TRANSPORTATION RESEARCH RECORD 730

Issues in Transportation Planning for Small and Medium-Sized Communities

TRANSPORTATION RESEARCH BOARD

COMMISSION ON SOCIOTECHNICAL SYSTEMS
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C. 1979

Transportation Research Record 730
Price $3.00
Edited for TRB by Mary McLaughlin

modes
1 highway transportation
2 public transit

subject areas
12 planning
13 forecasting

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

Notice
The papers in this Record have been reviewed by and accepted for publication by knowledgeable persons other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The views expressed in these papers are those of the authors and do not necessarily reflect those of the sponsoring committee, the Transportation Research Board, the National Academy of Sciences, or the sponsors of TRB activities.

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Issues in transportation planning for small and medium-sized communities.

(Transportation research record; 730)
Reports for the Transportation Research Board 58th annual meeting.
1. Urban transportation policy–United States–Addresses, essays, lectures. 2. Local transit–United States–Addresses, essays, lectures. 3. Transportation planning–United States–Addresses, essays, lectures. I. Title. II. Series.
TE7.H5 no. 730 [HE308] 380.5s [388.4'0973] 80-10611
ISBN 0-309-02984-8 ISSN 0361-1981

Sponsorship of the Papers in This Transportation Research Record

GROUP 1–TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Leon M. Cole, Library of Congress, chairman

Transportation Systems and Planning Innovations Section
Richard J. Bouchard, Daniel, Mann, Johnson and Mendenhall, chairman

Committee on Transportation Planning Needs and Requirements of Small and Medium-Sized Communities
Max R. Sproles, Harland Bartholomew and Associates, International, chairman
Edward G. Bates, Jr., Michael J. Demetsky, David A. Doctor, Gorman Gilbert, Thomas J. Hillegass, Antoine G. Hobeika, R. Keith Jones, Joseph C. McBride, Carl C. McChesney, Monty C. Murphy, Edward S. Neumann, Marion R. Poole, Donald B. Rhodes, Robert R. Roberts, George E. Schoener, Kenneth W. Shiatte, John L. Skinner, Jr., Vergil G. Stover, John H. Waggoner, C. Michael Walton, Bruce B. Wilson, Eugene M. Wilson

James A. Scott, Transportation Research Board staff

The organizational units and officers and members are as of December 31, 1978.

Contents

Estimation of Trip Tables from Observed Link Volumes

Mark Turnquist*, School of Civil and Environmental Engineering, Cornell University, Ithaca, New York
Yehuda Gur, Urban Systems, Inc., Chicago

Traffic assignment studies, which are instrumental in transportation plan evaluation, require an origin-destination trip table as input. A method of estimating the trip table from observed link volumes within the framework of small-area assignment is described. The technique is based on solution of an optimization problem and is compatible with equilibrium traffic assignment assumptions. It thus differs significantly from previous methods based on statistical estimation. Because it avoids expensive home-interview surveys and time-consuming work with trip generation and distribution models, the method provides an attractive and cost-effective alternative analysis procedure for small and medium-sized communities. It is likely to be particularly effective in the evaluation of short-term, low-capital improvements in the transportation network.

This paper describes a model for estimating a trip table, or origin-destination (O-D) matrix, based mainly on the knowledge of link volumes and possibly turning movements. This matrix can be used in a number of traffic assignments in which alternative network improvements are examined. A major intended application of the model is as an element in a process of developing and evaluating short-range, subregional traffic improvement plans. As such, this tool should prove very useful to transportation planners in small and medium-sized communities.

The proposed model provides a cost-effective alternative to available methods for constructing trip tables, methods that are dependent on expensive, time-consuming surveys and a chain of trip generation and distribution models. Such methods often require planning resources in excess of what is available in many smaller communities and metropolitan areas. The proposed model, on the other hand, is based mainly on count data, which are easily obtainable. It will be most effective in analyzing short-range problems in which no significant changes in land use are expected.

This paper discusses the general nature of the problem and summarizes previous attempts to solve it. It then describes the proposed model and an example of its application.

NATURE OF THE PROBLEM

Any solution procedure for determining an O-D matrix from observed link volumes must address itself to two major problems. The first of these is that the set of link volumes may be internally inconsistent; that is, there may be no trip table that, when assigned, exactly duplicates the observed flows. There are several possible reasons for this, some of which are described below. It is important that the solution procedure be capable of finding an approximate solution to the problem in this case.

A second major problem is that, if link flows are consistent, exactly the same set of flows might be generated by many, quite different trip tables. Because there are typically far fewer links (pieces of information) in a network than O-D interchanges (unknowns), the problem is underspecified. Thus, there will be a set of feasible trip tables (perhaps many), all of which produce the same link volumes. A major element in the solution procedure must be a mechanism for identifying a desired trip table from among this set.

Inconsistency of Input Data

There are a number of reasons why there may be no trip table that exactly reproduces the observed link flows. First, a coded network is an abstraction of the road network being represented. Vehicle counts from the real network must be converted into volumes on more abstract links of the coded network, and this process can give rise to internal inconsistencies. This is particularly the case for the ends of trips, since the coded network describes trips as beginning or ending in a limited number of "load nodes" whereas, in reality, these trips originate and terminate at many locations on the network.

A second source of inconsistency is that the logic of traffic assignment is always based on the assumption (expressed in various forms) that travel is made along minimum impedance paths between an origin and a destination. This assumption, although plausible, is not a completely accurate description of the real world. Thus, one should expect certain discrepancies between the real world and its description in the model.

Finally, inconsistencies may arise from errors in measurement and definition. Traffic counts will likely contain some errors, and the counts on various links in the network will often be from different days and times. The travel impedance that is minimized by travelers is likely to be incompletely defined, and its relation to measurable attributes, such as travel time and distance, is not completely understood. These problems, among others, are likely to result in a set of input link volumes that cannot be reproduced by assigning any trip table.

The approach taken in this research is to "smooth out" these inconsistencies so as to find a set of link volumes that are approximately the same as the observed volumes but that could be produced by assigning a trip table to the coded network. Once such a set of link volumes is found, the problem of underspecification can be addressed.

Underspecification of the Problem

Link volumes alone (with or without turn volumes) do not provide enough information to construct a unique trip table; the same link volumes might be generated by highly different trip tables. A simple example will serve to demonstrate this problem.

Consider the simple network shown in Figure 1, with travel times and volumes as indicated. Nodes 1 and 2 serve as origins, and nodes 3 and 4 are destinations. If Wardrop's first principle—i.e., that all paths used between a given origin and destination have equal impedance and no path that is not used has an impedance less than that of paths that are used (1)—is assumed to determine network equilibrium, the observed flows could have arisen from either of the O-D matrices described in Figure 2.

There is no way to determine from link volumes alone which of the alternative O-D matrices is the appropriate one. However, these two O-D matrices might produce substantially different flows as a result of a change in the characteristics of a network link. Since the objective of this method is to provide trip tables that can be used to evaluate such changes, attention must be given to the problem of distinguishing among the alternative solutions.

Previous Attempts to Solve the Problem

The principal previous work on determination of O-D matrices from observed link volumes has been done by Robillard (2, 3), Nguyen (4, 5), and Gur and others (6). Robillard's work represents the first major effort in this area. He studied only the class of proportional traffic assignment algorithms, namely algorithms that do not take link capacity or congestion into account, either explicitly or implicitly.

Robillard's method of determining the trip table is to minimize a measure of the difference between observed link flows and assigned flow. For the case of proportional assignment algorithms, the assigned flows can be written explicitly in terms of the O-D interchange volumes, since the proportion of each O-D volume that uses a given link is constant. Robillard's approach is to solve a linear regression problem to determine total originating and terminating trips for each zone. A generalized gravity model is then used to determine the trip table.

Robillard's work offers some useful insights into the problem. First, it focuses attention on the need to obtain estimates of total origins and total destinations at each node. Second, the need to bring to bear additional information on the nature of trip distribution is recognized. This is an important element in the construction of a solution. However, Robillard's proposed solution technique is quite crude.

Nguyen (4, 5) has presented an approach to estimating an O-D matrix based on the assumption that observed link flows represent network equilibrium in the sense of satisfying Waldrop's first principle. Nguyen's work is valuable because it demonstrates a formulation that is compatible with nonproportional assignment techniques. This is a significant generalization of Robillard's earlier work, which was limited to proportional assignment methods. Nguyen's formulation of the problem is described in more detail later in this paper.

Figure 1. Simple network with indicated link flows and travel times.

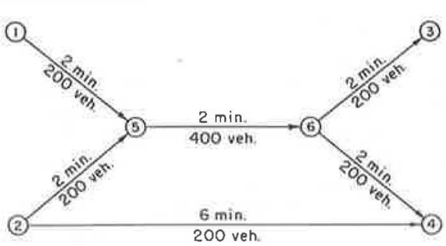


Figure 2. Possible trip tables for observed flows shown in Figure 1.

(a)

From \ To	3	4
1	0	200
2	200	200

(b)

From \ To	3	4
1	100	100
2	100	300

The approach Nguyen proposes for solving the problem suffers from a number of major deficiencies:

1. He does not address the problem of how to deal with the large number of potential solutions that reproduce the same link volumes.
2. He assumes that the observed flows are in equilibrium and does not address the problem of cases in which the observed volumes are inconsistent.
3. This proposed solution procedure, although theoretically correct, is extremely inefficient in application.

A third approach for solving the problem was formulated as a part of the present project (6). This approach used a linear programming formulation that specifically addressed the problems of volume inconsistencies and multiple solutions. The approach was formulated primarily for proportional assignment procedures. Although it was conceptually promising, it was not developed to its fullest extent. The major reasons for this were the need (as a part of project objectives) to address primarily nonproportional assignments and the need to invest heavily in software development before the approach could be fully developed and tested.

After the nature of the problem and possible solution procedures were studied, the direction chosen in this research was to use Nguyen's basic approach as a point of departure for developing an operational procedure.

APPROACH TO SOLVING THE PROBLEM

In general, the problem of finding an O-D matrix that, when assigned, would replicate observed link volumes can be stated formally as follows:

$$\min F = \sum_a \left[\int_0^{f_a} t_a(x)\,dx \right] - \sum_j \hat{u}_j T_j \tag{1}$$

subject to

$$T_j - \sum_k h_j^k = 0 \qquad \text{for each O-D pair } j \tag{2}$$

$$\sum_j \sum_r d_{ja}^k h_j^k = f_a \qquad \text{for each link } a \tag{3}$$

and

$$f_a,\ T_j,\ h_j^k > 0 \tag{4}$$

where

f_a = observed flow on link a,
$t_a(x)$ = impedance function for link a,
$\hat{u}_j$ = observed O-D impedance for interchange j,
T_j = trips for interchange j,
h_j^k = number of trips from interchange j using path k, and
$d_{ja}^k = \begin{cases} 1 & \text{if link a is in path k for interchange j} \\ 0 & \text{otherwise.} \end{cases}$

This is a nonlinear programming problem. The general form reflects the situation in which the observed link flows represent network equilibrium. In the special case in which no congestion effects are present, $t_a(x)$ is a constant and the first term of the objective function becomes simply $\Sigma_a t_a f_a$. This results in a linear programming problem.

In either event, the decision variables (or unknowns) are the T_j's and the h_j^k's, which represent trips on each interchange and allocation of these trips among paths, respectively. Note, however, that the h_j^k's are internal to the problem and can be computed indirectly. Nguyen (4) proved that the problem has a unique solution in link volumes (the observed link volumes); thus, the set of decision variables T_j (the output trip table) is a solution to the problem.

A solution to this mathematical programming problem is one trip table that could have resulted in the observed link flows. In light of the previous discussion, however, there are likely to be a number of alternative optima to this problem (i.e., alternative trip tables), each of which would produce the same link flows. The solution procedure, then, must be sensitive to this aspect of the problem and be capable of converging to a desired solution from among this set.

Since the optimization problem described above is very similar to the equilibrium traffic assignment problem, it is not surprising that solution algorithms for the two problems are quite similar. An iterative solution algorithm for this problem has been developed that requires the following basic steps:

1. Specify an initial trip table T^1 and a volume-delay (impedance) function for each link.
2. Find the $\hat{u}_j$ (skim trees) by using observed link impedances.
3. Assign T^1 to the unloaded network by using free-flow impedances to obtain a set of link volumes f^1. Denote this current solution as a vector (f^1, T^1).
4. Let $i = 1$.
5. Determine link impedances at the current volume f^i and again build minimum impedance trees. Denote the resulting values u^i.
6. Given T^i, $\hat{u}$, and u^i, find a correction trip table V that is closer to a solution.
7. Assign V^i to the trees built in step 5 to obtain correction link volumes s_i.
8. Find a weight r^i such that $0 \le r^i \le 1$ and the solution $[(f^{i+1}, T^{i+1}) = r^i(s^i, V^i) + (1 - r^i) \cdot (f^i, T^i)]$ minimizes the objective function F.
9. Check the convergence criterion. If it is met, stop; otherwise, set $i = i + 1$ and go to step 5.

Within the basic framework of this algorithm, there are a number of opportunities for variation. Specifically, one would expect the results to be sensitive to (a) choice of the initial trip table, (b) choice of link impedance functions, and (c) choice of procedure for computing V^i. The initial trip table is important because the problem generally has multiple optima. The algorithm will have a tendency to converge to a solution that is "nearest" to the initial solution. Link impedance functions are important because they play a central role in determining the order in which alternative paths for a given interchange are selected and thus the way in which equilibrium is approached. Obviously, the procedure for computing the correction trip table is a crucial element in the algorithm, since it determines the efficiency of the technique as well as the nature of the final solution.

A substantial battery of tests has been performed on this algorithm to determine its sensitivity to the choices indicated above under various network situations. The major findings are summarized here.

The theoretical constraints on the problem leave considerable latitude in selecting link impedance functions. Nguyen (5) has proved that the problem solution will replicate observed link flows as long as the impedance functions satisfy two simple criteria: (a) They must be increasing functions of volume, and (b) they must take on the value of observed impedance at the observed link volume.

Several forms of impedance functions have been tested, and the best results have been obtained by using a piecewise-linear form, as shown in Figure 3, in which t_0 is the observed impedance and f_0 is the observed volume. This function has provided superior results in terms of both the speed of convergence and the quality of the final solution.

A number of procedures for computing the correction trip table V^i are also possible. In his initial work on this problem, Nguyen (4) suggested a procedure in which a single interchange would be updated at each iteration. Although the convergence properties of such a technique can be established, it represents a very inefficient method for problems of realistic size. As a result, considerable effort has been devoted to developing more efficient procedures. To date, the procedures developed must be regarded as heuristics; i.e., there is no proof available that they will always converge to a solution. However, a number of these procedures have performed very well in empirical tests. The best correction technique identified thus far is as follows:

$$V_j^i = \begin{cases} T_j^i * 1 + 2 * [(\hat{u}_j - u_j^i)/(u_j^i - u_j^0)] & \text{if } \hat{u}_j > u_j^i \\ 0 & \text{if } \hat{u}_j \leqslant u_j^i \end{cases} \quad (5)$$

where u_j^o is the impedance for interchange j on the unloaded network (free flow). This correction is based on projecting the history of past corrections and has performed well in empirical tests. It must be emphasized, however, that it is a heuristic method.

As an example of the performance of the algorithm, Figure 4 shows how a solution is approached for a given network from three different starting points. The value on the ordinate is total volume error, computed as

$$\text{error} = \sum_a |\hat{f}_a - f_a| \quad (6)$$

where $\hat{f}_a$ is the observed volume on link a and f_a is the estimated volume from the algorithm. The network in question is shown in Figure 5, and the three initial trip tables are shown in Figure 6. Table 1 gives the observed volume and impedance attributes assumed for the network.

Trip table A represents essentially a "no-prior-information" starting point. All feasible interchanges are assigned equal numbers of trips, and the only information represented is that the matrix is scaled to match total observed vehicle hours traveled on the network. Trip table B represents a better starting solution but one that still contains major errors; i.e., this trip table, when assigned, would not produce the observed link volumes. Finally, trip table C represents a starting point that is quite close to an acceptable solution.

Figure 4 shows two important points about the algorithm. First, at least in terms of total volume error, the better the initial solution is, the better the computed solution will be after a given number of iterations. This is to be expected. However, the second point is that the procedure is capable of movement to a reasonable solution even when the starting point is far from a solution, as in the case of trip table A. This tends to increase our confidence in the convergence properties of the heuristic method.

It is also interesting to examine the estimated trip tables produced from the three different starting points. Figure 7 shows the solutions after 15 iterations. Note that the three solutions differ considerably from one another. This indicates the presence of alternative solutions, as discussed previously. Note also, however, that each solution has retained several important properties of its starting solution. These properties include the number of nonzero interchanges and variability in the relative size of interchange volumes.

This is very important, for it indicates the tendency of the algorithm to identify a solution that is close to the starting solution. This provides a mechanism for distinguishing among the alternative optimal solutions. By using as a starting point a "target" trip table that represents a number of properties we would like to see retained in the final trip table, we can generate a final solution that is close to this target, subject to the restriction that it (approximately) replicates observed link volumes.

Figure 3. Pseudo volume-delay function used in experiments.

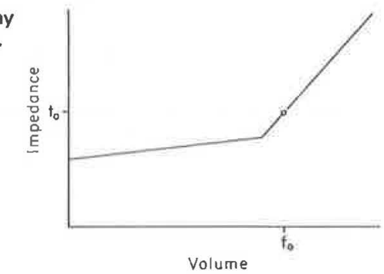


Figure 4. Total volume error in test network for three initial trip tables.

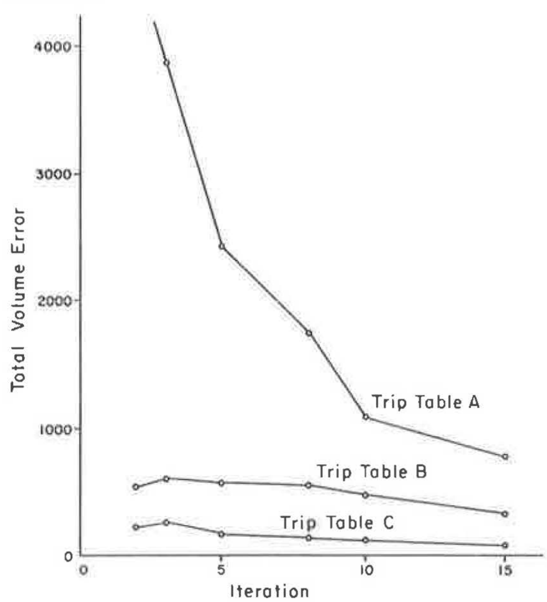
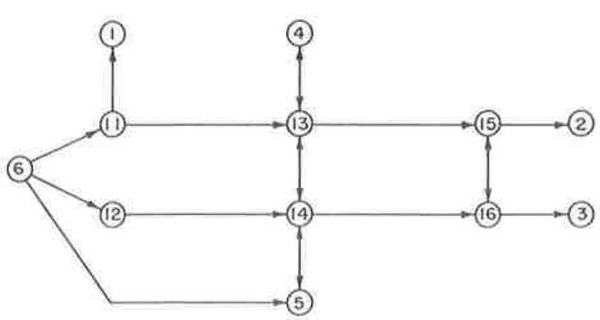


Figure 5. Test network.



This property of the algorithm has led us to devote substantial attention to the problem of constructing a reasonable target trip table. This was necessary because our major concern is small-area analysis and available trip distribution models deal primarily with large regions. A special trip distribution model—SMALD—has been designed for this purpose and is described in detail elsewhere (7).

If the solution trip table is judged by the analyst to be too far from the target table to be acceptable, postprocessing procedures are available to formally identify the trip table that will reproduce (approximately) observed link volumes and that lies closest to the target table. This procedure uses the basic information produced by the iterative algorithm in a second-stage mathematical programming problem. The technique is described by Gur and others (6). Experience to date has indicated, however, that this second-stage procedure is seldom necessary.

APPLICATION TO A REALISTIC NETWORK

To demonstrate the feasibility of the model in a realistic setting, it has been applied to a network that represents an area of about 15 km^2 (6 miles^2) in Hudson County, New

Figure 6. Initial trip tables for test network.

(A)

From \ To	1	2	3	4	5
4	-	983	983	-	983
5	-	983	983	983	-
6	983	983	983	983	983

(B)

From \ To	1	2	3	4	5
4		800	500	-	1100
5	-	1500	500	0	-
6	500	2000	0	2000	600

(C)

From \ To	1	2	3	4	5
4	-	600	300	-	1500
5	-	2000	3000	1500	-
6	500	4000	400	500	200

Table 1. Attributes of test network.

Link Origin Node	Link Destination Node	Loaded Impedance	Zero Volume Impedance	Observed Volume
4	13	10	7	2400
5	14	10	7	2000
6	5	40	30	100
6	11	10	7	5000
6	12	10	7	500
11	1	10	7	500
11	13	20	15	4500
12	14	20	15	500
13	4	10	7	2000
13	14	10	7	1500
13	15	20	15	4900
14	5	10	7	1600
14	13	10	7	1500
14	16	20	15	900
15	2	20	15	4800
15	16	10	7	300
16	3	20	15	1000
16	15	10	7	200

Jersey. The area includes 34 "boundary" load nodes and 24 internal load nodes. The coded network includes a total of 132 nodes and 369 links. The observed link volumes are based on adjusted ground counts collected from 1973 to 1975.

The initial trip table was estimated by using an early version of the SMALD model (7). This table contained a number of obvious imperfections, particularly total vehicle kilometers traveled, which was overstated by 14 percent. In the observed data, vehicle kilometers of travel = 1 482 304. The table below gives the results of the input trip table and the solution trip table for this measure (1 km = 0.62 mile):

Measure	Input Trip Table	Solution Trip Table
Vehicle kilometers of travel	1 729 567	1 482 169
Volume error (number of vehicles)		
With approach links	609 364	234 041
Without approach links	550 085	141 280
Root-mean-square error (%)		
With approach links	42.5	18.3
Without approach links	42.7	13.5

In assigning the inital trip table to the network by using five iterations of equilibrium assignment, a root-mean-square (RMS) error of 42.7 percent of the mean volume was found.

