*TRANSPORTATION RESEARCH RECORD* 677

# Transportation System Analysis

*TRANSPORTATION RESEARCH BOARD*

*COMMISSION ON SOCIOTECHNICAL SYSTEMS*
*NATIONAL RESEARCH COUNCIL*

*NATIONAL ACADEMY OF SCIENCES*
*WASHINGTON, D.C. 1978*

**Transportation Research Record 677**
Price $6.40

modes
1 highway transportation
2 public transit

subject areas
12 planning
13 forecasting
15 socioeconomics



**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.

To eliminate a backlog of publications and to make possible earlier, more timely publication of reports given at its meetings, the Transportation Research Board has, for a trial period, adopted less stringent editorial standards for certain classes of published material. The new standards apply only to papers and reports that are clearly attributed to specific authors and that have been accepted for publication after committee review for technical content. Within broad limits, the syntax and style of the published version of these reports are those of the author(s).

The papers in this Record were treated according to the new standards.

**Library of Congress Cataloging in Publication Data**
National Research Council. Transportation Research Board.
Transportation system analysis.

(Transportation research record; 677)
1. Transportation planning--Data processing--Congresses.
2. Urban transportation--Congresses. 3. System analysis--Congresses. I. Title. II. Series.
TE7.H5 no. 677 [HE193] 380.5'08s [380.5] 79-12500
ISBN 0-309-02822-1

**Sponsorship of the Papers in This Transportation Research Record**

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

**Transportation Forecasting Section**
*John R. Hamburg, John R. Hamburg and Associates, chairman*

Committee on Transportation Systems Design
*Gorman Gilbert, University of North Carolina, chairman*
*William G. Barker, Edward A. Beimborn, Melvyn Cheslow, Lewis P. Clopton, Robert T. Dunphy, Jerry L. Edwards, Gary R. Erenrich, Kevin E. Heanue, William S. Herald, James Dan Jones, Harold Kassoff, Robert L. Knight, Douglass B. Lee, Jr., Marvin L. Manheim, Charles William Ockert, Robert E. Paaswell, Richard H. Pratt, Sandra Rosenbloom, Larry H. Sams, William L. Smith, Darwin G. Stuart*

Committee on Transportation Information Systems and Data Requirements
*Joseph M. Manning, Urban Transportation System Associates, chairman*
*Robert E. Barraclough, Daniel E. Benson, H. W. Bruck, Kenneth J. Dueker, David J. Dunlap, Robert T. Dunphy, Christopher R. Fleet, Richard L. Forstall, Alexander French, Larry R. Goode, Yehuda Gur, Dayton P. Jorgenson, Roy E. Lave, Jr., William A. Lesansky, Alfred J. Marston, James J. McDonnell, Morton A. Meyer, Philip D. Patterson, Alan E. Pisarski, James E. Sale, John L. Skinner, Jr., Garrison P. Smith, Jr., Arthur B. Sosslau, Robert N. Tap, Ronald W. Tweedie*

Committee on Urban Activity Systems
*David E. Boyce, University of Illinois, Urbana-Champaign, chairman*
*Brian K. Barber, Alan Black, Leslie P. Boudrot, Jeffrey M. Bruggeman, John W. Dickey, Ralph Gakenheimer, Michael A. Goldberg, Peter Gordon, Richard L. Gustafson, John M. Hartwick, Gregory K. Ingram, Marc P. Kaplan, James B. Kau, Douglass B. Lee, Jr., Steven Richard Lerman, James T. McQueen, Will Terry Moore, Stephen H. Putman, Arlee T. Reno, Labh S. Sachdev, Kumares C. Sinha, Carl N. Swerdloff, H. C. W. L. Williams*

Committee on Computer Graphics and Interactive Graphics
*Jerry B. Schneider, University of Washington, chairman*
*William G. Barker, Dwight W. Briggs, Robert B. Cody, Robert B. Dial, Larry J. Feeser, William A. Fetter, Yehuda Gur, Hubert A. Henry, Robert H. Hinckley, Edgar M. Horwood, Larry L. Howson, Everett S. Joline, Donald C. Kendall, Gene Letendre, Paul Michel Lion, Peter S. Loubal, Patrick E. Mantey, Harold Moellering, Lloyd M. Pernela, Matthias H. Rapp*

**Transportation Systems and Planning Innovations Section**
*Charles V. Wootan, Texas A&M University, chairman*

Committee on Intergovernmental Relations
*Albert A. Grant, Metropolitan Washington Council of Governments, chairman*
*Lonnie E. Haefner, Washington University, secretary*
*Kurt W. Bauer, John J. Bosley, Richard J. Bouchard, William K. Brussat, Kathryn Cousins, John A. Dyer, Joel Ettinger, Stanley G. Feinsod, Douglas F. Haist, Edward M. Hall, Richard J. Huff, William E. Loftus, William L. Mertz, Thomas P. Messier, Clyde E. Pyers, Arthur D. Rogers, Joanne Vlecides, David B. Walker, James A. Walsh, Paul C. Watt, William S. Weber, Phillip L. Wilson*

Committee on Transportation Planning Needs and Requirements of Small and Medium-Sized Communities
*Max R. Sproles, Harland Bartholomew and Associates, International, chairman*
*J. Stephen Guhin, Federal Highway Administration, secretary*
*Edward G. Bates, Jr., David A. Doctor, Gorman Gilbert, George F. Hagenauer, Thomas J. Hillegass, R. Keith Jones, Monty C. Murphy, Edward S. Neumann, Marion R. Poole, Robert R. Roberts, Billy J. Sexton, Kenneth W. Shiatte, John L. Skinner, Jr., Vergil G. Stover, Robert D. Studer, John H. Waggoner, C. Michael Walton, Bruce B. Wilson, Eugene M. Wilson*

James A. Scott, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each report. The organizational units and officers and members are as of December 31, 1977.

# Contents

# Decision-Maker-Defined Cost-Effectiveness Framework for Highway Programming

David I. Wilson,* Department of Civil Engineering, University of Melbourne
Joseph L. Schofer, Department of Civil Engineering, Northwestern University

**Federal regulations that specify the development of the transportation improvement program charge metropolitan planning organizations with responsibility for establishing a forum for cooperative transportation decision making. This paper describes an effort by the Chicago Area Transportation Study to perform that role. It focuses on the development of a simple linear weighting scheme for use in ranking federal-aid urban system highway projects for inclusion in the transportation improvement program for DuPage County, Illinois. This approach was developed in close working relationship with local decision makers, who participated in the selection of measures of effectiveness, the weighting of those measures, and the definition of the overall scheme. The results of the method, which combined both traffic and environmental measures for two points in time, were provided to decision makers as a basis for choice rather than as a rule for decision making. The success of the effort was measured by the correspondence between the analytic ranking of projects and the transportation improvement program ultimately chosen by local officials.**

Recent federal regulations that define the requirement for the transportation improvement program (TIP) mandate metropolitan planning organizations (MPOs) to function as "forums for cooperative decision-making by principal elected officials. . . . " (1). The development of the transportation improvement program as a major product of the planning process focuses planning (and MPOs) on decision-making processes by calling for planning organizations to be more supportive of decision making than in the past when their products were plans with a long-range focus rather than immediate investment programs.

The role of the MPO is to guide the development of the processes of choice from those based largely on political negotiations to decision making based on objective consideration of needs, benefits, and costs while remaining politically responsive to the preferences and desires of citizens and interest groups. Working this closely with decision makers requires the development of interactive planning and analysis processes that are both sensitive to decision makers' needs and capabilities and technically sound. In developing such tools, compromises must be made, the ultimate objective being to provide technical support to an effective decision-making process. Thus, decisions and decision making must be the ultimate focus of the development of tools in the practical environment of the MPO (2).

This paper documents an evaluation procedure developed to select federal-aid urban system (FAUS) highway projects for inclusion in the annual element of the TIP in DuPage County, Illinois. This work was conducted by the Chicago Area Transportation Study (CATS) in cooperation with the Council of Mayors and City Managers of DuPage County, which is the principal political decision-making unit concerned with the investment of FAUS funds in the county. Although the methodology itself was tailored to meet the needs and desires of the DuPage County Transportation Subcommittee (TSC) of the subregional council, the approach itself may be useful in other programming contexts.

The charge to CATS in the development of this method was to create a procedure to produce a realistic ranking of highway projects selected from a larger list of locally generated projects, based on objective measures of the worthiness of each project, the readiness of each project for immediate implementation, and community values as reflected by the preferences of local officials. It was understood that this ranking procedure would be applied in support of decision-making processes and that the ultimate TIP would be chosen by decision makers; the ranking procedure would thus be an input to the overall decision-making process and might even help structure negotiations among local officials. It was not to be viewed as an inflexible decision rule, however.

Before the development and adoption of a relatively objective strategy of project evaluation, the TSC selected projects through a relatively unstructured process of negotiations. The first projects suggested for implementation were typically the first projects to be funded, and the readiness of a project for immediate implementation was sometimes more important in the choice process than the worthiness of the project itself. Such an approach was reasonable when the number of projects submitted was within the budget allocated for the FAUS program in the county. However, when the total costs of the suggested projects exceeded available capital, this unstructured negotiation process became more difficult to use. Therefore, the TSC sought a new evaluation and choice process based more directly on measures of project worthiness. It was important for the method to be objective to avoid claims of favoritism and provide clear justification for programming decisions.

## THE DECISION-MAKING ENVIRONMENT

The Council of Mayors and Managers of DuPage County is a subregional forum for decision making regarding FAUS funds allocated to the county level. This council retains decision-making authority, but the principal tasks associated with transportation investment choices are delegated to the TSC, which is made up of some city managers, public works directors, and principal traffic engineers from various communities. There are 12 members on this subcommittee. These individuals have a limited time budget and must annually examine 20 or more highway projects under this program and compare and evaluate them for the purpose of making a choice.

Subcommittee members have varied backgrounds, and all were not well versed in transportation planning. Each participant does have detailed knowledge of some of the candidate projects and at least a general knowledge of the others. This allows the application of personal experience and expertise in the choice of investment alternatives. Participants are generally aware of the unique aspects of proposed projects that are not easily accommodated in an objective evaluation process. For example, a bridge that has deteriorated to the point where it has become a safety problem can easily be accepted as a high priority for maintenance or improvement even without sophisticated benefit measures.

Still, given budget limitations, it is important to provide this decision-making group with some objective guidance for project ranking and selection of a TIP. The technique discussed below was developed to provide a high level of decision-maker participation in the evaluation process, to accommodate objective measures of project worth as well as other subjective and more political factors that were clearly important to the choice, and to encourage a detailed and structured discussion of alternatives within the decision-making body.

COST-EFFECTIVENESS EVALUATION

The approach adopted was based on the philosophy of cost-effectiveness analysis (3). The effectiveness of each alternative is characterized by a number of measures of effectiveness chosen as relevant to the local situation. An effectiveness index is related to project cost, and the result is used to create an initial ranking. The decision on what is cost-effective and which projects are to be included in the final improvement program is made by decision-makers and is clearly subjective and open to negotiation. The value of this general approach is in its ability to compare a variety of project types in a number of important dimensions and to provide guidance to decision makers while not making the decision itself.

The method begins with the selection of a set of measures of effectiveness or criteria that are used to characterize the alternative projects. These were generated through an interactive process that involved members of the TSC and the CATS technical staff. Discussions of issues and factors to be measured (and specific measures themselves) were held in small groups of three or four subcommittee members. The meetings were conducted both to identify the factors of importance to the subcommittee and to introduce participants to various measures of effectiveness that might prove to be relevant. The result of these small meetings was a broad set of measures, a number of which could not be put in operation because they were either ill-defined or too costly. In the second stage of the measure-development process, this broad set of measures was examined critically by the technical staff, and definitions were sharpened where necessary. The set of measures finally selected by the TSC was as follows:

1. Change in peak-hour travel time,
2. Change in equivalent property damage (EPD) rate of accidents,
3. Change in average daily congestion (volume/capacity ratio),
4. Change in off-peak daily travel time,
5. Change in noise pollution,
6. Change in air pollution, and
7. Number of dwelling units taken.

Change refers to the difference between the build and no-build cases for the measure. Effectiveness was thus characterized as reductions of various negative attributes. The measures were applied to two distinct time periods: the present (actually 1975 because of data limitations) and 1985. This provides an evaluation of both the immediate and medium-range effects of the proposed project. Decision makers chose to use both of these time points in their evaluation process; continuous time streams were not considered because of the problems associated both with treating them directly and with aggregating them to some "present value."

This set of measures is the result of interaction and compromise among decision makers and the technical staff. It is clear that there is a high level of intercorrelation among the measures, particularly measures 1 and 3 but also measures 2, 4, 5, and 6; this amounts to multiple counting of some effects of projects. A conceptually more sound approach might call for a reduced set of measures and an associated shift in the aggregation weights. But these measures reflect the impact categories of direct concern to the decision makers, and thus the choice was made to compromise in the direction of decision-maker interests to ensure their involvement with the expectation that a modest success at this level might provide an opportunity for further improvement of the methodology at a later date.

AGGREGATION OF MEASURES

The technical staff proposed to work with a matrix display format that illustrated the disaggregate effectiveness dimensions and costs of the alternatives and highlighted the trade-offs among them. Although aggregation of the effectiveness measures simplifies the choice process, it covers up important information about the specific attributes of alternatives, and it assumes that appropriate and stable weights can be derived.

The decision makers, however, expressed a preference for a simple linear weighting scheme to aggregate effectiveness. Arguments against weighting in general, and linear weighting in particular, were rejected in favor of the simpler approach. Since the decision makers were relatively familiar with the alternatives and their attributes and since the ultimate choice of projects for implementation was to be made by using the evaluation scheme along with other information and judgment, it was felt reasonable to adopt this simplified linear approach. The scheme was not a decision rule but merely a decision aid.

Weights or priorities associated with each of the measures were constructed from information provided by the participating decision makers by using a Delphi process (4). Delphic voting ensured that all members of the subcommittee had an equal opportunity to influence the outcome; in an open voting situation, it is more likely that dominant individuals would be able to have an inordinately large influence on the outcome. Delphi has also been shown to be an effective consensus-building technique within a group that has divergent interests (5). It tends either to move opinions toward a fairly strong consensus or to polarize them and thus to isolate areas of disagreement on which further, open discussion is warranted. Delphi also provided the possibility for individual participants to indicate a lack of understanding of one or more of the measures and to get an explanation to clarify those measures without having to admit to a problem within the group as a whole.

The Delphi weighting process was conducted over three rounds by using mail-back forms (Figure 1). A five-point semantic scale was used, and participants were provided with a graphical description of voting in the previous round as well as a personalized indication of how they had voted previously. The opportunity was then provided for a new, modified vote by each participant. Throughout this process, close personal contacts were maintained between the technical staff and the decision makers. This was accomplished by telephone calls and personal visits to clarify the procedure and the measures and to help participants respond to the questionnaires correctly. A secondary benefit of these close contacts was the establishment of a strong relation between the technical staff and the decision makers, which facilitated the successful implementation of the overall methodology.

A comparison of the group average weights for the various measures between rounds 1 and 4 indicates that there was a general trend toward convergence on a mean value for most measures. Average weights for each of the rounds on each of the measures are given in Table 1 along with the variance associated with each weight.

The data given in Table 1 show that the participants generally valued effectiveness in 1985 as more important than immediate effectiveness. That this is contrary to the economic theory of discounting was pointed out to members of the TSC (6). Their rationale was based on the observation that DuPage County is one of the most rapidly growing parts of the Chicago metropolitan area. They perceived their role as that of creating a transportation system to meet future needs when the population of the county would be larger and development more diverse.

For similar reasons, participants felt that environmental impact measures were generally not as important as traffic-related impacts. Therefore, they adopted a weighting structure that separated the two types of measures and placed a heavier weight on traffic measures with respect to environmental measures in a ratio of 5:2.

Thus, a simple, two-tiered, linear weighting system was used that partitioned the measures into traffic and environmental sets, both of which had their own measure-specific weights as produced in the Delphi process.

**Figure 1. Example of Delphi measure weighting response form.**

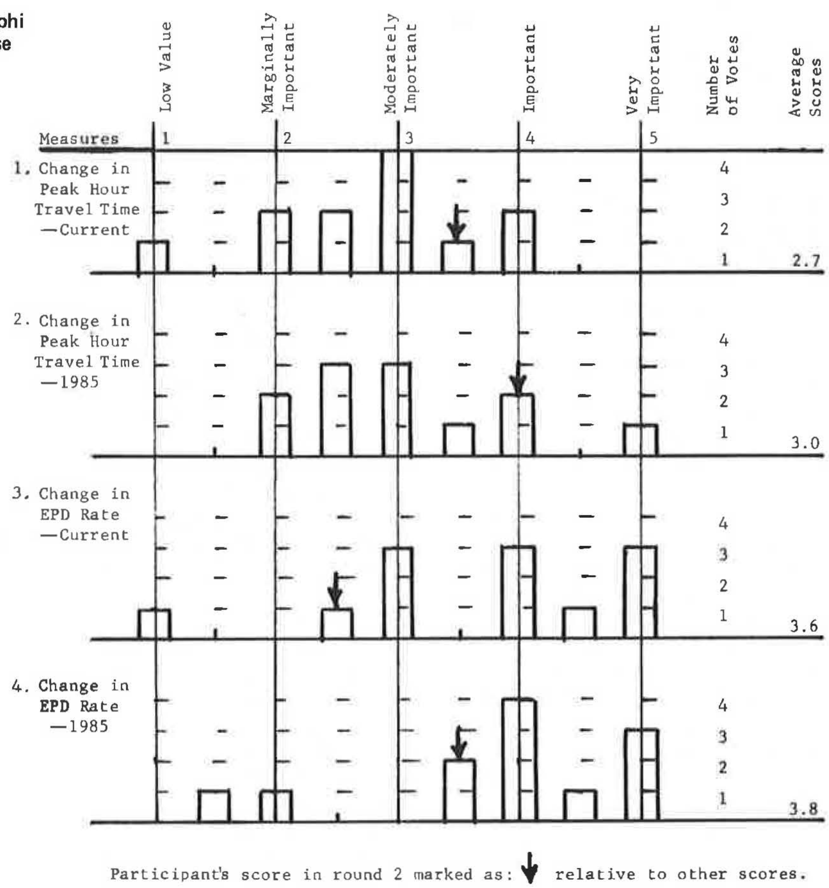



**Table 1. Means and variances of measure weights for three rounds of Delphi process.**

| Measure | Round 1 Average | Round 1 Variance | Round 2 Average | Round 2 Variance | Round 3 Average | Round 3 Variance |
|---|---|---|---|---|---|---|
| Peak-hour travel time | | | | | | |
| 1975 | 2.7 | 0.84 | 2.7 | 0.75 | 2.9 | 0.27 |
| 1985 | 3.0 | 0.61 | 3.0 | 0.81 | 2.9 | 0.97 |
| Equivalent property damage rate | | | | | | |
| 1975 | 3.7 | 1.79 | 3.6 | 1.47 | 3.8 | 0.79 |
| 1985 | 4.2 | 0.78 | 3.8 | 1.24 | 4.1 | 1.3 |
| Volume/capacity | | | | | | |
| 1975 | 3.5 | 0.91 | 3.5 | 1.2 | 3.7 | 0.46 |
| 1985 | 3.9 | 0.04 | 3.8 | 0.95 | 3.9 | 0.32 |
| Off-peak daily travel time | | | | | | |
| 1975 | 2.5 | 1.0 | 2.6 | 0.65 | 2.6 | 0.28 |
| 1985 | 3.2 | 1.31 | 2.8 | 1.19 | 2.8 | 0.62 |
| Environment, 1975 | | | | | | |
| Noise pollution | 2.7 | 0.39 | 2.7 | 0.61 | 2.6 | 0.45 |
| Air pollution | 2.5 | 0.79 | 2.5 | 0.94 | 2.5 | 0.8 |
| Dwellings | 3.4 | 2.17 | 2.8 | 1.86 | 3.1 | 0.95 |

Note: 1985 environmental variables have the same value as 1975 variables.

Table 2. Evaluation matrix.

| | Effectiveness of Performance Measures | | | | | | | | Effectiveness of Environmental Measures | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1975 Changes[a] | | | | 1985 Changes[a] | | | | 1975 Changes[b] | | 1985 Changes[b] | |
| Project | Peak Travel Time (s) | EPD Rate | Volume/ Capacity Ratio | Avg Daily Travel Time (s) | Peak Travel Time (s) | EPD Rate | Volume/ Capacity Ratio | Avg Daily Travel Time (s) | Noise | Air | Noise | Air |
| 1 | -6 | 1.45 | 0.17 | -3 | -8 | 1.45 | 0.17 | -3 | 0.5 | 0.5 | 0.5 | 0.5 |
| 2 | -11 | 0.28 | 0.06 | -4 | -10 | 0.28 | 0.06 | -3 | 0.5 | 0.5 | 0.5 | 1.0 |
| 3 | 228 | 4 | 0.12 | 228 | 228 | 4 | 0.24 | 228 | 0.5 | 0.5 | 0.5 | 0.5 |
| 4 | 21 | 2 | -0.05 | 16 | 21 | 2 | 0.08 | 16 | 0.5 | 0.5 | 0.5 | 0.5 |
| 5 | 0 | 2.4 | 0.2 | 12 | 108 | 2.4 | 0.22 | 42 | 0.5 | 0.5 | 0.5 | 0.5 |
| 6 | 18 | 3.5 | 0.06 | 18 | 18 | 3.5 | 0.05 | 18 | 0.5 | 0.5 | 0.5 | 0.5 |
| 7 | 0 | 3.9 | 0.13 | -2 | 11 | 3.9 | 0.17 | 12 | 0.5 | 0.5 | 0.5 | 0.5 |
| 8 | 192 | 0 | 0 | 193 | 192 | 0 | 0.01 | 193 | 0.5 | 0.5 | 0.5 | 0.5 |
| 9 | 17 | 5.1 | 0.24 | 15 | 70 | 5.1 | 0.28 | 10 | 0.5 | 1.0 | 0.5 | 1.0 |
| 10 | 0 | 2.7 | 0.15 | 0 | 0 | 2.7 | 0.34 | 0 | 0.5 | 0.5 | 0.5 | 0.5 |
| 11 | 12 | 7 | 0.18 | 0 | 348 | 7 | 0.22 | 0 | 0.5 | 0.5 | 0.5 | 1.0 |
| 12 | 5 | 0.38 | 0.22 | 0 | 45 | 0.38 | 0.28 | 0 | 0.5 | 0.5 | 0.5 | 0.5 |

[a] Negative changes are reductions in performance caused by the project.
[b] No project required taking dwelling units, and thus this measure was deleted; other environmental measures were scaled as described in the text.

## TECHNICAL ANALYSIS

The highway projects analyzed in this effort included those that dealt with intersections and those concerned with roadway sections: widenings, new signals, additional lanes, bridge replacement, and new portions of roadways. Standard forecasting tools were used to predict changes in traffic volumes, travel speeds, travel time, delays, accidents, and air quality for each of the proposed projects, both now and in 1985. Methods included studies of intersection delay and accident reduction, limited use of traffic assignment for analyzing larger projects, and simplified air-quality and noise-level forecasting tools.

Measures for the no-build alternatives were collected from several sources. On-site field observations were carried out to obtain operating speeds and delay times for peak and off-peak periods. Data collected in this way also permitted a more accurate calibration of the forecasting tools. Volume data were obtained from inventory files and from supporting documentation for each of the proposed projects. Forecast (1985) no-build volumes were synthesized from past, 1980, and 1990 assignments over the existing highway network.

In most cases, the nature of the projects allowed two simplifying assumptions to be made: the independence of the project and negligible traffic diversion caused by a project. Since the projects were local, relatively small in scale, and scattered throughout DuPage County, the first assumption—that of minimal project interaction—was reasonable. Before-and-after data on similar, previously implemented projects in DuPage County showed no clear pattern of traffic diversion. In all cases, natural growth could have accounted for any changes in volume that were observed. Most projects were therefore treated as though no traffic diversions would result. In a few proposals, the nature of the project suggested that diversions would occur; these were subjected to a more detailed analysis, which included a simplified traffic assignment.

## COST-EFFECTIVENESS ANALYSIS

Effectiveness measures for each of the projects for both 1975 and 1985 were assembled in an evaluation matrix in which the rows were projects and the columns were dimensions of effectiveness. Table 2 gives a simplified version of the actual matrix presented to decision makers, showing the trade-offs between projects for a single measure (cells in a single column) as well as the performance profile of each project across the measures (cells in a single row). This structured data set was presented to decision makers and remained available to them throughout the evaluation process. Their expressed preference, however, was to seek a collapsing of the data to fewer dimensions to facilitate understanding and choice.

It seemed apparent that two interrelated factors increased the willingness of decision makers to compromise the quantitative evaluation procedure in favor of simplicity. First, it became increasingly clear throughout the technical process that the TSC members reserved the right to make the final choice of a TIP independent of the results of the quantitative studies. If the studies proved instructive, they intended to use them; still, the intricacies of intercommunity negotiations were not to be given up for any formal evaluation tool, and thus the use of a simpler tool made sense. In addition, because of the relatively small scale of the entire set of projects and the established working relation among the decision makers, it was expected that considerable richness of information exchange would exist when any quantitative structure was used.

