients and adjusted group means could not be rejected at 95 percent level of confidence. Violations in the equality of variances assumption would lead to the same conclusion with even greater confidence, i. e., inasmuch as the computed F statistic which was calculated was lower than the F value required for rejection, an increase in the required F value would only add support to our conclusion. In the cases where floor area and truck trips are employed as independent variables, the computed F statistic is several times larger than that required for rejection, and the conclusions formed are assumed to remain unchanged, although less confidence can be placed in the rejection of the null hypothesis.

In any event, the best estimating equation of manufacturing trip generation for the data analyzed is given by

$$\begin{aligned}\text{Total person trips} &= 33 + 1.112\ (\text{emp.}) + 176\\ &= a + b\ (\text{emp.}) + (\text{avg. truck trips})\end{aligned}$$

This equation assumes the following:

Standard error of regression estimate = 270
Standard error of slope coefficient = 0.006
Coefficient of variation (slope coeff.) = 0.54 percent
Standard error of truck trip estimate = 376
Standard error/avg. truck trips = 214 percent

Because of the large standard error of estimate and the relatively small variation in the regression coefficient, it might be concluded that trip generation estimates based on employment would be more suitable for planning networks where traffic zones were the basic unit of employment estimates than for projecting traffic demand of a specific generator. Thus where trip generation is to be determined for a specific generator with little employment, a larger percent variation would be expected.

One would suspect that there existed some functional relationship between truck trips and employment intensity. Table 8 indicated very little correlation within two-digit groups (coefficient of determination was not significant in any group). But Table 9 did indicate some functional relationship between two-digit groups. Between groups there seems to be an inverse relation between average employment intensity and truck trips. This relationship does not appear linear, but some type of inverse relationship does exist. Unfortunately, excessive variation of employment intensity within groups limits a valid analysis of this relationship.

For the data analyzed, it appears that trip purpose is an important consideration in relation to trips generated by manufacturing land-use activities. Truck trips indicate to some extent the functional relationship of the activity type and its employment intensity. At the 80 percent confidence limit it has been established that there are differences in the average truck trips between groups. The most significant purpose of trip-making to industrial land is for the purpose of work. Thus, two primary purposes of trip-making are relevant to manufacturing trip generation: those trips made for purpose of work and those which are otherwise related to the functional operation of the land-use activity. Fortunately, in the model derived, manufacturing employment is the only variable which must be projected. Regional growth models make use of employment projections, and this would complement the determination of manufacturing generation for work trips.