The model was applied to the problem for 35 iterations. The model changed the trip table substantially; the final trip table, when assigned, showed an RMS error of 13.5 percent, as given above, and underestimated the vehicle kilometers by about 2 percent. (Successful capacity-restrained assignments usually approximate observed volumes with an RMS error of 25-35 percent.) Figure 8 shows a plot of the observed versus assigned volumes that demonstrates close replication of the observed volumes over the entire volume range.

In a detailed analysis of the results, the following findings were made:

1. Many of the significant residual volume errors can be traced to inconsistencies in the input data (e.g., the total observed volume entering a network node is different from the total volume leaving that node). The algorithm distributed those errors in an acceptable way.

2. The extent to which total productions and attractions are preserved for internal zones can be controlled by the user, through the slope of the pseudo-delay function. Changes in production-attraction rates can be traded off against lower residual errors in the observed volumes. The trade-off should be based on the user's knowledge of the quality of the data.

3. The algorithm preserves the original trip table to the extent possible, subject to the objective of approximating the observed link volumes.

Figure 7. Output trip tables resulting from various initial trip tables.

(A)

From \ To	1	2	3	4	5
4	0	884	483	0	1147
5	0	1248	313	415	0
6	478	2743	259	1636	520

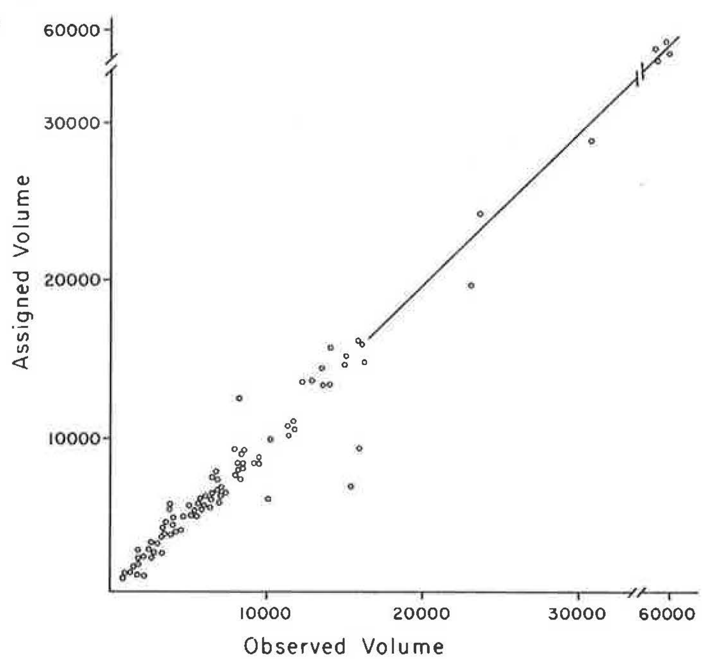
(B)

From \ To	1	2	3	4	5
4	0	725	554	0	1164
5	0	1547	453	67	0
6	495	2540	5	1955	590

(C)

From \ To	1	2	3	4	5
4	0	597	708	0	1096
5	0	1700	291	4	0
6	500	2503	0	1996	600

Figure 8. Observed versus assigned volume for Hudson County network.



4. Required computer resources are quite modest. The run required 65 s of central processing unit time and 120K bytes of core on an IBM 370/165.
5. Advance knowledge of total area vehicle kilometers of travel is not essential for a successful application of the model. Thus, the model can easily be enhanced to treat cases in which volumes are known for only part of the links.

The tests succeeded in demonstrating the applicability of the model to actual planning problems.

CONCLUSIONS

Traffic assignment is an important tool in the evaluation of potential changes in the transportation network. Such assignments require an O-D trip table. These tables have traditionally been estimated based on home-interview data and a complex chain of trip generation and distribution models. Careful analytic evaluation of transportation plans has thus often been beyond the financial capabilities of small and medium-sized communities. Even if the required data are available, in many cases they are outdated and inaccurate and thus result in assigned volumes that hardly resemble observed traffic. Such large discrepancies may cause the assignment analysis to be unreliable or even useless, especially for small-scale, short-range problems.

This paper documents a procedure by which an O-D trip table can be estimated based primarily on link-volume data. Most communities maintain an ongoing program for monitoring link volumes. Thus, they can easily provide a set of good ground counts to be used as a basis for constructing the trip table.

An important property of the problem in general is that there are usually alternative optimal solutions that correspond to different trip tables that produce the same link volumes when assigned. A characteristic of the algorithm developed is that it tends to converge to a solution that is close to the given starting point. This provides the motivation for creating a starting point (a target trip table) that represents desirable attributes. The algorithm will then modify this target table but only enough so as to result in approximate replication of observed link volumes.

Tests of the algorithm have indicated that it is relatively efficient from a computational standpoint and should thus represent a cost-effective tool for transportation planning in small and medium-sized communities. It should be particularly useful in evaluating short-term and/or low-capital improvements in the transportation network, which are not likely to result in significant changes in land-use or general trip-making patterns.

ACKNOWLEDGMENT

The work described in this paper was performed by John Hamburg and Associates under the sponsorship of the Office of Research, Federal Highway Administration, U.S. Department of Transportation. We wish to thank Steven Cohen, the contract manager, for his encouragement and helpful suggestions in all phases of the project. The advice and ideas of Morton Schneider and Larry Leblanc have also been of great value.

We are solely responsible for the findings and opinions expressed here.

REFERENCES

1. J. G. Wardrop. Some Theoretical Aspects of Road Traffic Research. Proc., Institute of Civil Engineers, Vol. 5, Part 2 (I), 1952, pp. 278-325.
2. P. Robillard. Estimating the O-D Matrix and Network Characteristics from Observed Link Volumes. Proc., 1st International Conference of Transportation Research, 1973, pp. 736-740.
3. P. Robillard. Estimating the O-D Matrix from Observed Link Volumes. Transportation Research, Vol. 9, 1975, pp. 123-128.
4. S. Nguyen. Estimating an O-D Matrix from Network Data: A Network Equilibrium Approach. Centre de Réchèrche sur les Transports, Univ. of Montreal, Rept. 60, Feb. 1977.
5. S. Nguyen. On the Estimation of an O-D Trip Table Matrix by Equilibrium Methods Using Pseudo Delay Functions. Centre de Réchèrche sur les Transports, Univ. of Montreal, draft paper, 1977.
6. Y. Gur, M. A. Turnquist, M. Schneider, and L. J. Leblanc. Determining an Origin-Destination Trip Table Based on Observed Volumes. Presented at Annual Meeting of Operations Research Society of America and the Institute of Management Science, New York, May 1978.
7. D. Kurth, M. Schneider, and Y. Gur. Small-Area Trip Distribution Model. TRB, Transportation Research Record 728, 1979, pp. 39-41.

**M. Turnquist was at Northwestern University when this research was performed.*

Equilibration of Supply and Demand in Designing Bus Routes for Small Urban Areas

Antoine G. Hobeika, Department of Civil Engineering, Virginia Polytechnic Institute and State University, Blacksburg
Chaushie Chu, Department of Civil Engineering, Northwestern University, Evanston, Illinois

A methodology that models the interactive relations between bus-system supply and demand and results in an optimal or near-optimal bus-route structure is described. On the supply side, the route structure is developed by using a heuristic algorithm called SWEEP, written in FORTRAN language. The algorithm partitions the total bus stops in the urban area into sectors and uses a three-optimum traveling-salesman algorithm or Hamiltonian-path algorithm to link these stops. The objective function of the algorithm is to minimize the total distance traveled by all buses, subject to the capacity and distance constraints on each bus. On the demand side, the program uses the already developed bus network to determine the percentage of total community travel that requires bus service. This is carried out by using a disaggregate mode-choice model that is based on the total time and cost difference between travel by automobile and travel by bus for each individual user. Costs of bus operations are calculated from a four-variable unit-cost model. An iterative, interactive feedback process is used to achieve the equilibrium state of the transportation market. Equilibrium is reached when the bus share of the transportation market cannot be increased by improving the bus network, under certain resource conditions and financial constraints. The program is tested in developing feeder bus routes to the proposed Glebe Metro station in Arlington County, Virginia.

The problems of bus transit planning in small urban areas can be boiled down to two major aspects—supply and demand. On the supply side, the development of route structure, the frequency of buses on each route, and the estimation of system operating costs and required subsidies are all bus functions that require improvement in the existing transit planning process. The system-wide configuration of bus routes, which is the working skeleton of the transit system and the medium of contact between the users and the bus company, is still developed by hand. Bus routes and frequencies are developed on a qualitative basis by using a number of routing and operating criteria to judge the route network (1, 2). This procedure limits the number of alternative configurations to be considered, does not have a defined objective function, and does not make use of the interactive relations between supply and demand. Expected operating costs and required subsidies cannot be treated independently of developed route configurations and equilibrium demand functions. On the other hand, demand is a function of the attractiveness of the supply system, which includes such factors as the characteristics of the bus system and its performance under specific physical and financial constraints. In short, a computerized methodology that will equilibrate the supply and demand functions where the supply functions are endogenous to the model is not available for bus-system planning in small urban areas.

Few mathematical models have been formulated to determine the equilibrium condition of the supply and demand functions in a transportation network. Equilibrium conditions in highway networks in which demand is elastic have been investigated by Florian and Nguyen (3), Martin and Manheim (4), Wilkie and Stefanek (5), and Wigan (6). A survey of the literature and possible approaches to the problem are presented by Ruiter (7). Kulash (8) has developed two simulation models for analyzing fixed-route bus systems. These models evaluate the quality of service that results from various operating policies and are used to predict the impacts of various operating decisions that are needed to improve route and schedule designs. Yet the route structure in Kulash's models is still developed manually, and no demand interaction is considered. Similarly, Rapp and Gehner (9) developed an interactive graphic computer system known as the Urban Transit Analysis System (UTRANS). UTRANS is used to evaluate different route and schedule policies based on quality of service. Frequency and route structure, again, are used as input data.

The methodology presented in this paper develops bus-route structure and computes equilibrium flows in a network in which demand is elastic.

METHODOLOGY

Equilibrating the supply and demand functions of the bus system requires the following models: a collection of supply models, a demand model, and an integrating model. The demand model represents the demand side of the system. It predicts the passenger demand for each bus stop in the system based on economic and population forecasts and on characteristics of the various transportation modes that serve the city. The supply models include the set of activities that represent the flow of passengers on bus routes. Two models represent the supply activities: the network development model and the cost model. The integrating model is a program that processes the supply and demand functions to determine an equilibrium of supply and demand quantities. The structure of the supply-demand equilibrating framework is shown in Figure 1.

It is useful to discuss the methodology in three sections: (a) supply functions, (b) the demand model, and (c) the integrating model.

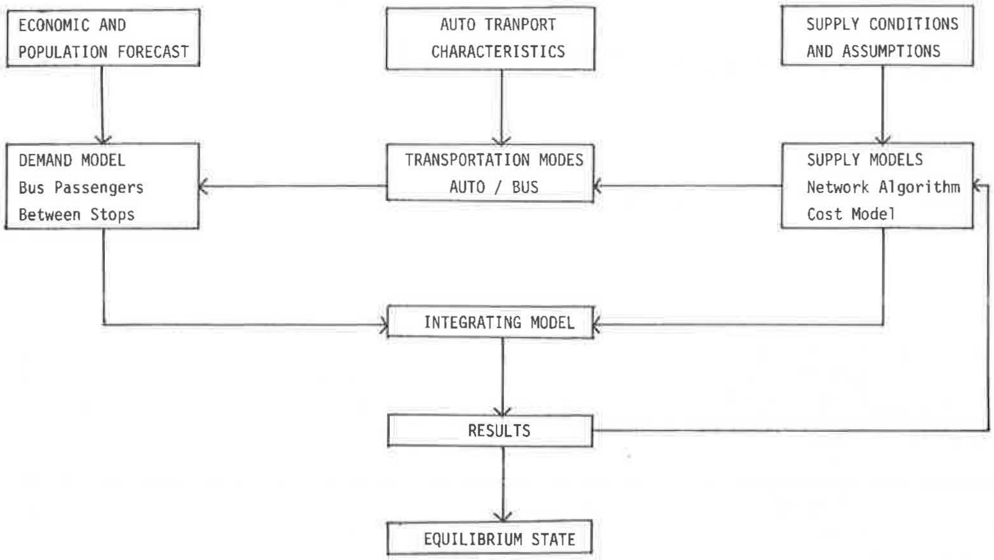
Supply Functions

Network Development Model

The model used to develop route structure, referred to as the SWEEP algorithm, was originally proposed by Gillet and Miller (10) as a solution to the general problem of vehicle dispatch. This algorithm was modified and applied to bus network design (11).

Optimal computer algorithms for development of bus routes had been investigated and applied in the past and found to be ineffective and to require an extensive amount of computer time. But, in light of the new generation of heuristic algorithms now available in the literature for solving vehicle-dispatching problems (12-19), these limitations are no longer valid. Near-optimal solutions to problems of large bus networks could be accomplished

Figure 1. Supply-demand equilibrating framework.



in a reasonable amount of computer time.

The objective of the SWEEP algorithm is to minimize the total distance or time traveled by all buses in satisfying demand at all bus stops, subject to the load and distance constraints on each bus. The problem is to determine the number of routes and the paths in each route according to the above objective. The coordinates of each bus stop and bus-stop demand are inputs to this algorithm. The following notation will aid in its explanation:

N = number of bus stops, including the transfer station (the transfer station is bus stop 1 and the dispatch point of all buses);
Q(I) = demand at bus stop I (I = 2, 3, ..., N);
X(i), Y(I) = rectangular coordinates of I bus stop (I = 1,2,3, ..., N);
C = capacity of each bus;
D = maximum distance each bus can travel;
A(I, J) = distance between stops I and J;
An(I) = polar coordinate angle of the Ith stop (I = 2,3, ..., N); and
R(I) = radius from transfer station to stop (I).

The constraints on the problem are as follows:

$$Q(I) \leq C \tag{1}$$

for all I;

$$A(I, J) > 0 \tag{2}$$

for all I ≠ J;

$$A(I, I) = W \tag{3}$$

where W is a constant that denotes extra distance or time per stop; and

$$A(I, 1) + A(1, I) \leq D \tag{4}$$

for all I.

The stops are renumbered according to the size of their polar-coordinate angle, and the transfer station is stop 1. The stops are partitioned into routes beginning with the stop that has the smallest angles—namely, stop 2. The first route, then, consists of stops 2, 3, ..., J, where J is the last stop that can be added without exceeding the vehicle capacity or distance constraint. The second route contains stops (J + 1, J + 2, ..., L), where L is the last stop that can be added to the second route without exceeding the constraints. The remaining routes are formed in the same manner.

The total distance or time traveled is just the sum of the distances for each route. An iterative procedure is then used to improve the total distance traveled by replacing one stop in route K with one or more stops in route (K + 1) for (K = 1, 2, ..., m - 1), where m is the number of routes formed.

The process of adding one or more stops to route K and deleting another stop continues until no improvements are found. The X and Y axes are then rotated counterclockwise so the first location becomes the last, the second becomes the first, and so forth.

This procedure of partitioning routes and interchanging locations between routes is then repeated until all possibilities have been exhausted and the minimum total distances are calculated. The smallest of these minimums provides a good heuristic solution for a bus-route network.

Each time a set of bus stops is considered for a given route, a "traveling-salesman" algorithm is solved to determine the minimum path to service each of the stops in the route. As a consequence, a loop route is formed. To modify the algorithm to form linear bus routes instead of loop routes, a Hamiltonian-path algorithm is developed to replace the traveling-salesman algorithm.

The solution to the well-known traveling-salesman problem, in which the origin and destination points coincide, represents a "Hamiltonian circuit". A Hamiltonian path represents a route in which the origin and destination vertices are at two distinct points. The linear bus route, on which the bus goes to the end of the route and returns along the same route, is mathematically equivalent to the determination of the shortest Hamiltonian path.

In the algorithm for determining the shortest Hamiltonian path, which is given below, the following notation is used: Edges are incident to or incident from a vertex; $v_1, v_2, \ldots, v_n$ = the set of vertices; and L_{ij} = distance or time for the edge incident from v_i and incident to v_j.

Figure 2. Sample of output of SWEEP algorithm.

```
Best solution is:
Route 1 has load 46.00 with distance 184.54 is
102  49  50  1  44  45  103  107  106  104  105
78  79  80  82  75
Route 2 has load 46.00 with distance 255.44 is
68  91  92  66  94  111  114  110  109  81  108
69  76  1  48  72  77  113  73  112  90
Route 3 has load 44.00 with distance 281.23 is
2  1  70  74  83  95  64  65  67  63  62  84  85
61  58  57  86  87  71
Route 4 has load 48.00 with distance 235.73 is
100  43  34  33  1  46  47  42  41  54  59  55
56  60  101
Route 5 has load 45.00 with distance 218.14 is
96  97  99  98  52  53  37  36  35  1  32  89
Route 6 has load 47.00 with distance 205.15 is
1  6  24  88  38  39  40  31  18  10
Route 7 has load 47.00 with distance 192.33 is
1  6  24  88  38  39  40  31  18  10
Route 8 has load 43.00 with distance 182.97 is
5  1  11  9  14  29  28  27  26  25  30  17  3
Route 8 has load 43.00 with distance 182.97 is
4  23  20  19  15  1  8  16  13  22  21
Route 9 has load 19.00 with distance 54.49 is
1  12  7  51
Total Distance Is. . . . .1810.01685  (Distances are in 100 feet)
```

The steps in the Hamiltonian path are as follows:

1. Define a T set of edges to be included in the Hamiltonian path. Initially, T is empty.
2. Determine the shortest L_{ij}. Include the edge (v_i - v_j) in T.
3. Merge the vertices v_i and v_j to form a single vertex. This would imply the revision of distance L_{ik} and L_{pj}, where v_k and v_p are typical vertices other than v_i and v_j.
4. Check as to whether the set T has (n - 1) edges. If yes, go to step 5; if no, go to step 2.
5. The set T defines the shortest Hamiltonian path.

The routes developed by the algorithm, whether one uses the traveling-salesman path or the Hamiltonian path, conform to the general guidelines for bus-route development (1) and yet minimize the total distance or time traveled by all buses. The computer output under each variation shows the number of routes, the path of each route, the distance traveled by each bus, and the total number of passengers carried on each route. A sample of the output of the SWEEP algorithm is shown in Figure 2.

The algorithm described above is version 1 of the SWEEP algorithm, which determines a bus network in which there is only one central transfer point in the service area. To overcome this weakness, version 2—development of which is under way—deals with multi-terminal network design. The basic concept of version 2 is the same as that of version 1. Its modifications can be briefly described as follows.

Assume M to be the number of assigned terminal stations, which usually are the major traffic attractors in the city. The algorithm first assigns each bus stop to its most appropriate terminal by a ratio scheme and then partitions the primary problem into M smaller subproblems. These individual subproblems are considered as problems of single-terminal network design and can be solved by using version 1 of the SWEEP algorithm. The so-called ratio scheme involves the following steps:

1. Initially, all bus stops are unassigned. For each bus stop i, find the closest terminal j1 and the second closest terminal j2 and compute the ratio

$$r(i) = d(i, j1)/d(i, j2) \tag{5}$$

where d(i, j) is the distance between nodes i and j.

2. Assign all bus stops i that have r(i) greater than 2 to their closest terminal.
3. For each unassigned bus stop i, find j*, k* such that

$$\text{Min}\,\{d(i, j) + d(i, k) + d(j, k)\} = d(i, j^*) + d(i, k^*) + d(j^*, k^*) \tag{6}$$

for all j, k where j and k are bus stops already assigned to the same terminal, either j1 or j2. Assign stop i to the terminal to which j* and k* are assigned.

In this procedure, if r(i) is large it means that stop i is relatively close to one terminal. All such stops are immediately assigned to the closest terminal. Stops that are more or less midway between two terminals are assigned more carefully. The minimization implies including stop i between j* and k* as linearly as possible.

Cost Model

The task of explaining total operational costs as a function of the output and characteristics of the system has proved in many cases to be very difficult. But this is not the case for the operational costs of bus systems.

Several different approaches have been taken to developing cost models for bus systems, but they are all basically single-equation expressions of cost as a function of the output of the system. These models can be primarily categorized into three types: the four-variable unit-cost model, the four-variable regression model, and the slowness function model. A review and comparison of these models can be found elsewhere (20). Hurley and Siegel (21) suggest that "the unit-cost method of determining parameters appears to be an accurate method when used to predict future costs for the same system" and "the four-variable model is equal to, and usually superior to, the slowness function." Therefore, the computer program uses the unit-cost model, under some reasonable assumptions, to generate operational cost estimates.

The four-variable unit model has the following general form (since the models presented in this paper were formulated in U.S. customary units of measurement, no SI equivalents are given):

$$OC = a{*}VM + b{*}VH + c{*}PV + d{*}RP \tag{7}$$

where

OC = annual operational costs,
VM = annual vehicle miles,
VH = annual vehicle hours,
PV = number of peak-hour vehicles,
RP = annual revenue passengers, and
a, b, c, d = unit costs for their corresponding variables.

According to data collected in 1970, the national averages for these unit-cost coefficients for public bus operations are as follows:

$$OC = 0.277\,VM + 5.700\,VH + 6527.480\,PV + 0.038\,RP \qquad (8)$$

These costs are based on the 1970 dollar value and are multiplied by the inflation factor to estimate design-year operational costs. In this case, the average inflation factor is considered to be 7 percent for each year since 1970.