The second factor of importance is that this effort represented the first experience most of the decision makers had had with a formal approach to transportation evaluation. This led to a preference to start slowly and carefully (from their perspective) and not to be swallowed up by technical detail. There may be logic in the decision on the part of the technical analyst to work with simplified tools at the start in order to get a foot in the door.

The traffic-related effectiveness measures were originally prepared on a per user (i.e., per vehicle) basis to account for the fact that some measures were daily and others were annual. Information in the evaluation matrix was then factored by the appropriate daily volume for each project to produce measures of effectiveness per day.

Because measures of effectiveness were in different units, there was the possibility that "large-unit" measures such as daily travel-time savings might dominate "small-unit" measures such as dwellings taken. But it was important to ensure that the priorities assigned to various measures were those selected by the decision-makers (through the Delphi process described above) and not some technical feature of the measures. Therefore, all measures were converted from an absolute to a relative basis by dividing each element in a given column by the largest element in that column. This scaling approach is arbitrary, and a number of other options might have been used. The method selected set 1.0 as the maximum value and related all other values to that level; no measure could be less than zero.

The environmental measures were treated separately from the performance measures. Air pollution and noise levels were predicted for each project, related subjectively to established standards, and scored directly on a 0 to 1.0 scale by using the following decision rule: significant environmental deterioration with the project—0; no change or slight improvement in environmental conditions—0.5; improvement in conditions (including cases where the project is necessary to meet standards)—1.0. A similar arbitrary scaling system was used for dwelling units taken: no dwelling units taken—1.0; one to five units taken—0.5; more than five units taken—0.

The arbitrary nature of the resulting scale cannot be overlooked. It implies that the highest observed measure is the best possible, which is not the case. It arrays alternatives on a short scale between 0 and 1.0 with a bias toward 1.0 Most significantly, this method for rescaling incommensurables to a common arbitrary dimension assumes the feature of linearity in unit conversion, which is not realistic. For example, it assumes that the common-dimension value of the gap between 2000 and 4000 in one measure is the same as that between 2 and 4 on another measure. Although the approach chosen was simpler, there are conceptual advantages associated with using more complicated scaling methods that allow for the possibility of nonlinearities in this transformation process (7).

These scaled effectiveness measures were then aggregated by using the Delphi weights. The result was a scalar effectiveness score for each project. These were reported to the decision makers separately and in the form of cost-effectiveness indexes by dividing the effectiveness of each project by its capital costs. Operating and maintenance costs were not treated because it appeared that the proposed projects would not impose much change in existing operating and maintenance costs.

## PROJECT READINESS

Because funds in the FAUS program are allocated on a yearly basis and in general cannot be carried over to subsequent years, it is important to the subregional evaluation and the decision-making process to ensure that selected projects can start in the year for which they are programmed. Thus, it is important to know when a project will be ready for implementation as well as to know its cost-effectiveness.

Therefore, a checklist of administrative steps necessary to implement federally funded projects was prepared. The list included the following items:

1. Passage of TSC resolution;
2. Decision on sources of local match;
3. Development and approval of environmental impact statement (EIS), if required;
4. Public hearings (dependent on outcome of item 3 above);
5. A95 and state review processes;
6. Approval of design report (a description of the extent and nature of the project, which may include an EIS);
7. Completion of a joint funding agreement (final funding commitment of local, state, and federal governments);
8. Right-of-way acquisition, if necessary;
9. Preparation of the construction plan; and
10. Target letting date.

By knowing the type of project and the implementation time historically required for similar projects, it was possible to estimate the expected time it would take a project to proceed through all required stages. Projects not likely to be ready for implementation in the programming year were tabled for later consideration.

## DEVELOPMENT OF PRIORITIES

The cost-effectiveness index was used to prioritize the remaining projects subject to the constraint of the available budget. That is, projects were listed by decreasing cost-effectiveness until the budget was consumed. Some shifting of marginal projects is usually necessary to fit proposed projects into the available budget.

If maximization of benefits were the objective, it would be correct to use incremental benefit-cost analysis, or net present worth analysis, and in more complex cases to apply mathematical programming tools (8) to this ranking task. Here, however, the concern was not to maximize countywide benefits but to ensure a fair share of funds for participating communities where that share supported good projects. In this case, cost-effectiveness is more credible as a ranking tool. In addition, local decision makers desired considerable flexibility of choice so that their actions could reflect the needs and desires of communities and interest groups. Such factors cannot be easily accommodated within a formal evaluation framework. Indeed, decision makers have the responsibility for bringing such factors into the choice process; thus, it is logical to leave them with enough slack in the list of recommended priorities so that they can satisfy the variety of needs they represent.

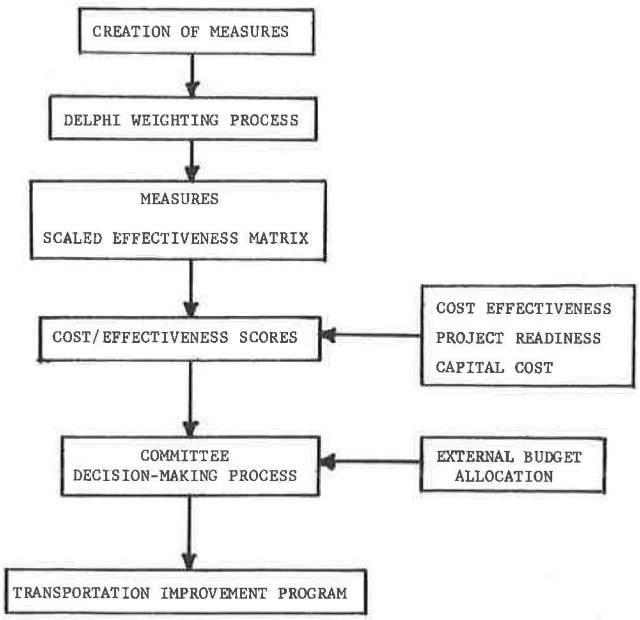
The tables below give a list of projects as they might be selected to be included in the annual element of the TIP—based on an attempt to maximize cost-effectiveness subject to the constraint on the budget—and the list of priority projects selected by local decision makers:

| Project | Cost-Effectiveness | Cost ($000 000) |
|---|---|---|
| A | 62.6 | 0.385 |
| B | 28.6 | 0.300 |
| C | 16.1 | 0.912 |
| D | 14.2 | 0.610 |
| E | 7.9 | 0.600 |
| F | 7.4 | 3.3 |
| H | 5.7 | 0.172 |
| I | 3.7 | 0.601 |
| Total | | 6.88 |

| Project | Cost-Effectiveness | Cost ($000 000) |
|---|---|---|
| A | 62.6 | 0.385 |
| C | 16.1 | 0.912 |
| F | 7.4 | 3.3 |
| G | 6.0 | 1.065 |
| I | 3.7 | 0.601 |
| J | 2.9 | 0.570 |
| Total | | 6.833 |

The correlation between the lists in these tables is apparent. Of the six projects selected by decision makers, all but one were selected by the cost-effectiveness method. Of the eight projects selected by the method, four were chosen by decision makers. The differences between these lists were easily explainable in terms of local history, needs, and compromises that

Figure 2. Overall evaluation and decision-making process.




are a natural part of the cost of a multiplicity of jurisdictions working together in planning and programming.

Although it can be argued that the differences between the recommended and adopted priority lists were substantial, given that this was the decision makers' first experience with using organized quantitative techniques for programming, it is not an unreasonable outcome. Even though the decision makers themselves called for the use of a more structured approach, they were not fully prepared to yield their negotiating positions totally to the results of the analysis. This is particularly the case because the analysis was performed by the MPO, which was to a certain extent viewed as an outside agency invited to assist in the programming process but not to seize it. Based on the reactions of decision makers, it is expected that their willingness to make better use of formal analysis has increased and, with proper technical support, will continue to expand in the coming years. But the forum for cooperative decision making had, in this case, begun to function.

## SUMMARY

The evaluation and decision process developed and tested in this effort is shown in Figure 2. It is nothing more than a simple linear weighting scheme, but its development in close working relationship with local decision makers makes it uniquely appropriate for supporting local transportation investment decisions. Decision makers initially sought out the assistance of such a scheme; they participated actively in the selection of the measures of effectiveness; and they worked together in a Delphi process to establish the weights associated with each of the measures. With the assistance of the technical staff, they studied the raw measures of effectiveness, and they carefully reviewed the products of the formal evaluation process. Although the process itself gave them a set of project rankings, they elected to implement a modified but quite similar ranking list in an effort to accommodate local goals.

The closeness of the decision makers' preferred project ranking to the product of the evaluation scheme suggests the merit of the general strategy if not of the evaluation scheme itself, not only in helping to justify a particular set of choices in this case but also as a potential tool for simplifying the choice process in future years. The differences between the rankings of the decision makers and those produced by the evaluation technique underscore the notion that the political process is still in control of investment decisions in the public sector and thus is still in a position to respond to the unique characteristics of the needs of individuals and groups.

## REFERENCES


1. Transportation Improvement Program. Federal Highway Administration and Urban Mass Transportation Administration, U.S. Department of Transportation, Federal Register, Sept. 17, 1975, p. 42977.
2. J. L. Schofer. Evaluating Transportation Alternatives. Proc., U.S. Department of Transportation Conference on Emerging Transportation Planning Methods, Daytona Beach, FL, Dec. 1976, in preparation.
3. E. N. Thomas and J. L. Schofer. Strategies for Evaluating Alternative Transportation Plans. NCHRP, Rept. 96, 1970.
4. G. S. Rutherford and others. An Experiment in Goal Formulation for Socio-Technical Systems. Journal of Urban Planning and Development Division, ASCE, Vol. 99, No. UP2, Sept. 1973.
5. N. Dalkey and O. Helmer. An Experimental Application of the Delphi Method to the Use of Experts. Management Science, Vol. 9, No. 3, April 1963.
6. T. Kuhn. Public Enterprise Economics and Transport Problems. Univ. of California Press, Berkeley, 1962.
7. Evaluation of Alternative Plans for the Indianapolis Region. Barton-Aschman Associates and City of Indianapolis, Feb. 1977.
8. J. G. Pigman and others. Optimal Highway Safety Improvement Investments by Dynamic Programming. TRB, Transportation Research Record 585, 1976, pp. 49-59.


*Publication of this paper sponsored by Committee on Transportation Systems Design.*

**Mr. Wilson and Mr. Schofer were with the Chicago Area Transportation Study when this research was performed.*

# Sensitivity Analysis of Selected Transportation Control Strategies

Robert J. Maxman and Darwin G. Stuart, Barton-Aschman Associates, Evanston, Illinois

**The relative potential of 13 different transportation control strategies for reducing projected regional vehicle kilometers of travel in the San Francisco Bay Area is analyzed. Through the use of a series of representative home-based work trips, it is possible to analyze mode-choice sensitivities directly by means of additional runs of the recently developed set of mode-split models for the region. These mode-split models are stratified by three automobile ownership categories for both primary and secondary workers. Ranges of potential mode-split shift for automobile, transit, and shared-ride modes are established for each transportation control strategy. Four combinations of strategies are examined. Graphs that help to depict relative mode shift potentials are developed by using stepwise incremental testing of various control measures. Mode-split sensitivities for representative trips are generalized to the regional level.**

Continuing concern about the air quality and energy consumption implications of current urban travel patterns has led to a growing interest in the potential effectiveness of transportation policies that might reduce overall vehicle travel. A wide range of such potential policies designed in one way or another to induce travelers to make greater use of multiple-occupancy vehicles (transit and car pooling) have been advanced. The general idea is to maintain current levels of personal mobility while reducing the number of vehicles and vehicle kilometers necessary to provide that mobility (7, 8).

A number of studies have been conducted in recent years to investigate the relative potential for different transportation control strategies to stimulate such mode shifts. A Los Angeles study (9) that examined three broad tactics—bus system improvements, car-pooling incentives, and economic disincentives—is most similar to the work reported here. Other efforts (1, 3, 4) have specifically emphasized car-pooling incentives and disincentives. Some studies have relied on more conventional zone-based mode-split models (9), and others have used recently developed disaggregate models calibrated at the household level (1, 3). Another promising approach has used a quantitative marketing model built around consumer preference surveys (4).

The purpose of this paper is to describe the results of a mode-split sensitivity analysis conducted for 13 different transportation control strategies in the San Francisco Bay Area. This work has been carried out for the Metropolitan Transportation Commission (MTC) as a part of the air quality maintenance plan being developed for the region (2). The different transportation control strategies were identified by MTC based on earlier transit, parking management, and air quality planning efforts.

## METHODOLOGY

Forecasts of 1985 mode-split changes that might be stimulated by the different strategies were prepared by using the new set of travel demand models recently developed by MTC (5). Logit-form models of mode choice have been calibrated on disaggregate (household level) data and are stratified by three income categories. Separate models were developed for the trip-making behavior of primary and secondary workers for the home-based work trip. For each of the three income levels, separate models for work-trip mode choice have been developed for each of three modes: transit, drive alone, and shared-ride automobile. Together with other trip generation and distribution models, the overall modeling system is fully compatible with the urban transportation planning system (UTPS) package. Twenty-one different models are included, and UTPS network assignment routes are used.

Only home-based work trips were examined to obtain a sample of origin-destination district interchanges across the region. Five different origin districts were identified, and trip interchanges between these districts and the San Francisco central business district (CBD) as well as an industrial area of Oakland were investigated. These two destinations are meant to generally represent CBD and non-CBD trips respectively. This district-based analysis of representative trips is intended to provide only an approximate picture of mode-choice sensitivities and does not make use of disaggregate, household-level travel data. A random sample enumeration method that uses a considerably larger representative sample of areawide households could provide a more statistically acceptable basis for analysis (1, 3).

Potential changes in mode split for automobile, transit, and shared-ride trips for each of the strategies tested were analyzed parametrically. Projected shifts in mode split for representative trips were converted to an estimated range of impacts at the regional level. This more generalized range of impacts covers work trips primarily oriented toward both CBD and non-CBD destinations. Conversion to shifts in vehicle kilometers of travel and estimates of changes in air pollutant emissions were made subsequently by MTC; a preliminary estimate of the impact of vehicle kilometers of travel is given here.

It was originally intended that the sensitivity analyses described here be carried out through the use of elasticities and cross elasticities (6). These mathematical expressions, which are derived for a specific mode-split forecast and for specific zone-pair interchanges (or household trip records), allow one to estimate what the percentage change in mode split would be for a 1 percent change in a given service characteristic such as travel time and travel cost.

But, because the control strategies tested amounted to major proportional changes in service characteristics (for example, some amounted to more than a 100 percent change in a component service feature such as walking time), and because elasticities must be computed for each of six different mathematical models (primary worker and secondary worker models, each stratified by three income categories), it was decided to develop procedures to compute mode-split changes directly. This more direct approach was made possible because of the flexibility of the new MTCFCAST computer system. Additional computer software was written that will permit a wide variety of additional sensitivity analyses to be conducted. This software is set within the UTPS UMODEL framework and focuses on the series of prespecified representative trips.

In the past, to test major changes in particular service variables, elasticities have been emphasized in sensitivity analyses of this type because of the expense and difficulty of complete runs of the travel demand model system. Elasticities were used instead of additional model runs. However, by using the new procedures developed here, it is possible to analyze mode-split sensitivities more easily and directly in terms of partial model runs (mode-split models only for representative trips only).

## REPRESENTATIVE TRIPS

To determine the effect on vehicle travel that might result from the various transportation control strategies proposed by MTC, the shift in mode split among transit, shared-ride, and single-driver automobile trips must be addressed in an efficient way. Efficiency in this case can be translated into the development of a set of representative trips that reflect the effect on regional travel in a prototypical fashion. The prototype regional trip chosen in terms of travel purpose was the home-based work trip, which represents the travel category most susceptible to modal service influences and also represents the trip purpose that can potentially yield the highest dividends in terms of improved air quality (because of the relation between the work trip and the morning and evening peak traffic hours).

To represent the region in a geographic sense, five origin districts were developed (Figure 1):

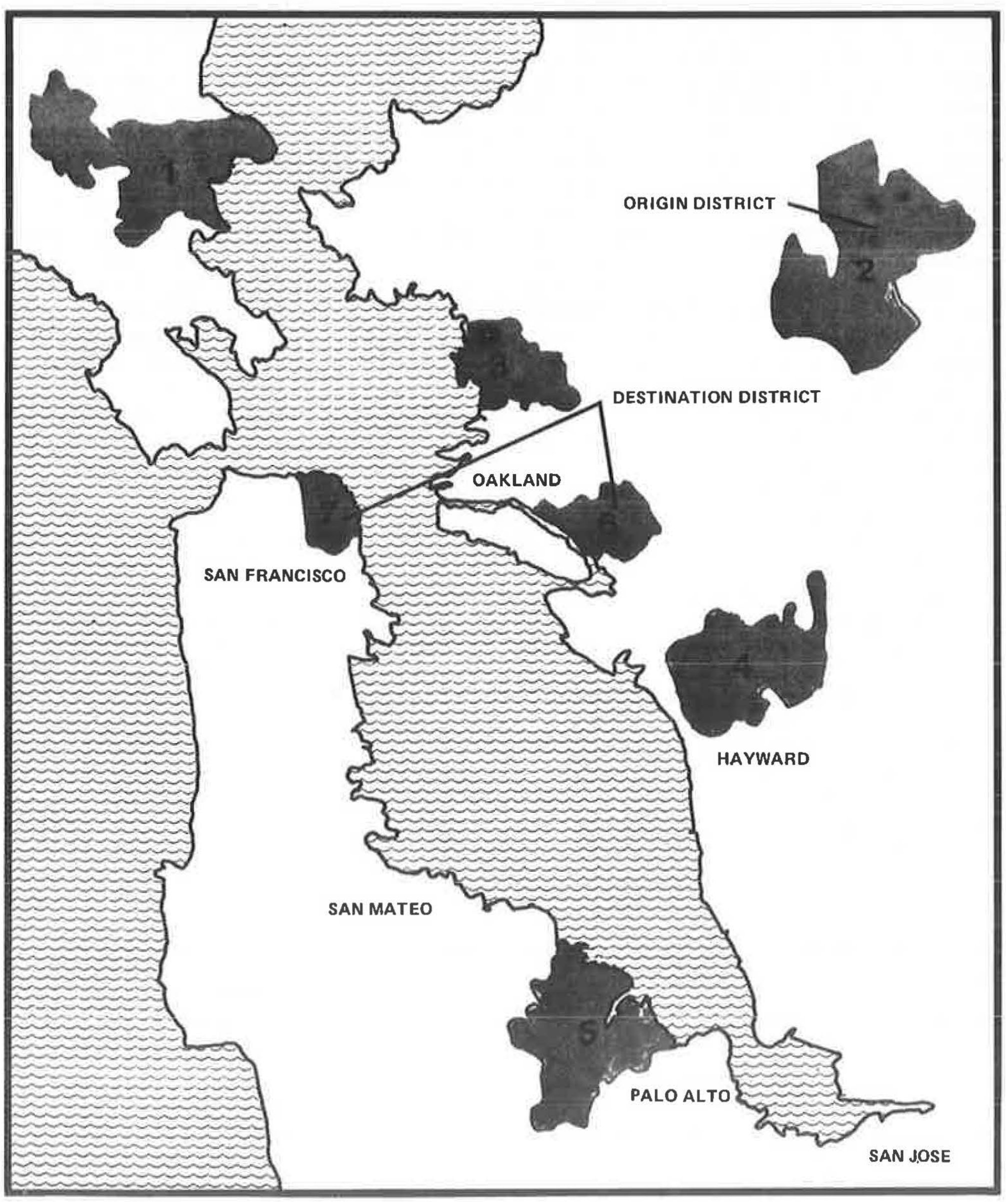
1. District 1: Larkspur, San Rafael—zones 8 through 15;
2. District 2: Concord, Walnut Creek—zones 96 through 104;
3. District 3: Berkeley—zones 125 through 132;
4. District 4: San Leandro, Castro Valley—zones 176 through 184; and
5. District 5: Redwood City—zones 319 through 324.

Two destination districts were developed to represent CBD and non-CBD travel (Figure 1):

1. District 6: South Oakland—zones 156 through 161 (non-CBD); and
2. District 7: San Francisco—zones 382, 383, and 417 through 437 (CBD).

The zonal information required to compute mode

**Figure 1. Representative Bay Area origin and destination districts.**



split was developed at the district level by computing the weighted average of the zonal data. The network-related data were computed on a district-to-district interchange basis based on the weighted aggregation of zone-based highway and transit network data. All weighting was done on the basis of total work trips per zone or trip interchanges per zone pair.

## TRANSPORTATION CONTROL STRATEGIES

The transportation control strategies that were tested and the means of representing each strategy are given in Table 1.

## RESULTS FOR SINGLE-STRATEGY TESTS

A computer program was written to print out the productions and attractions for each of the seven (five origin and two destination) districts for the two modes that relate directly to vehicle use: automobile driver and transit. In this case, automobile driver includes the drivers of the shared-ride mode as well as drive-alone drivers. By analyzing the modes in this way, the effect on the key factors that relate to air pollution and energy consumption (number of vehicles and vehicle type) can be directly determined.

**Table 1. Transportation control strategies.**

| Strategy | Means of Representing Strategy |
|---|---|
| Limit on number of parking spaces | Increase walking time at destination end of trip by 5, 10, and 15 min |
| Preferential parking for car pools | Increase walking time at destination end of trip by 5 min for drive-alone only |
| Automobile-free zone | Increase walking time by 5 min for automobile modes at destination end of trip for zones 419 through 423, 429, and 430 (minidistrict within San Francisco CBD) |
| Parking cost increases | Increase daily parking cost by $1.00 and $2.00/d |
| Limit on long-term parking | Increase daily parking cost by $3.00/d |
| Area license | Increase automobile operating cost by 50 and 100 percent |
| Gasoline tax | Increase automobile operating cost by 50 and 100 percent |
| Toll increase | Increase bridge tolls by 100 percent |
| Parking cost incentive for car pools | Use shared-ride parking cost of zero and shared-ride parking cost of zero with drive-alone parking cost increase of $2.00/d |
| Toll reduction for car pools | Use zero bridge toll for shared ride |
| Additional transit service | Double service (halved headways) by reducing initial wait and transfer wait by factor of 50 percent |
| Bus lanes with ramp metering | Reduce transit in-vehicle times by 10, 20, and 30 percent |
| Paratransit alternatives | Not tested because of extensive network modifications required |
| Car-pool lanes with ramp metering | Reduce highway travel time for shared-ride mode by 10, 20, and 30 percent |

**Table 2. Effect of strategies on mode choice.**

| | Change in CBD Trips (%) | | | Change in Non-CBD Trips (%) | | |
|---|---|---|---|---|---|---|
| Strategy | Number of Automobiles | Transit Ridership | Shared Ride | Number of Automobiles | Transit Ridership | Shared Ride |
| Limit on number of parking spaces | | | | | | |
| Walking time increased by 5 min | -7 | +23 | -6 | -3 | +25 | -3 |
| Walking time increased by 10 min | -14 | +47 | -13 | -7 | +56 | -6 |
| Walking time increased by 15 min | -22 | +75 | -21 | -12 | +104 | -18 |
| Preferential parking for car pools (drive-alone walking time increased by 5 min) | -9 | +10 | 14 | -6 | +18 | +19 |
| Automobile-free zone (walking time for affected zones increased by 5 min) | -3 | +12 | -4 | None | None | None |
| Increase in parking cost | | | | | | |
| $1.00/d | -5 | +7 | +5 | -3 | +14 | +5 |
| $2.00/d | -9 | +16 | +6 | -6 | +29 | +11 |
| Limit on long-term parking (parking cost increase of $3.00/d) | -13 | +28 | +5 | -9 | +47 | +14 |
| Area license and gasoline tax | | | | | | |
| Automobile operating cost increased by 50 percent | -1 | +1 | +1 | None | +2 | -1 |
| Automobile operating cost increased by 100 percent | -1 | +2 | +1 | -1 | +6 | -2 |
| Toll increase of 100 percent | -6 | +11 | +4 | None | None | None |
| Parking cost incentive for car pools | | | | | | |
| Zero shared-ride parking cost | -2 | -1 | +9 | None | None | None |
| Zero shared-ride parking cost and drive-alone parking cost increased by $2.00/d | -11 | +1 | +24 | -7 | +2 | +25 |
| Zero toll for car pools | -1 | -4 | +6 | None | None | None |
| Additional transit service (all headways cut in half and waiting time reduced by 50 percent) | -16 | +53 | -15 | -12 | +94 | -11 |
| Bus lanes with ramp metering | | | | | | |
| In-vehicle transit time reduced by 10 percent | -2 | +4 | None | None | +3 | -3 |
| In-vehicle transit time reduced by 20 percent | -3 | +9 | -1 | -1 | +8 | -2 |
| In-vehicle transit time reduced by 30 percent | -5 | +16 | -4 | -1 | +9 | None |
| Car-pool lanes with ramp metering | | | | | | |
| Shared-ride highway travel time reduced by 10 percent | -1 | -3 | +6 | None | -2 | +3 |
| Shared-ride highway travel time reduced by 20 percent | -2 | -6 | +12 | -1 | -4 | +4 |
| Shared-ride highway travel time reduced by 30 percent | -3 | -8 | +17 | -1 | -6 | +13 |

Person-trip attractions and productions were also printed for each district, which enabled the number of "shared riders" to be computed by subtracting transit and automobile driver from total person trips. For each of the 13 strategies that were individually tested, a base case and the modified situation as well as the differences in productions and attractions computed for each district are printed. Trip interchanges among the districts are also printed. All results are computed by summing the primary and secondary work trips for all three income levels. The actual equations used the number 18, i.e., 2 (mode choice models) × 3 (income levels) × 3 (modes) = 18.

The strategies described above range in individual (single-strategy) effect from reducing by less than 1 percent the number of automobiles used for work trips to a potential reduction in work-trip automobiles of greater than 20 percent. Under the general heading of incentives versus disincentives, the strategies can be categorized as follows:

1. Cost increases versus cost reductions,
2. Service improvements versus service reductions, and
3. Car-pool incentives versus transit incentives.

Exactly which strategy or combination of strategies or strategy category is eventually chosen will depend on a number of factors that include political and public acceptance and the degree of air quality improvement required. The purpose of this initial analysis was to point out the potential of each individual strategy for causing a shift in mode choice.

The range of potential decreases in the number of work-trip automobiles attributed to each strategy is given in the table below:

| | Reduction in Number of Work-Trip Automobiles (%) | |
|---|---|---|
| Strategy | CBD Travel | Non-CBD Travel |
| 1 | 7-22 | 3-12 |
| 2 | 9 | 6 |
| 3 | 3 | NA |
| 4 | 5-9 | 3-5 |
| 5 | ≥13 | ≥9 |
| 6 | ≥1 | 0-1 |
| 7 | ≥1 | 0-1 |
| 8 | ≥6 | No effect |

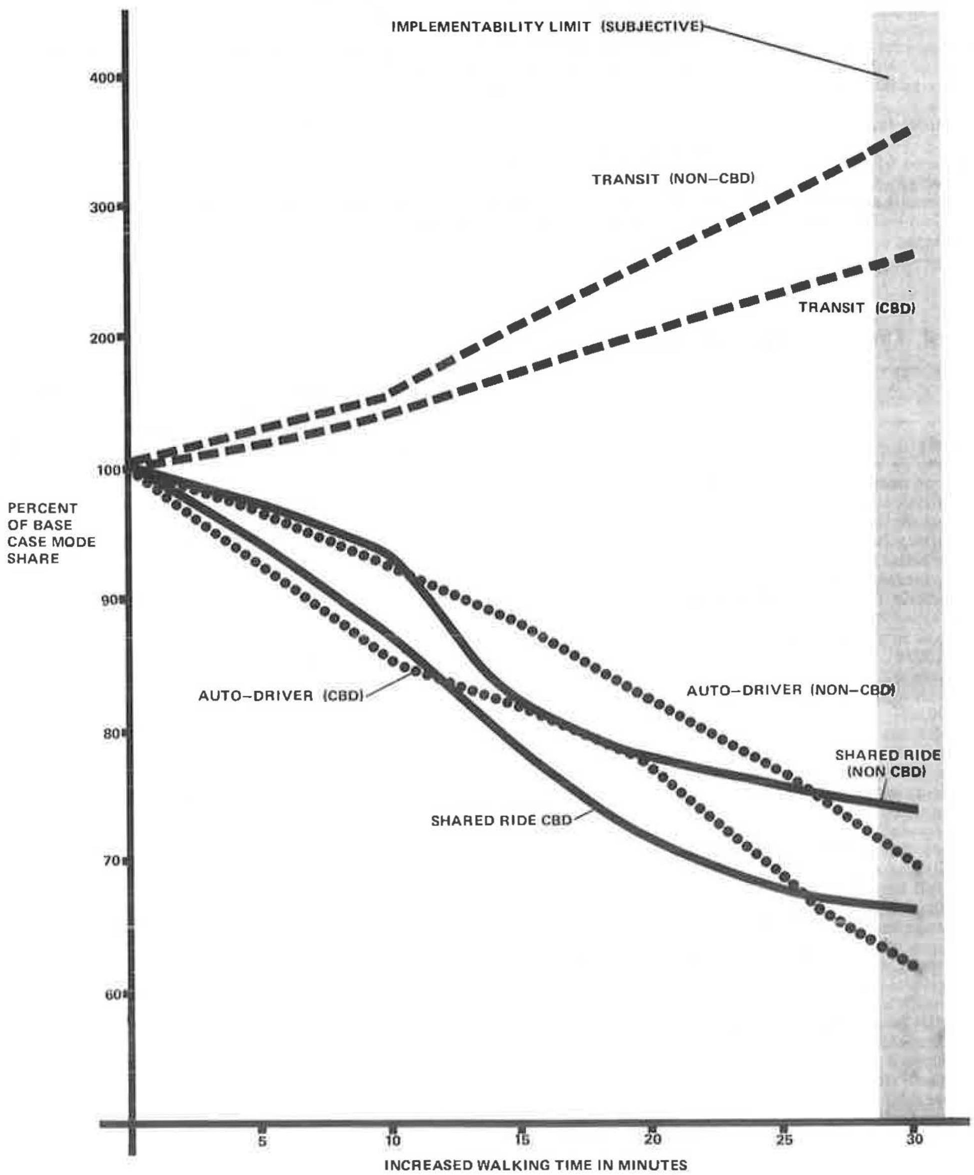



**Figure 2. Percentage mode shift versus reduced availability of parking (increased walking time).**

| Strategy | Reduction in Number of Work-Trip Automobiles (%) | |
|---|---|---|
| | CBD Travel | Non-CBD Travel |
| 9 | 2-11 | 0-7 |
| 10 | ≥1 | No effect |
| 11 | ≥16 | ≥12 |
| 12 | 2-5 | 0-1 |
| 13 | Not tested | Not tested |
| 14 | 1-3 | 0-1 |

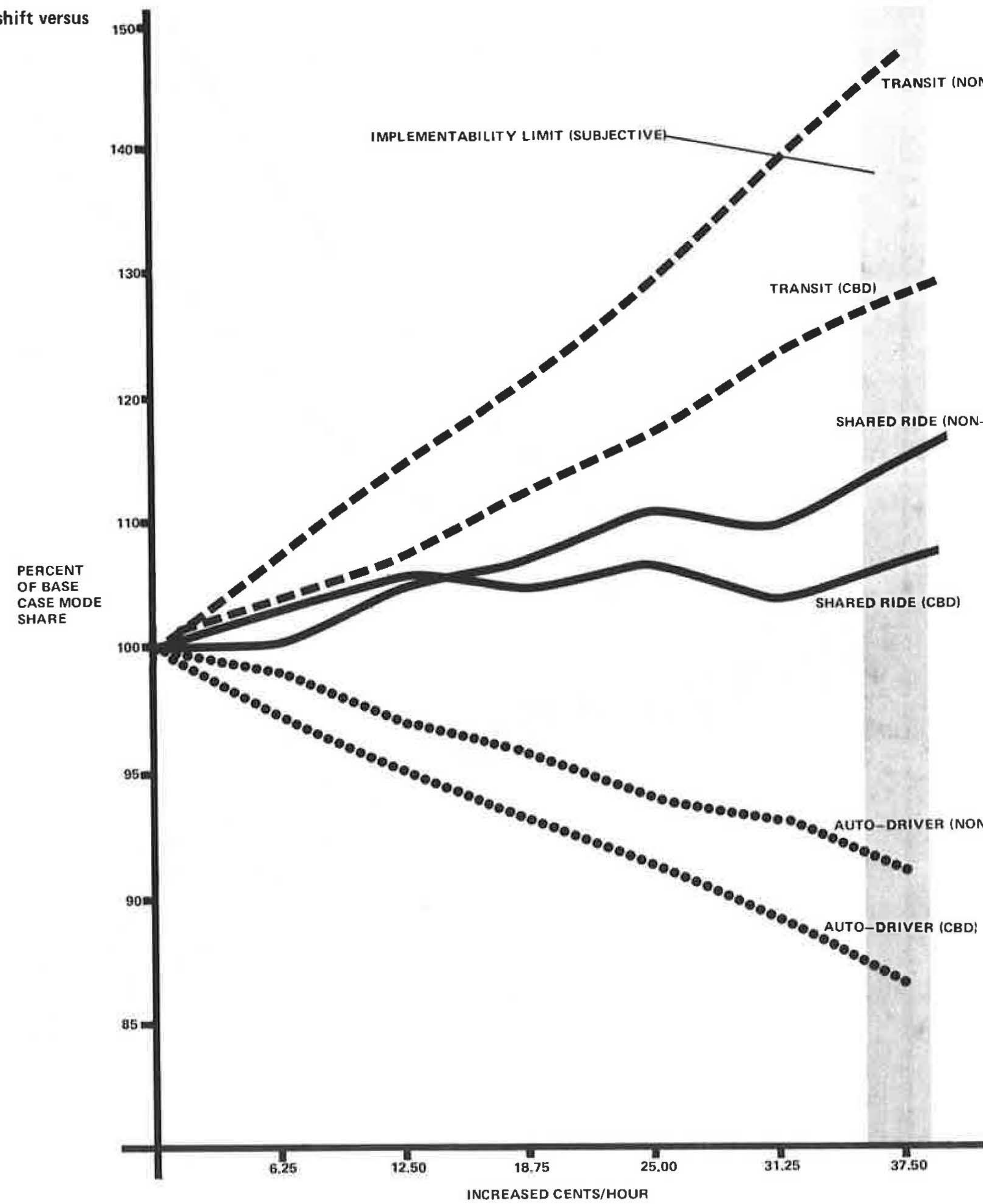
This table represents results based on the range of variable modifications used in the analysis; the exact value for each affected variable (e.g., walking time) will, of course, affect the shift in mode split. The ranges displayed therefore represent a comparative evaluation of those strategies that could potentially have the greatest effect on automobile use. The analysis is broken down by CBD and non-CBD trip orientation to show the varying effect of the strategies on different parts of the region.

For each individual transportation control strategy, the decrease in work-trip automobiles is reflected by an increase in transit or shared ride or both. Table 2 gives a detailed picture of the effect on each mode for each strategy and each policy level tested within that strategy.

In summary, the strategies that have the highest potential for reducing the number of work-trip automobiles making single-driver work trips are (a) limiting the number of parking spaces, (b) increasing transit service, and (c) increasing parking cost. The strategies that have the greatest potential for increasing transit use are the same because the competitive position of transit is strengthened the most.

The strategies that have the greatest potential for increasing car pooling are (a) reducing shared-ride parking cost and increasing the parking cost for driving alone, (b) reducing shared-ride travel time through exclusive car-pool lanes and ramp metering, and (c) preferential parking for car pools.

Figure 3. Percentage mode shift versus increased parking cost.



COMPARISON OF STRATEGIES

The computer program written for these analyses has the capability to systematically test prespecified incremental shifts in each control strategy (e.g., successive 3-, 4-, or 5-min increases in walking time from parking facilities). By using this capability, it was possible to test a wider range and number of policy levels for each strategy beyond the specific levels for which testing was requested. For 9 of the 13 control strategies, these systematic incremental variations were examined and plotted in the form of a graph. Figures 2, 3, and 4 show the percentage changes in mode share obtained for each of the three modes and for CBD and non-CBD work trips for the three most promising strategies: limiting the number of parking spaces, increasing transit service, and increasing parking cost.

The maximum level for each control strategy shown in Figures 2, 3, and 4 was treated as an "implementability limit," defined purely for further analysis purposes. For eight of the nine control strategies examined in this way, this implementability limit is on the order of 1.5 to 2.0 times greater than the most ambitious policy levels tested in Table 2. The exception is increased parking costs, for which it was felt that previously tested levels were probably already at maximum feasible limits. Because all of the remaining implementability limits would pose serious implementation problems from a political standpoint, these broader

Figure 4. Percentage mode shift versus reduced transit headways.

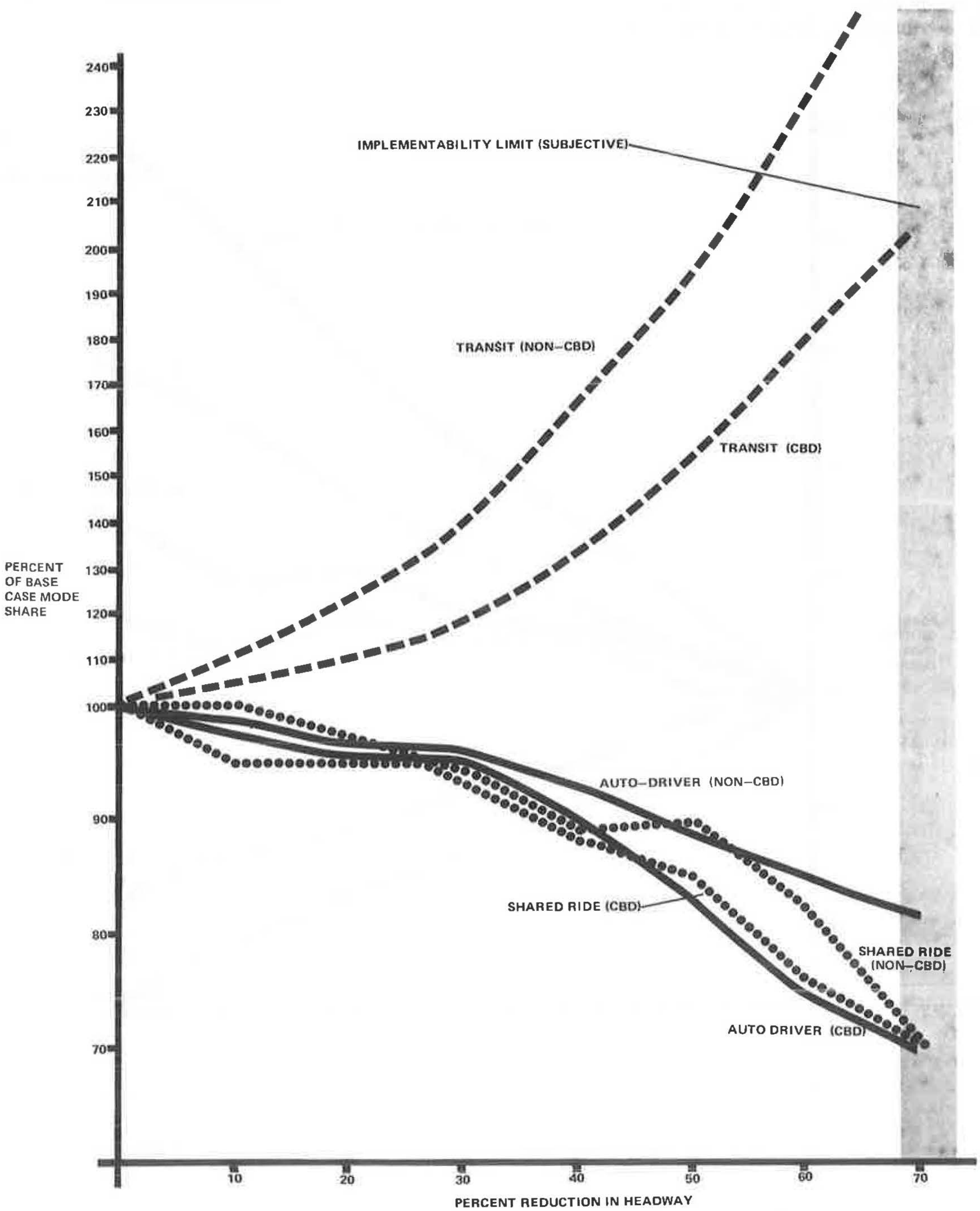


Table 3. Comparative index of effectiveness of strategies.

| Strategy | CBD-Oriented Work Trips: Automobile Driver | CBD-Oriented Work Trips: Transit | CBD-Oriented Work Trips: Shared Ride | Non-CBD-Oriented Work Trips: Automobile Driver | Non-CBD-Oriented Work Trips: Transit | Non-CBD-Oriented Work Trips: Shared Ride |
|---|---|---|---|---|---|---|
| Reduce parking availability | -0.38 | +1.60 | -0.33 | -0.30 | +2.50 | -0.24 |
| Preferential parking for shared ride | -0.18 | +0.20 | +0.24 | -0.13 | +0.34 | +0.42 |
| Increase parking cost | -0.14 | +0.27 | +0.06 | -0.08 | +0.46 | +0.14 |
| Increase automobile operating cost | -0.01 | +0.06 | +0.05 | -0.02 | +0.09 | +0.04 |
| Toll increase | -0.19 | +0.18 | +0.03 | - | - | - |
| Parking cost differential | -0.13 | +0.13 | +0.35 | -0.15 | +0.38 | +0.51 |
| Reduce transit headways | -0.32 | +1.05 | -0.29 | -0.20 | +1.50 | -0.28 |
| Improve transit operating speed | -0.18 | +0.30 | -0.11 | -0.03 | +0.20 | -0.02 |
| Shared-ride lanes and ramp metering | -0.06 | -0.20 | +0.41 | -0.03 | -0.15 | +0.39 |

feasibility ranges are estimated in a purely technical or engineering sense.

By establishing judgmentally this broader theoretical range for each variable tested for the various control strategies, it was possible to compare them more consistently. These larger ranges were assumed to be approximately equal in terms of technical implementability, and a normalized index of potential impact was computed. This index reflects the percentage of change in mode use for an equivalent "1 percent of implementability" change in the variable being tested. This admittedly judgmental index number allows a generalized view of strategy effectiveness to be developed that would not be possible otherwise because the range of previously tested values for one variable may involve a 200 percent change (e.g., parking cost) or only a 20 percent change (e.g., reduction in transit travel time). By normalizing over the implementable range, a more accurate (though generalized) comparison can be made. Table 3 gives these index values for each combination of mode and strategy.

For CBD travel, the control strategy that will decrease single-driver automobiles to the greatest extent is still to reduce the number of parking spaces available and thereby increase walking distance and hence walking time for automobile drivers and passengers. The second most effective control strategy in reducing the number of single-driver automobiles for CBD travel is to improve transit service by cutting transit headways. It is significant, however, that, because of major implementation difficulties that have emerged in the development of possible parking management plans for different portions of the region, increasing parking cost shows much less potential effectiveness than was initially observed. The third most effective means of reducing the number of automobiles in the downtown area is reflected by either of two individual strategies: (a) providing preferential parking for shared-ride vehicles or (b) improving transit service by providing preferential treatment for transit vehicles, which would reduce in-vehicle transit travel time.

For non-CBD travel, the strategies that will reduce the number of work-trip automobiles are essentially the same: (a) reduce parking, (b) improve transit service by decreasing headways, (c) provide a preferential parking cost for shared ride, and (d) provide preferential parking locations for shared-ride vehicles.

## RESULTS FOR COMBINATION STRATEGY TESTS

Four combination strategies were developed and tested for their effectiveness in terms of realizing potential mode shifts and hence reductions in vehicle kilometers of travel in the region. The four combination transportation control strategies are given in Table 4. Table 5 indicates the relative change in mode use that could result from each of these combination strategies.

The four strategies were developed to represent a generally realistic set of transportation control options. Results given in Table 5 suggest that the option that combines factors of both travel time and cost has an expected cumulative effect in terms of reducing the number of automobiles. That is, the reduction in single-driver automobiles for the time-and-cost option is very nearly the sum of the reductions in the number of automobiles for the strategies that emphasize travel time and cost. This additive relation does not hold, however, for transit ridership or for car pooling. Significantly higher increases in car pooling could be achieved under strategy 3, whereas increases in transit ridership under this strategy are only slightly greater than those obtained under the travel-time strategy alone.

Table 5 also indicates that very little if any improvement results for CBD-oriented trips from implementing the maximum-effort strategy versus the combination time-and-cost strategy. This suggests that the assumed policy levels that go into the time-and-cost transportation control strategy already realize the upper limit on mode-split shifts that can be expected for CBD-oriented travel. For non-CBD trips, however, the maximum effort does realize greater decreases in the number of automobiles than does the combination time-and-cost control strategy, which indicates that room still exists for further reduction in vehicle travel to non-CBD areas perhaps beyond that achieved by the time-and-cost and maximum-effort strategies.

For CBD travel, a reduction of 27 percent in the number of automobiles used for the work trip could potentially be accomplished by a realistic transportation control strategy that involved both time and cost factors. In terms of impact on total regional travel, it has been found that the work trip accounts for approximately 23 percent of the total trips made in the region but involves approximately 33 percent of the vehicle kilometers of travel. Specifically, the CBD work trip represents about 15 percent of total work trips. Therefore, a 27 percent reduction in the number of automobiles for CBD-oriented work trips would result in a commensurate decrease of 1 percent in regional vehicle kilometers of travel. If similar computational logic is used for the non-CBD-oriented work trip, a reduction in regional vehicle kilometers of travel of 2.5 percent could potentially be realized.

Therefore, transportation control strategies aimed at the work trip, which is, of course, a primary target for such control measures, could realize a total potential decrease in regional vehicle kilometers of travel of approximately 3.5 percent. This impact is applicable for a combined control strategy that involves both time and cost factors over what is considered to be initially a relatively realistic range of control. Greater decreases in vehicle kilometers of travel may be realized by developing other combination strategies based on the sensitivity analyses (and further sensitivity analysis capabilities) developed in the project.

Table 4. Combination transportation control strategies.

| Strategy | Type | Combination of Measures |
|---|---|---|
| 1 | Travel time | Increase drive-alone travel time by 20 percent<br>Decrease transit in-vehicle time by 10 percent<br>Decrease shared-ride travel time by 10 percent<br>Increase walking time at destination for drive alone by 200 percent<br>Reduce initial transfer and transit wait time by 10 percent |
| 2 | Travel cost | Increase drive-alone parking cost by $1.00/d<br>Decrease shared-ride parking cost by 80 percent<br>Reduce shared-ride toll cost to zero<br>Increase drive-alone toll by 150 percent |
| 3 | Time and cost | Combine all policy levels used in strategies 1 and 2 |
| 4 | Maximum effort | Decrease transit in-vehicle travel time by 20 percent<br>Decrease shared-ride travel time by 20 percent<br>Increase drive-alone travel time by 20 percent<br>Increase walking time at destination for drive alone by 500 percent<br>Decrease initial and transfer transit waiting time by 20 percent<br>Decrease shared-ride parking cost by 80 percent and increase drive-alone parking cost by $2.00<br>Reduce shared-ride toll to zero<br>Increase drive-alone toll to $2.00 |

Table 5. Potential effect of combination strategies on mode choice.

| Combination Strategy | Change in CBD Trips (%) Number of Automobiles | Transit Ridership | Shared Ride | Change in Non-CBD Trips (%) Number of Automobiles | Transit Ridership | Shared Ride |
|---|---|---|---|---|---|---|
| Travel time | -14 | +22 | +14 | -5 | +19 | +14 |
| Travel cost | -13 | +11 | +20 | -3 | +10 | +12 |
| Time and cost | -27 | +25 | +42 | -9 | +32 | +24 |
| Maximum effort | -27 | +28 | +39 | -18 | +68 | +48 |

The transportation control strategies examined here will, of course, also have some impact on nonwork travel. Though nonwork travel was not examined explicitly in the study, other work (1) suggests that those policies that tend to reduce the number of single-driver work trips also tend to increase the number of nonwork trips. In the short run, at least, this is because the automobile normally used for the single-driver work trip (for those who shift to transit or car pooling) is now available at home for use by other family members. These other family members tend to make additional discretionary or nonwork trips.

For example, in Washington, D.C. (1), it was found that a daily parking cost increase of $3.00 would reduce the number of drive-alone work trips by 15.6 percent and the number of work-trip vehicle kilometers by 10.2 percent. However, nonwork vehicle kilometers would increase by 2.3 percent for a combined net impact on total vehicle-kilometer reductions of 2.5 percent. Similar balancing impacts might be expected in other urban areas so that the 3.5 percent reduction in Bay Area regional vehicle kilometers of travel because of work-trip mode shifts might be offset somewhat by accompanying increases in nonwork travel.

## CONCLUSIONS

Several general conclusions can be drawn from these analyses:

1. Although the various transportation control strategies examined here, either singly or in combination, could reduce the number of automobiles used for work trips by about 20 percent or more, the potential impact on total regional vehicle kilometers of travel is considerably lower—perhaps a 5 percent reduction or less. This is consistent with the findings of other studies.

2. The most promising single transportation control strategy involves limiting the number of close-in parking spaces available for work trips at both CBD and non-CBD destinations. Such a strategy could significantly increase peak-hour transit ridership although, without selective treatment for car pooling, car pooling might also be reduced.

3. The second most promising single transportation control strategy involves additional improvements in transit service, which are reflected in major reductions in waiting time. For significant impact, such a strategy would call for major capital and operating investments in additional transit equipment, labor, and operating schedules.

4. Although increased parking cost also shows significant potential for inducing mode shifts, the levels of increase necessary to achieve a major impact on work trips appear to face serious difficulties of implementation. When these difficulties are considered, this control strategy appears to be of less potential effectiveness.