In the absence of particular specifications, the magnitude of these four variables is calculated by using the following relations:

VM = total bus-route miles * service frequency in vehicles per hour * operating hours per day * operating days per year,
VH = vehicle miles/bus average speed in miles per hour,
PV = number of vehicles in peak-hour operation, and
RP = total number of passengers on all routes for each bus trip during peak hours (capacity) * service frequency in vehicles per hour * [peak-hour operation per day + (operation per day - peak hours per day)/2] * operating days per year.

Demand Model

Travel demand is divided into captive and choice riders. The captive riders are further subdivided into automobile captives and bus captives. The division between choice and captive riders is assumed to be known a priori, through survey or other data.

The demand model for choice riders used in this research is an individual mode-choice model based on logit functions. The mathematical expression of this model is

$$P_{i(b)} = \exp(B_i) \Big/ \sum_{j=1}^{n} \exp(J_i) \qquad (9)$$

This equation states that the probability of a passenger i taking the bus travel mode (b) is the exponential of the bus model utility (B_i) divided by the sum of the total exponentials of all modal utilities $\sum_{j=1}^{n} \exp(J_i)$ in the market, where n is the number of total available modes. In a small urban area in which there are only two major modes available, the equation is

$$P_{i(b)} = \exp(B_i)/\exp(A_i) + \exp(B_i) \qquad (10)$$

where A_i and B_i are the utilities of automobile mode and bus mode for an individual traveler i.

The development and use of these types of models, also referred to as disaggregate travel behavior models, are fully described and discussed elsewhere (22-26).

The above equation is simplified by dividing both the numerator and the denominator by $\exp(B_i)$. The following equation is obtained:

$$P_{i(b)} = 1/[1 + \exp(Z_i)] \qquad (11)$$

where Z_i is the difference in utility function between bus and automobile.

There are many exogenous variables that can be considered in calibrating the difference utility function Z_i. Because of unavailable data, a model calibrated for Schaumburg/Hoffman Estates Transit demand prediction (23) is used here. The difference utility function is

$$Z_i = -1.37 + 0.0544(T_a - T_b) + 0.0021(C_a - C_b) \qquad (12)$$

where

T_a = total travel time by automobile,
T_b = total travel time by bus,
C_a = total travel cost by automobile, and
C_b = total travel cost by bus.

Integrating Model

The integrating model is a simulation model that uses the route network and the operating costs developed by the supply functions and the demand quantities from the demand model to determine the equilibrium state of bus-system supply and demand, under specific operational and financial management policies such as bus fare, the boundaries of the service area, and bus headways and capacity. The input data for the integrating model include

1. The bus network and its associated operating cost developed from the supply model;
2. The O-D matrix for the system service area;
3. The split for automobile-captive riders, choice riders, and transit-captive riders; and
4. All values of travel time and cost parameters used in the demand model.

If the split for automobile captives, choice riders, and transit captives is not available for the service area, it is assumed that 75 percent of total trip makers are choice riders and 25 percent are captives. Among the captives, 25 percent are assumed to be transit captives and 75 percent automobile captives. In other words, 75 percent of total trips are sensitive to the level of service of the bus system and applicable to the demand model.

The simulator is based on the Monte Carlo technique. It first calculates the probability of a trip origin for each bus route and each bus stop. The individual route probability is the initial total demand on the route divided by the total bus demand over the whole service area. The nodal probability is the amount of initial demand at the particular bus stop divided by the total amount of initial demand along the route. Uniformly distributed random numbers between zero and one are generated and fed into the foregoing cumulative probabilities for the routes and stops, respectively. Comparisons with the calculated route and nodal probabilities determine the location of the trip origin. The location of the trip destination is determined by repeating the same process based on the number of passengers getting off at each bus stop.

The travel times for this trip are calculated separately for the two travel modes, private automobile and public bus. Direct travel times are determined based on the available street network and the first developed bus network. Automobile access time and bus walking time, waiting time, boarding and departing time, and transfer time are all determined from appropriate distributions and appropriate boundary values (8, 27, 28) that correspond to the system under consideration. The total travel times for each individual trip, plus other utility variables determined or obtained from supply functions and operational policies (such as automobile travel cost, bus fare, and automobile parking fare), are fed into the demand model to obtain the individual probability of choice rider i traveling by bus, referred to here as the bus system's attractiveness to choice rider i.

This simulation process is repeated for a large sample of individual trips. The sample size is chosen to be statistically acceptable. Then statistics on all of these individual probabilities (bus attractiveness) are collected according to classes of trip length. The mean of total probabilities—the bus share of choice riders in the transportation market—is assigned the variable name PROB2, which implies the "current" attractiveness of the bus system, in contrast to PROB1, the "previous" attractiveness of the bus system.

PROB1 is among the initial input data, which are determined by either demand survey or analogy. PROB1 can be calculated as follows:

$$\text{PROB 1} = (\text{total bus riders} - \text{bus captives})/[\text{total travelers} - (\text{bus captives} + \text{automobile captives})] \quad (13)$$

If the numbers of automobile captives and bus captives are not available, the model determines PROB1 by the previously assumed division between captive and choice riders:

$$\text{PROB 1} = (\text{total bus riders} - \tfrac{1}{16}\ \text{total travelers}) \div (\text{total travelers} - \tfrac{1}{4}\ \text{total travelers}) = (\text{bus choice riders}/\text{total choice travelers}) \quad (14)$$

The values of PROB1 and PROB2 are compared. If they are different, the market is in an unstable condition and demand is subject to change. Demand at each bus stop (one of the initial input data) is modified by the following relation:

$$Q(I) = Q(I)\,\{[1 - (\text{bus captives}/\text{total choice riders})]\,(\text{PROB 2} \div \text{PROB 1}) + (\text{bus captives}/\text{total choice riders})\} \quad (15)$$

or, for the assumed division,

$$Q(I) = Q(I)[(11/12) \times (\text{PROB 2}/\text{PROB 1}) + (1/12)] \quad (16)$$

where bus captives/total choice riders = (1/16)/(3/4) = 1/12. The previous system attractiveness is then dropped and the current attractiveness is substituted; i.e., PROB1 = PROB2.

New demands [Q(I)] at each bus stop are fed back to the SWEEP algorithm, and the whole process is performed again. Current system attractiveness (PROB2) is once again generated from the integrating model. PROB1 and PROB2 are compared. If they are different, the whole process is repeated. The termination of the iterative process occurs only if PROB2 falls within a 95 percent confidence interval of PROB1. At this stage, it is assumed that the equilibrium state is reached—i.e., that the supply and demand sides are in stable condition. The equilibrating strategy is shown in Figure 3.

As Figure 3 shows, the equilibrium state or condition changes with system management policies. Different sets of policies will produce different levels of bus-system attractiveness. Under a specific set of policies, maximum local attractiveness is obtained at the equilibrium state. When different policies are evaluated by sensitivity analysis, the one that generates global maximum attractiveness is the most preferable set of policies if maximization of attractiveness is the management objective.

APPLICATION

The methodology described above is now applied to the development of feeder bus routes to the proposed Glebe Station of the Washington, D.C., Metro rail rapid transit system. The study area is located in the midwestern section of Arlington County, Virginia. It surrounds the last Metro station on the Rosslyn-Ballston Corridor and acts as a catchment area for the station. The boundaries of the designated area, determined by using a conservative radius from the transit station, extend 3.2 km (2 miles) to the north, south, and west of the station. The area to the east is assumed to be serviced by the previous station on the Metro line.

The locations of bus stops are determined by land-use patterns, the major attractors, the road network, and passenger walking time. Candidate streets for bus routes are initially identified. These streets fall into the categories of major and minor arterials. Bus stops are then located on these streets based on land-use patterns, the major generating and attracting points of the study area, and the consideration that passenger walking distance to a bus stop should not exceed 0.4 km (0.25 mile). Because 70 percent of morning-peak work trips in the study area are bound for Washington, D.C., the Metro station can be considered as the major traffic attractor in the area and version 1 of the SWEEP algorithm appears to be applicable. The Metro station is considered to be bus stop 1 and the dispatch point of all buses. One hundred and thirteen bus stops, excluding the Metro station, were developed for the study area. The study area and its boundaries, as well as the location of bus stops, are shown in Figure 4.

Initial demand for each bus stop is calculated from the 1970 census-tract data for bus ridership in the study area.

Various computer runs were conducted by using different input parameters. On the supply side, bus capacity, maximum allowable travel distance for each bus, and the schedule of bus headways are all input variables. On the demand side, bus-fare structure, automobile parking fare, and automobile cost per mile are also input parameters that affect the measure of bus-system attractiveness. All of these variables are tested in order to investigate how alternative policies will affect the attractiveness of bus to the community.

The findings can be summarized as follows:

1. Capacity and distance constraints versus operational cost—Two types of buses were chosen for computer runs: buses with 32 seats and a maximum allowable travel distance of 9 km (5.68 miles) and buses with 50 seats and a maximum allowable travel distance of 13 km (7.97 miles). The results show that the system with the larger buses will operate at lower cost.
2. Capacity and distance constraints versus attractiveness—From intuitive judgment, a decrement in attractiveness should result if buses of greater capacity are used. The bus with greater capacity can serve longer routes, which will consume longer travel time. Not too surprisingly, however, bus capacity does not significantly affect system attractiveness in this case study. The variation ranged between 0.005 and 0.01. Such a small range of variation can be attributed to the structure of the behavioral model.
3. Fare structure and parking price versus attractiveness—Four levels of bus-fare policies were tested: \$0.25, \$0.30, and \$0.50/ride and free fare. Three levels of parking-price policies were tested: \$1, \$2, and \$3/trip. The results show that higher bus fares suppress ridership and higher parking prices increase ridership. Increasing parking price will be a more effective means of increasing ridership than decreasing bus fare.
4. Trip length versus attractiveness—The simulation output stratifies the data on trip length. The results show that the range of maximum distance for bus use is approximately 4.8-6.4 km (3-4 miles). When trip length is longer, the attractiveness of bus decreases rapidly.

Other findings related to the supply side are that the

Figure 3. Flow diagram of equilibrating strategy.

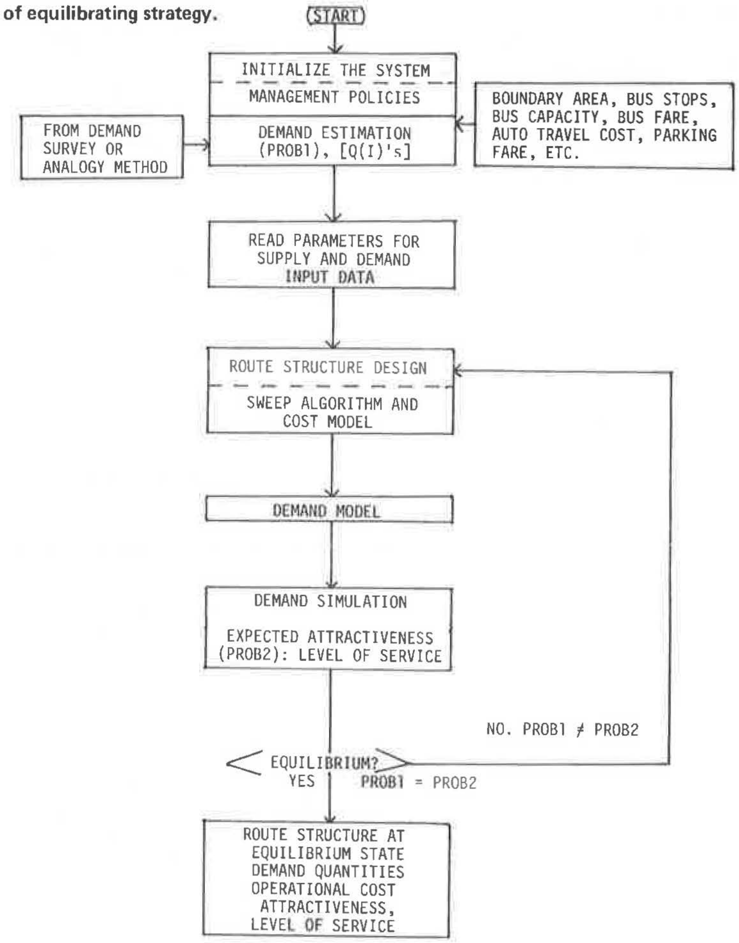


Figure 4. Selected bus-stop locations in the study area.

linear routes developed by the Hamiltonian-path algorithm, in comparison with the loop-routes developed by the traveling-salesman algorithm, saved 5-10 percent of the total distance traveled by all buses under the same conditions of bus capacity and travel-distance constraint. Although both algorithms are computationally efficient, the Hamiltonian-path algorithm is far superior for this specific problem. Each computer run for the Hamiltonian path took an average of 1 min of central processing unit (CPU) time versus 9.5 min for each run with the traveling-salesman algorithm. In both cases, computer time increased linearly with the total number of bus stops and quadratically with the number of stops per route.

Based on the sensitivity experiments, the following observations are made:

1. The larger the constraint is on the distance traveled by each bus, the fewer routes will be formed. This yields a lower distance traveled for all buses.
2. The greater bus capacity is, the fewer routes will be formed. This yields a lower total distance traveled.

CONCLUSIONS

The methodology described in this paper is a comprehensive computer model and an efficient tool for analyzing and evaluating bus systems in small to medium-sized urban areas. It determines the impacts of different policies and conditions on the bus system. It can be useful to small bus companies in evaluating and designing their systems and to local planning commissions in planning bus systems in their communities. The model is easy to use and provides quick answers to many of the decision maker's questions. In addition, it is a strong educational tool by means of which the user can easily learn the interactions among the different elements of the bus system.

However, additional improvements and modifications in bus-network development in the model are being sought, such as (a) inclusion of the actual street network and variable travel times on each link of the network and (b) a built-in capability in the algorithm to change the locations of bus stops and their spacing according to demand.

REFERENCES

1. A. G. Hobeika and V. K. Walker. Developing the Fixed Routes of a Bus System in Small Urban Areas. Modeling and Simulation, Proc., 6th Annual Conference, Pittsburgh, 1975.
2. G. M. Smerk and others. Mass Transit Management: A Handbook for Small Cities. Institute for Urban Transportation, Indiana Univ., Bloomington, Feb. 1971.
3. M. Florian and S. Nguyen. A Method for Computing Network Equilibrium with Elastic Demands. Transportation Science, Vol. 8, 1974, pp. 321-332.
4. B. V. Martin and M. L. Manheim. A Research Program for Comparison of Traffic Assignment Techniques. HRB, Highway Research Record 88, 1965, pp. 69-84.
5. D. F. Wilkie and R. G. Stefanek. Precise Determination of Equilibrium in Travel Forecasting Problems Using Numerical Optimization Techniques. HRB, Highway Research Record 369, 1971, pp. 239-252.
6. M. R. Wigan. Benefit Assessment for Network Traffic Models and Application to Road Pricing. British Road Research Laboratory, Crowthorne, Berkshire, England, Rept. LT417, 1971.
7. E. R. Ruiter. The Prediction of Network Equilibrium: The State of the Art. Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, June 1973.
8. D. J. Kulash. Routing and Scheduling in Public Transportation Systems. Massachusetts Institute of Technology, Cambridge, Ph.D. thesis, March 1971.
9. M. H. Rapp and C. D. Gehner. Criteria for Bus Rapid Transit Systems in Urban Corridors: Some Experiments with an Interactive Graphic Design System. Urban Transportation Program, Univ. of Washington, Seattle, Res. Rept. 72-1, April 1973.
10. B. E. Gillett and L. R. Miller. A Heuristic Algorithm for the Vehicle Dispatch Problem. Operations Research, No. 22, 1971, pp. 340-349.
11. A. G. Hobeika and V. K. Walker. A Heuristic Algorithm for Developing Feeder Bus Routes to a Metro Station. Modeling and Simulation, Proc., 8th Annual Conference, Pittsburgh, 1977.
12. H. Christofides and S. Ellen. An Algorithm for the Vehicle Dispatching Problem. Operations Research Quarterly, 1969, pp. 309-318.
13. M. Iri, T. Morigaki, S. Moriguti, and M. Yabe. A Network Model for Commuter Transportation. Journal of Operations Research Society of Japan, Vol. 10, Nos. 3-4, June 1968, pp. 108-124.
14. W. S. Jewell. Allocation of Route Service in a Transportation Network. Presented at 4th International Symposium on the Theory of Traffic Flow, Karlsruhe, Federal Republic of West Germany, June 17-20, 1968.
15. R. L. Karg and G. L. Thompson. A Heuristic Approach to Solving Traveling Salesman Problems. Management Science, 1964, pp. 225-247.
16. W. Lampkin and P. D. Saalmans. The Design of Routes, Service Frequencies, and Schedules for a Municipal Bus Undertaking: A Case Study. Operations Research Quarterly, Vol. 18, No. 4, Dec. 1967, pp. 375-397.
17. S. Lin and B. W. Kernighav. An Effective Algorithm for the Traveling Salesman Problem. Bell Laboratories, Murray Hill, NJ, Computer Science Tech. Rept. 1, April 1972.
18. J. D. Little, K. G. Murty, D. W. Sweeney, and C. Karel. An Algorithm for the Traveling Salesman Problem. Operations Research, No. 11, 1963, pp. 972-989.
19. R. Seshagiri, R. Harasimhan, S. L. Mehndiratta, and B. K. Chanda. Computer-Generated Time Table and Bus Schedules for a Large Bus Transport Network. Transportation Science, Vol. 3, No. 1, Feb. 1969, pp. 69-85.
20. J. H. Miller and J. C. Rea. Comparison of Cost Models for Urban Transit. HRB, Highway Research Record 435, 1973, pp. 11-20.
21. J. W. Hurley and R. L. Siegel. A Cost Forecasting Model for Bus Transit Operation. Modeling and Simulation, Proc., 7th Annual Conference, Pittsburgh, 1976, pp. 427-432.
22. M. G. Richards and M. E. Ben-Akiva. A Disaggregate Travel Demand Model. Saxon House, D.C. Heath, Ltd., England, 1975.
23. B. D. Spear. Application of New Travel Demand Forecasting Techniques to Transportation Planning: A Study of Individual Choice Models. Federal Highway Administration, U.S. Department of Transportation, March 1977.
24. P. R. Stopher. Derivation of Values of Time from Travel Demand Models. TRB, Transportation Research Record 587, 1976, pp. 12-18.
25. T. C. Thomas and G. I. Thompson. The Value of Time for Community Motorists as a Function of Their Income Level and Amount of Time Saved. HRB, Highway Research Record 314, 1970, pp. 1-19.
26. S. L. Warner. Stochastic Choice of Mode in Urban Travel: A Survey of Binary Choice. Northwestern Univ. Press, Evanston, IL, 1962.
27. J. J. Bakker. Transit Operating Strategies and Levels of Service. TRB, Transportation Research Record 606, 1976, pp. 1-5.
28. V. G. Hurdle. Minimum Cost Schedules for a Public Transportation Route. Transportation Science, Vol. 7, No. 2, pp. 109-157.

Demand-Estimating Model for Transit Route and System Planning in Small Urban Areas

Marvin Golenberg, SG Associates, Inc., Boston
Steven Pernaw, Alan M. Voorhees and Associates, Inc., Boston

A simplified model for directly estimating transit route and system patronage for small urban areas is presented. A category approach is used to determine basic transit trip generation by automobile-ownership classification. The basic rate is then modified by a series of adjustment relations for trip length, walking distance, and service frequency to arrive at an estimate of patronage for the service alternative under study. The model can be manually applied and used to assess new service, extension of existing service, or improvements in the existing level of service.

A principal component of transit planning and development studies centers on procedures for estimating demand and patronage. These procedures are important because they are the main means of assessing (a) benefits of new service or modifications to existing service; (b) financial feasibility of new or modified service; (c) mobility impacts on population that result from service changes; (d) potential impacts on areas served by new or modified transit services, areas such as central business districts (CBDs) and public service complexes; and (e) potential impacts on other transportation modes or transportation-related social, environmental, and economic factors.

Patronage-estimating procedures are well developed for large urban areas, particularly for long-range, capital-intensive transit improvements. These techniques have proved relatively successful because of a combination of one or more of the following factors characteristic of transit improvements in larger urban areas: (a) major system changes; (b) large travel markets in which transit service can have significant impacts; and (c) dominance in systems or corridor studies of travel that is diverted to transit rather than captive travel or latent transit demand. In addition, large urban areas generally have available a wealth of travel data for both highway and transit travel, data that represent a fairly wide range of conditions and permit the development of reasonably stable and statistically accurate forecasting relations.

Smaller urban areas present somewhat different problems in transit planning, both in the scale of proposed transit development and in the data base from which to derive forecasting relations. Transit improvements in smaller areas are less dramatic than those in larger areas, both in level-of-service changes and in overall impact on the total transportation system. Adding a new local bus route in a corridor or increasing service frequency on a local bus route—each an example of typical small-area transit improvements—is far less dramatic than building a rapid transit line in a corridor. In small urban areas, the smaller-scale changes in transit service, the relatively small travel markets, the lower potential for diversion of travel to transit, and the comparative dominance of the captive-rider and latent-demand market generally cause changes in the level of demand that cannot be satisfactorily addressed by competitive mode-forecasting techniques.

Techniques for direct estimation of transit patronage are more appropriate for smaller urban areas and also for comparatively minor service changes, for the following reasons:

1. Transit service improvements in small urban areas and minor changes in large urban areas generally have a greater impact on the generation of new travel than on diverting travel; new demand is mainly generated travel and only a minor amount is diverted travel.
2. The amount of new patronage, although significant in transit planning (mainly because of low existing patronage), does not have a significant effect on highway traffic volumes in relation to highway planning decisions.
3. Mode-share models are generally unsatisfactory for estimating the relatively small changes in patronage that occur as a result of the level-of-service changes most common in small-area and short-term transit improvements.