5. An upper limit on potential mode-shift impact for CBD work trips for various combination transportation control strategies may lie in the area of 25 to 27 percent reduction in single-driver automobile use. For non-CBD work trips, the existence of such an upper limit is less clear; it could exceed an 18 to 20 percent reduction in single-driver automobile use.

6. A series of representative CBD and non-CBD work trips does provide an efficient method for analyzing the sensitivity of a wide range of transportation control strategies. Emphasizing work-trip analysis appears to be an adequate means for drawing general conclusions regarding the impact on overall regional travel and potential reductions in vehicle kilometers of travel.

7. The additional computer software developed in association with the UTPS UMODEL program provides considerable flexibility for further sensitivity analyses and permits a wide variety of combination control strategies to be tested. This software has been incorporated in the MTC travel demand forecasting system.

## ACKNOWLEDGMENTS

Significant contributions to the analyses reported here were made by Bill Davidson, Marty Inouye, Larry Lavery, and Byron Riffe, all colleagues at Barton-Aschman Associates. The assistance of Hanna Kollo and Vince Petrites of MTC is also gratefully acknowledged.

REFERENCES

1. T. J. Atherton, J. H. Suhrbier, and W. A. Jessiman. Use of Disaggregate Travel Demand Models to Analyze Car-Pooling Policy Incentives. TRB, Transportation Research Record 599, 1976, pp. 35-40.
2. Sensitivity Analysis of Selected Transportation Control Measures: Potential Reductions in Regional Vehicle Miles of Travel. Barton-Aschman Associates and Metropolitan Transportation Commission, Berkeley, CA, July-Aug. 1977.
3. M. Ben-Akiva and T. J. Atherton. Methodology for Short-Range Travel Demand Predictions: Analysis of Car Pooling Incentives. Journal of Transport Economics and Policy, Jan. 1977.
4. J. M. Bruggerman, R. B. Rubin, and F. Griffiths. Findings of a Study to Estimate the Effectiveness of Proposed Car-Pool-Incentive Policies. TRB, Transportation Research Record 650, 1977, pp. 36-42.
5. MTCFCAST Travel Demand Forecasting System. Comsis Corporation, Cambridge Systematics, and Barton-Aschman Associates; Metropolitan Transportation Commission, Berkeley, CA, Aug. 1977.
6. T. A. Domencich, G. Kraft, and J. P. Valette. Estimation of Urban Passenger Travel Behavior: An Economic Demand Model. HRB, Highway Research Record 238, 1968, pp. 64-78.
7. G. K. Ingram. Reductions in Automobile Use in Four Major Cities as a Result of Car Pooling and Improved Transit. TRB, Transportation Research Record 650, 1977, pp. 59-65.
8. J. B. Margolin, M. R. Misch, and R. D. Dobson. Incentives and Disincentives to Ride-Sharing Behavior: A Progress Report. TRB, Transportation Research Record 592, 1976, pp. 41-44.
9. W. T. Mikolowsky, W. L. Stanley, and B. F. Goeller. Effectiveness of Near-Term Tactics in Reducing Vehicle Miles (Kilometers) of Travel: Case Study of Los Angeles Region. TRB, Transportation Research Record 574, 1976, pp. 1-14.

*Publication of this paper sponsored by Committee on Transportation Systems Design.*

# Evaluation of Road and Transit System Requirements for Alternative Urban Forms

R. G. Rice, Department of Civil Engineering, University of Toronto

**Research was performed for the purpose of evaluating the road and transit system requirements of a range of cities that have different density and spatial patterns and thereby assessing the effects of varying urban forms on transportation investment and service measures. The assessment is conducted in the context of a proposed policy evaluation framework that uses the end-state transportation and land-use plan for policy guidance and the time stream of benefits and costs as the object of evaluation. For the analysis of the transportation implications of a number of urban forms, a two-mode network generation model is developed and applied to six hypothetical city types of 2 million population. The comparison of the transportation requirements for these urban forms indicates a range of transit use among the city types of from 8 to 34 percent and wide differences in the need for high-capacity service routes. In terms of person hours of travel and mean trip length, the multicentered city in particular and the centrally oriented cities in general have the lowest requirements. These conclusions have important implications for the use of horizon-year transportation and land-use plans within the proposed framework of dynamic evaluation.**

The current emphasis in urban planning on a more comprehensive and open process and an orientation to strategic choices among a wider range of alternative policies has placed new demands on the transportation planner. In spite of the ability of the transportation planner to simulate travel demands on a complex multimodal network, there is a notable lack of success in responding to currently relevant planning issues and alternatives. In general, existing transportation planning models are expensive and cumbersome to use and difficult to interpret, and their attention to the detail of system performance is too restrictive in scope (1, 2).

Perhaps the best way to exemplify the deficiencies of current transportation modeling is to define a comprehensive three-level hierarchical structure of the planning process against which present achievements can be compared (3, 4). The three components of the structure are defined as follows:

1. Policy planning—Policy planning is a political exercise concerned with the broad issues of urban development and the authority and constraints relating to the resolution of these issues within a public forum. This process provides the contextual setting within within which transportation system alternatives may then be assessed.
2. Systems planning—For transportation planners, this process is concerned with the analysis and evaluation of multimodal networks that consist of major transportation facilities and associated terminals from the point of view of location, operation, and regulation.
3. Project planning and programming—The third level of the structure involves those activities required to design and implement a particular component or link of the system plan and includes engineering design, right-of-way acquisition, and capital programming and budgeting.

The recent primary efforts of the transportation researcher have been to refine the planning model programs and packages available at the second level, presumably to improve the reliability of the model output for subsequent use in design at the third level. As a consequence, the linkage between the second and third level is very satisfactory, but the linkage upward to the highest level is almost totally neglected. Transportation system models are unresponsive to policy issues that relate to alternatives for land-use development, environmental and socioeconomic impacts, energy consumption, and effects of income distribution. This results in a relation between policy and systems planning that tends to be so ad hoc that systems planning often completely fails to reflect political and public views.

In summary, there is a need for analytical procedures or models at the policy planning level that can both provide a set of appropriate constraints for input to the systems planning process and in return accept corresponding system performance measures. Such a policy model would have to be easy to use and interpret, have the ability to accept a comprehensive set of development alternatives, and be capable of estimating a broad set of interrelated measures and indicators.

It is the purpose of this paper to describe a proposed framework for policy planning and to present a series of analyses that represent an important initial component of this framework. The paper also briefly defines the proposed framework in abstract terms and then describes a numerical experiment to assess the effects of varying urban forms on transportation investment and service requirements. Finally, the relevance of the experiment to the policy framework is discussed.

## PROPOSED FRAMEWORK

One of the essential differences between the policy planning and systems planning levels of the hierarchical structure already defined is the relevant time dimension. Although the transportation planner is concerned with a 20- to 25-year planning horizon, political decision makers at the policy level are primarily interested in short-term issues. The task of the urban transportation planner, then, might be defined as that of providing technical advice on short-range problems in such a way that consistency is maintained with regard to long-range intentions. Somehow the planner must understand the relation between immediate choices and evolving end-state alternatives. This relation then becomes a central element in establishing an appropriate linkage between policy and systems planning and a central element of the policy framework to be proposed.

A second characteristic that influences the proposed framework has to do with the approach the transportation planner follows in determining the most appropriate end-state plan. Typically, a static comparison is made of a series of horizon-year plans, and a "best" plan is chosen on the basis of the accommodation of future travel demand at an acceptable level of service. This is a rather limited perspective, however, since the interrelationship between transportation service and land use (an important policy issue) is a very dynamic one. Whether particular route links are built early or late in the planning period has much to do with the location and rate of urban growth, and whether the improvement of the road network is emphasized before or after the transit system is completed influences both travel demand and automobile ownership levels as well as location of residence and employment.

The implications for both policy and systems planning, then, lie in the requirement for a process that makes the time stream of benefits and costs rather than the end-state condition achieved the object of evaluation (5, 6, 7). The end-state transportation plan may serve as a general policy guide, but it is evaluated only in the light of the sequence of actions that it takes to achieve that state (if, in fact, the state is achievable or feasible).

Interestingly enough, this need for a dynamic evaluation structure parallels and satisfies the first requirement for a linkage between short-range decisions and long-range alternatives (1). To define this proposed structure or framework, let us assume initially that a statement of objectives is available so that the basis for the policy evaluation structure is given. These objectives may be translated into a list of measures of effectiveness that may be defined as $x_i$, where i is an index $i = 1, 2, \ldots, r$. Since the time stream of these r measures is of central interest, a vector $X^t$ must be defined at any time t:

$$x^t = (x_i^t) \quad i = 1,2,\ldots,r \tag{1}$$

and a matrix X must be defined for the full planning period:

$$X = (x_i^t) \quad i = 1,2,\ldots,r \text{ and } t = 1,2,\ldots,T$$

$$= \begin{bmatrix} x_1^1 & x_1^2 & \cdots & x_1^T \\ x_2^1 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ x_r^1 & \cdots & \cdots & x_r^T \end{bmatrix} \tag{2}$$

In essence, this matrix becomes the focus for the dynamic policy evaluation process. It is perhaps of some value at this point to compare this evaluation base with that typically used by transportation planners. Existing systems procedures that relate to both modeling and evaluation are concerned with the end-state condition:

$$X^T = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_r^T \end{bmatrix} \tag{3}$$

and the comparison of the estimated measures of performance for a number of these conditions: $X^{T'}$, $X^{T''}$, $X^{T'''}$, and so on. These alternatives might represent a road-oriented network, a rail transit-oriented network, and a surface transit-oriented network, and the measures of effectiveness $(x_i^T)$ might be average travel time by mode and socioeconomic group, average travel cost, and volume-capacity ratios by link. These measures are estimated as a function of socioeconomic variables, travel behavior parameters, and transportation system characteristics at time T, in accordance with the calibration of an identical function at the existing point in time.

The proposal for a policy evaluation process, on the other hand, requires that the full X matrix be estimated and evaluated. Since the number of measures of effectiveness may be expected to be greater, reflecting the increased scope of policy issues, the requirements for modeling will be both different and more demanding than those described for systems planning. Not only is it necessary to relate the performance and impact measures to socioeconomic factors, travel behavior, and transportation system variables but also the relations between the dependent measures, or $x_i$'s (such as transportation service and land-use change) become very important. Thus, consideration of the requirements of the policy planning level in the hierarchical structure proposed above results in more than a simple extension of the techniques currently available at the systems planning level. Although the primary difference has to do with the incorporation of the time dimension, the implications of this change for both the type of forecasting model and the mode of application of the model are obviously substantial.

In summary, policy planning is very much an exercise in evaluating alternative paths or time sequences of actions that lead to a desirable end-state condition. From an efficiency point of view, it is likely that the evaluation of planning actions through time will still require the definition of an end-state or boundary condition. As Rice and Nowlan (1, p. 102) have observed,

> It will always be computationally efficient to pre-select one or more specific terminal-year structural configurations, so that feasible paths of urban change will have a beginning point in the present structure and a terminal point or points in the pre-determined terminal-year structure. It is in general possible to optimize an objective function without specifying a terminal-year structure, but the increased computation ... makes the exercise much messier. More importantly, however, the terminal-year structure, because it functions both as legacy and target, is interesting in its own right and worthy of separate analysis.

It is the primary purpose of the following section of this paper to initiate this type of separate analysis by assessing the two-mode transportation service and investment requirements of a series of alternative static terminal-year urban development conditions. This is very much a policy planning exercise since the terminal-year or end-state structure comes very close to representing a "policy statement." No attempt is made here to assess alternative plans or sequences of actions that would achieve this policy statement or target configuration (6). Instead, an approach for defining the general implications of alternative policy states is described (1, p. 103):

> Its purpose would be to make numerically specific the rather vague notions of policy that are part of current politics. Thus, different policy positions could be specified in quantitative terms. Such competing concepts as intensified central business district development versus nodal business district development, denser suburban populations versus stabilization at present densities, public transit transportation modes versus expressway development may all be defined by means of feasible terminal-year positions.

## RESEARCH PROCEDURE AND ANALYSIS

The objective of the research described here is to define the notion of policy alternatives in terms of their transportation and land-use implications: That is, how might we determine the most effective combinations of road and transit systems to serve a defined number of cities that have different density and spatial patterns? Given the perspective offered in the literature (8, 9), it is apparent that any effort in this area should allow first for full modal interdependence, permitting shifts in travel mode with changes in land-use and socioeconomic characteristics as well as transportation level of service, and second for a greater number and broader type of output indicators. In essence, the estimation of modal travel demands for each urban form must be sensitive to both the level of service supplied and the space- and density-related pattern of land-use activities. To permit the investigation of transportation system characteristics for a range of urban patterns, there is a strong need for an analytical base that allows for the comparison of transportation measures between land-use plans. That is, the procedure for developing the transportation network and mode combination for each urban form must be consistent and not unduly bias any particular city patterns. This, in fact, is the central issue of the research procedure.

The description of the research method and results has been divided into three separate stages: The first deals with the definition of the range of hypothetical urban forms and the estimation of basic peak-hour travel demands, the second with the generation of optimal two-mode transportation networks to accommodate the estimated travel demands in each city, and the third with the comparative analysis of transportation system requirements by type of city.

### Definition of Urban Form and Travel Demand Estimation

Because of the complexities associated with transportation and land-use interrelationships, hypothetical rather than actual urban forms were generated for testing in the research project. This process was quite complex and has been described in detail elsewhere (8, 9). Very simply, it was necessary that the range of hypothetical city types be defined so that they resulted in a significantly broad range of transportation conditions. In total, population and employment characteristics were developed for six different urban forms of 2 million population: (a) central core, (b) homogeneous (uniform density), (c) multicentered, (d) radial corridor, (e) linear, and (f) satellite. These forms were developed through the application of a series of realistic and empirically derived constraints on density variation, socioeconomic characteristics, and relative population and employment location. The six urban forms are shown in Figure 1, and their summary characteristics are given in Table 1.

Definition of the characteristics of urban form made it possible to estimate travel demands by trip purpose. The demand estimation procedure involved the use of the conventional four-stage process of applying zonal trip generation and attraction equations and the gravity model to produce origin-destination trip matrixes for each of the six urban forms. To permit the development and use of unique functions of travel impedance for each city type, reference was made to previous research that related the distribution of work opportunities to average trip length (10).

Finally, work-trip origin-destination matrixes were assigned to spider networks for each type of city. The rest of the analysis of the transportation implications of urban forms was restricted to the work trip on the basis that this trip purpose set the condition for the design of the transportation networks. The results of the spider, or desire-line, assignments for the work trip are given in Table 2. The advantage of the desire-line assignment is that the volume flow condition that has been estimated is not constrained by the form or characteristics (capacity and mode) of the spider network (Figure 2). It therefore provides a relatively objective and consistent base for deriving more comprehensive two-mode transportation networks.

### Generation of Two-Mode Network

As indicated earlier, the critical phase of the research requires that a procedure be developed for generating a unique, two-mode, capacity-restrained transportation network for each city type to permit a realistic and unbiased comparison of mode performance for the six urban forms. In basic terms, this procedure is dependent on a definition of modal balance, which might be stated as the condition in which both mode subsystems are used effectively in and of themselves and in such a manner as to produce collectively optimal total system performance.

The network generation procedure that was applied is a two-stage heuristic process that depends on an initial division of modal service for each network link (the supply equilibrium cycle) and the refinement of modal volumes in accordance with mode and route choice (the demand equilibrium cycle). This procedure is described in Figure 3 in the form of a flow chart.

The supply equilibrium phase starts with the desire-line volumes from the spider assignment for each type of city and defines an initial two-mode transportation network that can accommodate the expected demand.

**Figure 1. Six selected urban forms.**

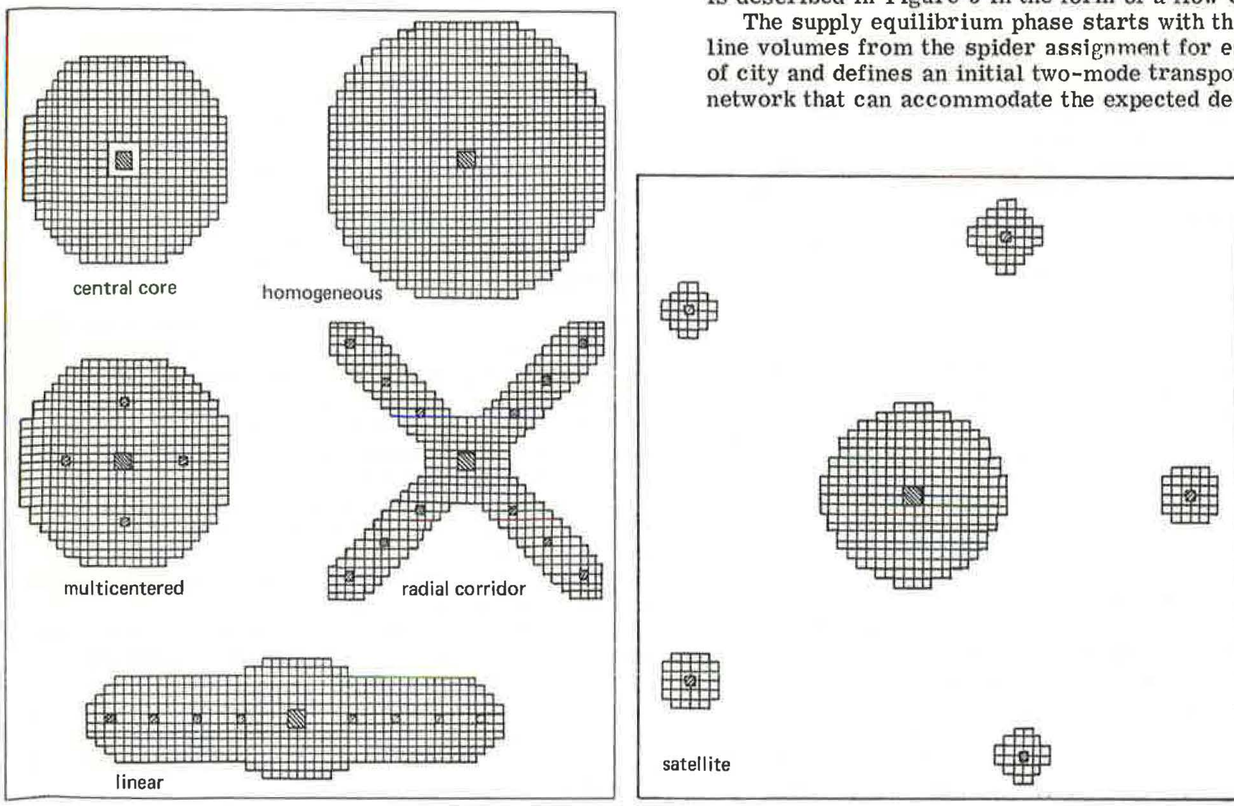



**Table 1. General characteristics of urban forms.**

| Urban Form | Designation | Employment | Developed Area ($km^2$) | Gross Population Density (persons/$km^2$) | Average Net Residential Density (persons/$km^2$) | Number of Dwelling Units | Average Number of Persons per Dwelling Unit |
|---|---|---|---|---|---|---|---|
| Central | 1A | 800 000 | 1240 | 1610 | 11 200 | 565 300 | 3.54 |
| Homogeneous | 1B | 803 000 | 2190 | 910 | 5 200 | 534 300 | 3.74 |
| Multicentered | 2 | 804 700 | 1240 | 1610 | 12 000 | 567 000 | 3.52 |
| Radial corridor | 3 | 802 200 | 1045 | 1920 | 13 200 | 567 000 | 3.52 |
| Linear | 4 | 803 000 | 1270 | 1580 | 11 500 | 566 100 | 3.54 |
| Satellite | 5 | 800 300 | 1140 | 1760 | 12 600 | 567 300 | 3.53 |

Notes: 1 $km^2$ = 0.386 $mile^2$.
All urban forms are for population of 2 million.

**Table 2. Desire-line assignment.**

| | | | | Total Trips | | | | Work Trips | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Urban Form | Total Length of Network (km) | Average Length of Links (km) | Number of Two-Way Links | Average Volume per Link | Maximum Link Volume | Total Person Kilometers | Total Person Hours | Average Volume per Link | Maximum Link Volume | Total Person Kilometers | Total Person Hours |
| 1A | 602.2 | 5.8 | 104 | 32 710 | 104 500 | 32 880 000 | 342 470 | 5465 | 34 600 | 5 440 000 | 56 700 |
| 1B | 739.8 | 6.7 | 108 | 31 360 | 103 000 | 43 250 000 | 450 230 | 4900 | 24 000 | 6 640 000 | 69 200 |
| 2 | 610.2 | 5.1 | 120 | 29 390 | 80 100 | 32 530 000 | 338 000 | 4270 | 21 000 | 4 670 000 | 48 600 |
| 3 | 463.8 | 5.6 | 82 | 50 800 | 170 500 | 37 920 000 | 395 000 | 8530 | 47 100 | 5 970 000 | 62 200 |
| 4 | 604.2 | 5.3 | 114 | 37 970 | 169 000 | 41 650 000 | 433 760 | 6470 | 57 300 | 7 060 000 | 73 500 |
| 5 | 955.2 | 7.8 | 122 | 30 520 | 107 300 | 48 540 000 | 505 640 | 5130 | 28 800 | 9 180 000 | 95 700 |

Notes: 1 km = 0.62 mile.
For the desire-line analysis, a speed of 95 km/h (60 mph) was assumed.

Figure 2. Spider network for satellite city.

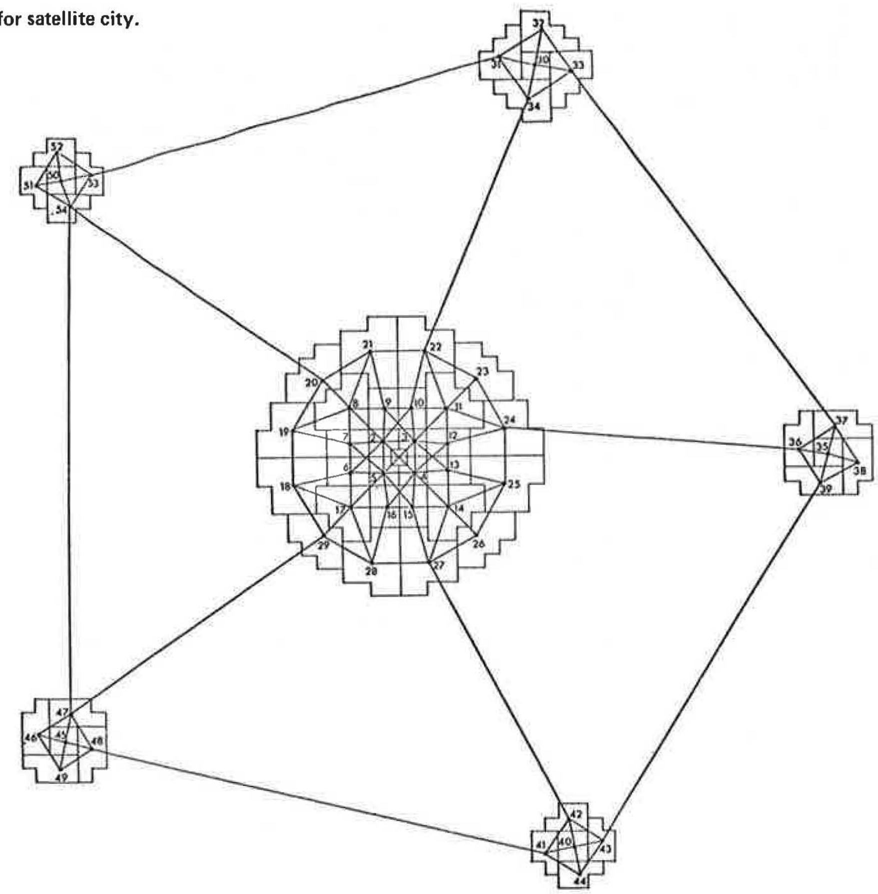
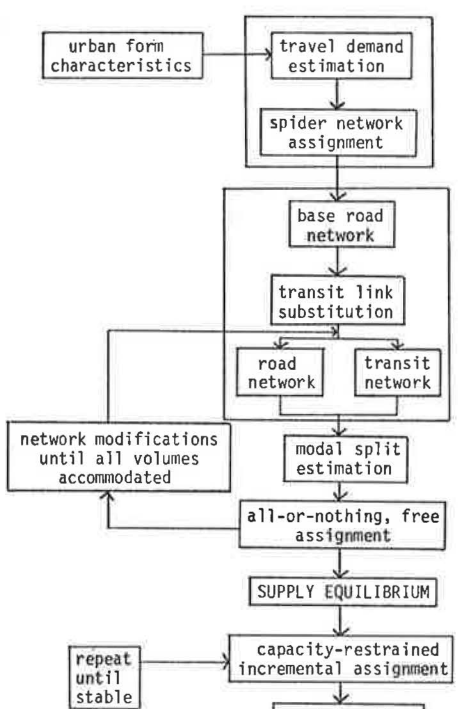


Figure 3. Procedure for two-mode network generation.




This was achieved by first designing a base road network with four possible link types in such a way that it would just carry the estimated desire-line volumes. Modifications were then made to this initial feasible solution by substituting one of eight levels of transit service on a link-by-link basis so that a trade-off function between the consumption of space for the transportation facilities and user travel time was always satisfied. This was achieved by calculating the ratio of the change in travel time to the change in transportation facility space for each feasible transit substitution per link. The transit service for which this benefit ratio was closest to the mean value of the ratios for all feasible substitutions for the link was chosen for implementation. This, in effect, means that the trade-off is directly related to total link volume and higher volumes are more willing to accept larger increases in travel time per unit of facility space gained. This seemed plausible, and a testing of the decision rule indicated that higher capacity transit service does replace road capacity on high-volume links.

It is obvious that the introduction of transit service will have a substantial effect on both mode and route choice. It was necessary, therefore, to reestimate modal split and trip assignment after the initial round of transit substitution. When this was done, however, it was found that the new routes selected took advantage of the links with the higher level of service so that there was a natural aggregation of trip movements into specific modal corridors. In a similar manner, other links were deleted in trip volume; this modified the service available on these links in the next round of transit service substitution. In an iterative sequence, then, a process of network rationalization takes place that involves both the route and mode choices of the traveler so that natural corridors of travel demand build up in accordance with network geometry and demand orientation (11).

The process of network rationalization is most easily demonstrated by the diagrams shown in Figures 4 and 5, which represent the results for the road and transit networks of city type 1A (central core) for three phases of the iterative supply cycle. These figures represent only the high-capacity links in the modal networks, but it is apparent that the road network expands and the transit network contracts. This is obviously a function of the modal-split submodel, but this phenomenon did occur for all six urban forms.

The supply equilibrium cycle was repeated until no further changes were required in each of the modal network links to accommodate the travel volumes estimated in the previous iteration of the cycle. It should be noted that the trip-assignment component of the supply cycle is a free or desire-line assignment because the objective of the network synthesis is to develop a natural expression of the required transportation system. In other words, network rationalization must be unconstrained by physical limitations; modal capacity is simply provided in accordance with traveler demand. This resulted, therefore, in the need for the demand equilibrium cycle, which reestimates modal split and route assignment in an iterative sequence under assumptions of capacity-restrained flow on all links. This cycle is also indicated in the flow diagram shown in Figure 3 and completes the procedure of network generation.

## Comparative Analysis

The results of the procedure of network generation are most easily demonstrated by the total system output measures given in Table 3 (the $X^T$ vectors defined previously). The differences in travel conditions for the six cities are obviously quite significant. For two cities (the homogeneous and the multicentered), no rail transit links were generated. For both total person hours and mean trip length, urban forms 1A, 1B, and 2 have lower conditions than the other three forms. This is not a function of mode use because modal split varies from 8 to 34 percent and this variation occurs in both groups of cities.

The comparison of mode use for the range of cities

Figure 4. Transitional sequence for road network for central-core city.

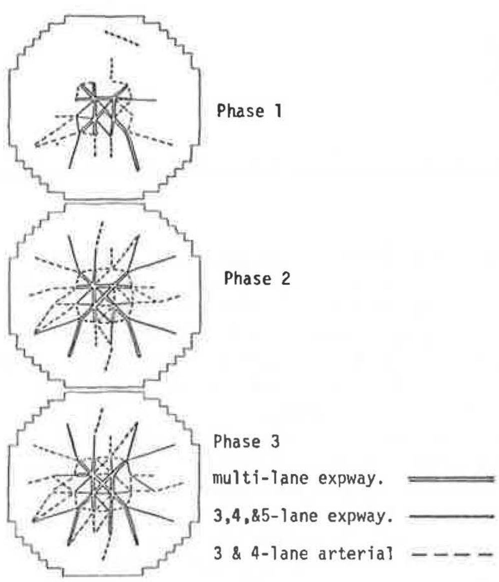



Figure 5. Transitional sequence for transit network for central-core city.

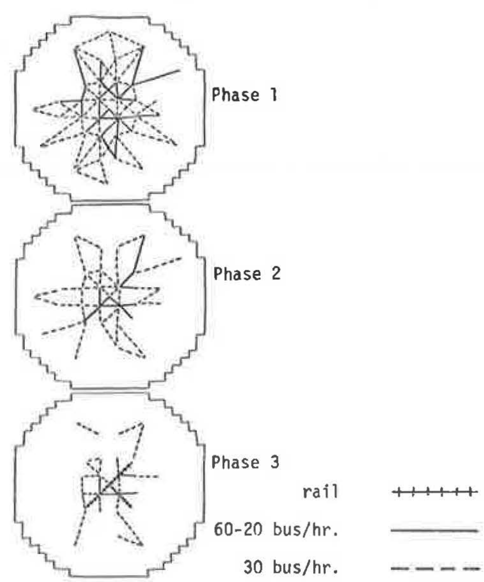



Table 3. Final characteristics of travel demand by urban form.

| | Urban Form | | | | | |
|---|---|---|---|---|---|---|
| Output Measure | 1A | 1B | 2 | 3 | 4 | 5 |
| Total system | | | | | | |
| Total work trips | 346 500 | 341 380 | 326 900 | 334 700 | 375 100 | 387 950 |
| Person hours | $7.00 \times 10^4$ | $7.71 \times 10^4$ | $5.00 \times 10^4$ | $9.65 \times 10^4$ | $9.51 \times 10^4$ | $17.14 \times 10^4$ |
| Mean trip length, min | 11.68 | 13.1 | 9.07 | 17.01 | 14.49 | 26.0 |
| Percentage transit | 18.2 | 8.0 | 10.3 | 33.8 | 19.9 | 8.0 |
| Road network | | | | | | |
| Work trips | 293 100 | 313 900 | 293 500 | 221 200 | 300 350 | 357 950 |
| Percentage on expressway | 63 | 55 | 59 | 85 | 75 | 90 |
| Trip length, min | | | | | | |
| Mean | 11.78 | 12.16 | 8.30 | 17.2 | 13.58 | 26.66 |
| Standard deviation | 10.88 | 7.64 | 4.99 | 18.8 | 12.54 | 40.8 |
| Transit network | | | | | | |
| Work trips | 53 400 | 27 480 | 33 400 | 113 500 | 74 750 | 30 550 |
| Percentage on rail | 30 | 0 | 0 | 51 | 37 | 26 |
| Trip length, min | | | | | | |
| Mean | 11.22 | 24.06 | 15.82 | 16.62 | 18.82 | 18.39 |
| Standard deviation | 8.27 | 11.02 | 8.98 | 13.16 | 13.94 | 18.34 |

is shown in Figures 6 and 7. The percentage of trips that use transit corresponds closely with the number of train kilometers supplied except in the case of the satellite city where the existence of rail transit has little effect on the use of public transit. With regard to the percentage of person hours on high-speed service links (expressway and rail), the two corridor cities (3 and 4) and the satellite city (5) depend most on facilities that have high levels of service, as would be expected. Although the cost of transportation investment has not been estimated directly, it is possible to form some general conclusions on capital costs from the amount of high-service facilities required in each of the six city types. The two corridor plans and the satellite plan are dependent on high-service facilities for both modes and hence will require high capital investment. The satellite plan easily claims the position of the most expensive form even though its rail service requirements are not the largest. The radial-corridor plan is likely to be the cheapest of the corridor plans. The remaining three cities (central core, homogeneous, and multicentered) require the lowest transportation investment; the multicentered city requires the absolute minimum.

Finally, in order to measure the relative efficiencies of the high-service links in the modal networks, intercity comparisons were made between person hours per expressway lane kilometer and person hours per rail transit train kilometer. For expressway efficiency, the satellite city ranked highest, and the multicentered city and the linear city performed rather poorly. For rail transit, however, the networks of the radial-corridor city and the linear city performed well and those of the satellite city very poorly. With regard to the efficiency of the total networks, including all link types, it may generally be concluded that it is more difficult to achieve an efficient transit network than it is to achieve an efficient road network. The satellite form is the primary example of this disparity, but it is demonstrated in the other urban forms as well. In terms of overall efficiency, the radial-corridor city rates best overall. This might be expected, but the fact that the homogeneous city ranks second certainly is not. Even though the homogeneous city has minimal transit service, what is available is effectively used.

Figure 6. Use of automobile mode and expressway.

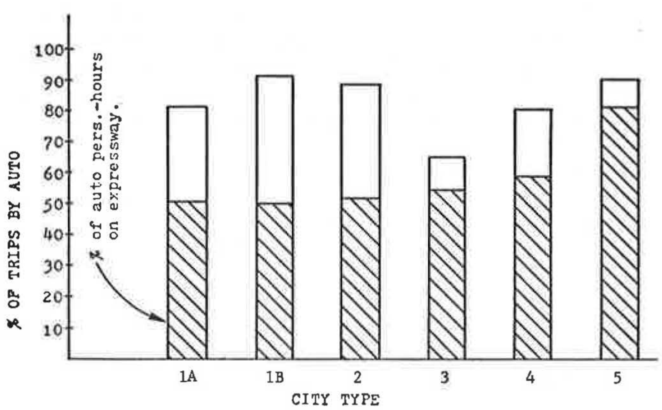
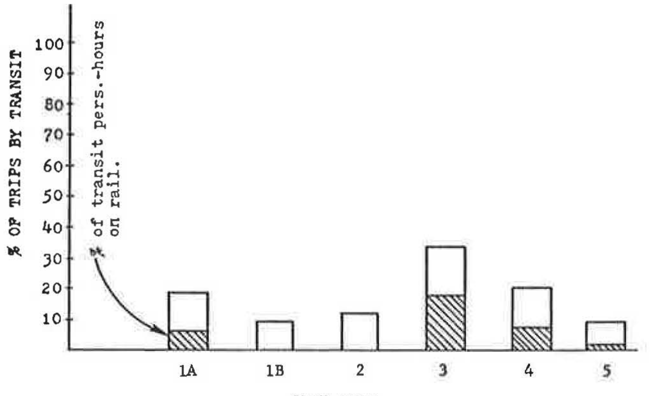



Figure 7. Use of transit mode and rail transit.




## CONCLUSIONS

This paper has been concerned with refining the transportation planning process so that short-term policy issues might be incorporated more effectively. As part of the proposed framework of dynamic policy evaluation, an assessment of the importance and relevance of transportation and land-use alternatives was undertaken. Conclusions that relate to the framework and the static analysis can be summarized as follows:

1. Given the wide range of transportation service and investment requirements that resulted from the different urban forms, the horizon-year policy states are definitely essential inputs to the proposed dynamic evaluation framework. From this analysis, it would appear that their use within the framework will probably improve the computational efficiency of the method of selecting optimal time paths or plans for improvements through time, which is aimed at achieving the selected end-state condition.

2. The transportation requirements of any particular urban form are uniquely defined. The average work-trip lengths for the six cities that were analyzed differed by a factor of almost three. Also, the investment cost implications would appear to indicate the relative inexpensiveness of the centrally oriented city types (central core, homogeneous, and multicentered) relative to the two corridor plans and the satellite plan. This conclusion is verified by the requirement for a large percentage of high-service links in the satellite cities. This conclusion generally supports similar research by Balkus (12) but runs counter to the conclusions of Zupan (13) and Hemmens (14). With regard to specific urban forms, the research results are confirmed by Voorhees, Barnes, and Coleman (15), who conclude that the existence of subcenters reduces average trip lengths, but conflict with Jamieson and others (16), who contend that the linear form rather than the radial-corridor plan is most efficient. With regard to the Metropolitan Toronto Transportation Plan Review (17), the research supports the contention that a nucleated pattern is preferable to a single-core plan in terms of all transportation measures.

3. With regard to the effect that alternative urban forms have on specific modal requirements and use, the analyses indicate that substantial variability in mode use (8 to 34 percent for transit) may be expected among the six city types. In addition, the submode balance (rail-bus and expressway-arterial) is also dramatically different: Two urban forms—homogeneous and multicentered—have no rail service at all.

4. The implications of the modal differences for the urban forms may also be used to indicate differences in user travel costs and facility investment costs. Using either average trip length or number of person hours of travel as a proxy for user costs results in the following ordering of city types: multicentered, central core,

homogeneous, linear, radial corridor, and satellite. Using lane kilometers of high-service facilities as a measure of investment cost results in the same ordering from low to high cost for the first three urban forms and a reversal in order for the linear and radial-corridor plan in the last three cities. The consideration of transportation efficiency (output per unit of input, such as person hours per lane kilometer or per bus kilometer) is even more instructive, resulting in the choice of the radial-corridor plan as the most efficient plan in the use of capital funds, followed by the homogeneous and the multicentered forms.

The research results described here are preliminary in several respects. First, a number of internal modifications and checks should be applied to the transportation modeling procedure. These include the incorporation of transportation investment and user cost functions, the empirical verification of results, the inclusion of new transportation systems, and the application of the model procedure to cities of different population sizes. The most important direction, however, has to do with the further elaboration of the proposed policy evaluation framework and the construction of a policy model that will permit the evaluation of the most efficient path of investment over time in such a way that a desirable end-state plan may be achieved. Only in this way will the transportation system plan be relevant to policy issues.

## ACKNOWLEDGMENTS

Although I accept all responsibility for the statements made here, personal acknowledgments are extended to Richard M. Soberman and John Dakin for their encouragement and guidance and to D. N. Dewees, D. M. Nowlan, M. J. M. Posner, and A. J. Scott of the Urban Transportation Impact Study Group of the University of Toronto-York University Joint Program in Transportation for their collective and individual contributions. The financial assistance offered by the Central Mortgage and Housing Corporation; the Ontario Department of Treasury, Economics, and Intergovernmental Affairs; and the Transport Canada Research and Development Centre is gratefully acknowledged.

## REFERENCES

1. R. G. Rice and D. M. Nowlan. Toward a Dynamic Framework for Urban Transportation and Land-Use Planning. Proc., 16th Annual Meeting, Canadian Transportation Research Forum; Transportation Research Forum, Vol. 16, No. 1, 1975, pp. 91-106.
2. F. C. Dunbar. Quick Policy Evaluation With Behavioral Demand Models. TRB, Transportation Research Record 610, 1976, pp. 47-49.
3. B. G. Hutchinson. Principles of Urban Transport Systems Planning. McGraw-Hill, New York, 1974.
4. L. A. Neumann and others. Integrating System and Project Planning for Effective Statewide Programming of Investments. TRB, Transportation Research Record 499, 1974, pp. 83-93.
5. K. V. Smith. A Mix-of-Modes Evaluation Model for Transportation Systems. Rand Corp., Santa Monica, CA, Rept. P-4059, May 1969.
6. R. G. Rice. Transportation Plan Evaluation as a Continuous Process: A Proposed Framework and Case Study. Paper prepared for Annual Conference of Roads and Transportation Association of Canada, Oct. 11-14, 1977.
7. L. A. Neumann and W. M. Pecknold. Application of the Time-Staged Strategic Approach to System Planning. HRB, Highway Research Record 435, 1973, pp. 20-31.
8. R. G. Rice. Performance Characteristics of Two-Mode Transportation Systems for Varying Urban Form. Department of Civil Engineering, Univ. of Toronto, PhD thesis, 1975.
9. R. G. Rice. The Effect of Alternative Urban Forms on Two-Mode Transportation System Requirements. Paper presented at World Conference on Transport Research, Rotterdam, the Netherlands, April 26-28, 1977.
10. Alan M. Voorhees and Associates. Factors and Trends in Trip Lengths. NCHRP, Rept. 48, 1968.
11. J. C. Rea. Designing Urban Transit Systems: An Approach to the Route-Technology Selection Problem. HRB, Highway Research Record 417, 1972, pp. 48-59.
12. K. Balkus. Transportation Implications of Alternative Sketch Plans. HRB, Highway Research Record 180, 1967, pp. 52-70.
13. J. M. Zupan. Social Implications of Land Use and Mode Choice. Transportation Engineering Journal, ASCE, Vol. 99, No. TE2, May 1973, pp. 383-391.
14. G. C. Hemmens. Experiments in Urban Form and Structure. HRB, Highway Research Record 207, 1967, pp. 32-41.
15. A. M. Voorhees, C. F. Barnes, and F. W. Coleman. Traffic Patterns and Land-Use Alternatives. HRB, Bulletin 347, 1962, pp. 1-9.
16. G. B. Jamieson and others. Transportation and Land-Use Structures. Urban Studies, Vol. 4, No. 3, Nov. 1967, pp. 201-217.
17. Development of Land Use and Transportation Alternatives. Metropolitan Toronto Transportation Plan Review, Rept. 62, Toronto, May 1974.

*Publication of this paper sponsored by Committee on Transportation Systems Design.*

# Modelers, Muddlers, and Multitudes: Establishing a Balanced Transportation Planning Process

Michael A. Goldberg, Faculty of Commerce and Business Administration, University of British Columbia

**Urban model builders and policy makers are turning toward a more flexible approach to planning and implementing urban transportation investments. This paper seeks to extend these recent efforts in unexplored planning processes to enable the newly emerging transportation planning process to better cope with the uncertainties of complex urban and regional systems. The discussion is conceptual and builds on current literature and trends. Diagrammatic representation of "old" and "new" approaches to transportation planning are set out to clarify the nature of the emerging processes. The paper concludes that the trend away from highly structured analytical methods of planning toward more synthetic and open-ended approaches is worthwhile but should not be overdone. What is most needed is a delicate balance between rigorous analytic techniques and less rigorous synthetic and qualitative ones. It is through such balance that technicians (analysts) will work closely with politicians and other policy makers (synthesists) to provide flexible, responsive, and carefully thought-out urban transportation planning.**

As the inherent uncertainty of our various environments has made itself apparent (often painfully so), planning interventions into these environments have become increasingly flexible and open-ended to accommodate the unpredictable. There are many combinations of existing, extinct, or innovative configurations of land use and transportation to cushion such uncertainties. The techniques available for identifying, choosing, and implementing suitable combinations are also many. But technical knowledge is not sufficient. What is needed is a planning process appropriate to an uncertain and dynamic urban environment.

Significant strides have recently been made toward developing planning processes capable of handling the multidimensional complexity of urban systems. The emergence of such processes has been particularly noteworthy in transportation planning, which is the focus of this paper. A report from a recent Philadelphia workshop on communication among planning professionals and researchers explored the differences between this emerging transportation planning process and more traditional approaches (1, p. 6):

> The "old" process in this somewhat overdrawn dichotomy can be described as long range, comprehensive, top-down, end state, closed option planning, based on the engineer-architectonic approach that requires a detailed, fixed end product from which everything else is subsequently determined, the whole predicated on the belief that it is possible to forecast future events. The alternative, or the emerging "new" process, is characterized as short range, incremental, politically open, and multi-optioned in the sense of narrowing but not eliminating choice. Methodologies and techniques for the emerging paradigm have not been settled upon, but the intent of sketch planning and quick response analytic procedures is in this direction. The shift, technically, is clearly well underway, but there is still a long way to go.

Central to the "new" process is the need for information and communication among those involved. Much of the burden of information generation and processing has fallen on urban simulation models of transportation and land use. With the shift from long-range to short-range priorities and from single-purpose comprehensive plans to multiple incremental policies, the demands on technologies for generating and processing information are significant.

It is the goal of this paper to explore

1. The utility of existing models in light of these changes in the transportation planning process,
2. The demands imposed on modelers and future models to meet the needs of this evolving process, and
3. Some approaches to model building that are compatible with these evolving needs.

To accomplish these tasks, the paper first briefly explores traditional transportation planning and then examines the new process in contrast with traditional approaches. Next, a framework for discussion is set out to examine the present and the future of models and model building in the context of the planning process. Finally, ways are suggested to enable the model-building process to complement the planning process.

## THE OLD PROCESS

In reading any of the standard references on transportation planning, one is immediately struck by the order and neatness of the process. It is well structured and highly rational. It is designed like any other production process, the output being a comprehensive transportation plan rather than an automobile or a dishwasher. The similarity to processes designed by engineers is quite reasonable given the origins of urban transportation planning in civil engineering. As such, the "old" process reflects its roots and also its formative era—the 1950s, when the world's problems seemed capable of straightforward (if large-scale) engineering and design solutions.

With such a history, the striving for analytic rigor in the planning process and technical efficiency in the end product is to be expected. Urban and regional planning generally followed similar paths during the 1960s and early 1970s. These and other related academic disciplines are built implicitly on assumptions of order and predictability in the area being studied. Such assumptions, although analytically tidy and emotionally comforting, do not bear up well under "battle conditions" (i.e., actual as opposed to hoped-for or assumed conditions).

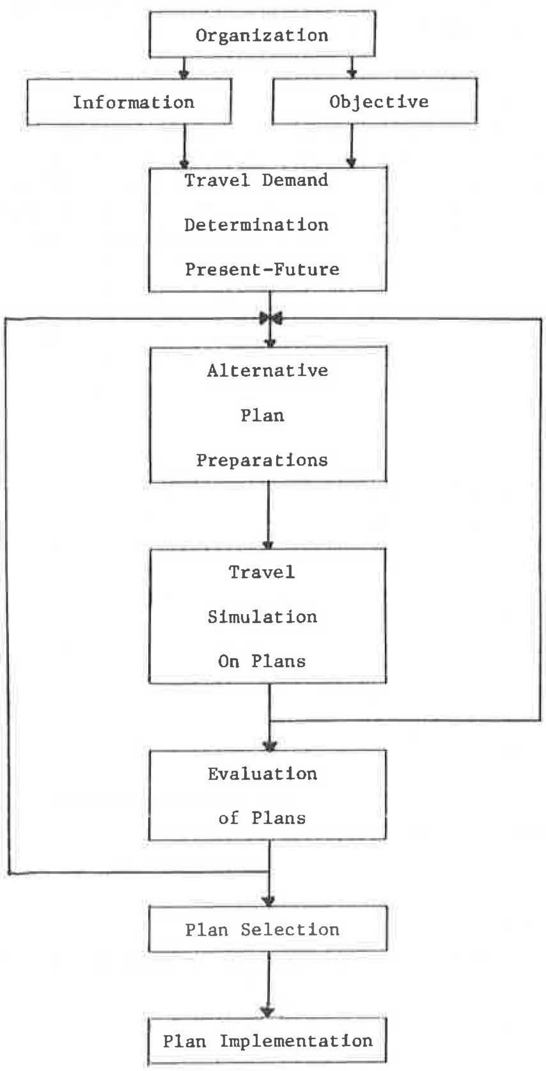
It is inevitable therefore that the old must give way to a newer approach. The rationality, rigor, efficiency, scale, and order of the metropolitan transportation master plan must pass into history.

## THE NEW PROCESS

In a world drifting toward disorder, the quest for order is laudable but fraught with difficulties. Just as our own biophysical environment eludes the chaos of entropy through the openness of the system (open here with respect to energy inputs from outside), so must planning processes become open-ended to avoid the chaos that results from imposing order on inherently uncertain and dynamic urban systems. The rigidities of large-scale high-rise public housing and urban renewal schemes, urban freeways, rail rapid transit, green-belt girdles, and single-use zoning must be replaced by more flexible approaches to perceived problems. To complement analytic skills, synthetic ones need to be sought and refined. The restraints imposed by the narrowness of technically efficient evaluative criteria must be expanded to acknowledge political realities. Transportation planning—indeed, all planning and public decision making—is, when stripped to its essence, a political process. The politics of the process need and deserve to be placed in their proper perspective and context.

These are the kinds of demands that are currently being placed on transportation planning. (It should be stressed that what is said here about transportation planning applies to other forms of urban, regional, and even development planning as well.) To accommodate these demands, the new process must have many of the attributes that Lee, the chronicler of the Philadelphia workshop, ascribes to it (1, pp. 4-12). It must also include other features. For example, the emerging process should foster an appreciation of political necessities in the minds of professional and technical participants and also inform politicians and citizens of technical requirements and constraints. As the Lee report stresses (1, pp. 13-17), communication is a key ingredient. Communication of the sort required can only be built on mutual respect among participants in the process.

Figure 1. Traditional transportation planning process.




Perhaps the most important new element in the emerging process is politics. Politics, politicians, and political considerations are no longer viewed as antithetical to "good" planning but are acknowledged to be fundamental. Indeed, a strong case could be made that the short-range, incremental, and multioption features of the process are a direct result of politicizing it. Thus, Lee emphasizes the changing attitudes of professionals toward politicians and directs much of the text of his workshop report to the increased need for better communication among politicians, citizens, professionals, and technical people involved in transportation (1, pp. 5, 9, 13-16).

Significant changes are already well under way. Compare, for example, Figures 1 and 2. Figure 1 is taken from a 1968 overview paper on transportation planning (2, p. 154). Though it is barely 10 years old, it exemplifies the highly rational structure of traditional approaches to transportation planning. Information flows are unidirectional in general, and a nice, crisp organizational structure permeates the whole. In contrast, Figure 2 (3, p. 8) shows a more recent schematic view. Here we see that the political process is an integral and pervasive element in the planning process. The dramatic difference reflects the distance that has been covered in the very recent past and points the way toward equally dramatic, though less obvious, directions for the future (4).