There is a need for a technique of estimating demand and patronage that is appropriate to the characteristics and commensurate with the requirements of small areas and comparatively minor transit service improvements. As with more traditional mode-share techniques, this technique must be sensitive to the policy-related factors inherent in transit planning: frequency, coverage, fares, and travel time. In addition, the comparative size of existing travel markets, characteristics of potential trip makers, and latent demand generation potential should be accounted for. Other considerations are the desire to simplify application of the technique to increase its utility as a planning tool by reducing the data requirements, the application effort, and reliance on senior professional staff.

A desirable technique is one that

1. Is responsive to all major policy issues;
2. Is sensitive to trip-maker and level-of-service characteristics;
3. Is intuitively simple;
4. Has minimal data requirements (can be used with census data and planning descriptions of transit service);
5. Can be applied efficiently as a manual technique for route or small-system planning but, if need be, can be computerized to simplify bookkeeping for repetitive or larger-system applications; and
6. Is intuitively correct (e.g., patronage changes are intuitively consistent with the direction of change in a particular service characteristic).

PAST EXPERIENCE

A number of approaches to the estimation of demand and patronage have been developed for small-area transit planning. These have generally been of two types: estimation of areawide system patronage and route-corridor estimation. Techniques for both have ranged from simple (standard productivity rates with specified route mileage and hours) to complex [behavioral system models developed by the New York State Department of Transportation (1-3)]. A number of the more widely used approaches or variations are briefly described here.

One simplified aggregate systemwide approach is based on use of an annual per-capita transit ridership that corresponds to a typical average, overall transit operation for a small urban area. The per-capita trip rate is modified for variations in systemwide frequency and fare from the per-capita-trip-rate reference condition. Data are empirically derived from a number of smaller urban areas. Route planning cannot be done by using this technique.

The approach developed by Hillegass (4) is a simplified technique based on major corridor or route structure for travel to the CBD. The premise of the approach is that in smaller urban areas the predominant type of transit trips is trips to the CBD, which are largely work trips. A generalized relation between mode split and automobile-occupancy and income and automobile ownership, developed from national statistics, is used to estimate systemwide (and corridor) travel to the CBD for a given estimate of CBD person work travel for the urban area. The procedure explicitly addresses the area of transit route coverage and user characteristics but does not contain specific relations to reflect frequency and fare variations.

Procedures developed in Massachusetts—in the Merrimac Valley (5) and Northern Middlesex (6) transit development programs (TDPs)—are both variations and extensions of the approach advanced by Hillegass. The Merrimac Valley approach is a corridor technique that implies a radial CBD-dominant transit system. Basic transit trip rates within a 0.8-km (0.5-mile) coverage band for separate automobile-ownership categories are used to estimate a base route demand. Relations for service frequency and fare changes developed in other studies are used to modify the base route demand. The Northern Middlesex technique is similar but uses a trip rate based on income. Adjustments for frequency and fare variations are again based on relations developed in national studies.

All of the above techniques stress simplification in application and a complexity in balance with the demand-forecasting problem. The four approaches have similar inherent assumptions that are not always explicitly presented. All are based on a radial route structure that focuses on the CBD. Transit travel is predominantly home-based travel to the CBD; there is little crosstown or non-CBD corridor travel. Routes are fairly short in length and generally do not extend beyond the older, denser residential core; few routes extend to newer, low-density residential areas. Consequently, these techniques are intended to be used primarily for direct travel to a single dominant activity center over fairly short [6.4-km (4-mile)] maximum travel distances.

Each of the approaches is intuitively acceptable, and each contributes to the state of the art. Collectively, they form a good basis for further extension of a simplified forecasting technique.

CONCEPTUAL ASPECTS OF MODEL DEVELOPMENT

A number of conceptual hypotheses are presented that establish the structure for model development. These are based partly on previous work, on general findings from analysis of transit data, and from intuitively derived relations based on observation of transit and travel data.

The nature of transit trip making in smaller urban areas and, to a degree, new trip making associated with transportation system management (TSM) type of improvements to bus service in any area argues strongly for a demand-estimating technique that emphasizes generation of trips rather than mode splitting of existing demand. This establishes the approach for model development.

Specification Guidelines

The objective of the process of model development is to produce a set of relations that can be applied at the planning level to determine potential ridership levels for variations in service as well as service to activity centers of different sizes. Conceptually, the variations that should be addressed are route relocation, route extension, frequency changes, route speed changes, changes in fare levels, route coverage, trip-maker characteristics, and activity location and size. It is desirable to structure these relations so that they can quickly and easily be used to evaluate transit service improvements. Graphic and tabular representations, rather than mathematical equations, are desirable. Model development is directed toward this format.

The model is in the form of direct transit trip generation and is structured as a set of separate but integrated relations. The components are

1. A basic transit trip generation rate by socioeconomic category and
2. Relations for modifying trip generation by (a) variation in walking distance to service, (b) distance from the attraction center, (c) change in service frequency, (d) change in fare levels, (e) change in schedule running speed, (f) size of activity center, and (g) size of urban area.

Model development does have some basic constraints. These are primarily dictated by available empirical data. The source of empirical data on ridership is on-board travel surveys for conventional fixed-route, radially oriented surface bus systems. The basic data limitations are that (a) the model is for conventional transit service and should not be extended to paratransit services and (b) the model is for trips to an activity center at the focus of radially oriented transit service. The second limitation reflects the CBD data bias. However, introduction of concepts of trip-rate adjustment, based on both the absolute and the relative size of an activity center, and use of the principle of superposition allow CBD-based data to be extended and used for transit planning in areas that contain multiple activity centers.

Basic Trip Generation Rates

Transit trip generation rates have been shown to be related to such socioeconomic characteristics of trip makers as income and automobile ownership. Both variables have been shown to be highly correlated. Since information on automobile ownership taken from survey data is more reliable than data on income, it is selected as the basic variable. In addition, automobile-ownership distributions are readily available from census data, there are fewer variable stratifications, and impacts of energy crises will be more readily reflected in automobile ownership than in income.

The basic trip rate should be for each category of automobile ownership: households with no automobile, households with one automobile, and households with two or more automobiles. Trip generation is based on trips per household (since this is a short-term forecasting technique, the apparent trend toward lower household size can be ignored). To ensure that trip generation reflects "effective" transit service, only data from covered areas are used in developing trip rates. Each trip-rate value will inherently reflect "averages" of the other parametric values, such as trip length and distance from a route. A basic trip-rate table is shown in Figure 1.

Trip Length

Transit trip generation rates should vary with distance from the CBD. This reflects two phenomena: trip distribution and mode share. Both concepts are described below.

A hypothetical radial travel corridor to the CBD is used to show the effect of trip distribution. The corridor consists of four zones of equal length and width; all zones have identical socioeconomic and trip generation characteristics. The width of the corridor is taken as being equal to the coverage of a transit route—approximately 0.8 km (0.5 mile)—as shown in Figure 2.

The generalized distribution of trip length and frequency for each zone can be estimated by using trip distribution theory and empirically based observation. In Figure 3, zone 1 sends a trips to the CBD, zone 2 sends b trips, zone 3 sends c trips, and zone 4 sends d trips, where $a > b > c > d$. If trips to the CBD from the corridor were plotted by distance from the CBD, a trip-length distribution of CBD-oriented trips would result. For example, Figure 4 shows that travel from a zone to the CBD decreases as the distance from that zone to the CBD increases.

The implications of the relations shown in Figures 3 and 4 for procedures for forecasting the transit trip rate are significant. For example, if the transit trip rate were taken as a constant value, it would imply either increasing mode share or greater latent demand generation or both. This is contrary to empirical evidence.

Generalized mode-share relations along the corridor can be hypothesized from mode-share theory and empirical evidence. A generalized mode-share profile is shown in Figure 5. Transit is not an attractive mode for short trips, primarily because of the relatively high waiting times; walking and the automobile are more attractive, and hence mode split or transit trip generation would be lower for short trips. At the other extreme—long trips—transit begins to lose its attractiveness as line-haul time and cost begin to favor the automobile; the mode share for transit then decreases.

A conceptual relation of variation in transit trip generation rates along a transit service corridor can be derived from the above characteristics. This relation should also show differences for different strata of automobile ownership. Total person-trip-length distribution by distance will vary by category of automobile ownership because of the comparative difficulty in reaching the same spatial opportunities in the same travel time for each category. Two-automobile households have a superior mode available for trip making and can "cover more ground" in the same time as zero-automobile households, which are more dependent on "inferior" modes such as transit, taxi, and shared ride. Generalized profiles of trip-length distribution by category of automobile ownership in a trip production zone are shown in Figure 6. The relation between variation in transit trip generation along a corridor and distance from the CBD is shown in Figure 7.

All of these concepts can be applied to any trip attraction subarea, such as shopping centers and public service complexes.

Trip Frequency

Empirical observation, elasticity studies of transit system characteristics, and research based on behavioral mode-share models have all shown mode split to be sensitive to frequency of service. The relation between transit trip rate and headway, based on disutility mode-split findings, is shown in Figure 8. Separate response surfaces are indicated for each category of automobile ownership.

Fare

Transit trip generation (mode share) has been shown to vary with the fare charged. The relation between transit trip generation rate and fare, derived from existing mode-share and elasticity research, is shown in Figure 9.

Figure 1. Basic trip rate.

0 Auto	1 Auto	2+Auto
x.x	y.y	z.z

Figure 2. Prototypical CBD travel corridor.

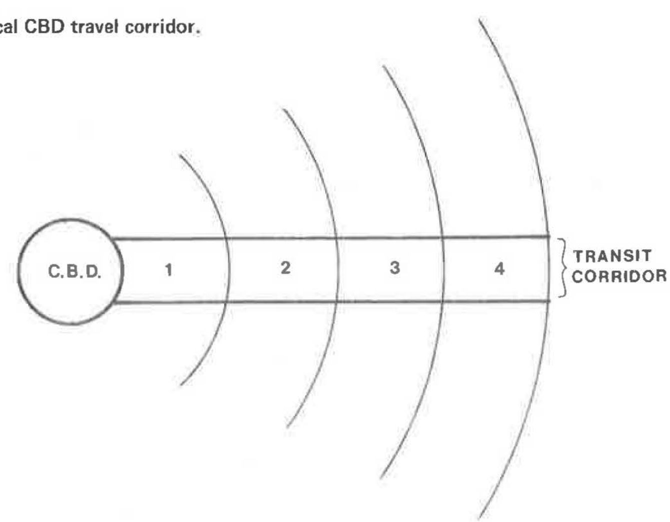


Figure 3. Generalized distribution of trips for an origin zone.

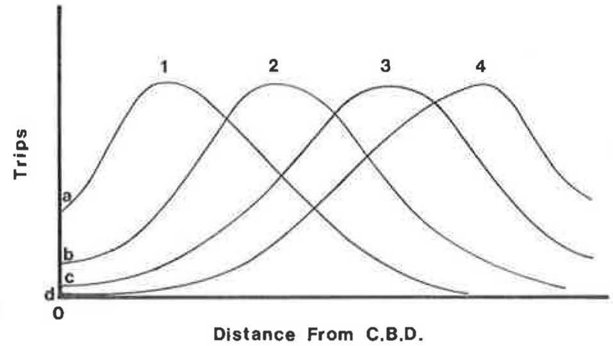


Figure 4. Trips to the CBD by distance of origin from the CBD.

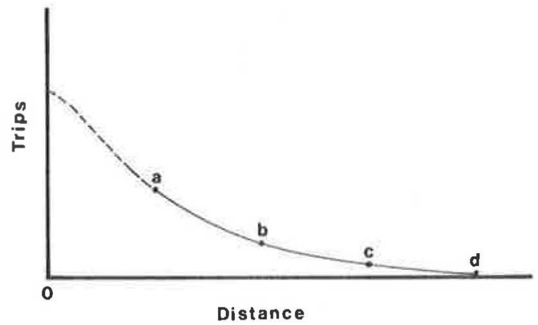
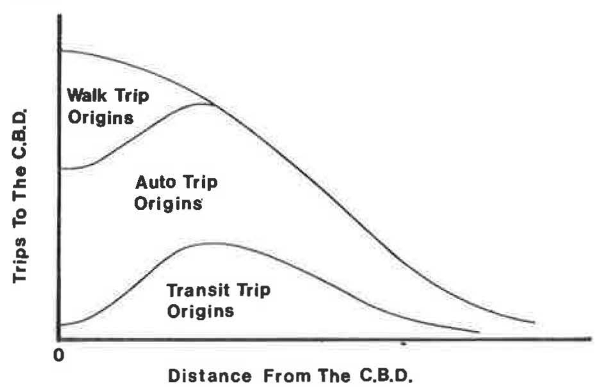


Figure 5. Generalized distribution of trip origins for trips to the CBD.



Walking Distance

Analysis of empirical data has shown that the walking distance between a potential trip origin and a transit stop has an effect on the rate of transit ridership. A generalized form of the relation between walking distance and transit trip rate is shown in Figure 10.

Trip Line-Haul Speed

Mode-split model analysis has shown transit demand to vary with changes in line-haul transit service time, all other factors remaining constant. Generally, as travel time by transit improves, the trip rate increases. A

Figure 6. Generalized distribution of trips from a zone by automobile ownership.

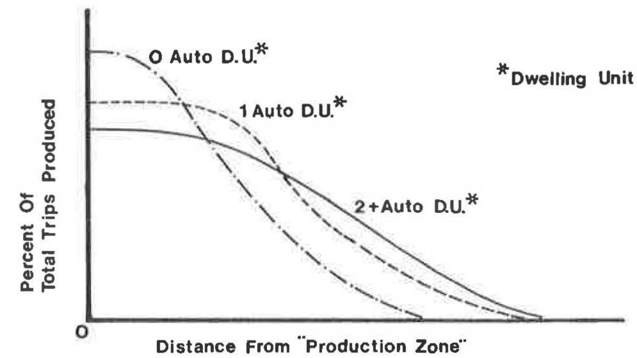


Figure 7. Relative transit trip generation rates for travel to the CBD by distance from the CBD.

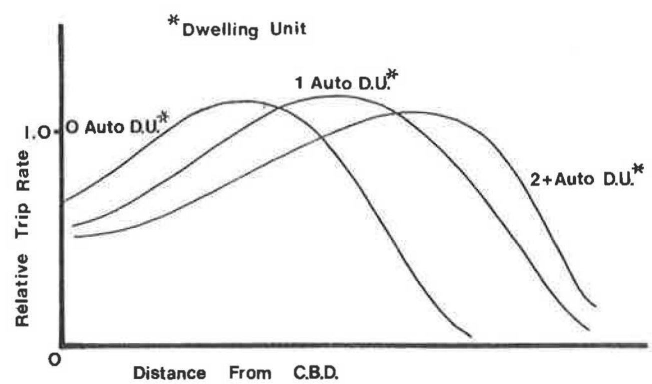
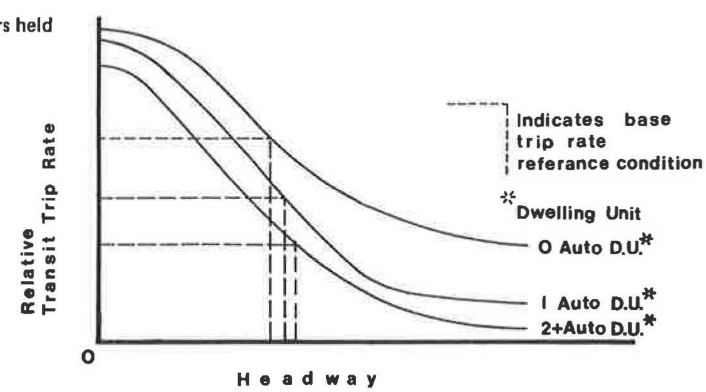


Figure 8. Transit trip rate versus headway (all other factors held constant).

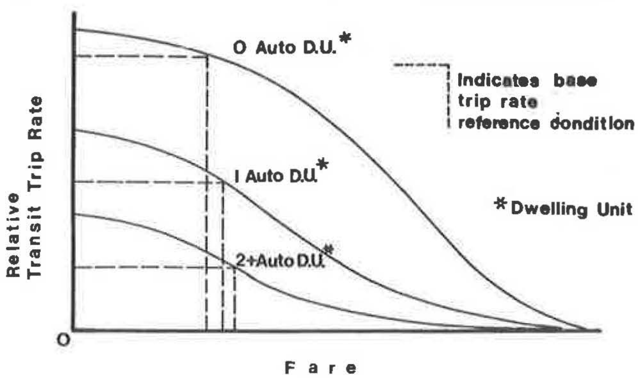


hypothesized relation between speed and trip rate for a constant person-trip length is shown in Figure 11.

Size of Urban Area

Data on transit trip rates from several urban areas of different size but with approximately the same quality of transit service indicate variation in trip rates. An a priori hypothesis is that city size may be a factor in transit trip generation rates. This may be a result of a number of factors that become more pronounced as city size increases, such as increases in traffic congestion and parking cost and decreases in parking space and in walking as a primary mode of travel. Although many of these factors may be directly or indirectly accounted for in other hypotheses, a general conceptual relation between transit trip rate and size of the urban area is shown in Figure 12.

Figure 9. Fare versus transit trip rate (all other factors held constant).



Size of Attraction Subarea

The size of the attraction subarea measured in the number of trip attractions and in the proportion of total urban-area trip attractions should have an impact on transit trip generation rates. These measures reflect trip distribution, the size of the trip market, and traffic congestion as well as the cost and difficulty of parking in the subarea. It is hypothesized that, as both the percentage of area attractions and the absolute number of attractions in a subarea increase, the transit trip rate should increase, transit service parameters remaining constant. Figure 13 shows the relation between transit trip rate and urban-area attraction activity.

This hypothesis has added significance. The specification for the proposed model is for the estimation of transit trips to a single attraction subarea (the CBD). With the exception of the trip-length adjustment, the estimate of transit trip generation is independent of transit travel to any other location. It is therefore possible, and conceptually valid, to develop a number of separate estimates that correspond to transit service to other specific attraction subareas and combine them in an additive manner to yield route and system estimates for areawide travel by transit. An adjustment factor to scale the basic trip rate according to relative subarea activity would permit this.

Principle of Superposition

Superposition occurs when events taking place in the same environment are independent of one another in their effect on the environment. Impacts of each event are additive, having a linear cumulative effect. The model specification is defined to take advantage of superposition to simplify use of the model. The discussion above on estimating transit demand for more than one attraction area is an example of superposition.

Figure 10. Relative transit trip rate versus walking distance from transit route.

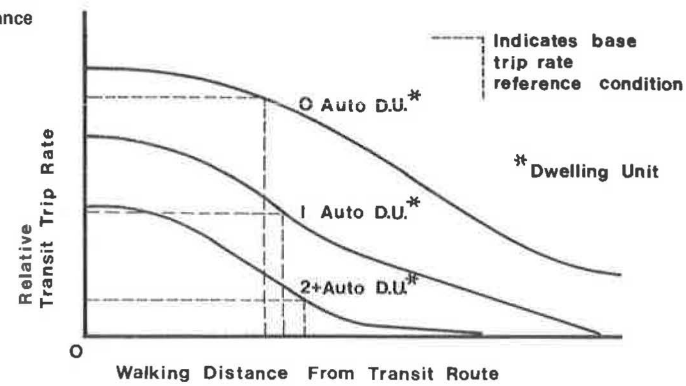


Figure 11. Route line-haul speed for a person trip of constant distance versus relative transit trip rate.

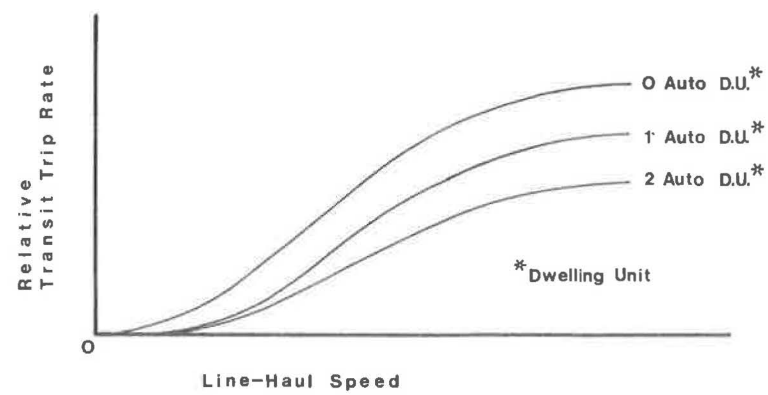


Figure 12. Relative transit trip rate versus size of urban area.

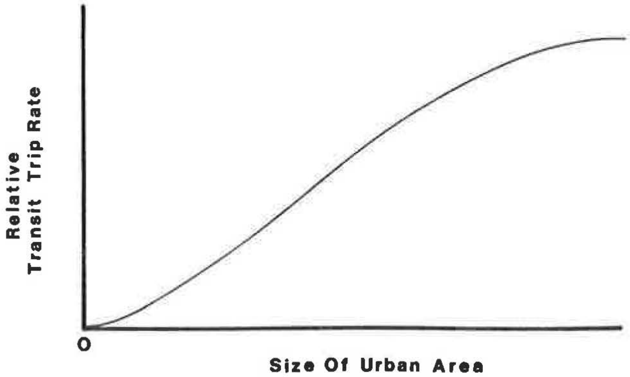


Figure 13. Relative transit trip rate versus urban-area attraction activity.