To accommodate more heterogeneous participants, the process must employ increasingly simpler methodologies. To accommodate the political process in its various forms, research must increasingly focus on policy issues and researchers must become more accessible to the users of their work. Finally, the analytic skills that predominated to date must be balanced by synthetic skills to bring disparate analyses successfully to bear on pressing short-range policy issues. This lengthy agenda is already being whittled away.

In sum, the emerging transportation planning process is different. It has the capacity to overcome many of the shortcomings of past approaches but also promises to create a whole host of new shortcomings.

Figure 2. Changing structure of the transportation planning process.

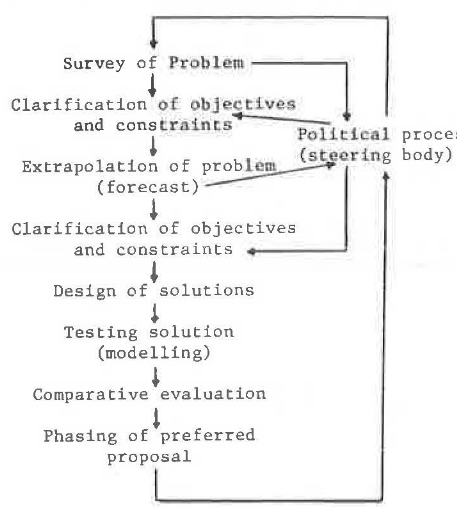
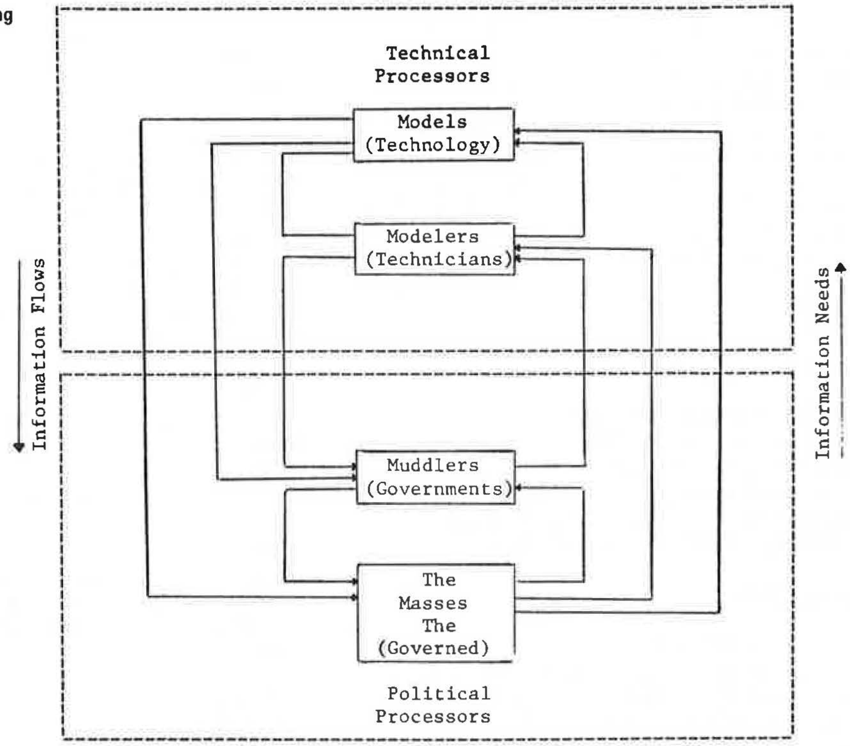


Figure 3. Participants in the emerging transportation planning process.




## The Cast of Characters

To make sense of the emerging process and of its implications for its own further evolution and for the future planning of urban areas, it is worthwhile to look at the various actors in the process and to develop an appropriate framework within which they might effectively interact.

Making plans for a system as complex and diverse as an urban region requires, among other things, an extraordinary amount of information and people to generate and digest the information. Organizing such planning processes even schematically is not a trivial task. The schematic variants rapidly approach in complexity and number the complexity and diversity of the urban region itself. Several different views of planning processes and model-building processes have been set out diagrammatically elsewhere (5, pp. 150-151). Lee offers yet another complementary visual aid to understanding and therefore potentially aiding the planning process (1, pp. 1-3). Given the changing nature of the transportation planning process, it is useful to offer another diagrammatic description to provide a framework for discussion and analysis (Figure 3).

There are four groups of information processors and generators:

1. Models—The models of concern here are conceptual constructs and their computerized analogs that provide simulated policy impacts as outputs. Models are generators of simulated information and also processors of raw data and policy variables. They represent the product of the model-building process at any one point in time.

2. Modelers—Models are designed, debugged, calibrated, and run by technically trained people. The modelers receive information inputs from their clients (politicians, bureaucrats, and citizens) and from the models themselves. They also generate information to develop, refine, and run the models as well as provide technical guidance to model users.

3. "Muddlers"—Muddlers are bureaucrats, politicians, and all servants of the citizenry at large who receive information both from the models and the modelers relating to the simulated effects of contemplated policies. They also receive information from their clients, the polity. Muddlers also provide information: For modelers they provide goals and objectives for model construction and policy, and for models they provide policy inputs and needed data. To the people go their policy decisions and their rationales.

4. The masses—Ultimately, in a self-governing democratic society, all this planning and policy making is done because "the people" want to improve the quality of their society and view such efforts as essential if needed and perceived improvements are to occur. Ideally, the masses receive information from their representatives (the muddlers), from their representatives' representatives (the modelers), and from the representatives' representatives' technologies and expertise (the models). The masses in their turn provide the essential political context within which all this information generation and processing takes place.

Of course, complex categories such as those chosen above begin to blur at the edges. It is my intent here only to sketch the principals in the process and their interactions in the broadest terms. This is intended as a didactic and not a definitive exercise.

## Defining Appropriate Roles

By fixing the boundaries of responsibility of the interacting elements in the planning process, we can begin to ensure that each will operate within its appropriate sphere and not dominate so that the process does not become subservient to any one or more of the elements of which it is composed. The first two groups are essentially processors of technical information. The last two are more concerned with political issues surrounding the establishment of weights for decision making and of goals and objectives for the transportation planning process in the larger sense.

This process can also be viewed as one of information flowing down through the various interacting elements and requests for information flowing upward. As we move down through Figure 3, we move from agents to clients until we reach the fundamental client in a democratic society—the masses. When the process is viewed in this oversimplified context, it should be obvious that anything that impedes the flow of either information or requests for information (i.e., anything that in essence cuts one of the lines of communication among the actors) can destroy, or seriously damage, the entire process.

Putting Models in Their Place

Information and the meaningful communication of information occupy center stage in this schematic. Models must serve their masters. By placing models (and modelers) in this framework, the challenges that face future modeling efforts as well as some of the problems that confront existing models become clearer.

Utility of Existing Models

During the past two decades, a great deal of progress has been made in the development of urban simulation models, mostly as an adjunct to the transportation planning process. A number of benefits that are valuable for the future have been derived from these past activities and represent a positive legacy of past and current modeling work. Experience with models has also uncovered some serious deficiencies.

The usable features of existing models include the following:

1. Technically sound models have resulted from past modeling efforts. The entropy-maximizing models of Wilson (6, 7) and others establish an analytically neat, computationally efficient, and usable framework. The econometric models of metropolitan development by Kain and Quigley (8) and Straszheim (9) of the National Bureau of Economic Research represent significant strides in the application of economics to modeling. Putman's work on interactive transport and land-use models (10) has moved us ahead by directly linking access and land-use activities. There are also numerous bibliographies on the subject (2, 10, 11).

2. A fund of model-building expertise grew out of past modeling activities. Communication among these technical experts is excellent, resulting in a well-informed, relatively closely knit group of researchers armed with powerful technical skills that can accommodate the changing demands being placed on models and modeling. A fund of model-using expertise is also developing (12).

3. Large-scale data bases and data-handling systems have come into being to cope with the ravenous data requirements of most models. Since future models are likely to be more modest, past gains in the field of large-scale urban data collection, manipulation, storage, and retrieval should be more than adequate to meet expected future demands.

4. Large, fast, interactive, and graphic computer hardware and software enable existing models to be run economically and be potentially widely available. Present hardware and software capabilities appear adequate to meet any unforeseeable future demands.

Given technical considerations alone, current models and modeling expertise provide some essential preconditions for the next generation of models that are likely to be needed by the more open-ended and incremental emerging process described above.

The deficiencies of existing models include the following:

1. Until very recently, high cost and limited output characteristics typified most existing models, severely restricting their usefulness.

2. Comprehensive incomprehensibility characterizes a large portion of the current stock of models. Their comprehensiveness makes them incomprehensible to all but a very few highly trained technical people who have the opportunity to work directly with the models. It is not surprising that potential users are discouraged at the outset from using tools that they cannot understand.

3. Limited access to models, model builders, and computer machinery also restricts the utility of current models. This, combined with technical complexity, further constrains the potential success of existing modeling.

4. Lack of policy inputs and outputs hinders many current models. Without relevant policy variables, it is perfectly understandable why policy makers and their clients—the public—have shown little interest in making better use of models.

Despite technical achievements, current models have not realized their full potential, often because of the technical rigor and awesomeness of modern computing machinery. Much greater emphasis must be placed on making models useful and usable if they are really to become an integral part of the more broadly based emerging transportation planning process (5, 13, 14, ).

## FUTURE DEMANDS ON MODELS AND MODEL BUILDING

Weaknesses in current modeling efforts point the way toward future needs (5). The overriding need is to fully appreciate the context within which modeling takes place. In this context, a number of specific issues come to mind:

1. Smaller, less ambitious models would appear to be an obvious and direct consequence of the deficiencies of current modeling. Such models should be much more easily understood, dramatically less expensive to operate and refine, easier to program for different hardware and software configurations, and, as a result, markedly easier to use.

2. Special-purpose models such as Huff's 1962 retail model (15) are needed. Simple in design and structure and directed toward well-defined transportation and land-use elements, such models have the potential for ready acceptance and use because of their high degree of specificity and singleness of purpose. Models for specific public facilities; for specific recreational uses such as local parks, golf courses, fitness activities, and playgrounds; for high-rise and low-rise offices; and for other quite narrowly defined land uses and transportation activities would be consistent with both the needs of the emerging process and the foregoing comments on smallness. During the past half-dozen years, this has begun to happen (16, 17).

3. Modular submodels easily combined to form more comprehensive, but still relatively simple, models are also in the offing. By building larger, if still simple, models from well-defined, well-designed, and well-used submodels, economies in operating characteristics can be achieved while the compound models are kept sufficiently elemental to allow different kinds of users to

understand their structure, logic, and use.

4. Interactive and graphics capabilities of the present generation of computers should be exploited to the utmost. Current hardware and software packages enable programs to be used at numerous locations by diverse users and provide graphic output that can be designed to increase understanding and use of models. The computer technology exists to make models more accessible, comprehensible, and thus usable to ever wider audiences. Every effort should be made to take advantage of these technological achievements.

Simpler, special-purpose modules can potentially overcome the weaknesses of existing models, exploit computing technology, and assist the planning process.

## ISSUES OF MODEL-BUILDING STRATEGIES AND PROCESSES

Just as transportation planning takes place in the confines of a larger transportation planning process, so does model building go on within a larger model-building process. If that process is to achieve goals that are consistent with the new transportation planning process, a number of strategic points are worthy of consideration before new modeling activities begin (5, 13):

1. Interinstitutional modeling teams hold the promise of providing technical modelers with real-world inputs, constraints, and uses for models. Such a blending of policy-making skills opens the possibility of achieving the best of all possible worlds through choosing the most appropriate technical elements and bringing them to bear on the most important practical issues (18, pp. 629-34).
2. The evolution of models should be allowed for. Placing modeling in an evolutionary framework highlights the process nature of model building as opposed to a pure production orientation designed to produce models. Models of dynamic systems must themselves be capable of change. Keeping sight of the model-building process allows for such evolutionary change.
3. Demystification of models is also a high priority. It would be helpful to remind unsuspecting users that, after all, "models are only human." They have weaknesses and are far from infallible. Only through sufficient attention to weaknesses and clear and simple elucidation of the structure and function of any given model can the user (ultimately the public) be protected from inappropriate application and use of modeling technology in relation to pressing policy questions.
4. Disposable institutions should be found to house models, and the life of both the models and the institutions should be gracefully ended when they are no longer useful. The North American urban landscape is dotted with formerly useful institutions that have taken on lives of their own quite independent of their original purpose.

Model-building processes are subordinate to the higher order transportation planning processes discussed earlier. If models and modelers develop in this larger context, they are likely to continue to play important roles in the emerging transportation planning process. Otherwise, I suspect they will be returned to the academy from whence they came.

## CONCLUSIONS

In any range of disciplines, a shift similar to that under way in transportation is noticeable. The almost obsessive concern with rigor, analysis, and precision that has come to typify inquiry is being replaced by calls for more relevant integrative approaches (18). "Value-free" economics, sociology, anthropology, history, and so on are being challenged by unabashedly value-based research. Awareness of the need for synthetic skills, for sound processes as well as sound products, and for generalists as well as specialists illustrates this point. Lee directs attention to many of these issues (1, p. 20) as I have done (5). More frightening than the continuation of analysis uber alles is the possibility that professionals will respond to these exhortations and abandon sound analytic tools in favor of synthetic ones alone.

The point is that both sets of skills are essential. In a strict sense, there can be no good analysis without some sound, previously synthesized hypotheses. Similarly, without analytic evidence there is nothing to synthesize.

Accordingly, if the new emerging transportation planning process degenerates into just another (albeit synthetic and process-oriented) technology, I would anticipate that its impact will be of short duration. If, on the other hand, the emerging process acts to bring together varied and needed skills of analysis and synthesis in an evolutionary and dynamic setting, then its promise for significantly improving transportation and, more generally, urban and regional planning is great.

Balance is called for between the paired elements of product and process, analysis and synthesis, individuals and societies. This is the real challenge facing planning and decision-making processes in our societies. This is where better communication can have its most telling impact. By providing bridges across the gaps between elements, communication can begin to engender some sense of the total effort required to plan and administer our urban and rural environments. By engendering respect for specifics among generalists and respect for generalities among specialists, communication can help those who formerly held a dichotomous view of the world appreciate that at best they represent only half the picture. Technicians without policy makers to implement technically based suggestions are likely to be as helpless as policy makers who face technically based decisions in the total absence of knowledge. It takes both engineers and politicians to build highways, subways, and city streets.

Open-minded and cooperative participation in the process is a necessary condition for success. Attitudes change slowly, usually for good reason. For attitudes to change, they must, among other things, be shown to be inadequate to current needs; simultaneously, it must be shown that there exists an alternative set of values (attitudes) that are more appropriate. The new transportation planning process does have the flexibility and breadth to foster diversity, to bring differing attitudes in contact with each other, and ultimately to provide for the evolution of attitudes that are needed to complement the evolution of the process itself.

## REFERENCES


1. D. B. Lee, Jr. Improving Communication Among Researchers, Professionals and Policy Makers in Land Use and Transportation Planning. U.S. Department of Transportation, 1977.
2. F. W. Memmott and C. Guinn. Transportation Planning: Streets, Highways and Mass Transportation. In Principles and Practice of Urban Planning (W. I. Goodman and E. F. Freund, eds.), International City Managers' Association, Washington, DC, 1968, pp. 138-163.

3. D. N. M. Starkie. Transportation Planning, Policy and Analysis. Pergamon Press, Oxford, England, 1976.
4. R. A. Gakenheimer. Urban Transportation Planning: An Overview. In Taming Megalopolis (H. W. Eldridge, ed.), Anchor Books, Garden City, NY, Vol. 1, 1968, pp. 392-411.
5. M. A. Goldberg. Simulating Cities: Process, Product and Prognosis. Journal of American Institute of Planners, Vol. 43, No. 2, 1977, pp. 148-157.
6. A. G. Wilson. Entropy in Urban and Regional Modeling. Pion, London, 1970.
7. A. G. Wilson. Urban and Regional Models in Geography and Planning. Wiley, London, 1974.
8. J. F. Kain and J. M. Quigley. Housing Markets and Racial Discrimination. National Bureau of Economic Research, New York City, 1975.
9. M. Straszheim. An Econometric Analysis of the Urban Housing Market. National Bureau of Economic Research, New York City, 1975.
10. S. H. Putman. Urban Land Use and Transportation Models: A State-of-the-Art Summary. Transportation Research, Vol. 9, 1975, pp. 187-202.
11. F. J. Cesario. A Primer on Entropy Modeling. Journal of American Institute of Planners, Vol. 41, No. 1, 1975, pp. 40-48.
12. J. R. Pack. The Use of Urban Models: Report on a Survey of Planning Organizations. Journal of American Institute of Planners, Vol. 41, No. 3, 1975, pp. 191-199.
13. D. B. Lee, Jr. Requiem for Large-Scale Models. Journal of American Institute of Planners, Vol. 39, No. 3, 1973, pp. 163-178.
14. A. H. Voelker. Some Pitfalls of Land-Use Model Building. Oak Ridge National Laboratory, 1975.
15. D. L. Huff. Determination of Intraurban Retail Trade Areas. Real Estate Research Program, Univ. of California, Los Angeles, 1962.
16. G. M. Carter and E. Ignall. A Simulation Model of Fire Department Operations: Design and Preliminary Results. Rand Corp., Santa Monica, CA, 1970.
17. R. C. Larson and K. A. Stevenson. On Insensitivities in Urban Redistricting and Facility Location. Rand Corp., Santa Monica, CA, 1971.
18. M. A. Goldberg. Simulation, Synthesis and Urban Public Decision-Making. Management Science, Vol. 20, No. 4, Part 2, 1973, pp. 629-643.
19. M. A. Goldberg. On the Inefficiency of Being Efficient. Environment and Planning A, Vol. 7, 1975, pp. 921-939.

*Publication of this paper sponsored by Committee on Urban Activity Systems.*

# How to Do a Transit Station Land-Use Impact Study

Douglass B. Lee, Jr., Institute of Urban and Regional Research, University of Iowa

**Several improvements in the conceptual basis and methodology for studies of land-use impacts have occurred over the past two decades, but the framework is still incomplete because the need to incorporate the policy context into the study design has not been fully recognized. A revised model for impact studies is proposed, and the approach is illustrated by a case study of a planned rail rapid transit station. One of the major differences between this and previous methods is that the method described in this paper acknowledges several possible outcomes or impacts as a function of alternative public policies in addition to the transit station itself. Five categories of impacts are evaluated: public facilities, environment, market, neighborhood, and costs and revenues.**

The purpose in asking the positive question, What are the land-use impacts of a major transportation project?, is to evaluate better the feasibility and desirability of such projects, and the answer to the question depends heavily on public policies other than the project itself. The theory and case study presented here are an attempt to construct a workable framework for executing land-use impact studies of major transportation investment projects from a planning- or policy-oriented perspective.

## IMPACT MODEL

Refinements in the before-and-after and the more recent with-and-without impact methodologies have advanced the state of the art (1, 4), but the model, derived from experimental design in the physical and natural sciences, is still incomplete. Figure 1 shows schematically an extension of the with-and-without model in which the comparison is made between two sets of outcomes ("options" because they are a consequence of conscious policy choices) that result from the decision to build or not build the project. State-of-the-world assumptions are things that are held constant for comparative purposes: regional population and employment growth, aggregate travel demand, and the rest of the transportation system. Policy assumptions, in contrast, are specific to each option: For example, policy assumptions associated with intensive redevelopment are different from those associated with neighborhood preservation. The impact of the project is the difference between (a) the options available without the project and (b) the options available with the project.

Previous impact studies and the proposed model can be distinguished, in part, by the way the question is asked. In relation to the case study of the Metro transit station in Vienna, Virginia, the old research question is, What will happen if a transit station is placed at I-66 and Nutley Road? The policy research question is, What will be the differences between the choices available if a transit station is or is not placed at I-66 and Nutley Road?

CASE STUDY

Vienna, the town after which the proposed station is named, lies just to the north of the station site in the Virginia suburbs of Washington, D.C. The Vienna Metro station is the terminus of the Vienna line of Washington's Metro rail rapid transit system. The immediate station area, shown in Figure 2, is largely vacant now, and the station itself is located in the median of I-66 just west of Nutley Road.

Specifically, the question being asked in relation to this station is the following impact question: Given that a transit station is located at I-66 and Nutley Road, what will its impact be? Alternative questions that are not addressed include the following:

1. Given the locations of all other transit stations and lines, what are the impacts of locating the Vienna station at I-66 and Nutley Road versus other possible locations?
2. Given the general configurations of the line, what are the impacts of alternative numbers and locations of stations?
3. Given the existence of a system, what are the impacts of alternative line locations and lengths of extensions?
4. What are the impacts of a rail rapid transit system on the Washington, D.C., metropolitan area?

Figure 1. Proposed land-use impact model.

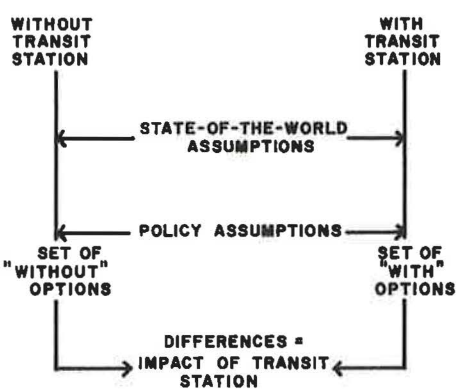
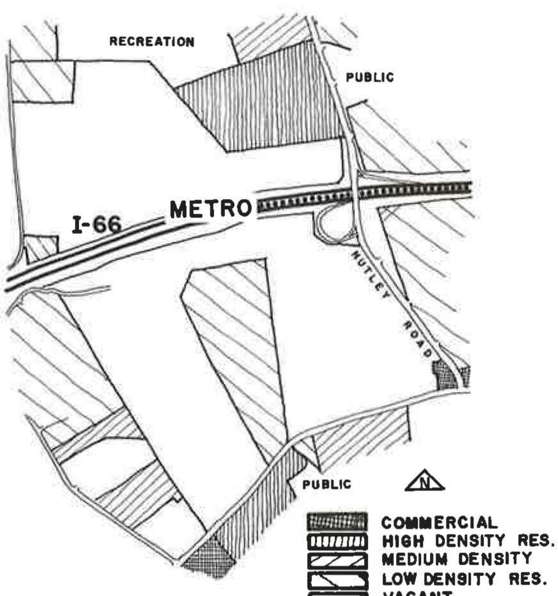



Figure 2. Existing land use and anticipated development at Vienna station site.




These questions represent respectively the station location, route decision, corridor decision, and build-no-build decision questions. Each is a separate question and must be addressed within a separate and suitable analytic framework. Most notably, it is not possible to add the pieces together to get the whole; the answer to the macroquestion is not the summation of the answers to the microscopic questions.

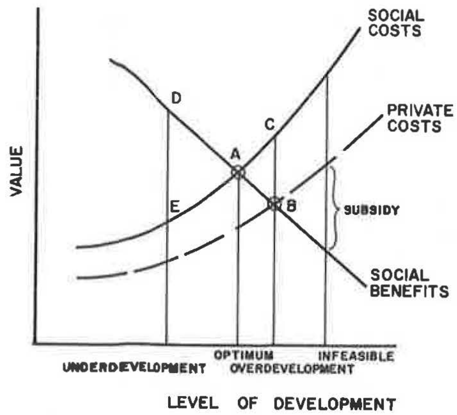
DESIGN AND SELECTION OF OPTIONS

It is important to emphasize that the options are discovered rather than invented although a good deal of creativity is often required to ferret out the real options that exist. Because the process of discovering options is largely heuristic and judgmental, it is misleading to break the process into separate steps. A working approximation might include the following:

1. List all possible alternatives for future development in the vicinity of the station. Clearly, it is not possible to carry this out to the letter, but it is not necessary to list most of the implausible alternatives because they will be eliminated in the next step.
2. Delete infeasible alternatives. Feasibility will, of course, be one of the judgmental determinations, but a key component will be market demand for various land uses at the particular site. Techniques for market studies are well-known applications of macroeconomic concepts (3, 5).
3. Group options into categories. The categories used for the case study are based on levels of development or development intensity, and this might be a dimension suitable to many impact studies although other dimensions can be used.
4. Rank the options within the categories according to normative objectives. These objectives are specific to each of the five impact categories and are described below in the context of the land-use impacts.
5. Evaluate the preferred option or options within each category. Impacts are estimated for each type, and results are tabulated as to costs, benefits, or residual impacts (those that are of interest but cannot be aggregated as either costs or benefits).
6. Revise options and categories as appropriate. Steps 3 through 6 can then be repeated until a stable set of options is generated.

The desired result of the option design effort is a limited number of real choices that can be reviewed from both technical and political perspectives. Thus, the impact study is also, not surprisingly, a planning study in that it provides information that will aid in resolving a problem of social choice. The choice among options, represented in abstract form in Figure 3, is an attempt to find a balance between social costs and social benefits. On the benefit side, demand is reflected in the prices consumers are willing to pay for such items as housing, personal services, retail goods, and hotel rooms; these benefits are transformed into demand for land development through entrepreneurs and lenders who are able to perceive the demand and willing to invest in the development. On the cost side, the supply curve represents the opportunity costs of resources foregone by both the private and public sectors to achieve different levels of development. The optimum

Figure 3. Social costs and benefits of development options.