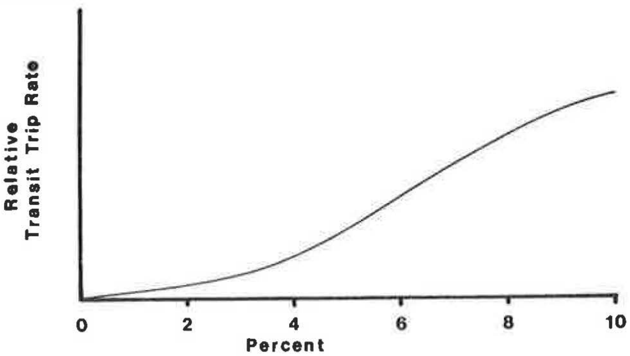
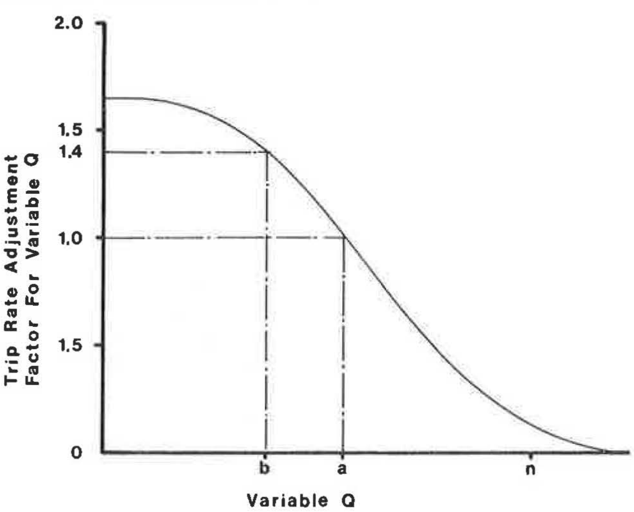


Figure 14. Trip-rate adjustment factor versus variable Q.



Application Concept

The model is intended to be applied on a route or corridor basis. It can be applied on an aggregate system basis if average route characteristics and areawide socioeconomic characteristics and percentage coverage are used.

In using the model to estimate route demand, the basic trip generation rate for each category of automobile ownership is successively modified by an adjustment factor that reflects the route characteristics, the spatial relation of the trip origin zone to the attraction subarea, and the size of the urban area and the attraction subarea. This is expressed as follows:

$$T_k = \sum_{i=1}^{n} \sum_{j=1}^{m} T_{ijk}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} (H_{ij})(R_j)(F_j)(W_{ij})(Q_j)(D_{ij})(SP_{ij})(A_k)(U) \qquad (1)$$

where

T_k = total trips generated on the route to attraction k,
T_{ijk} = trips from zone i by trip-maker category j to attraction k,
H_{ij} = number of households of type j in zone i,
R_j = basic transit trip generation rate for trip-maker category j,
F_j = fare adjustment factor for the route for trip-maker category j,
W_{ij} = walk-distance adjustment factor for trip-maker category j in zone i,
Q_j = frequency adjustment factor for trip-maker category j,
D_{ij} = distance adjustment factor for trip-maker category j in zone i,
SP_{ij} = route-speed adjustment factor for trip-maker category j in zone i,
A_k = adjustment factor for subarea size and concentration, and
U = adjustment factor for urban-area size.

As can be seen from this expression, the application is very similar to Highway Capacity Manual procedures for calculating intersection capacity (7).

The basis of this approach is that each adjustment factor is referenced to the value each variable had for calculation of the basic trip rate. This is accomplished by normalizing each of the relations by dividing trip rates by the average basic trip rate. The value of the variable at a normalized trip rate of 1.0 is the reference condition. A generalized curve for the relation between the variable and the trip-rate adjustment factor is shown in Figure 14. (In the figure, a is the value of variable Q, corresponding to the base trip generation rate; thus, the adjustment factor for a is 1.0. If the proposed service improvement resulted in a value of b for variable Q, the base trip generation rate would be multiplied by an adjustment factor of 1.4.)

Use of the procedure implies measuring the transit system variables as the trip maker sees them. When a zone is served by only one route, there is no measurement problem; characteristics of only that route are used. However, when a zone is served by two or more routes for travel to the attraction subarea, the effective combined service characteristics must be used. This will almost always be limited to the frequency variable. As an example, a zone with two 30-min services is treated as having one 15-min service.

MODEL DEVELOPMENT

Data Base

The data base for model development consisted of an on-board transit O-D survey of approximately 1000 interviews, a description of the transit system, and socioeconomic census data from the Montachusett, Massachusetts, regional planning agency (RPA).

Figure 15. Walking-distance adjustment factor.

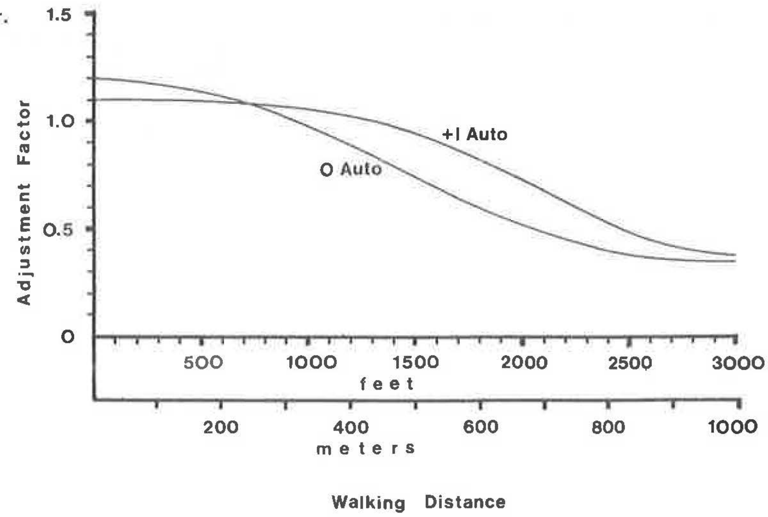


Figure 16. Trip-length adjustment factor.

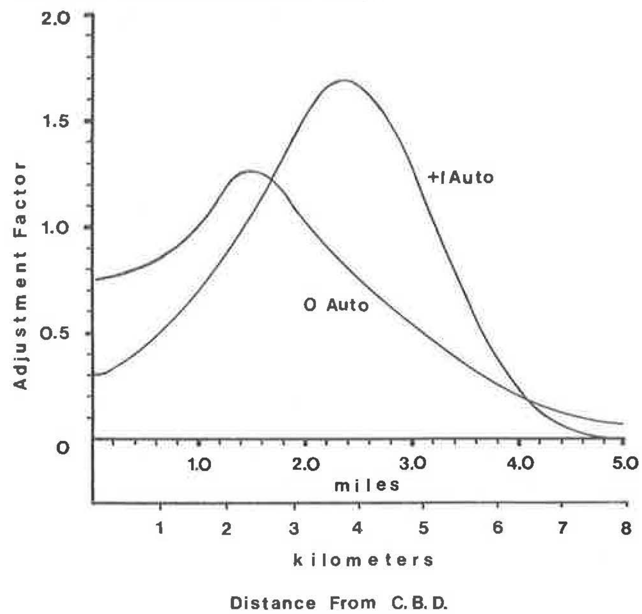
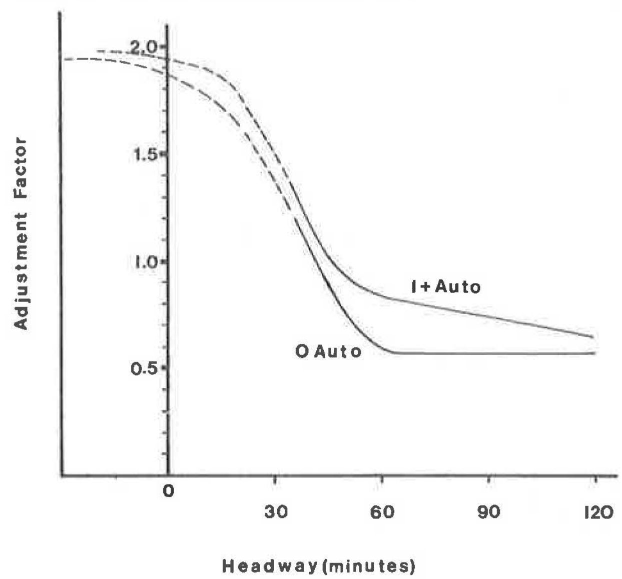


Figure 17. Trip-frequency adjustment factor.



Relations Investigated

Limitations imposed by the data restricted direct analysis to basic trip rate, trip length, walking distance, and service frequency. Fare-change relations were approached by using findings from other studies. Analysis of the effect of route speed was not possible because of a lack of suitable observations. An attempt was made to study subarea relations by using the Fitchburg and Leominster CBDs within the RPA, but this analysis was inconclusive because of problems encountered in structuring the analysis. Because of the single-area data set, the effect of city size could not be investigated.

Data Definition, Preparation, and Analysis

Trip generation was defined on standard gravity model notation, home-based and non-home-based. Only home-based trip productions were used in the model. Non-home-based trips were not included because of difficulty in associating causative factors and because these trips represented a small proportion of total transit trips. Trip attractions were not explicitly addressed; consideration of only the CBD as a trip attraction area and use of only trips to this subarea implied both distribution and balanced trip generation. This was also necessary to conform to the single-attraction focus of the model specification.

Home-based trips were not stratified by purpose, primarily because of the thin data base. Use of a single, combined home-based purpose appears sufficient for estimating transit patronage for CBD travel, but models by purpose should be more useful, particularly for estimating trips to more homogeneous subareas such as shopping centers, medical/health-care complexes, and large industrial parks.

Use of the home-based production definition produces a round-trip estimate that results in the estimate for a route in nondirectional total passengers. Directional loads and load profiles are estimated by splitting total trips equally into boardings and alightings and loading these on the route. Non-home-based trips are accounted for by factoring the home-based trips by the ratio of non-home-based to home-based trips developed from the survey data. This is an approximation but provides a working estimate that is sufficient for planning.

Analysis was done by using aggregate and semi-disaggregate techniques. Trip and socioeconomic data were referenced to traffic zones. This was done for two reasons: to provide a convenient reference for measuring transit system characteristics and to modify trip rates for each automobile-ownership category to reflect non-transit-trip-making households in that category. Within each zone, the data were treated in a semidisaggregate manner. Only households in the zone within "coverage" walking distance were included in the analysis. By using census data and other socioeconomic data prepared by the RPA, an estimate was made of the number of zero-, one-, and two-automobile households in each of the traffic zones.

Transit system characteristics were estimated for each zone by using the transit route map and timetable. Distance to each zone from the CBD was measured along the route from the center of the CBD to a midpoint location in the zone. Frequency of service for the zone was taken as the combined effective frequency of all service between the zone and the CBD. A parallel resistance formula was used for the calculation. Walking distance to the route, as measured, was used to determine the limit of route coverage and was the criterion for the inclusion of data in the analysis. Walking distance as reported in the survey was used in the analysis of the effect of walking distance. Fare was constant.

Basic Trip Rate

Zonal average trip generation rates for each category of automobile ownership were graphically analyzed. Trip rates were taken as the total for the automobile-ownership categories in the zone, and there was no additional cross classification. The analysis indicated a distinct difference in trip generation rates between households with no automobile and those with one or more automobiles. Trip rates for households with two or more automobiles appeared not to be statistically reliable because of the small number of responses in that category. For this reason, households with one automobile and households with two or more automobiles were combined. The resulting basic trip generation rates are given below:

Number of Automobiles per Household	Daily Trip Rate per Household
0	0.21
≥1	0.04

Walking Distance

Trip-rate data within each of the two automobile-ownership categories were cross-tabulated by reported walking distance. Hand-fitted curves were normalized by dividing the trip rates by the basic trip rate for each automobile-ownership category; these curves are shown in Figure 15. The apparent inconsistency in the curves is produced by the normalization procedure. Normalization is within the strata, and the factor indicates the relation to the basic trip rate. Apparent inconsistency results from plotting both against internal relative scales. If the curves were each multiplied by their respective basic trip generation rates and replotted on a trip-generation-rate scale, the inconsistency would disappear.

Trip Distance

Trip-rate data by each category of automobile ownership were cross-tabulated by distance from the CBD. Curves were hand-fitted and then normalized by dividing by the respective basic trip generation rates (see Figure 16). The apparent inconsistency in the curves is attributable to the normalization approach.

Service Frequency

Automobile ownership was cross-tabulated by service-frequency trip rates to develop the normalized service-frequency curves, which are shown in Figure 17.

Validation

The initial basic trip rates and adjustment relations were applied in the Fitchburg-Leominster area to test if the model could estimate base-condition transit travel. Input data were the number of households by category of automobile ownership in each zone and measures of transit service to that zone. Estimation errors were found in the initial test, and modifications were made to the factor adjustment curves. Two additional interactions of testing and revision were required before all model components were judged to be acceptable for planning application. All information given in the text table above and in Figures 15-17 are for the final relations. The prediction accuracy of the model for zonal trip productions is shown in Figure 18.

Ancillary Models

Use of a semidisaggregate or fully disaggregate model requires specific estimates of the independent variables by the same discrete classification as that used in model development. For the Fitchburg-Leominster model, the only ancillary model required was one to estimate automobile-ownership strata from estimates of average zonal automobile ownership. This was developed from census data. An example of this relation is shown in Figure 19.

Applications

The model was applied in the Montachusett regional TDP to estimate patronage for various service improvements. In application, the model produced rational demand estimates for the proposed types of improvements. For one alternative—restoration of service to previous levels—the estimated demand was similar to actual patronage levels experienced when the service was still in effect.

The model was also used to estimate patronage for the Midstate Connecticut TDP. There was no existing transit service, and therefore patronage estimation had to be developed from "borrowed" relations. Application of this model produced demand estimates that were typical of patronage levels found in cities of a similar size and a similar level of transit service.

APPLICATION PROCEDURE

The steps followed in making an estimate for a transit line that serves an attraction subarea are presented below. Where a second or third line also provides service to zones served by the subject line, all services must be taken into account when the trip estimate for the zone is calculated; the zone trips are then split equally between each of the lines that serve the zone. If an estimate is to be made for more than one attrac-

Figure 18. Observed versus estimated zonal trips.

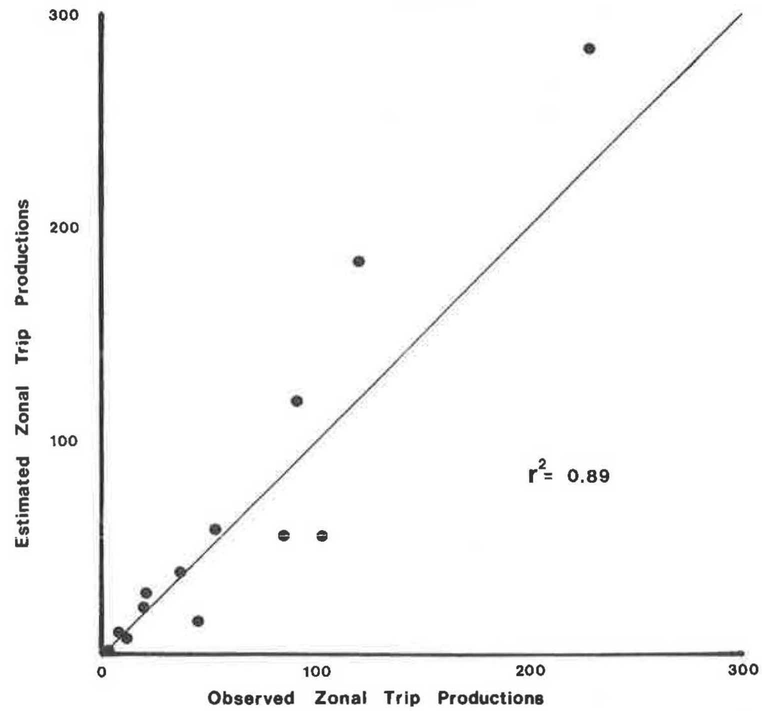
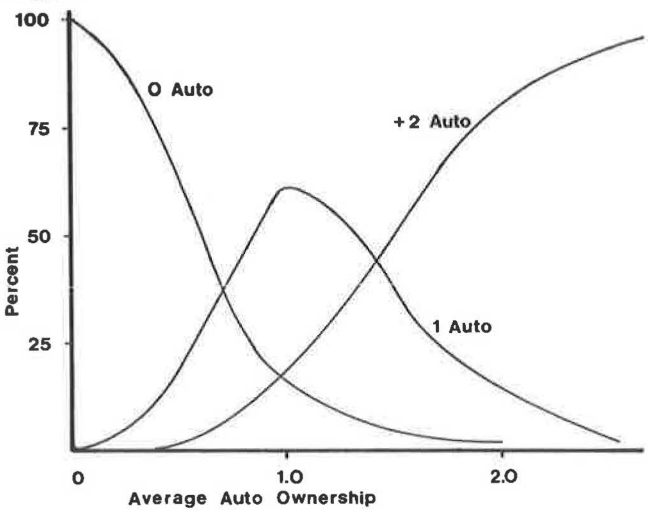


Figure 19. Average automobile ownership by automobile-ownership categories.



tion subarea served by the line, separate estimates are made for each and added together. The procedural steps are these:

1. Lay out the subject transit line on a zone map.
2. Define the coverage band of the line, approximately 0.8-1.2 km (0.5-0.75 mile) on each side of the line.
3. Note each of the zones included within the coverage band.
4. Estimate the number of households with no automobile and one or more automobiles in each zone in the covered area; this can be done by simply using a dwelling-unit density factor for the zone and a representative average automobile ownership derived from a census-tract-level estimate.
5. Estimate the average walking distance to the line; for small zones a single value is probably sufficient, but for larger zones it may be necessary to subdivide the zone into walking-distance bands. The adjustment factor for each band by automobile ownership for each zone is determined from the adjustment curve.
6. Measure the distance along the route from the attraction subarea to approximately a midpoint location of the route in the zone; select the appropriate adjustment factor for distance for each category of automobile ownership in the zone.
7. Estimate the service frequency from the zone to the attraction subarea, not the presence of multiple-line service. Select the appropriate automobile-ownership-category adjustment factors.
8. Calculate the transit trip estimate for each zone. This is done by multiplying the basic trip rate by the number of dwelling units and each of the factors. The line estimate is the sum of all individual "served" zone trip estimates, with allowance for multiple served zones.
9. Estimate total line demand by adding all activity center estimates and then factoring for non-home-based trips.
10. Estimate, if desired, route loading profiles, by converting the trip production estimate for each zone into two trips, one from the zone to the attraction subarea and one from the attraction subarea to the zone. Trips are manually assigned to produce the load profile.

EXTENSIONS AND REFINEMENTS

The results of the development, validation, and application of the model indicate that the conceptual framework and specification for the demand-estimation approach appear to be reasonable. Based on this, it is proposed that the findings be further tested by following this same approach in one or more other small urban areas. If the same results are found, the next steps should be to pursue other concepts that could not be addressed in this study: fare adjustment, urban-area size, route speed, and relative size of the attraction subarea. In addition, stratification of separate work, shopping, and other trip purposes might warrant further study.

ACKNOWLEDGMENT

The model described in this paper was developed under our direction as part of the Montachusett, Massachusetts, TDP study conducted by Alan M. Voorhees and Associates.

REFERENCES

1. D. T. Hartgen and C. A. Keck. Forecasting Dial-a-Bus Ridership in Small Urban Areas. Planning and Research Bureau, New York State Department of Transportation, Preliminary Res. Rept. 60, April 1974.
2. D. T. Hartgen. Modal Split in Small Urban Areas. Planning and Research Bureau, New York State Department of Transportation, Preliminary Res. Rept. PRR-15, July 1969.
3. D. T. Hartgen, G. H. Tanner, and R. A. Martino. An Empirical Approach to Estimating Mode Split in Small Urban Areas. Planning and Research Bureau, New York State Department of Transportation, April 1971.
4. T. Hillegass. Transit Travel Estimation for Smaller Urbanized Areas. Planning Methodology and Technical Support Division, Urban Mass Transportation Administration, Sept. 1975.
5. Ridership Estimating for the Merrimac Valley Area Transit Study. Merrimac Valley Planning Commission, Lawrence-Haverhill, MA, Staff Paper 7, March 23, 1976.
6. Brockton Area Transit Study: Estimating Ridership for Alternatives. Old Colony Planning Council, Lowell, MA, Staff Rept. 4, May 1975.
7. Highway Capacity Manual. HRB, Special Rept. 87, 1965.

Simulation of Travel Patterns for Small Urban Areas

Jerry G. Pigman and Robert C. Deen, Bureau of Highways, Kentucky Department of Transportation, Lexington

A study conducted to simulate travel patterns in small urban areas is reported. The purpose of the study was to develop models that would simulate internal-external trips and external-external (through) trips. Regression analysis and cross classification of data were tested in an attempt to predict the number of internal-external trips and the percentage of through trips. Regression analysis was used in the development of a through-trip distribution model. Grouping data for analysis created some problems; however, trial-and-error evaluation enabled the selection of variables that produced reasonable results. The variables found to be most significant in the development of internal-external-trip models are population and employment. For through-trip models, the variables used are population, functional classification, average annual daily traffic at the external station, and percentage of trucks. In developing through-trip distribution models, the variables of significance are average annual daily traffic at the destination station, percentage of trucks at the destination station, percentage of through trips at the destination station, and ratio of destination average annual daily traffic to total average annual daily traffic at all stations (value squared). Overall, for ease of application and accuracy, the models developed appear to be adequate for planning purposes.

Agencies responsible for determining when and where to construct new urban highways and streets, or to improve existing ones, must consider many factors in the decision-making process. One such factor is the purpose and volume of the traffic that can be expected to use the facilities in the future. Estimates of future traffic patterns are made by various traffic simulation models, usually some mathematical expression with parameters and constants to simulate traffic flow. Alternative transportation systems can be evaluated in terms of costs and benefits by entering socioeconomic descriptors into a simulation model to determine traffic patterns and volumes.

Travel patterns within an urban area are divided into three categories:

1. External-external or through trips—trips that originate and terminate outside the area,
2. Internal-external trips—trips that originate inside the area and terminate outside the study area or vice versa, and
3. Internal-internal trips—trips that originate and terminate within the area.