(A) is the point at which marginal social cost equals marginal social benefit.

In practice, a number of varieties of market failure distort resource allocation from the optimum. Only one variety is of concern here: negative externalities, which, in the form of noise, dust, disruption, and environmental degradation, allow some of the social cost to be exported by the private market. Decision-makers in the private sector consider only those costs represented by the dashed line in Figure 3 and choose a level of development (B) that is higher than optimal; the area BAC represents the loss to society from this overdevelopment.

Normative objectives, which are described below, are explicitly intended to place the full burden of costs on those who derive the benefits and constitute, for each option, a vertical movement from the private to the social cost curve. If costs are fully internalized, then a suboptimal level of development (D) results in a social opportunity loss equal to ADE—undesirable but perhaps preferable if the negative externalities cannot be controlled. A level of development higher than B would require a private market subsidy (even if external costs were ignored) and is, by our definition, infeasible. Two additional points should be made about negative externalities: First, if not controlled they may have the effect of reducing benefits (a lower social benefit curve), which would further reduce the optimum level of development; second, they amount to transfers of income from those who suffer the externalities to those who create them.

Only the end product is presented for the specific case of the Vienna station so that the options—low, medium, and high—embody the best mix of development at each level and negative externalities are assumed to be largely controlled as a result of specified public policies. Policy makers must then make their own assessments of whether the mixes are desirable and to what extent they are willing and able to impose regulations that will reduce the negative externalities. The three types of options can be generally described as follows:

1. Low—This includes a mix of residential and commercial units, but the largest single land use would be single-family residential. This would have the effect of extending the existing neighborhood into the area around the station, thereby providing a transition and a buffer against the station and its ancillary activities. Arrivals at the station would be predominantly by bus, kiss-and-ride, and park-and-ride.

2. Medium—Slightly more emphasis is placed on commercial development and considerably more on multifamily residential units. Some clustering of structures could be accomplished, and most of the land not covered would be in public areas such as those around garden apartments.

3. High—More emphasis on commercial development and multistory apartments, lower land coverage, and more clustering would characterize this option. Pedestrians would form a relatively high proportion of the trips to and from the station.

Specific requirements in units and space for the three development options (2) are given below (1 $hm^2$ = 2.5 acres and 1 $m^2$ = 10.76 $ft^2$):

| Land Use | Low | Medium | High |
|---|---|---|---|
| Residential | | | |
| Single family | | | |
| Units | 364 | 0 | 0 |
| Space, $hm^2$ | 37 | | |
| Townhouse | | | |
| Units | 600 | | |
| Space, $hm^2$ | 24 | | |
| Garden apartment | | | |
| Units | 825 | 1620 | 1830 |
| Space, $hm^2$ | 22 | 44 | 49 |
| Elevator apartment | | | |
| Units | 1250 | 1850 | 3250 |
| Space, $hm^2$ | 18 | 24 | 38 |
| Total units | 3039 | 4300 | 5420 |
| Office space, $m^2$ | 22 300 | 33 400 | 65 000 |
| Retail space, $m^2$ | 4600 | 14 000 | 23 000 |
| Hotel rooms | 100 | 200 | 300 |

Much of the substantive information presented here for the case study is taken from a study of three stations on the Vienna line (2), and these market forecasts project an adequate demand for any of the three options.

It is the conclusion of this study that the high option comes the closest to constituting the optimum (A in Figure 3). But this result depends on the many policy assumptions and other assumptions discussed below, and no implication that high levels of development are generally suitable for transit stations is intended. The Vienna station was selected in part because it is illustrative of situations in which a range of options are available and hence the impact of the station is not uniquely predetermined.

## EVALUATION OF LAND-USE IMPACTS

Impacts are grouped in five categories—public facilities, environment, market, neighborhood, and costs and revenues—on the basis of policy treatment and underlying assumptions. Table 1 gives three aspects of each type of impact: normative (ideal policy) objectives, the nature and measurement of impacts, and the evaluation of impacts. Evaluation concerns the extent to which the impacts can be entered and aggregated in a cost-benefit framework as well as the extent to which the assignment of values is inherently political. Impacts of each option are summarized, evaluated, and compared with the options that are available without the station.

### Public Facilities

Services provided by public facilities can be roughly separated into those that create direct benefits for the consumer (e.g., travel, water, and waste disposal) and those that create general benefits for the community as a whole (e.g., government and primary education). For facilities that benefit consumers, costs should be paid either through direct user charges such as parking fees and hookup charges or through development charges such

Table 1. Evaluation of land-use impact categories.

| Category | Normative Objective | Impact Measurement | Evaluation |
|---|---|---|---|
| Public facilities | Costs of all facilities and services that create benefits that occur directly to the user should be paid for with suitable user charges; capacity of public facilities should be adequate to provide for expected demand | Measure (a) drawdown in capacity of existing facilities resulting from each option, and (b) extent to which demand has been anticipated and capacity programmed to meet demand | Value can be attached to the consumption of capacity only when the demand created by land-use development could not reasonably be foreseen and constraints such as long lead time and bonding limits exist on providing adequate capacity (this condition is, by definition, temporary) |
| Environment | Environmental resources should be protected by suitable constraints on development | Measure residual changes in environmental characteristics | Values to be placed on net changes in environmental variables can only be assessed through the political process because normal market mechanisms undervalue most environmental resources |
| Market | Resources such as labor and materials exchanged in private markets related to station development are properly valued in those markets (i.e., there are no significant externalities, inefficiencies, or market imperfections) | Estimate changes in market activities (employment, housing mix, land use), including those indirectly related to the existence of the transit station | No costs or benefits can be attached to market impacts except in cases where (a) there is specific evidence of significant market failure (public-sector imposition of D in Figure 3, for example) or (b) there are expressed community goals that pertain to certain market impacts |
| Neighborhood | Existing and constructed neighborhood resources should be protected by suitable constraints on development or compensation should be paid to affected parties | Measure residual changes in neighborhood characteristics | Inadequately compensated changes to existing neighborhoods should be considered costs; other changes are a matter of individual taste and perhaps political choice |
| Costs and revenues | Same as for public facilities | Estimate changes in annual revenues and expenditures for affected municipal budgets | Changes in general revenue patterns should be noted and corrective measures taken if problems appear; underpayments by users and direct beneficiaries of facilities should be regarded as costs of development to be minimized as much as possible |

as fees or in-kind contributions from developers. Facilities that create general benefits can be financed from general revenues such as property, sales, and income taxes. If these policies are adhered to, the infrastructure required by development is paid for by those who benefit, and general facilities are supported by the community in proportion to ability to pay.

Facilities required to support development at the Vienna transit station are listed below:

1. Nutley Road should be widened from two to six lanes, and a number of similar improvements should be undertaken to increase the capacity of vehicle access to the station. All three options require these road expenditures.
2. The road and parking area immediately adjacent to the station needs to be redesigned in order to better facilitate pedestrian arrivals. This is especially needed to support the high development option. The present design requires pedestrians to cross a large parking area in order to reach the station.
3. Pedestrian walkways throughout the station area, public squares and furniture, landscaping, shelters, and other items should be constructed at the expense of developers. More amenities can be obtained under the high option because of higher intensity of use and economies from clustering structures.
4. Capital facilities needed to support each development option should be provided and financed in accordance with the guidelines given above. More recreation area and open space are needed for the high option than for the low, for example, and should be provided by developers.

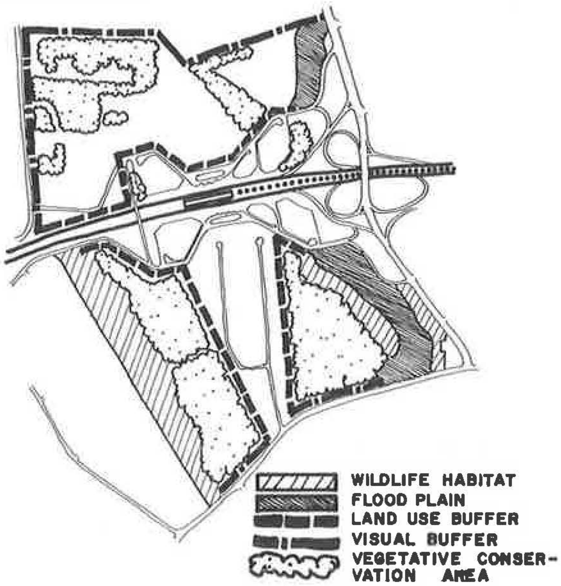
## Environment

Clearly, some changes in the natural environment will occur if any development at all takes place; minor reductions in environmental quality may be offset by the absence of such reduction elsewhere. The first component of the problem is to determine which changes are acceptable, which changes are acceptable if minimized, and which are unacceptable. The second component is the design of standards or other methods to achieve only acceptable changes. The environmental controls required are given below:

1. Several notable stream valleys traverse the site, and these are generally wooded. No development should be permitted in any 100-year floodplains or within 30 m (100 ft) of a stream bed.
2. Portions of the stream valleys have been identified by Fairfax County as wildlife habitats. These should be protected by a minimum of 76 m (250 ft) of natural buffer on either side of the stream.
3. Because of the high clay content of the soil and the frequency of sudden, hard rains, water quantity must be explicitly controlled. Natural vegetation should be retained as much as possible, especially on slopes, and retention facilities should be required for all development so that the natural drainage capacity will not be overloaded.
4. Slippage-prone soils should be identified by the developer, and measures should be taken to ensure stability or to avoid the problem of slippage.
5. The county has delineated "environmental quality corridors" that are designed to create a network of open space and also protect stream valleys and other environmental resources. A portion of the site for the Metro station is included in this network.

Degradation of most aspects of the environment can be kept to tolerable levels by appropriate policies and attendant costs without detracting from development potential. Because the high-density option emphasizes clustering of structures and lower coverage, environmental resources such as open space, stream valleys, and quality and quantity of water are actually more easily protected under high development than otherwise.

Figure 4. Effect of environmental and development constraints on site planning.




Market

For the most part, market impacts are simply redistributive, spatially or among sectors; i.e., the activities (such as employment and housing) would have occurred somewhere, perhaps in a different form; the differences may be of interest, but it is seldom a matter of new jobs or net increases in land value. Changes that occur as a result of properly functioning markets can be legitimately diverted only by "buying out" responsible property rights, e.g., through acquisition of land for parks instead of development. The market impacts of the Vienna Metro station include the following:

1. Private market land-use changes resulting from the presence of the station could range from minor to major depending on public policies. If low-density development were the option followed, land-use changes would largely be limited to those involving vehicular access to the station. High-density development, however, would result in substantial changes in land use. Hence, market impacts of the station depend heavily on development policies and not solely on the presence or absence of the station.

2. Which development option was chosen would have only a small effect on the number of housing units constructed in the region, but location of the units within the corridor and perhaps within the region would be altered. High development would shift the mix of structure types away from single-family units toward townhouses and apartments and would allow (with suitable policies) more moderate- and low-income units to be constructed.

3. More specialized commercial activities would also be likely under the high-density option in comparison with the highway and shopping center development that would take place under the low- and medium-density options.

Neighborhood

Neighborhood quality depends on many factors. The group of factors that land-use controls attempt to ameliorate are those negative externalities, or "nuisances," created by land-use interactions. Protecting neighborhood resources means preventing the negative impacts of new development on existing neighborhoods as well as ensuring compatibility within new development. The potential impacts of the Vienna station on neighborhood quality are given below:

1. The neighborhood surrounding the station area is generally low-density residential so noise levels should be compatible: moderately low during the day and quiet at night. Potential sources of noise are traffic (especially trucks and motorcycles), loading and unloading of trucks, garbage containers, power equipment, stereos, parties, and discotheques. The source of most objectionable noise in the station area is motor vehicles, and the most efficient protection is design standards for buffering development from trafficways.

2. High-density land uses are visually incompatible with low-density neighborhoods, but the impacts can be almost fully eliminated by means of three measures: (a) placing the largest structures closest to the station and reducing the intensity of use outward, down to garden apartments and townhouses; (b) using vegetative buffers between different intensities of use that are incompatible; and (c) imposing a height restriction on structures of 12 m (40 ft) above the highest local grade to ensure that structures blend in rather than stick up (taller buildings would be permitted on lower grades). Because of the existing vegetation and the topography of the site, both of the last two measures would be very effective in this instance. Figure 4 shows the combined effects of environmental constraints and buffering requirements on site planning.

3. Low-density development will maintain the age, family structure, and income mix that already exist in the area, whereas high-density development would also allow the elderly, single people, young couples, and moderate-income households to join the community.

4. Dust, fumes, loss of important architectural sites or historic sites, vibration, and flooding can also reduce neighborhood quality; under the stated policy conditions, problems with these impacts are not expected. Suitable access control should make construction impacts on the largely vacant site minimal.

Costs and Revenues

Calculations of costs and revenues typically reflect little more than the number of school children that will be brought in by new development. Preferably, each direct-benefit government function, such as utilities, should be balanced separately, and user fees should be distinct from general revenues. Road users do not pay property taxes on the right-of-way, sales tax on gasoline (they pay an excise tax), or a share of construction, maintenance, and administrative costs; hence, any increase in highway capacity implies an increased and continuing transfer of advantage from general taxpayers to highway users. Unfortunately, this inefficiency cannot be corrected at the local level although the costs of some kinds of facilities can be levied on developers on the assumption that the costs will be passed on to those who create the need for the facilities. Several fiscal viewpoints are needed, including those of the county, the town of Vienna, Metro, and the highway department.

Summary

The impacts of the presence of the Vienna Metro station depend to a large extent on public policies that affect the amount of development that takes place in the immediate vicinity and the regulatory constraints placed on that development. Comparisons can be made among the three development options by using a cost-benefit framework and a tabulation of residual impacts.

Costs and Benefits

Given the assumption set forth, there should be no uncompensated costs of high-level versus low-level development. One possible exception would be traffic. To the extent that the high option generated more total trips than it substituted walking for automobile trips, there would be some negative neighborhood effects; one estimate is that there would be 1400 additional vehicles in the peak hour (2). If this factor is considered, the benefits of high-level over low- or medium-level development (area ADE in Figure 3) are as listed below:

1. Desirability of integrated, mixed land uses, housing types and price ranges, population ages and incomes, and commercial enterprises, as reflected by what consumers would be willing to pay in the market;
2. Additional public facilities and amenities that can be provided (instead of savings in the cost of public facilities caused by clustering or higher profits to private entrepreneurs);
3. Better use of the rail transit system (if other facilities would be needed for highway travel while there is excess capacity on Metro, the benefit is the savings in the cost of new facilities); and
4. Greater retention of existing vegetation and protection of environmental resources.

Because the low-density option is similar to what will occur without the station, the costs and benefits of the transit station under high-density development (relative to no station) are similar to the comparison between options. The major differences are in the road improvements and traffic impacts since these will occur under any development option.

Residual Impacts

For the Vienna site, the location of a transit station offers opportunities for development that would not be available without the station but will not necessarily occur with the station. In fact, rather stringent policy assumptions (the normative objectives) are required to realize the full potential of the opportunities; if these assumptions are generally not followed in implementation, the resulting impacts would be different from those stated. Assuming that a high level of development and the corresponding constraints are implemented, the remaining impacts would be limited to the following:

1. Impacts listed above as benefits;
2. A change in the character of the neighborhood from suburban to low-density residential with a small semiurban neighborhood core;
3. Impacts of increased traffic volumes in the neighborhood to the extent that these are not buffered (primarily in comparison with no station at all);
4. Some reduction in open space and vegetation (relative to no development) but an increase in public open space;
5. Somewhat higher ambient levels of noise, particulates, and air pollution in the immediate environs but less in the aggregate;
6. Increases in land value in the area immediately adjacent to the station but dampened increases because of the requirements for public amenities, facilities, and environmental controls.

Finally, although there has been little mention of citizen participation in the decision-making process, the structure of the impact analysis and evaluation is designed to be able to maintain (even depend on) a continuous dialogue between the technical and political sides of the process. Various groups—neighbors, developers, investors, residents, taxpayers, and modal lobbies—have both positive and negative considerations at stake in the outcome, and they should be encouraged to participate actively in the many choices to be made. The impact evaluation framework provides them with a solid yet flexible basis for debate.

## CONCLUSIONS

An extension of the with-and-without impact methodology framework has been proposed and demonstrated in conjunction with a case study of a rail rapid transit station. The primary intent was to incorporate the policy context as a part of the impact study, and the result was to generate a range of possible outcomes rather than a single impact, each outcome being associated with a matched set of policy conditions. The impact of the station is then the difference between the options available with the station and those available without the station. Although the extended impact framework is still incomplete, it is offered as a step toward improved evaluation of major transit or transportation projects through the analysis of land-use impacts.

## ACKNOWLEDGMENTS

Many helpful suggestions and comments on a previous draft were provided by Kevin Laverty and several anonymous reviewers, and their efforts are greatly appreciated.

## REFERENCES

1. D. E. Boyce. Notes on the Methodology of Urban Transportation Impact Analysis. HRB, Special Rept. 111, 1970, pp. 41-44.
2. A Study of Land Use Planning Alternatives and Recommendations in the Vicinities of West Falls Church, Dunn Loring and Vienna Metro Stations. Kamstra, Dickersen Associates and Office of Comprehensive Planning, Fairfax County, VA, June 6, 1975.
3. W. Z. Hirsch. Urban Economic Analysis. McGraw-Hill, New York, 1973.
4. R. L. Knight and L. T. Trygg. Land Use Implications of Rapid Transit: Implications of Recent Experience. U.S. Department of Transportation, Aug. 1977.
5. S. C. Wheelwright and S. Makridakis. Forecasting Methods for Management. Wiley, New York, 2nd Ed., 1971.

*Publication of this paper sponsored by Committee on Urban Activity Systems.*

# Models for Predicting the Impact of Transportation Policies on Retail Activity

Clifford R. Kern, State University of New York at Binghamton and Cambridge Systematics, Inc.
Steven R. Lerman, Massachusetts Institute of Technology and Cambridge Systematics, Inc.

**Comprehensive urban land-use models designed in the past to predict the effects of large, capital-intensive transportation facilities on the spatial distribution of urban activities are not well suited for predicting the impacts of newer policies to control and manage existing facilities. This paper describes a case study that develops two alternative models with a much sharper, policy-oriented focus and substantially reduced requirements for data and computational resources. The case selected for study involves the hypothetical adoption of transportation control measures to improve air quality in the Denver central business district and the potential impact of controls on retail activity. The two models are a cross-section, lagged-adjustment regression that identifies determinants of aggregate sales at any location and a set of disaggregate travel demand models that predicts the equilibrium between shopping trips and retail activity. The forecasts of both models are consistent in predicting substantial declines in retail activity in response to restrictions on automobile access and negligible offsetting effects of improvements in transit service. It is concluded that compensatory nontransportation measures that enhance downtown amenities or the uniqueness of downtown retail opportunities may offset the negative influence of reduced accessibility.**

With the increasing importance of transportation system management and transportation control plan strategies, transportation planners have been called on to forecast the impacts of new policy options that existing planning tools are ill-suited to simulate. Comprehensive urban activity, or land-use, models present an obvious example of this deficiency. Because they attempt to forecast the spatial distribution of all urban activities for the entire metropolitan area, these models invariably require a large data base, a major calibration effort, and substantial computational resources. Moreover, because of their generality, they often do not specify carefully the behavioral structure that lies behind observed location patterns. As a result, important determinants of location decisions are omitted from the models, and policies that affect these determinants cannot be accurately represented.

This paper demonstrates an alternative approach to activity system modeling that focuses more narrowly on a specific set of policies and a specific activity of interest. The policy selected for study is the adoption of transportation control measures to improve air quality in the central business district (CBD). The impact of interest is the response of retail activity in the CBD, an issue that has raised widespread concern among retailers who fear that transportation controls will undermine their competitive position. The potential impact of transportation controls on retail activity is examined in a case study of Denver, Colorado. [A more detailed discussion of this issue and of the Denver case is available elsewhere (1).] By sacrificing comprehensiveness and focusing on a single policy issue, the modeling strategy described in this paper requires substantially fewer data and resources, incorporates a much richer description of the determinants of retail activity, and forecasts the impacts of a wider range of policies than do comprehensive activity allocation models.

Two separate models have been developed to illustrate this approach. The first, termed the aggregate model, is a cross-section, lagged-adjustment regression that identifies determinants of aggregate sales at any location. It was constructed to make use of statistical skills and data sources that should be commonly available to local planning agencies. The second, the disaggregate model, predicts the destination, mode, and frequency of individual shopping trips and uses these predictions to determine retail activity at any site. Its purpose is to provide a particularly detailed representation of the behavior that underlies shopping travel. The models are first described in detail, and then model predictions for the Denver case are presented.

## AGGREGATE MODEL

The aggregate model is a cross-section regression that specifies retail sales at different locations as a function of market characteristics such as access to households, household income, and noon-hour shopping by nearby workers. An important objective of this methodology is the use of data that are readily available to the local planning agency. In Denver, the best such data are those originally assembled for the calibration of the Denver EMPIRIC activity allocation model. All data used to estimate the aggregate model have therefore been taken from this source. Each observation in the regression model is one of the geographic zones for which the EMPIRIC data are reported. The level of retail sales in each zone, which is not reported directly in the EMPIRIC data set, has been estimated from EMPIRIC information on retail employment by applying sales per employee ratios computed from the Census of Retailing. In cities where information equivalent to the Denver EMPIRIC data has not already been collected, all the information required for the aggregate model could be assembled from readily available sources such as census publications and local transportation network information.

The focus of the model is on estimating the sensitivity of sales to the access of stores to customers. For each retail zone, accessibility to households throughout the metropolitan area is defined for the EMPIRIC model as

$$A_i = \sum_{j=1}^{N} H_j \cdot f(t_{ij}) \qquad (1)$$

where

$A_i$ = accessibility of stores in zone i to households,
$N$ = total number of zones,
$H_j$ = number of households in zone j,
$t_{ij}$ = travel time between i and j, and
$f(\,)$ = a travel impedance function.

The form of the impedance function is the gamma function,

$$f(t_{ij}) = (b^a/\gamma)t_{ij}^{a-1}e^{-bt_{ij}} \qquad (2)$$

where a, b, and $\gamma$ are empirical parameters that have been estimated for the Denver EMPIRIC model to be $a = 3.434$, $b = 0.314$, and $\gamma = 3.0922$.

The stores in zone i do not operate in isolation but in competition with retailers in all other zones in the metropolitan region. Their competitive position depends not only on their accessibility to potential customers but also on the access of competitors to this market. It is useful, therefore, to introduce a slightly modified accessibility measure for retailers in each zone. This modified measure, which has been named competitive accessibility, is simply the access to customers of stores in zone i divided by the average of access measures for all competitive zones:

$$A_i/\overline{A}_j = \left\{A_i/[1/(N-1)] \sum_{\substack{j=1 \\ i \neq j}}^{N} A_j\right\} \qquad (3)$$

Because of arbitrary variations in zonal retail sales caused by variations in zone size, the dependent variable has been defined as sales per zone acre (the models described here were calibrated in U.S. customary units of measurement, and therefore no SI units are given). The estimated relation between sales and competitive access takes the form

$$S_i^* = \alpha_0 (A_i/\overline{A}_j)^{\alpha_1} \qquad (4)$$

where $S_i^*$ is sales in zone i per zone acre. If all zones were equally accessible, sales in each would be $\alpha_0$. The coefficient $\alpha_1$ measures the percentage change in sales with respect to a 1 percent change in competitive accessibility.

Equation 4 implicitly assumes full adjustment of sales to current levels of accessibility. To satisfy this assumption, retail centers must expand in areas where access to households has recently increased and contract where access has recently declined. Households must adjust shopping patterns in favor of centers that have become more accessible to them and those whose growth has created greater shopping opportunities. Such adjustments typically take many years. New stores are not opened immediately nor are existing ones closed in response to changes in demand. Households continue to be governed by shopping habits acquired in the past. These lags in adjustment can be explicitly modeled so that

$$S_i^t = (S_i^*/S_i^{t-1})^{\Theta} \cdot S_i^{t-1} \qquad (5)$$

where

$S_i^t$ = current sales,
$S_i^*$ = fully adjusted sales, and
$S_i^{t-1}$ = sales in the past period.

Current sales depend on the relative size of fully adjusted sales and last-period sales and on $\theta$, which measures the fraction of the disparity that is closed in the current period. If, for example, $\theta = 0.4$ and fully adjusted sales are 10 percent above those achieved in the past, current sales will rise by 4 percent.

Substituting Equation 4 into Equation 5 and taking natural logs yield

$$\ln S_i^t = \Theta \ln\alpha_0 + \Theta\alpha_1 \ln(A_i/\overline{A}_j) + (1-\Theta)\ln S_i^{t-1} \qquad (6)$$

Equation 6 shows the basic form of the regression model, which can be easily estimated by linear regression techniques. In the central role accorded accessibility as a determinant of retail sales, this model resembles earlier work by Lakshmanan (2) and Lakshmanan and Hansen (3, 4). However, the specification employed here is quite dissimilar.

The model actually estimated for Denver is somewhat more complex. To measure the separate effects of automobile and transit accessibility on retail sales, measures of each are included separately as explanatory variables in the regression equation. Both of the accessibility variables are of the form just described. They differ from one another in numerical value because of differences in travel times by automobile and by bus.