Historically, travel data for these three types of trips have been obtained from origin-destination (O-D) surveys. The external O-D survey, in which drivers of vehicles are interviewed at the study area boundary, provides data for internal-external and external-external trips. Internal-internal trip data are generally obtained by using home-interview, truck, and taxi surveys. The collecting, coding, editing, processing, and summarizing of these data often represent a major portion of the time and cost of conducting a transportation study. However, a review of completed studies has indicated that there are many similarities in models developed for trip generation and trip distribution that involve internal-internal trips, and this makes it possible to synthesize internal-internal trips by modeling. Many similarities are also apparent in internal-external and external-external trips. Synthesis of the trips involves applying values from O-D studies to other urban areas that have similar population and socioeconomic characteristics.

The models discussed in this paper were developed for simulating internal-external and external-external trips by emphasizing previously tested procedures and by selecting variables that characterize small urban areas in Kentucky.

REVIEW OF THE LITERATURE

The differences between large and small urban areas are apparently significant enough to make it necessary to separate them in traffic modeling. Most planners categorize areas that have less than 50 000 population as small urban areas.

Initial work in North Carolina was directed toward simulating internal travel by using trip generation data either from a small sample of home interviews or from data obtained from another similar urban area. By 1970, a procedure for synthesizing internal travel had been perfected to the extent that its use had become standard operating procedure (1). In 1970 and 1971, Modlin (2), working with the North Carolina Department of Transportation, was successful in synthesizing internal, external, and through travel for small urban areas.

The estimating procedure for through trips has consisted of three models (3). The first dealt with estimating the percentage of through trips from each external station given the functional classification of the facility external to the cordon, the current average annual daily traffic (AADT), the percentage of the facility external to the cordon, the percentage of panel and pickup trucks, and the population of the urban area. The second was a composite model composed of distribution models for each functional classification, which produced a triangular through-trip table. A third model estimated the percentage of total external trips by vehicles garaged inside the cordon as a function of employment available in the urban area.

In another study (4), previously developed corridor growth-factor models for developing future estimates of internal traffic in small urban areas were tested and modified. Regression equations were developed to provide data that are usually obtained from external cordon surveys. Alternative procedures for providing external survey information, based on historical data, were also developed. The completed procedure provided traffic volumes within the accuracy necessary for planning major thoroughfares in small urban areas.

Most studies of trip generation undertaken in the 1960s relied heavily on regression analyses. But a recent study sponsored by the Federal Highway Administration indicates that a combination of cross-classification and rate analysis was a more efficient and straightforward procedure for forecasting trip generation (5). Some of the advantages of using combined cross-classification and rate analysis are that the procedure is easy to understand, uses the data efficiently, and is easy to update.

DEVELOPMENT OF MODELS

Transportation studies of 20 cities scattered throughout Kentucky that have populations ranging from 6000 to 50 000 were the primary source of data for the analyses. As is the case with most prediction models, the procedure followed was a trial-and-error process of selecting independent variables that were easy to predict, met the test of reasonableness, and produced statistically sound results. Model formulation was confined to regression analyses and cross-classification techniques.

Internal-External Model

Inspection of internal-external equations developed in urban-area transportation studies reveals the types and the combinations of independent variables that were used to predict internal-external trips. The dependent variable (internal-external trips) and independent variables (various planning and socioeconomic factors) were the best combination of variables to represent base-year conditions and to predict future trip generation. Internal-external trips were obtained from O-D surveys. Population and employment data were available from censuses, and projections of these variables were considered good predictors of conditions at some point in the future. The study areas were grouped according to population.

Regression Analysis

Data on dwelling units, population, various types of employment, and internal-external trip attractions by zone were collected, tabulated, keypunched, and coded for computer analyses. Linear regression was the first type of analysis performed to derive a prediction model. Several combinations of independent variables were tested by using available data from the 20 Kentucky cities. Each internal zone was considered to be a separate set of data so that a total of 816 sets of data were available. The data sets were reduced from 816 to 762 because some exhibited unusually large, or small, numbers of internal-external trips.

An attempt was made to perform a regression analysis by using the complete data. The result was an inaccurate and unresponsive prediction equation. A second regression analysis was performed by using the zones within each study area as a data set. These equations characterized individual areas well, but the equations were not applicable to predicting trips in other areas. It became apparent that the study areas should be combined into population groups. Regression analyses were performed in which five population groups were used. The resultant equations are given in Table 1. The

Table 1. Internal-external trip prediction models: regression analysis.

Number of Study Areas	Population Group	Equation
6	5 000- 9 999	$Y = 10.25 + 0.53P + 5.41C + 0.81E + 0.57I$
7	10 000-14 999	$Y = 123.45 + 0.15P + 2.73C + 3.20E + 0.80I$
2	15 000-19 999	$Y = -28.41 + 0.38P + 2.72C + 3.28E + 0.69K$
3	20 000-29 999	$Y = 1.78 + 0.30P + 1.87C + 1.64E + 0.53I$
2	30 000-49 999	$Y = 60.76 + 0.05P + 1.26C + 0.30E + 0.051I$

Table 2. Cross-classification prediction of internal-external trips per internal zone.

	0-150 Population		151-500 Population		>500 Population	
Total Employment	Trips	Data Entries per Cell	Trips	Data Entries per Cell	Trips	Data Entries per Cell
0-5	59	87	87	51	317	8
6-50	154	46	185	73	340	63
51-100	179	22	222	39	485	52
101-300	436	30	464	70	610	87
>300	945	42	1150	43	1309	49

Table 3. External-external trip models: cross classification.

Functional Classification	AADT	Trucks in AADT (%)	Through Trips (%)	Entries per Cell
Primary arterial		0-5	12	2
	0-2500	6-10	31	3
		>10	41	6
		0-5	39	2
	2501-5000	6-10	31	7
		>10	49	15
		0-5	24	2
	>5000	6-10	49	10
		>10	64	15
Minor arterial		0-5	16	17
	0-2500	6-10	20	30
		>10	15	8
		0-5	28	9
	2501-5000	6-10	20	8
		>10	36	18
		0-5	10	2
	>5000	6-10	32	4
		>10	40	5
Collector	All	All	25	11
Local	All	All	19	3

following notation is used in the equations:

Y = internal-external trips by zone,
P = population of internal zone,
C = commercial employment by zone,
E = public employment by zone, and
I = industrial employment by zone.

Cross-Classification Analysis

The second type of analysis used to obtain internal-external prediction models was cross classification of data. The independent variables used for this analysis were zone population, total employment by zone, and dwellings by zone. The first cross-classification matrices were developed by using large numbers of categories for each variable. It was found that the number of entries per cell was not sufficient to give significance to this high degree of stratification because only 816 zones constituted the data base. From regression analyses, it was found that dwellings and population exhibited characteristics of collinearity; therefore, one or the other had to be dropped from the regression equations. Since both variables relied on the same characteristics of the urban area for prediction purposes, dwellings were omitted from the cross-classification analysis.

The resulting model, in its final form, is given in Table 2. Total employment by zone and population by zone are stratified into five and three groups, respectively. Because of the unusual nature of the attractors (businesses and institutions) previously mentioned, only 762 of the 816 internal zones were used for the final cross-classification analysis. The number of entries per cell in the matrix is also given in Table 2. A report on trip generation analysis by the Federal Highway Administration (5) suggests that at least 25 observations be accumulated for each cell. Only 2 of the 15 cells had fewer than 25 observations.

External-External Model

Percentage Through Trips

Regression Analysis

A North Carolina study (3) was used as a guide in testing a model with several independent variables to evaluate the percentage of through trips in the AADT at external stations. The independent variables in the regression analysis were AADT at the external station, percentage of trucks, population, functional classification of the highway at the external station, and employment. The same areas used in developing models to predict the percentage of through trips were used in developing internal-external trip models. There were 20 urban areas and a total of 177 external stations.

Of the 177 external stations, four functional classifications were represented. There were 61 external stations on primary arterials, 102 on minor arterials, 11 on collectors, and 3 on local routes. In the North Carolina study (3), functional classification was used as a dummy variable. The method of dummy variables involves coding the data in such a manner that only selected classifications would be entered into the regression equation; others would be omitted. Functional classification, however, yielded no improvement in the statistical values for the equation. Functional classifications were also considered in an equation for each class, but this also proved unsuccessful. Employment data did not significantly improve the predictive ability of the equation. Generally, it is best that prediction equations have relatively small constants; however, equations that were forced to have smaller constants were not acceptable because predictions were less accurate. After several attempts to segregate the data, the simplest equation that represented all functional classifications and gave the best predicting ability was developed (see Table 3):

$$Y = 0.003A + 1.49T - 0.0007P + 17.43 \qquad (1)$$

where

Y = percentage of through trips of AADT at external station,
A = AADT at external station,
T = percentage of trucks of AADT at external station, and
P = population of urban area.

Cross-Classification Analysis

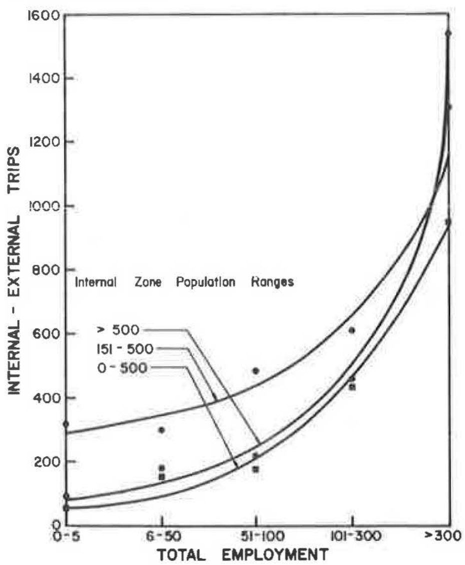
Recent work with cross-classification models has increased the confidence in this type of model for prediction purposes. Here, the first attempts to predict percentages of through trips by using cross classification were generally unsuccessful because too many variables and too much stratification were used. The population of the study area, the functional classification of the route at the external station, the AADT of the route at the external station, and the percentage of trucks in the AADT were the variables first considered. Area population was dropped first because too many blanks appeared in the cross-classification matrix. Functional classification, which was not a significant variable when it was entered into the regression equation, was found to be a practical means of segregating data for cross-classification analysis. Cross-classification models were developed for primary arterial and minor arterial functional classifications, but insufficient data were available to develop models for collector and local routes. The average percentage of through trips for the 11 collector routes and 3 local routes was considered to be representative of the 20 urban areas analyzed in this study.

After several attempts, the final cross-classification model used only three groups of AADT data and three groups of truck percentages for each AADT group. Therefore, there were nine cells in each of the models that represented primary arterials and minor arterials. These models and the average percentages of through

Table 4. External-external trip distribution models.

Functional Classification	Equation
Primary arterial	Y = 0.0001A + 0.11T + 0.22TT + 385.83R - 2.58
Minor arterial	Y = 0.0008A - 0.08T - 0.03TT + 228.14R + 6.20
Collector	Y = -0.000 01A + 0.11T + 0.05TT + 295.06R + 3.10
Local	Y = -0.01A - 0.03T + 0.83TT + 2704.73R + 1.95

Figure 1. Relation between internal-external trip attractions and total employment for various population ranges.



trips that represent collector and local routes are given in Table 3.

Distribution of External-External (Through) Trip Ends

The distribution of external-external (through) trip ends was accomplished by developing regression equations for each of the four functional classifications so that trip ends were distributed from each functional classification to all other functional classifications. External-external trip data were available for only 17 of the 20 urban areas used in the development of the other models in the study. A total of 1332 combinations of trip interchange data were available for use in the analyses.

External-external trip data had to be balanced and then doubled before being input into the distribution models. This was necessary to make the distribution of trips from one external station to all other stations equal to 100 percent. For example, if the balanced number of trips from external station A to external station B is 10 and the number from B to A is 10, then the total number of trips between the two external stations is 20. Handling the trip tables in this manner, the volumes at the external stations represent two-way traffic.

Of the 14 independent variables used in an attempt to predict the distribution of through trips, only 4 were considered significant enough to be included in the final model. To adequately represent two-way trips, it was felt that some function of both origin station and destination station should be included in the model. Results from the regression analysis, however, indicated that the variables that represented the origin station were relatively insignificant, and thus they were omitted from the equation. One variable—the ratio of the destination station AADT to the combined AADT at all external stations—did represent the origin station in an indirect way. The other three independent variables were the AADT, the percentage of trucks, and the percentage of through trips at the destination station. The models in their final form are given in Table 4. The following notation is used in the equations:

Y = percentage of trip ends from origin station distributed to each of the other functional classifications,
A = AADT at the destination station,
T = percentage of trucks in AADT at destination station,
TT = percentage of through trips in AADT at destination station, and
R = square of ratio of destination AADT to total AADT.

RESULTS

Internal-External Trip Models

Regression equations for internal-external trips are given in Table 1. In the equations, internal-external trip attractions are a function of the population of the internal zone, commercial employment, public employment, and industrial employment by zone. Table 2 summarizes the internal-external cross-classification model. In this model, internal-external trip attractions are a function of employment by zone and population by zone. Figure 1 was prepared as a graphical representation of internal-external trip attractions as a function of employment and population by internal zone. For all three population ranges, the number of internal-external trip attractions increases with increasing total employment.

Several statistical values were used to evaluate the accuracy and reliability of the internal-external trip models. For the regression analyses, the statistical values were the squared correlation coefficient, the standard error of estimate, the mean of the dependent variable, and the coefficient of variation. These values for each study and each group of studies are given in Table 5. As should be expected, the statistical results for the individual study areas were better than the results for the combination of studies.

Table 6 gives data on the predictive abilities of internal-external regression models and internal-external cross-classification models for each of the study areas based on the group equations. Included in the table are the number of zones used, actual and predicted trips, and root-mean-square errors for each of the 20 study areas. Root-mean-square errors were used as a means of comparing the predicted values calculated from the regression equations and the actual data obtained from O-D surveys. Two-thirds of the time, the predicted values will deviate from the observed values by an amount no greater than the root-mean-square error.

It is obvious that considerably better predictions were achieved by using the model developed from regression analysis than by using the model developed by the cross-classification analysis. As the data given in Table 6 show, the root-mean-square errors were significantly less for the regression model in all but one (Berea) of the 20 studies in which combined equations were used to generate predictions. Results also indicated that, when the study areas were grouped by population, greater accuracy was achieved by using the regression model. The large root-mean-square errors associated with some of the predictions can be explained in some cases by the unusually large or unique producers and attractors of trips. For example, the Murray area was examined from the standpoint of eliminating unique zones to see how the error of prediction was affected (6). Three zones that had employment three times greater than the average were discarded. The change in the root-mean-square error was from 346 to 249 for the regression model and from 693 to 238 for the cross-classification model. This indicated that the decision to discard some of the zones was very critical to the outcome of the prediction model. If some zones were discarded in the development of the general prediction model, it would be necessary to estimate the internal-external trip attractions by some other means. The most valid estimates are based on data from past studies that involve similar trip producers and attractors.

External-External Trip Models for Percentage of Through Trips

As Table 3 indicates, the regression equation developed

Table 5. Statistical comparison for each study area: internal-external regression equations.

Study Area	Study Year Population	Number of Internal-External Zones	R	Standard Error	Mean of Dependent Variable	Coefficient of Variation
Franklin	7 898	28	0.91	195	370	53
Cynthiana	6 700	20	0.98	138	563	25
Hazard	6 145	15	0.97	243	906	27
Mount Sterling	7 695	19	0.90	293	771	38
Nicholasville	7 464	24	0.95	234	646	36
Berea	9 210	24	0.81	120	331	36
Combined group		130	0.81	353	564	63
Murray	14 713	20	0.95	240	970	25
Glasgow	12 979	32	0.96	190	473	40
Somerset	14 031	20	0.87	383	1188	32
Elizabethtown	12 300	45	0.94	195	488	40
Danville	12 755	30	0.86	472	706	67
Corbin	11 430	31	0.95	135	426	32
Mayfield	13 436	25	0.90	289	1016	28
Combined group		203	0.79	404	690	59
Madisonville	18 224	48	0.96	147	411	36
Winchester	16 205	30	0.95	179	627	29
Combined group		78	0.94	171	494	35
Henderson	24 965	77	0.70	153	289	53
Hopkinsville	26 647	74	0.84	93	224	42
Richmond	23 477	31	0.87	356	793	45
Combined group		182	0.78	229	348	66
Paducah	50 000	95	0.58	133	212	63
Bowling Green	36 553	74	0.79	153	309	50
Combined group		169	0.71	143	255	56

Table 6. Internal-external trip predictions: comparison of regression analysis and cross classification.

Study Area	Internal-External Zones Used in Model	Actual Internal-External Average Trips per Zone	Cross-Classification Prediction (average trips per zone)	Cross-Classification Root-Mean-Square Error	Regression Prediction (average trips per zone)	Regression Root-Mean-Square Error
Franklin	28	370	311	459	418	203
Cynthiana	20	563	507	678	588	228
Hazard	15	849	533	809	990	280
Mount Sterling	19	771	449	684	641	298
Nicholasville	24	645	271	777	397	555
Berea	24	331	369	274	532	316
Combined group	130	564	488	621	563	339
Murray	20	970	652	693	910	347
Glasgow	32	472	481	640	536	330
Somerset	20	1187	704	882	900	459
Elizabethtown	45	488	395	554	534	254
Danville	30	706	543	959	950	686
Corbin	31	406	371	292	414	188
Mayfield	25	1016	677	564	840	367
Combined group	203	687	520	670	689	386
Madisonville	48	411	476	626	445	156
Winchester	30	627	533	551	580	186
Combined group	78	494	498	598	498	168
Henderson	77	289	418	281	298	177
Hopkinsville	74	224	421	437	298	147
Richmond	31	793	589	673	597	413
Combined group	182	348	458	439	349	226
Paducah	95	213	450	329	212	136
Bowling Green	74	309	585	435	285	171
Combined group	169	255	509	380	244	153

to predict the percentage of external-external trips was a function of AADT at the external station, the percentage of trucks, and population. The statistical accuracy of this equation was reasonable: The standard error was 15.53, the multiple correlation coefficient (R^2) was 0.53, and the coefficient of variation was 49.

Table 3 gives the final cross-classification model used to predict the percentage of external-external trips at an external station. This model was also a function of AADT at the external station and the percentage of trucks in the AADT at the external station, but the matrix did not include population. Functional classification was another means of segregating the data for the cross-classification analysis.

A comparison of the predictive abilities of external-external trip models is given in Table 7. Included in the table are the number of external stations used, the actual and predicted trips, and the root-mean-square errors for each of the 20 urban areas. The accuracy of the two models was approximately equal, but the number of entries per cell in the cross-classification matrix was so small that the reliability of the results must be questioned.

External-External Trip Distribution Models

As a result of exhaustive regression analyses, equations for each of the four functional classifications were developed (Table 4). Each of the equations was a function of AADT at the destination station, the percentage of trucks and the percentage of through trips at the destination station, and the ratio of the AADT at the des-

Table 7. External-external trip predictions: comparison of regression analysis and cross classification.

Study Area	Number of Stations	Actual Percentage of Through Trips (average per station)	Cross-Classification Prediction (average percentage of trips per station)	Cross-Classification Root-Mean-Square Error	Regression Prediction (average percentage of trips per station)	Regression Root-Mean-Square Error
Franklin	6	25.3	23.5	8.5	33.3	11.6
Cynthiana	6	30.5	31.2	15.1	34.0	11.9
Hazard	4	17.7	37.0	25.8	41.5	24.8
Mount Sterling	7	42.3	29.4	18.0	39.6	13.9
Nicholasville	7	41.8	34.4	12.7	36.0	9.9
Berea	8	20.0	22.8	6.1	29.5	14.6
Murray	9	18.5	30.0	12.5	33.1	16.1
Glasgow	8	33.5	36.1	15.1	35.4	14.0
Somerset	9	43.1	27.2	19.2	20.3	26.4
Elizabethtown	12	49.3	35.6	24.0	37.6	26.3
Danville	8	28.1	34.9	11.5	31.6	7.7
Corbin	7	44.1	42.1	15.3	31.6	28.0
Mayfield	12	31.7	30.2	17.4	30.3	15.7
Madisonville	8	30.2	29.7	10.3	33.4	8.8
Winchester	11	34.0	26.9	17.5	22.4	24.4
Henderson	10	47.2	34.9	19.4	27.4	25.8
Hopkinsville	12	25.9	34.3	18.0	26.9	15.1
Richmond	7	28.0	30.8	13.4	23.1	8.7
Paducah	15	18.1	29.2	16.5	22.8	9.3
Bowling Green	11	23.3	31.0	18.2	19.0	14.4
All areas	177	31.7	31.6	16.7	29.4	17.8

Table 8. Statistical results for external-external trip distribution models.

Functional Classification	Total Observations (functional class at origin)	Mean of Dependent Variable	R^2	Standard Error	Coefficient of Variation
Primary arterial	478	11.60	0.54	12.85	111
Minor arterial	733	11.61	0.43	11.13	97
Collector	79	11.74	0.35	12.64	108
Local	42	7.03	0.63	7.93	113

Note: Regression equations given in Table 4.

Table 9. Independent input variables.

Model	Regression Equation	Cross Classification
Prediction of internal-external trips	Population of internal zone Commercial employment by zone Public employment by zone Industrial employment by zone	Population of internal zone Total employment by zone
Prediction of external-external trips	AADT at external station Percentage of trucks in AADT at external station Population of urban area	Functional classification at external station AADT at external station
Distribution of external-external trips	AADT at destination station Percentage of trucks in AADT at destination station Percentage of through trips in AADT at destination station Square of ratio of destination station AADT to combined AADT at all external stations	

tination station to the combined AADT at all external stations.

Statistical results that show the accuracy of the models are given in Table 8. Although some statistical measures appear to produce inaccurate predictions, it is generally assumed that reasonably high standard errors exist with these prediction models. Results from these four distribution models compare favorably with results obtained by others (2, 3). Overall, the models appear to be adequately reliable for planning purposes, especially in relation to ease of application and accuracy.

SUMMARY AND CONCLUSIONS

In this research, three prediction models were developed: a model to predict the number of internal-external trips, a model to predict the percentage of external-external trips, and a model to distribute external-external trips. Both regression analysis and cross-classification techniques were tested in the development of the first two models, but only regression analysis was used to predict the distribution of through trips. Segregation of data into groups suitable for analysis did create some problems, but a method of trial-and-error evaluation enabled selection of the best combination of variables. The independent variables required as input into the two internal-external models, the two external-external (through) models, and the through-trip distribution models are summarized in Table 9. These independent variables were selected from data that were readily available, easy to forecast, and easy to monitor.

Population was the most significant variable that affected the outcome of the internal-external trip regression model. As previously noted, there were five population groups. These were found to be the most distinctive means of separating the study areas for analysis. Many of the small urban areas in Kentucky were found to have travel patterns very similar to those of other towns of comparable population. Although it is not verified here, other studies have shown that geographical distribution has considerable influence on travel patterns, as does the proximity of the town to Interstate, parkway, or other major routes. The socioeconomic characteristics of small urban areas also play a significant role in determining travel patterns.

For predictions of internal-external trips, the regression equations given in Table 1 should be used. These equations are categorized into five groups according to population of the urban area, and predictions of internal-external trips by zone are functions of zonal population and employment. The cross-classification prediction presented in Table 2 may have useful application if considerable care is taken to identify unique producers and attractors of trips and if special procedures for handling these trips are developed.

For predictions of the percentage of external-external (through) trips, the regression equation presented in Table 3, which is representative of all cases, should be used. The model for cross classification is also presented in Table 3, but its utility is questionable because of the small number of entries in each cell in the matrix.

It was necessary to develop an external-external trip distribution model to implement results from development of a percentage-through-trip model. Results from the percentage-through-trip model can be input directly into one of the four distribution models presented in Table 4. This will enable the user to determine the percentage of through trips at a particular external station and then to distribute these trips to the other external stations within the study area. The final results will be an external-external triangular trip table.

Overall, the models developed in this study appear to be appropriate for planning purposes, especially in their ease of application and accuracy.

REFERENCES

1. Procedure for Synthesizing Internal Traffic Movements. North Carolina Department of Transportation, Raleigh, 1976.
2. S. Khasnabis and M. R. Poole. Synthesizing Travel Patterns for a Small Urban Area. Traffic Engineering, Aug. 1975.
3. D. G. Modlin, Jr. Synthesis of Through Trip Patterns in Small Urban Areas. Department of Civil Engineering, North Carolina State Univ., Raleigh, 1971.
4. A. D. Jones. A Simplified Procedure for Major Thoroughfare Planning in Small Urban Areas. Purdue Univ., West Lafayette, IN, and Indiana State Highway Commission, Indianapolis, Joint Highway Research Project Rept. 18, July 1972.
5. Trip Generation Analysis. Federal Highway Administration, U.S. Department of Transportation, Aug. 1975.
6. Mel Conner and Associates. Murray, Kentucky, Urban Area Transportation Study. Kentucky Department of Highways, Lexington, Sept. 1968.

Framework for Transferring Travel Characteristics of Small Urban Areas

Hani S. Mahmassani*, Massachusetts Institute of Technology, Cambridge
Oreste M. Bevilacqua*, De Leuw, Cather and Company, San Francisco
Kumares C. Sinha, School of Civil Engineering, Purdue University, West Lafayette, Indiana

It is felt that smaller urban areas (less than 250 000 population) can benefit significantly from the transportation studies that have already been conducted in other, similar urban areas. The results of a study of the spatial transferability of various urban travel characteristics are examined. Such characteristics for the small metropolitan areas of Indiana, as well as other selected midwestern communities, are compiled and critically analyzed and compared with each other and with other local and national characteristics and trends. Trip frequency (generation) is

examined at three levels of aggregation: areawide, zonal, and household. A framework is then provided for the transferability of trip-generation parameters of concern to planners.

Initiating a continuing, comprehensive, and coordinated transportation planning process can represent a formidable task for small metropolitan areas as their population reaches 50 000. This process usually follows the pattern set by the larger metropolitan areas in the 1950s, with its high requirements in terms of data collection (by origin-destination survey), technical complexity, and financial resources. However, the problems faced by smaller areas are usually different in nature, in magnitude, and in context from those in larger areas. Moreover, the changing emphasis in urban transportation planning (1) from long-range, large-system planning to short-range improvements aimed at making better use of existing facilities, coupled with public concern over environmental and energy issues, has led to a reassessment of program priorities. As a consequence of this change in planning orientation, the development and application of "conventional", full-scale, comprehensive transportation studies is becoming less appropriate. It is therefore necessary to develop and implement simplified alternative planning approaches that can reduce the time and cost required by the transportation planning process, thus saving resources that can be redirected toward program implementation and the resolution of other issues.

One approach to simplifying the transportation planning process is to eliminate the need for a full-scale origin-destination (O-D) survey, which is by far the most costly, time-consuming, and time-delaying element of the conventional process. This can be achieved by reproducing the travel patterns in the urban area under study, based on the socioeconomic and physical characteristics of that area, and using parameters and relations developed and calibrated in other areas where comprehensive O-D surveys have been conducted. In such an approach—referred to as synthetic travel demand modeling—the information traditionally obtained from the O-D survey is fabricated or synthesized by using parameters "borrowed" from "similar" areas. Questions arise, however, as to which parameters and relations can be transferred and which areas can be used as a source of such parameters.

This paper addresses the above questions for parameters and models that characterize the trip-generation (frequency) aspect of trip making. It presents a critical appraisal of some of the suggested synthetic modeling techniques and develops a framework for transferring trip-frequency parameters for three levels of aggregation: areawide, zonal, and household. This theoretical framework is supported by empirical evidence derived from the comparison of parameters obtained from various study areas for each of the relevant levels of aggregation.

SCOPE OF THE STUDY

Urban areas in Indiana that have populations between 50 000 and 250 000 were studied. The following Indiana urban areas were examined: Anderson, Evansville, Fort Wayne, Lafayette, Muncie, South Bend, and Terre Haute. The general characteristics of these areas are given in Table 1. The purpose of the study was to make more efficient use of the information made available by the full-scale transportation studies conducted in these areas. Its primary objective was to determine the extent to which this information could be used to develop "universal" travel parameters that could be applied in other, comparable urban areas and to provide the framework for such use. In other words, it would assess the transferability of parameters and models calibrated in certain areas to other areas and relate this transferability to characteristics of the urban areas (socioeconomic in the case of trip-frequency parameters).

Secondary objectives included the development of a data base that could be used for the following purposes:

1. Cross-checking of the output of the planning process in a given area by comparing key travel parameters with available information from other areas for the purpose of assessing its reasonableness (2), and
2. Input for quick-estimation ("quick-response") techniques that might be needed for rapid evaluation of policy alternatives (3).

Although this study focuses on urban areas in Indiana, it is anticipated that the results and procedures would be directly applicable to other midwestern communities that fall into the same size group. Because of the general nature of the results, their validity is by no means limited to the state or regional level.

DISTRIBUTION OF TRIPS BY PURPOSE

Variation Between Urban Areas

Transportation studies have usually classified internal trips by as many as seven purposes. In the interest of simplification, it is recommended that fewer trip purposes be used in the demand modeling process (4). For small urban areas, the following three trip purposes are usually adequate: home-based work (HBW), home-based other (HBO), and non-home-based (NHB).

Trip distribution by purpose (for internal vehicle trips only) for each of the study areas is given in Table 2. For those areas in which more than three purposes were used in the transportation study, the trips were combined accordingly into the three categories. Since the number of truck trips was relatively small compared with the total number of trips in these areas, truck trips were combined with non-home-based trips (standard procedure used in most small-area transportation studies).

Wilson and Kristoffersen (5) have shown that trip distribution by purpose is independent of city size for smaller and medium-sized cities. However, this does not necessarily imply that this distribution is identical for all cities. The hypothesis of the independence of the distribution of trips among purposes from the various factors that differentiate cities is tested in this section. A chi-square test was used for this purpose (6).

To use this test, the number of trips by purpose and by urban area can be arranged in a contingency table, with urban area as the row factor and purpose as the column factor. However, the actual frequencies (number of trips) obtained from the O-D survey (before expansion) should be used instead of the numbers or percentages given in Table 2 because the test statistic is sensitive to sample size. The hypothesis of independence was rejected with 99.5 percent confidence (χ^2_{data} = 5596 versus $\chi^2_{12,0.005}$ = 28.299). The results of this test, therefore, demonstrated that the distribution by purpose of internal vehicle trips was not the same for all small urban areas in Indiana.

Pairwise comparisons of each of the urban areas by use of chi-square tests also provided statistical evidence for the rejection of the hypothesis that trip distribution by purpose is identical in any two of the urban areas studied. However, from a practical standpoint (as opposed to a purely statistical one), and considering the

Table 1. General characteristics of study areas.

Characteristic	Anderson	Evansville	Fort Wayne	Lafayette	Muncie	South Bend	Terre Haute
Survey year	1971	1970	1961	1970	1971	1967	1971
Population	90 338	175 514	203 861	101 125	100 056	219 018	102 729
Occupied dwelling units	29 808	60 500	64 780	29 758	31 015	69 091	41 418
Automobiles per household							
Census[a]	1.30	1.26	1.36	1.27	1.30	1.30	1.18
Transportation study reports[b]	1.30	1.30	1.16	1.54	1.52	1.25	1.00
Persons per household	3.03	2.90	3.15	3.40	3.33	3.17	2.48

[a]From 1970 data for the SMSA.
[b]From reports for the study area (within the cordon area).

Table 2. Total trips by purpose.

Urban Area	HBW		HBO		NHB[a]		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Anderson	47 105	16.2	145 958	50.3	67 139	33.5	290 202	100
Evansville	74 586	16.5	188 285	41.6	189 923	41.9	452 794	100
Fort Wayne	90 203	23.3	146 990	38.0	150 029	38.7	387 222	100
Lafayette	44 337	15.0	154 067	52.1	97 306	32.9	295 710	100
Muncie	38 591	13.4	159 017	55.2	90 641	31.4	288 249	100
South Bend	79 672	15.1	208 452	39.3	241 896	45.6	530 020	100
Terre Haute	36 745	15.7	121 836	52.1	75 130	32.2	233 711	100

[a]Including truck trips.

Table 3. Distribution of household trips by purpose for city of Evansville.

Number of Automobiles Owned	HBW		HBO		NHB		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
0	6	22.2	15	55.6	6	22.2	27	100
1	2351	23.5	4996	49.9	2661	26.6	100 08	100
2	4000	25.2	6997	44.1	4872	30.7	158 69	100
≥3	1117	25.0	1974	44.1	1381	30.9	44 72	100

relatively large uncertainty accepted in transportation studies because of the very nature of the issues addressed (especially in the context of smaller urban areas, where the decision may concern the number of lanes of a certain facility), the percentage distributions of trips by purpose obtained for the study areas and given in Table 2 can be useful. The average (unweighted) distribution for these areas is as follows: HBW, 16.5; HBO, 46.9; and NHB, 36.6. These percentages can be used to develop gross estimations or as an initial assumption (to be adjusted later in the process) in a synthetic modeling effort (7).

Variation Within Urban Areas

In order to understand some of the factors behind the variation of the trip-purpose distribution between urban areas, this distribution was investigated within urban areas. Trips made by individual households are considered in this case. Separate trip-purpose distributions were developed for levels of household automobile ownership by using raw household O-D survey data collected in Evansville. The numbers and percentages of trips observed for each purpose and level (category) of automobile ownership are summarized in Table 3.

To test the hypothesis that the distribution by purpose of trips made by households is independent of the socioeconomic characteristics of households (as reflected in automobile-ownership status), a chi-square test similar to the one described earlier was used. Once again, test results led to the rejection of the independence hypothesis; i.e., it was found that the distribution of trips by purpose is related to the socioeconomic characteristics of the household.

No direct testing of whether the distribution of trips by purpose for each socioeconomic category is identical among urban areas was conducted. Later, however, this is done indirectly in relation to disaggregate household trip rates, where it is demonstrated that the average number of trips for each purpose made by individual households of a certain socioeconomic category does not differ significantly between urban areas. This conclusion, coupled with the conclusion that trip distribution by purpose is significantly different for each socioeconomic group within an urban area, appears to indicate that the variation of the overall (areawide) trip distribution by purpose between urban areas can, at least in part, be attributed to the variation of the socioeconomic mix (distribution of households among socioeconomic groups) between these areas.

The practical implications of the above are that areawide trip distributions by purpose should not be transferred indiscriminately between urban areas. Only if the socioeconomic mix is identical (which is not easy to prove in practice) can such a transfer be justified.

TRIP FREQUENCY (GENERATION)

Areawide Trip Rates

On an areawide basis, trip frequencies are most commonly expressed in terms of average number of trips per household and average number of trips per person, which are calculated from the overall number of trips and the overall number of households (or area population). These rates are useful in predicting total travel in an area or as a criterion for checking the reasonableness of survey results or model outputs. Vehicle trips per household as well as vehicle trips per person for the study areas are given in Table 4. These rates exhibit a fairly wide range of variation: 5.64-9.94 trips/household and 1.90-3.22 trips/person.

Table 4. Areawide vehicle trip rates.

Urban Area	HBW Per Household	HBW Per Capita	HBO Per Household	HBO Per Capita	NHB[a] Per Household	NHB[a] Per Capita	All Per Household	All Per Capita
Anderson	1.58	0.52	4.90	1.62	3.26	1.08	9.74	3.22
Evansville	1.23	0.42	3.11	1.07	3.14	1.08	7.48	2.57
Fort Wayne	1.39	0.44	2.27	0.72	2.32	0.74	5.98	1.90
Lafayette	1.49	0.44	5.18	1.52	3.27	0.96	9.94	2.92
Muncie	1.17	0.39	4.83	1.59	2.92	0.91	8.92	2.89
South Bend	1.15	0.36	3.02	0.95	3.50	1.11	7.67	2.42
Terre Haute	0.89	0.36	2.94	1.19	1.81	0.73	5.64	2.28

[a]Including truck trips.

Table 5. Cross classification of HBW vehicle trip production for Lafayette and Evansville.

City	Number of Automobiles Owned	Trip Rate by Household Size: One Member	Two Members	Three Members	Four Members	Five or More Members
Lafayette	0	0.004	0.004	0.091	0	0
	1	0.572	0.868	1.118	1.371	1.456
	2	0.867	1.870	2.266	2.280	2.231
	3+	-	2.0	2.704	2.232	2.888
Evansville	0	0.006	0.008	0	0.034	0
	1	0.829	0.910	1.179	1.375	1.345
	2	1.0	1.982	1.894	2.044	2.023
	3+	-	1.929	2.812	2.967	2.973

The factors behind this variation in aggregate areawide trip-frequency parameters were investigated by Chan in an effort to develop models for predicting areawide trip frequency (8). It was shown in that study that, for urban areas that had populations in the range of 50 000-800 000, only socioeconomic factors contributed to the differences in areawide trip-frequency parameters; urban form and structure were not differentiating factors in trip-frequency prediction. That same study found that average automobile ownership per dwelling unit (over the whole study area) showed the highest significant correlation with trip rates.

The regression equation developed by Chan for urban areas with populations of 50 000-800 000 is given below:

$$\text{Average number of person trips/household} = 1.262 + 6.591 \times (\text{average automobile ownership/household}) \quad R^2 = 0.412 \qquad (1)$$

Testing this equation for the Indiana study areas showed that the equation does not lead to very accurate predictions. However, any areawide aggregate regression model is likely to suffer averaging biases inasmuch as average automobile ownership per household might not be a good representative of the distribution of automobile ownership for each household in the urban area.

To test whether the distribution of automobile ownership is independent of the urban area (i.e., similar for all areas), a chi-square test was again used. The statistical test in fact substantiated that the distribution of households by automobile ownership is not the same for all of these urban areas. Using this technique for testing the similarity between household automobile-ownership distributions of pairs of urban areas (2x4 contingency tables) shows that only two of the urban areas, Anderson and Muncie, have similar household distributions by automobile ownership. All of the others are different with respect to each other. This socioeconomic similarity between Muncie and Anderson is also reflected in the areawide trip rates. For Muncie and Anderson, respectively, vehicle trips per household equal 8.92 and 9.74 and person trips per household equal 12.95 and 13.18.

Automobile ownership was used in the above tests only as an indicator of socioeconomic characteristics. This does not mean that it is the only important factor in determining areawide travel frequency. The finer is the categorization of households by various socioeconomic characteristics in urban areas, the more accurate is the comparison between these areas and the more reliable are the results obtained by using borrowed parameters. This concept is described in greater detail in subsequent sections of this paper.

TRIP-GENERATION ANALYSIS AT THE ZONAL LEVEL

Most transportation studies use the zone as the basic geographic unit of analysis. Multiple-regression techniques are generally used to relate zonal trips (or average zonal trip rates) to zonal socioeconomic and land-use characteristics. This is a very familiar technique, and it will not be discussed here.

A feeling exists among many planners that zonal regression equations developed in a certain area could be transferred and used to model travel in a different area, especially in the case of small urban areas (5, 9, 10). Along the same line, equations based on pooled data from different cities have been developed and recommended for use in synthetic trip-generation analysis (9). This direct borrowing of zonal regression equations was tested for some of the urban areas included in this study by using the equations for internal vehicle trip production for each of the three purposes described earlier. This test demonstrated that aggregate models cannot be reliably transferred between different urban areas. To overcome the disadvantages of using aggregate models, it was suggested that regression equations with trips per household as the dependent variable be calibrated at the household level so that the household is considered as the basic unit of analysis (11). Trip-generation rates at the household level are discussed in the following section.

DISAGGREGATE HOUSEHOLD TRIP RATES

Two types of disaggregate household trip-generation models can be used: regression models calibrated at the household level or cross-classification (category analysis) models. Cross-classification models stratify households according to their socioeconomic character-

Table 6. Summary of t-test results: cell-by-cell comparison of Lafayette and Evansville trip rates.

Cell: Automobiles	Cell: Members of Household	t^a	HBW[b]	Are Means Different?	t^a	HBO[b]	Are Means Different?	t^a	NHB[b]	Are Means Different?
0	1	0.3152	1.960	No	3.5734	1.965	Yes	2.9024	1.965	Yes
0	2	0.4713	1.970	No	1.7257	1.960	No	1.2866	1.960	No
1	1	3.379	1.960	Yes	4.5151	1.960	Yes	0.4817	1.960	No
1	2	0.709	1.960	No	4.6519	1.960	Yes	3.5254	1.960	Yes
1	3	0.5929	1.960	No	0.648	1.960	No	0.5013	1.960	No
1	4	0.0331	1.967	No	1.939	1.967	No	2.3513	1.968	Yes
1	5	0.8887	1.970	No	0.1375	1.968	No	1.8203	1.970	No
2	2	1.0944	1.960	No	2.3101	1.960	Yes	0.3085	1.960	No
2	3	2.8879	1.967	Yes	1.1887	1.966	No	1.3155	1.965	No
2	4	1.9060	1.967	No	1.0002	1.967	No	1.9298	1.967	No
2	5	1.8103	1.965	No	1.8753	1.960	No	2.2254	1.960	Yes
3	3	0.2959	1.960	No	0.6842	1.99	No	1.0664	1.981	No
3	4	2.4935	1.974	Yes	1.9689	1.974	No	1.0842	1.979	No
3	5	0.2833	1.974	No	3.459	1.978	Yes	3.1343	1.980	Yes

[a] t-statistic computed from the data.
[b] $t_{d,l,0.025}$, t-statistic from standard distribution.

istics and provide estimates of trip rates for each of the household categories. Both types of models predict equally well over the full range of households and appear to be indistinguishable with respect to sample-size sensitivity (12).

Disaggregate household models have been shown to be superior to aggregate models in various respects, such as yielding better estimates of zonal totals and the mean trip rate (11). Disaggregate models are also more data efficient than aggregate models, requiring fewer data for their calibration (10, 13). Moreover, because of their behavioral nature it is claimed that they are stable temporally as well as spatially (10, 13-16).

The spatial stability of household trip rates and their consequent transferability between small urban areas are investigated in this section. Cross-classification models for Lafayette and Evansville have been developed for this purpose. Category analysis has been used instead of the household regression technique because of its ability to express nonlinear relations, its inherent distribution-free characteristic (statistical distribution of the trip rates within each cell need not be assumed), and its ease of understanding and application. The independent variables used to classify households are automobile ownership and number of household members, both of which have been shown to be the major determinants of household trip generation (17). For illustrative purposes, Table 5 gives the average HBW vehicle trip rates per household category for each of the two test areas.

To compare trip rates, the use of a test statistic comparable to the chi-square distribution has been suggested (14). This test provides an overall comparison of all trip rates from two separate trip tables. It is, however, very sensitive to differences in individual-cell mean pairs. In other words, if the rates for one of the cells are somewhat different for the two tables, the test might lead to the conclusion that the two tables are significantly different. Since only some cells might have different average trip rates (often because of bad data points or slight deviations in each of the cells), trip tables should be compared on a cell-by-cell basis. A t-test can be used for this purpose. The test statistic used in this situation involves the comparison of two means for two cells, each of which contains a sample from two independent random variables, where the respective sample sizes are unequal and greater than 30 and the respective population variances are unknown and not necessarily equal. The results of these tests, summarized in Table 6, lead to the conclusion that, overall, trip rates are not significantly different for the two areas for all three trip purposes. Only a few cells show significant differences. These might be caused by the inaccuracy of some of the observations, which might lead to erroneous rates.

This comparison strengthens the belief that disaggregate household models are transferable from one area to another, especially in the case of small urban areas, regardless of the socioeconomic differences between them. This conclusion is also consistent with the finding, demonstrated in an earlier section of the paper, that the distribution of trips by purpose is related to household characteristics. In spite of its simplicity and seeming lack of sophistication, cross classification is nevertheless the most appropriate technique for trip-generation analysis within the structure of the conventional urban transportation planning process.

CONCLUSIONS

It has been shown in this paper that overall trip distribution by purpose is not similar between urban areas. Within an urban area, trip distribution by purpose varies among socioeconomic groups.

In a parallel way, it has been shown that the variation of areawide trip-frequency rates among urban areas reflects the variation of the socioeconomic distribution of households between these areas. Areawide frequency parameters would therefore be transferable between urban areas only when these areas have a similar socioeconomic distribution of households.

To transfer trip-frequency relations (number of trips as a function of other variables), these relations should relate trip making to its basic socioeconomic determinants at the household, or disaggregate, level. Aggregate equations tend to mask the causal aspect of the relations, and care must be exercised in borrowing such equations. Empirical evidence to that effect has been presented: Zonal regression equations developed in the study areas led to erroneous results when they were used in areas other than the ones for which they were developed. However, unlike aggregate models, disaggregate household models demonstrate the highest potential as well as the strongest theoretical justification for being transferred between small urban areas. Therefore, it is felt that trip-generation rates computed in one urban area whose population is in the 50 000-250 000 range can be successfully applied to synthesize travel in another urban area of comparable size, if the trip rates are derived from household-level data.

ACKNOWLEDGMENT

This paper summarizes the results of a research project conducted by the Joint Highway Research Project staff at Purdue University. The project is funded by the Federal Highway Administration, U.S. Department of Transportation, through Highway Planning and Research funds administered by the Indiana State Highway Commission.

We, the authors, accept sole responsibility for the contents of this paper.

REFERENCES

1. K. E. Heanue. Changing Emphasis in Urban Transportation Planning. Presented at the 56th Annual Meeting, TRB, 1977.
2. H. S. Levinson. Characteristics of Urban Transportation Demand: A New Data Bank. TRB, Transportation Research Record 673, 1978, pp. 53-59.
3. A. B. Sosslau, A. B. Hassam, and M. M. Carter. Manual Techniques and Transferable Parameters for Urban Transportation Planning. TRB, Transportation Research Record 673, 1978, pp. 32-40.
4. Transportation Planning in Small Urban Areas: The Appropriate Level of Effort. Federal Highway Administration, U.S. Department of Transportation, 1977.
5. S. Kristoffersen and E. M. Wilson. Trip-Generation Synthesis for Small and Medium-Sized Cities. TRB, Transportation Research Record 638, 1977, pp. 18-21.
6. I. W. Burr. Applied Statistical Methods. Academic Press, New York, 1974, p. 419.
7. Traffic Simulation Model Procedures Report for Columbus: Bartholomew County Transportation Study. Vogt, Sage, and Pflum Consultants, Indianapolis, June 1976.
8. Y. Chan. Review and Compilation of Demand Forecasting Experiences: An Aggregation of Estimation Procedures. Pennsylvania Transportation Institute, Pennsylvania State Univ., University Park, June 1977.
9. J. Bates. Development and Testing of Synthetic Generation and Distribution Models for Urban Transportation Studies. State Highway Department of Georgia, Atlanta, 1971.
10. W. L. Grecco and others. Transportation Planning for Small Urban Areas. NCHRP, Rept. 167, 1976.
11. H. Kassoff and H. D. Deutschman. Trip Generation: A Critical Appraisal. HRB, Highway Research Record 297, 1969, pp. 15-30.
12. R. Dobson and W. E. McGarvey. An Empirical Comparison of Disaggregate Category and Regression Trip Generation Analysis Techniques. Transportation, Vol. 6, 1977, pp. 287-307.
13. Trip Generation Analysis. Federal Highway Administration, U.S. Department of Transportation, Aug. 1975.
14. A. Chatterjee, D. R. Martinson, and K. C. Sinha: Trip Generation Analysis for Regional Studies. Transportation Engineering Journal, ASCE, Vol. 103, Nov. 1977.
15. E. J. Kannel and K. W. Heathington. Temporal Stability of Trip Generation Relations. HRB, Highway Research Record 472, 1973, pp. 17-27.
16. H. F. Wooton and G. W. Pick. A Model for Trips Generated by Households. Journal of Transport Economics and Policy, Vol. 1, May 1967, pp. 137-153.
17. W. Y. Oi and P. W. Shuldiner. An Analysis of Urban Travel Demands. Northwestern Univ. Press, Evanston, IL, 1962.

**H. S. Mahmassani and O. M. Bevilacqua were at Purdue University when this research was performed.*

Land-Use-Allocation Model for Small and Medium-Sized Cities

C. J. Khisty, Department of Civil and Environmental Engineering, Washington State University, Pullman

A residential land-use-allocation model most suitable for use in small and medium-sized cities is described. It can also be used in large metropolitan areas to serve as a check or backup method on the reasonableness of forecasts produced by more sophisticated models. The model makes use of Gompertz curves and the concept of holding capacity to allocate regional totals to planning areas. Residential development factors are then used to further distribute these planning-area totals to small areas such as census tracts or traffic zones. In an ex post facto test of this model in which the U-statistic was used as a measure of performance, the accuracy of the method was found to be excellent in comparison with that of sophisticated, computer-oriented urban development models. Use of the procedure will save money, time, and personnel, all of which are important considerations for planning organizations that work under a fixed budget.

Land-use-allocation models fuel the typical four-step sequential transportation models. The general land-use model used in this process takes areawide forecasts of several socioeconomic variables as control totals and uses some procedure to allocate them to small areas, usually traffic analysis zones. The allocation procedures currently used by transportation planning agencies range from traditional "manual" techniques to sophisticated urban development models such as the Projective Land-Use Model (PLUM). Many small and medium-sized cities do not have the expertise, time, or money to run these large-scale models but prefer to rely on simple, less expensive, and more transparent models. Such methods, however, have not been generally developed and validated.

This paper describes a simple method of land-use allocation for small areas, in which the concepts of holding capacity, Gompertz curves, rates of land consumption, and residential development factors are used.

Although many transportation study areas (1-3) have in the past used some of the concepts mentioned above in distributing areawide totals of population and other socioeconomic variables to small areas (such as planning areas, census tracts, and traffic analysis zones), these concepts have not been collectively used and tested in any one study.

CURRENT PRACTICES AND PROBLEMS

Land-use models are concerned with providing small-area forecasts of population and employment in a suitable form for input to a trip-generation analysis. Beginning in the 1950s with metropolitan transportation studies, numerous attempts have been made to forecast these phenomena by using varying degrees of complexity and with widely varying degrees of success. By the mid-1960s, a number of rather large efforts to forecast metropolitan dynamics by the use of computer simulation had been undertaken.

One general approach to land-use forecasting might be called the "planned requirement approach." A traditional method, it derives from a precomputer technology. The most complete and widely used formulation of this approach is that of Chapin (4). The main analytic components of this procedure are, for each land-use category, a set of location requirements and a set of space requirements. Specific rules for the resolution of conflicts among land uses competing for a site are not defined in this approach. These judgments must be made by the analyst, based on given principles and standards, special knowledge of local conditions, and what is considered to be in the best interest of the public.

The other and more sophisticated approach to land-use forecasting is the "market simulation approach". The archetype of this system of models was developed in the early 1960s by Lowry of the Rand Corporation (5). The general structure of the Lowry model, which has since been imitated, altered, and expanded, has been used in large metropolitan studies in recent years. Partly because of the overoptimistic outlook of their creators and partly because of the unrealistic expectations of their potential users, many of these efforts have been partial failures (6).

Most important, such models make heavy demands on the expertise and time of a metropolitan-area study staff and on the study budget. Most metropolitan-area studies would, for these reasons, like to rely on simple, less expensive, and far more transparent urban models. The model described here is one that meets this description.

METHODOLOGY

In this procedure, a thorough knowledge of the local area is assumed as a prerequisite for analysis and planning. It is assumed that the planner will use prescriptive planning dictated by the goals and objectives of the region—tempered, if necessary, by trend analysis—as opposed to purely predictive planning. The sequence of operations is as follows.

Surveys, Regional Totals, and Patterns of Land Consumption

A traditional survey of all existing land uses in the region is required. Additional studies such as a "land capability study" for the region can be very useful (7). In this study, factors such as soil, slopes, floodplains, woodlots, noise hazards, access, and utilities are incorporated into a rational quantitative and inductive system for determining the "ideal" feasible use of land over the long-term future for currently undeveloped or developing areas.

Inventories and areawide forecasts of economic activity and population are, of course, necessary. The land-use inventories should include historic development trends; topographic and physical constraints on development; square hectometers of land in urban use; square hectometers of vacant land, classified as unusable or usable and as publicly or privately owned; location of major travel generators; identification of neighborhood and community boundaries; and nature of land-use controls (8).

The rate at which vacant land is being absorbed into urban use is important in land allocation. This is determined by establishing current rates of land-use consumption. These rates are defined as the amounts of land (measured in square hectometers) brought into urban use by a one-person increase in population or employment.

With the help of regional base-year totals for population and employment, rates of land-use consumption are established for the base year for at least the following categories: residential, commercial, industrial and utility, recreational-institutional, roadways and agricultural, and vacant. Projecting this rate of land-use consumption for a future horizon year depends for the most part on the size, density, and locational preference of the population.

Goals, Objectives, Regional Totals, and Conceptual Land-Use Plans

Through public participation, a survey of community attitudes and preferences is made with regard to housing, shopping, public transportation, neighborhood characteristics, mobility, recreation, regional environment, urban services, and public expenditures. These preferences are reflected in a set of goals, objectives, and policies and also in a set of conceptual land-use plans.

There are several organizational congeries in the land market, and it is best to involve them in the final structuring of the concept plans (9). The first and most important of these congeries is the real estate and building business. The second is made up of larger industries, businesses, and utilities (although they may not consume the greatest quantities of land, they do purchase the largest and most strategic parcels). Individual homeowners and other small consumers of land form the third social constellation. The fourth organizational complex is composed of the many local government agencies that deal with land, such as zoning boards, planning commissions, school boards, traffic commissions, and other agencies. It may be best to meet individually with each of these groups to develop concept plans and then, if necessary, to blend them all together.

Land-Use and Socioeconomic Distribution to Planning Areas

Suitable planning areas, which consist of census tracts that have similar socioeconomic characteristics, are demarcated. The distribution of residential capacity to these planning areas is now undertaken. The concept of full development, or "holding capacity", is used in performing this distribution, for which the maximum amount of developable land in each planning area is established. The holding capacity of an area is the existing population plus the product of vacant, available, suitable land and the expected density. Dwelling units can be substituted for population in calculating holding capacity. Thus, the ratio of dwelling units to holding capacity in a planning area determines what stage in the development cycle the planning area has reached in the base year. It also provides the basis for estimating at which stage the area will be at a future date—say, the horizon year.

The question of establishing the development cycle is taken up next. The time required for each planning area to move from its current state to full development varies by the size of the area, its distance from existing urbanization, and the relative attractiveness of the area for development. This progression of development can be represented by a set of Gompertz curves (also known as logistic curves) to show typical patterns of growth for different planning areas (10). The general form of the Gompertz curve is given by the expression

$$P_t = L/(1 + P_0 e^{-bt}) \tag{1}$$

where

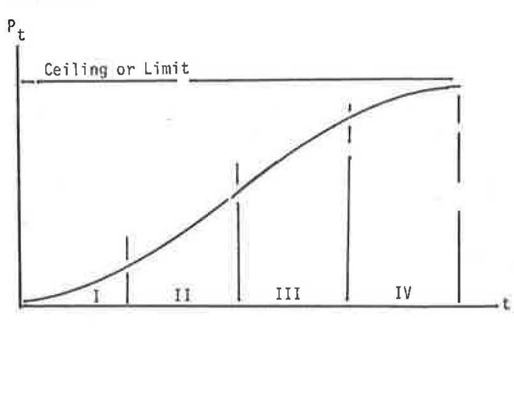
P_t = population at time period t,
L = some estimated maximum population (holding capacity of the area),
P_o = population at an arbitrary starting point in time,
b = rate at which population increases in time, and
t = time or some index of time.

The setting up of these working curves would need some historical sample data and calibration for establishing the value of parameter b.

Generally, four distinct stages of growth can be identified (see Figure 1). For example, first comes a very slow period of development in which areas move from totally rural to rural plus nonfarm. This stage is followed by a period of slow growth, often without total public utilities. Then comes a rapid-growth, or "boom", period, often heralded by, say, the installation of public sewers. Finally, a slow-growth period sets in as total capacity is reached. These Gompertz curves provide the rates of growth for individual planning areas. Future dwelling units for each planning area are then calculated based on projected average densities for both single- and multiple-family units. Multiplying the projected number of dwelling units by persons per household gives the future population.

The planning-area totals of population are then further distributed to census tracts and traffic zones by making use of residential development factors. These factors are dependent on several criteria, such as accessibility; water and sewer facilities; proximity to schools, employment centers, and shopping centers; and existing land use. Thus, the residential development factor for a

Figure 1. Gompertz (logistic) curve showing typical stages of growth.



given census tract provides a measure of the strength of the development potential and is used to allocate the planning-area total to the various census tracts. This process of activity allocation, like most manual distribution techniques, depends on judgment based on the understanding of the factors that promote growth. A typical set of residential development factors is given in Table 1. Zoning plans, aerial photographs, and the results of a land capability study can also be extremely helpful.

ILLUSTRATIVE EXAMPLE

A four-celled region will help to illustrate how the areawide totals of population are allocated to planning areas. Assume that the base year is 1978 and that the 1990 total population has been projected exogenously to be 72 000. Table 2 gives the basic information for each planning area. The additional land for residential use shown in this table can be calculated based on a set of assumptions. A typical set is shown below only for the purpose of this example; more elaborate and extensive assumptions may have to be established in real-world situations:

1. A deduction is made that reflects the fact that only 95 percent of the total land can be developed because of factors such as parcel shape, size, and ownership.
2. Physically constrained land is removed from the potential residential land.
3. Vacant land "committed" to nonresidential uses—such as industry, commerce, or major institutions—is subtracted. The committed uses are determined from

Table 1. Typical residential development factors.

Factor	Points
Community facilities	
Central sewer system service	
Existing (1965)	20
Planned to be in operation by 1980	15
Planned to be in operation by 1990	10
Central water service	
Existing (1965)	20
Planned to be in operation by 1980	15
Planned to be in operation by 1990	10
School, elementary school within 0.8-km radius	5
Accessibility	
Central business district	
0-5 min	5
6-10 min	3
≥11 min	0
Major shopping center	
0-5.5 km	5
5.6-10.8 km	3
≥10.9 km	0
Within census tract	0
Major employment center	
0-5 min	5
6-10 min	3
≥11 min	0
Within census tract	0
Highway system, census tract within 2.4 km of major arterial or freeway interchange	10
Mass transit system, established bus route within 0.4-0.8 km of census tract	5
Activity pattern	
Existing land use	
Industrial park	0
Subdivision	4
Commercial center	0
Population change, 1978-1990	
>25 percent	4
10-25 percent	2
<10 percent	0
Major recreational center	
Park within 3.2 km of census tract	4
Park within 8.0 km of census tract	2
Park available beyond 8.0 km of census tract	0

Note: 1 km = 0.62 mile.

Table 2. Capacity land-use projections by planning areas.

Planning Area	Total Land (hm^2)	1978 Residential Land (hm^2)	Additional Land for Residential Use (hm^2)	Maximum Land for Residential Use (hm^2)	1978 Dwelling Units	Dwelling Units at Capacity	1978 Dwelling Units as Percentage of Capacity
A	1056.1	233.6	110.4	344.0	5127	7 665	66.9
B	1462.4	538.0	362.8	900.7	6212	10 479	59.3
C	2133.0	548.1	521.1	1069.2	5536	11 930	46.4
D	1124.0	90.4	202.0	292.4	797	2 605	30.6

Note: 1 hm^2 = 2.47 acres.

Table 3. Development cycle: estimated years required to achieve given state of development.

Type of Growth	Percentage of Capacity Developed	Annual Growth Rate (%)	Approximate Number of Years in Stage
Very slow	0-10	1	10
Slow	11-20	2	5
Moderate	21-40	3	7
Boom	41-60	4	5
Moderate	61-80	3	7
Leveling off			
Fast	81-90	2	5
Slow	91-100	1	10
Total			50

existing zoning plans, local master plans, and announced major developments.

4. In addition to subtracting physically constrained land and land committed to major nonresidential activities, a final reduction may represent, say, a 23 percent roadway component and a 12-17 percent figure for residential services such as schools, parks, and commercial areas.

5. Finally, the net vacant residential land is translated into the number of dwelling units of the housing type and density that can be built for the area.

For planning areas that are almost all developed, the calculations of future growth simply reflect the potential development underzoning on the remaining vacant parcels. The typical development density and not the maximum permitted is used. For areas that are one-third to one-half developed, there is usually enough momentum in development trends to give a good idea of ultimate density. In currently rural areas, however, some judgment has to be used to project the ultimate urbanized density of these areas at some point in the distant future when all land use in the region is urbanized. The density to use is determined subjectively based on current densities, zoning on undeveloped land, and the socioeconomic character of the area.

Once the capacity residential land area is estimated, the next step in projecting future development is to assign total estimated dwelling units to the land. This is termed holding capacity and has already been described.

As mentioned before, development cycles are established for each planning area. Sample neighborhoods in these planning areas can be investigated for estimating the average number of years required to achieve a given stage of development. Table 3 gives a typical development cycle of 50 years. In actual practice, several such development cycles would be needed. It may also be noted that the last column in Table 2 provides the ratio of dwelling units to holding capacity in a planning area. This percentage is crucial because it determines the stage at which the area will be in, say, 1990.

The development cycle is then applied to the planning areas to estimate the percentage growth of dwelling units in each planning area, depending on the current stage of development of each area. This growth is given in Table 4, which also gives the forecast dwelling units for 1990. The forecast population for each planning area is given in Table 5. The vacancy and occupancy rates adopted in Table 5 can be derived by straight-line projections from historic data. The only information available is the total 1990 population of 72 000, obtained exogenously. It will be noticed that, since this total does not quite match the raw forecast total of 71 780, each planning-area total is prorated. The distribution to census tracts of 1990 forecast population by planning areas can then be performed by making use of the residential distribution factor analysis, described previously (Table 1).

After distributing the planning-area population to census tracts, it may be necessary to ensure that zoning ordinances applicable to different areas and census tracts are not violated. In this example, a one-shot estimate and distribution have been demonstrated. In the real-world situation, several trial cycles would have to be performed to match the regional control total of population.

MODEL PERFORMANCE

The reliability of the model was tested by using the U-test (11). In an ex post facto test of this method in which the U-statistic was used as a measure of performance, the accuracy of the method was found to be good (12).

CONCLUSIONS

As a viable alternative to the more costly, data-intensive computer models, the "manual" method described in this paper provides a simple, easy-to-understand, efficient procedure for developing land-

Table 4. Forecast growth of dwelling units in planning areas from 1978 to 1990.

Planning Area	Dwelling Units in 1978	Percentage of Growth by Annual Growth Rate[a]							1990 Forecast Dwelling Units
		1 Percent	2 Percent	3 Percent	4 Percent	3 Percent	2 Percent	1 Percent	
A	5127					3/4	2/5	1/3	6564
B	6212				4/1	3/7	2/4		8601
C	5536				4/3	3/7	2/2		7968
D	797			3/3	4/5	3/4			1193

[a] Percentage of growth per year/number of years.

Table 5. Forecast and actual population for 1990.

Planning Area	1990 Forecast Dwelling Units	Vacancy Rate (%)	Occupancy (persons per dwelling unit)	1990 Forecast Population	
				Raw	Final
A	6564	3.0	3.0	19 101	19 160
B	8601	2.6	3.1	25 970	26 050
C	7968	2.7	3.0	23 259	23 330
D	1193	3.6	3.0	3 450	3 460
Total				71 780	72 000[a]

[a]Control total derived exogenously.

use plans. The method forces the analyst to become intimately familiar with the study area and its zoning ordinances, physical characteristics, and growth trends before attempting to forecast. This is considered a positive feature of this method. Given a fixed budget of time, money, and personnel, simplification of currently used procedures is a prerequisite to the analysis of more alternatives or the incorporation of more impact analysis. Among the important features of this simplified model is the ability to communicate with those responsible for policy and administration and also with the public, both during and after the analysis.

REFERENCES

1. Land-Use and Development Forecast, 1970-2000: Technical Report. Akron Metropolitan Area Transportation Study, Akron, OH, 1976.
2. Population and Economics Study: Small Units Forecasts. Stark County Area Transportation Study, Canton, OH, 1977.
3. 1965 Regional Land-Use Inventory-Analysis and 1985 Forecasts. Toledo Regional Area Plan for Action, Toledo, OH, 1973.
4. S. F. Chapin. Urban Land-Use Planning. Univ. of Illinois Press, Urbana, 1965.
5. I. S. Lowry. A Model of Metropolis. Rand Corp., Santa Monica, CA, 1964.
6. D. B. Lee. Requiem for Large-Scale Models. Journal of American Institute of Planners, May 1973, pp. 163-178.
7. A Manual Land Capability Evaluation and Scoring System for Five Development-Related Uses. Ohio Department of Natural Resources, Columbus, Tech. Rept. 6, 1975.
8. D. K. Witheford. Urban Transportation Planning. In Transportation and Traffic Engineering Handbook, Prentice-Hall, Englewood Cliffs, NJ, 1976.
9. W. H. Form. The Place of Social Structure in the Determination of Land-Use: Some Implications for a Theory of Urban Ecology. Social Forces, Vol. 32, No. 4, May 1954.
10. M. Yeates. An Introduction to Quantitative Analysis in Human Geography. McGraw-Hill, New York, 1974.
11. Institute of Transportation Engineers. Informational Report: Land-Use and Demography—Growth Versus Forecast. Traffic Engineering, March 1977.
12. C. J. Khisty. An Evaluation of Alternative Manual Land-Use Forecasting Methods Used in Transportation Planning. Department of Civil Engineering, Ohio State Univ., Columbus, Ph.D. dissertation, 1977.