As indicated at the beginning of this section, several variables other than access to households are likely to be important in explaining variations in retail sales. These variables are included in the regression to control roughly for forces that operate simultaneously with accessibility to determine retail sales rather than to provide precise measures of their impacts. Therefore, the form of their appearance in the regression has been determined by the requirement that they be easily available from the EMPIRIC data set and that they be compatible with the functional form in Equation 6. Careful specification of functional form is not as important for these variables as for accessibility, for which precise measures of impact are the principal objective of the study.

Because sales depend not only on accessibility to households but also on levels of income in the most accessible markets, income in these markets is the first control variable. For each retail zone, it is measured by the fraction of households in the zone with incomes in the lowest 15 percent of the regional income distribution. The second control, which accounts for the impact of noon-hour shopping by white-collar workers, is the number of service, government, finance, insurance, and real estate workers in the zone. Because this number can vary arbitrarily with zone size, the variable has been defined as white-collar workers per zone acre.

The sales equation used for empirical analysis is

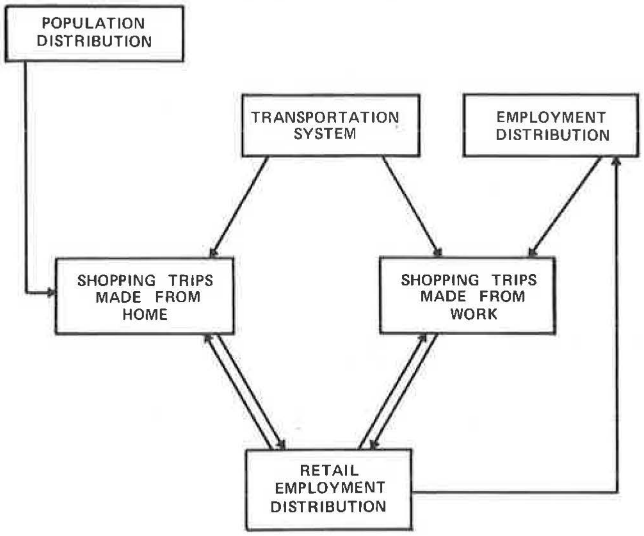
$$\ln S_i^t = \Theta\ln\alpha_0 + \Theta\alpha_1\ln(C_i/\overline{C}_j) + \Theta\alpha_2\ln(T_i/\overline{T}_j) + \delta_1\ln Y_i + \delta_2\ln E_i + (1-\Theta)\ln S_i^{t-1} + e \qquad (7)$$

where

$C_i/\overline{C}_j$ = competitive access to shoppers traveling by automobile,
$T_i/\overline{T}_j$ = competitive transit access,
$Y_i$ = the proportion of households in i with incomes in the lowest 15 percent of the regional income distribution,
$E_i$ = number of white-collar workers in i per zone acre,
e = a random, standard normal error, and
$\alpha_1$, $\delta$, and $\theta$ = parameters.

Empirical results are given in the table below. All coefficients are significantly different from zero (0.01 level) with the expected signs except for transit access, which is both insignificant and incorrect in sign. Because there are no plausible circumstances in which improved transit access should reduce retail sales, the combination of incorrect sign and statistical insignificance argues strongly that the true effect of transit on sales is zero. Therefore, the equation has been re-estimated without transit access as an explanatory vari-

Figure 1. Structure of disaggregate model system.




able, a procedure that constrains its coefficient to be zero. These results are also given in the table:

| | With Transit Access Variable | | Without Transit Access Variable | |
|---|---|---|---|---|
| Variable | Coefficient | t-Ratio | Coefficient | t-Ratio |
| Automobile access (1970) | 0.327 | 3.09 | 0.271 | 2.82 |
| Transit access (1970) | -0.081 | -1.25 | — | — |
| Low-income residents (1970) | -0.330 | -2.92 | -0.351 | -3.15 |
| White-collar workers (1970) | 0.209 | 3.82 | 0.200 | 3.68 |
| Last-period sales (1960) | 0.598 | 13.34 | 0.585 | 13.38 |
| Constant | 3.424 | — | 3.238 | — |

$R^2$ for the model results with and without the transit access variable was 0.834 and 0.833 respectively. The number of observations was 153 in both cases.

The reestimated results can be used to measure the responsiveness of sales in any zone to changes in the travel time for shoppers arriving by automobile. First, note from Equation 7 that the estimated speed of adjustment ($\theta$) can be determined by subtracting the coefficient of last period sales from 1. According to the reestimated estimates given in the table, therefore, if current market conditions imply equilibrium sales 100 percent greater than those of 10 years earlier, 41.5 percent of the full adjustment will have occurred over the decade. This observation implies that the measured coefficient of automobile access, as given in the reestimated results in the table above, represents only 41.5 percent of the true responsiveness of sales to access ($\alpha_1$), a conclusion confirmed by Equation 7. Therefore, a 100 percent increase in automobile access leads to a fully adjusted rise in retail sales of 0.271/0.415 = 0.653 or 65.3 percent. Finally, by using the relation between access and travel time shown in Equations 1 and 2, substituting into Equation 2 the values of the a, b, and $\gamma$ parameters estimated for the Denver EMPIRIC model, and using the actual 1970 travel times of shoppers driving to the CBD, it can be shown that an increase in travel time of 10 percent leads to a 15.6 percent reduction in the automobile access variable. This relation in turn implies that a 10 percent increase in travel time to the CBD will cause an ultimate decline in CBD sales of 0.156 × 0.653 = 0.102 or 10.2 percent. Analogous methods can be used to estimate the sales impact of any policy that restricts automobile access to the CBD.

## DISAGGREGATE MODEL SYSTEM

The disaggregate model system focuses on individual shopping trips and treats retail activity as simultaneously determined with people's shopping choices. It is derived from some of the basic Lowry model concepts but is based on a more sophisticated set of models than has been used previously. The number of shopping trips to each destination is predicted by separate submodels for home-based shoppers and for workers who shop during the noon hour. Shopping trips are linked to aggregate retail activity by two intuitively plausible relations. First, shoppers' choices of shopping destinations are influenced in part by the scale of retail activity at each destination as well as the level of service provided by the transportation system; and, second, the scale of retail activity at any location expands or contracts in accordance with the number of people who choose to shop there. The entire model system is shown in Figure 1.

### Home-Related Shopping Submodel

The home-related submodel used in the disaggregate model system is based on the short-range generalized policy (SRGP) model developed by Cambridge Systematics for the Metropolitan Transportation Commission in San Francisco. This model accepts as input a sample of households and predicts for each household (among other things) the expected frequency of shopping trips and the probability that each trip will go to every potential destination by both automobile and transit. These predictions are then appropriately factored to represent the population as a whole.

The destination-mode choice predictions are produced by a disaggregate choice model. Because such models are based on the decisions of individual households or travelers, they eliminate the need for aggregating various segments of the population either geographically or demographically. Model parameters, therefore, are in theory not subject to aggregation bias. Because they can be estimated by using very small samples, disaggregate choice models also offer the potential for significantly reducing the costs of data collection. However, most importantly, disaggregate choice models are based on a clear, credible, and consistent theory of how decision makers choose among available alternatives.

Choice theory is concerned with the behavior of an individual decision maker confronted with a mutually exclusive set of alternatives from which one and only one can be selected. The individual decision maker n associates some level of utility with each available alternative i. Denote this utility as $U_{in}$, and denote the set of alternatives available to individual n as $D_n$.

According to Lancaster (5), each alternative and decision maker can be characterized by a set of attributes. Thus, the utility of the ith feasible alternative to decision maker n can be expressed as follows:

$$U_{in} = U_{in}(V_i, W_n) \tag{8}$$

where

$U_{in}$ = utility of alternative i to individual n,
$V_i$ = a vector of attributes describing alternative i, and
$W_n$ = a vector of attributes describing decision maker n.

A more convenient expression for the utility function can be developed by defining a vector $Z_{in} = g(V_i, W_n)$, where g is some vector-valued function. Thus, $U_{in} = U_{in}(Z_{in})$.

Each decision maker is assumed to evaluate the attributes of every alternative and select the one yielding the greatest utility. However, since some of the attributes are unobserved, variables are improperly measured, or utility relations are misspecified, it is generally impossible for an observer to determine precisely which alternative any decision maker will select. However, by making suitable assumptions about the distribution of the unobserved elements in the utility function, it is possible to predict the probability with which any alternative will be selected. When each utility is a random variable, the probability that alternative i is selected from any set of alternatives $D_n$ is

$$\Pr(i \mid D_n) = \Pr[U_{in}(Z_{in}) \geq U_{jn}(Z_{jn}) \text{ for all } j \epsilon D_n] \quad (9)$$

Within the class of random utility model forms, the most generally applicable have been what Manski (6) defines as linear in the parameters with additive disturbances (LPAD). In this case, it is assumed that

$$U_{in} = \beta Z_{in} + \epsilon_{in} \quad (10)$$

where $\beta$ is a vector of parameters and $\epsilon_{in}$ is a random variable.

The LPAD form used in this study is the multinomial logit model. This model was chosen for a variety of practical and theoretical reasons including the lack of alternative methods for modeling decision problems with large choice sets and the substantial existing base of successful prior applications. The logit model relies on the assumption that the $\epsilon_{in}$'s are independently and identically distributed with the Gumbel distribution; i.e.,

$$\Pr(\epsilon \leq \omega) = \exp\left\{-[e^{-(\alpha + \omega)}]\right\} \quad (11)$$

By using this distribution, McFadden (7) demonstrates that

$$\Pr(i \mid D_n) = \left(e^{\beta Z_{in}} / \sum_{j \epsilon D_n} e^{\beta Z_{jn}}\right) \quad (12)$$

The parameters of this model can be estimated by maximum likelihood. Such estimates are consistent, asymptotically normal, and asymptotically efficient. McFadden also demonstrates that under relatively weak conditions such estimates exist with probability approaching unity and are unique.

Note that the set of available alternatives $D_n$ can vary from decision maker to decision maker. For example, a traveler without a driver's license or an available automobile would not generally be viewed as having the alternative of driving alone.

In the destination-mode choice component of the home-related shopping model, each feasible mode and shopping destination in the metropolitan area is an alternative available to the household. The utility of each alternative i and the probability of its being selected are determined by the attributes given in the table below. The coefficients given in the table are estimates of the $\beta$'s in the utility function and show how a unit of each attribute affects the utility of any alternative i. Therefore, for example, the utility of any destination-mode alternative is reduced by the travel time and cost that it entails and increased by the scale of retail activity as measured by total retail employment and retail employment density.

| Variable in Logit Model | Estimated Coefficient | t-Ratio |
|---|---|---|
| Constant for automobile mode | 0.797 | — |
| Constant for CBD destinations | 1.184 | 2.21 |
| Dummy for CBD destinations in automobile utility function | -0.946 | -1.74 |
| Number of automobiles divided by household size | 5.330 | 5.95 |
| LN[total travel time (min) x income ($)] | -0.130 | -10.57 |
| Out-of-pocket cost for automobile in cents per mile | -0.021 | -10.08 |
| Transit cost (cents) x household size | -0.022 | -4.55 |
| Density of retail and service employment per acre in destination zone | 0.006 | 4.22 |
| LN (retail and service employment in destination zone) | 0.494 | 13.70 |

The estimates in the table have been derived from data on San Francisco shoppers and applied to predict shopping travel in Denver. The constant for the automobile mode, however, has been recalibrated for Denver by a nonstatistical procedure described below. Therefore, there is no t-ratio for that coefficient. Summary statistics for the model as estimated from the San Francisco data include: $L^*(0)$ (the value of the log likelihood function when all parameters are zero, i.e., when every alternative has the same probability) = -2477; $L^*(\hat{\beta})$ (the value of the log likelihood function at the maximum likelihood coefficient values) = -1610; $\chi^2$ = 1733 [this statistic is equal to $-2\,[L^*(0) - L^*(\hat{\beta})]$, asymptotically distributed as chi square with the number of degrees of freedom equal to the number of parameters estimated, and provides a test against the null hypothesis that all parameters are zero]; NOBS (the number of households in the sample) = 572; NCASES (the number of available alternatives in excess of one per household used in the estimation) = 43 846; and the percentage of households for which the alternative with the highest nonstochastic component of utility was actually selected = 14.3.

The coefficients given in the table above are estimates of the $\beta$'s in the utility function and show how a unit of each attribute affects the utility of any alternative i. Therefore, for example, the utility of any destination-mode alternative is reduced by the travel time and cost that it entails and increased by the scale of retail activity as measured by total retail employment and retail employment density.

In addition to the destination-mode choice predictions, a second component of the home-related shopping submodel predicts the expected frequency of shopping trips from home. The frequency model is a single nonlinear equation in which the daily total of household home-to-shop and shop-to-home vehicle trips (Q) is a function of household size ($X_1$), household income ($X_2$), home-zone retail employment density ($X_3$), and the expected utility of a shopping trip ($X_4$). The functional form of the model, together with the coefficients estimated by nonlinear least squares for the SRGP model, is given by

$$Q = 0.609/\left\{0.0737 + \exp[-0.342X_1 - 0.515X_2 + 0.115(\ln X_3) - 0.527X_4]\right\} \quad (13)$$

Most of the relations described in Equation 13 are self-explanatory. For example, the total number of daily shopping trips increases where household income is high but decreases where the residential zone of the household is characterized by a high density of retail employment. A high level of home-zone retail employment presumably leads households to substitute short shopping trips on foot for the journeys by transit and automobile that are predicted by the model.

The expected utility of a shopping trip ($X_4$), which is positively related to the number of trips, requires further definition. It measures the expected value of the utility produced when a shopper makes a trip to the mode-destination alternative that yields the highest possible satisfaction. Mathematically,

$$X_4 = E[\underset{i \epsilon D_n}{\mathrm{Max}}(U_{in})] \tag{14}$$

It can be shown (8) that, whenever the utility of each individual alternative is a random variable characterized by the Gumbel distribution, as is true in the logit model, the expected utility of a shopping trip can be expressed as

$$X_4 = E[\underset{i \epsilon D_n}{\mathrm{Max}}(U_{in})] = \ln \sum_{i \epsilon D_n} e^{\beta Z_{in}} \tag{15}$$

All the information required to calculate $X_4$ is, therefore, available from the destination-mode component of the home-related shopping submodel. The variable $X_4$ is the critical link between the frequency component and the destination-mode component. Since it includes all of the level-of-service and socioeconomic variables given in the previous table, its inclusion in the frequency model makes the frequency of shopping travel responsive to transportation policy.

To predict home-related shopping travel in Denver, the home-related shopping submodel has been run by using 253 households sampled from the 1971 Denver home interview survey. Household residential locations and alternative shopping destinations are described by a total of 274 separate zones, and forecasting results are summarized at the level of 10 superzones, one of which is the CBD. Sample results have been appropriately factored to represent the entire population.

Two modeling issues must be addressed here. First is the issue of transferability. Can models estimated for San Francisco legitimately be used to predict shopping travel in Denver? Since these models are based on the decisions of individual households, it should be possible to estimate a model in one city and use at least some of the estimated parameters in another as long as the cities have populations with similar tastes. Atherton (9), Atherton and Ben-Akiva (10), and Pecknold and Suhrbier (11) all present evidence that most parameters are transferable among U.S. cities as diverse as Boston, Milwaukee, New Bedford, Philadelphia, San Francisco, and Washington. The only exceptions are constant terms that determine mode shares and total daily shopping trips. Therefore, the automobile constant in the preceding table and the two constant terms in the frequency model have been adjusted so that the aggregate mode shares and the total daily shopping trips predicted for the Denver region match reported Denver values derived from the 1971 home interview survey. For the logit model, this procedure chooses a value for the constant that is the maximum likelihood estimate for the Denver sample, conditional on the values of the transferred parameters.

A second issue arises because shopping destinations specified in the destination-mode model are not the individual shopping centers available to households but groups of shopping centers that represent the sum of shopping opportunities in each destination zone. Estimation of a disaggregate behavioral model is possible when alternatives are grouped as long as the model is structured to guarantee the following property: When any two destinations with identical characteristics are combined, the resulting probability of choosing the combined zone is equal to the sum of the two probabilities for the destinations treated separately. Lerman (12) has shown that this property, termed homogeneity, can be guaranteed in a logit model if the model includes a variable that represents the natural log of group size and if its coefficient is constrained to unity. In the original SRGP destination-mode model, the natural log of retail employment was chosen to measure group size. This choice, however, is unsatisfactory for the home-related shopping model. In the home-related shopping model, retail employment in any zone is determined endogenously by a simultaneous process in which shoppers are attracted to places with high retail employment and varied shopping opportunities and retail employment grows or declines in step with consumer demand. Constraining the retail employment coefficient to unity in the destination-model artificially inflates the number of shoppers who choose destinations where retail employment is high and tends to preclude equilibrium levels of zonal retail employment at values other than zero or infinity. Therefore, for use in the home-related shopping model, the original SRGP destination-mode model has been reestimated without the constraint. It is the set of coefficients from the reestimated model that is shown in the preceding table. Had it been possible to completely respecify the destination-mode model, homogeneity could have been preserved by choosing another measure of group size, such as acreage of the destination zone. Budget limitations precluded this option.

## Work-Related Shopping Trip Submodel

The second source of shopping trips represented in the disaggregate model is noon-hour shopping by workers. To forecast noon-hour trips, a model of work-related shopping trips originally developed for a study of the Bunker Hill area of Los Angeles has been adapted (13). Because the noon-hour model was originally estimated only for workers in the CBD, the work-based shopping model has been applied only to Denver's downtown zones. Noon-hour shopping is generally considered a substantial fraction of total sales only in the CBD.

Like the destination-mode component of the home-related submodel, the work-related model is a joint multinomial logit model. The probability that a worker will select any one of several mode-destination alternatives depends on the utility yielded by the attributes of each. Relevant attributes and coefficients that measure their contribution to utility are given below [the coefficients are taken from the Los Angeles travel demand model (13)]:

| Variable in Logit Model | Estimated Coefficient |
|---|---|
| Automobile mode constant | -0.592 |
| Walk mode constant | 0.115 |
| Minibus mode constant (free fare) | -2.376 |
| Regional bus mode constant ($0.25 fare) | -2.434 |
| Total travel time in minutes | -0.052 |
| Out-of-pocket travel cost for automobile in cents per mile | -0.008 |
| Trip attraction density per acre at destination | 0.032 |
| Employment per acre at origin zone (zero-frequency alternative only) | 0.008 |
| Zero-frequency constant | 8.578 |
| LN (area of destination zone in acres) | 1.000 |

Unlike the home-related shopping analysis, the destination, mode, and frequency of work-related shopping are predicted by a single choice model. During any noon-hour period, each worker may make a single trip, described by a particular mode-destination combination, or may choose not to travel at all. The alternative of no travel (zero frequency) is explicitly included in the set of alternatives.

The work-related model has been applied to each of five zones in the Denver CBD. These zones provide the geographic detail that describes both the location of workplaces and the alternative shopping destinations. Note in the table that this model makes no use of the socioeconomic characteristics of workers; such data are not likely to be readily available for forecasting. For this reason, the model predicts travel behavior for a representative worker at each workplace zone rather than operating on a sample of travelers as in the home-related model. For each workplace zone, it predicts the share of workers who choose each mode-destination alternative and the share who choose to make no trip at all. Total daily shopping travel is obtained by multiplying these shares times the total number of workers at each workplace zone.

Two adjustments have been made in applying the Los Angeles model to Denver. In the Los Angeles model, the attractiveness of each shopping destination is determined by its trip attraction density, which predicts, for example, that the number of shopping trips to any zone depends on the amount of retail floorspace per zone acre. For Denver zones, we have employment figures but no data for floorspace. For use in the model, employment figures have been converted to floorspace by using typical ratios of CBD floorspace per employee estimated from Boston data.

Second, the coefficients of the Los Angeles model imply an unrealistically low value of time for noon-hour shoppers: $0.15/h. This low value resulted from Los Angeles data in which money costs were constant for all trips by each mode except automobile. As a result of the high correlation between cost and mode, coefficients of the cost variable and of the constants representing different modes are poorly specified, although their combined effects are correctly estimated. Evidence on work trips for several cities indicates that workers often value travel time in the range of $3.00 to $5.00/h (11). We have therefore readjusted the coefficients of cost and of each modal constant so that the combined effect of the constant plus cost remains identical to that estimated by the Los Angeles model but the cost coefficient is consistent with a value of time of $3.85/h. This adjustment is embodied in the results given in the table above.

## Tests of Policy and Equilibration Procedure

Before the impacts of new transportation policies can be estimated, both the home-related and work-related shopping models must be run to predict a base number of trips to each destination zone given existing policy. Because both models predict the number of trips rather than the number of shopping visits per trip tour, some average number of shopping visits per trip must be assumed. For work-related shopping, evidence from Los Angeles suggests an average of 1.25 shopping trip ends per work-based shopping tour. To estimate the number of trip ends per home-based trip, we relied on estimates by downtown Denver merchants that the number of shoppers visiting each store is evenly divided between home-based shoppers and workers on noon-hour trips. The disaggregate model system produces this result for CBD destinations if an assumption of two shopping trip ends per home-based trip is made. This assumption has been adopted for all home-related trips. For both kinds of trips, the extra shopping visits that make up the tour are assumed to take place in the same superzone as the original home-based or work-based link of the trip.

To test any policy, the first step of the equilibration procedure is to modify level-of-service variables to reflect the effect of the policy on the Denver transportation system. Then both the home-related and work-related models are run to forecast total shopping trip ends for each zone. These results are aggregated to the 10-superzone level.

The change in shopping trip ends in each superzone is then computed as a percentage of base-year values. The percentage change in zonal retail employment is assumed to be equal to the percentage change in shopping trip ends. This procedure generates new levels of both retail and total employment in each zone. Both the home- and work-related submodels are run again with the revised employment levels as inputs, and the resulting trip-end forecasts are used to further revise retail and total employment. The process is repeated until the changes in trip ends and retail employment in each zone are relatively small.

Trial runs of several iterations of the model have indicated that, for many different policy options, the change in trip ends at each iteration is a constant fraction of the change at the previous iteration. Given this pattern of convergence and an initial change of $\Delta J_i$ in the number of trip ends in zone i, the total change can be approximated by

$$\Delta J_i + \rho\Delta J_i + \rho^2\Delta J_i + \rho^3\Delta J_i + \ldots = \Delta J_i \sum_{j=0}^{\infty} \rho^j = \Delta J_i/(1-\rho) \qquad (16)$$

where $\rho$ is the fraction that relates changes in trips at successive iterations. In accordance with a previously specified assumption, retail employment in any zone is expected to change ultimately by a percentage equal to the total change in shopping trip ends. From Equation 16, this percentage is given by

$$[\Delta J_i/(1-\rho)]/J_i \qquad (17)$$

where $J_i$ is the number of trip ends before the policy change.

This approximation greatly reduces the number of iterations required to operate the entire model system because $\Delta J_i$ is computed at the first iteration and a reasonable estimate of $\rho$ can be obtained in two or three iterations. This approximation has been used for all policy analyses in this paper.

It must be noted that the exact mathematical properties of the iterative equilibration procedure are still unknown and that the limited computational experience obtained in this study is insufficient to reach any definitive, empirically based conclusion about the stability of the model. Trial runs that led to the approximation adopted above indicated some instability after the first several iterations, and changes in the definition of zones and superzones affected the pattern of convergence. The conceptual appeal of the disaggregate system argues strongly for further research into its convergence properties and further refinement of its specification.

## TRANSPORTATION CONTROL POLICIES AND ESTIMATED IMPACTS

Both the aggregate and disaggregate models have been used to estimate the impact of several hypothetical transportation control measures on retail sales in the Denver CBD. These measures include

1. Implementation of an automobile-free zone or elimination of convenient parking spaces so that a shopper arriving by automobile must walk an extra 2.5 min after parking,
2. A similar policy that leads to an extra 5 min of walking after parking, and
3. A 5-min reduction in waiting time for a transit

Table 1. Tests of policy impacts: aggregate and disaggregate models.

| Policy | Total Ultimate Percentage Change in Sales (aggregate model) | Ultimate Percentage Change in Trip Ends and Retail Employment (disaggregate model): Response to Change in Level of Service | Response to Change in Retail Center Size | Total |
|---|---|---|---|---|
| Increase in one-way walk time for parkers | | | | |
| 2.5 min | -17.5 | -5.5 | -24.3 | -29.8 |
| 5.0 min | -32.7 | -9.9 | -32.7 | -42.6 |
| Reduction of 5.0 min in one-way wait time for transit riders | ±0.0 | +0.8 | +4.3 | +5.1 |

trip to or from the CBD (for noon-hour bus trips within the CBD, for which average wait times are currently estimated at 5 min, a reduction of 2.5 min has been assumed).

The results, which are given in Table 1, raise several issues worthy of brief discussion.

Both models forecast a substantial sensitivity of retail sales to reductions in access for shoppers who travel by automobile and negligible impacts for improvements in transit service. When a seemingly insignificant 2.5 min is added to automobile trips to the CBD, the forecasted drop in retail activity ranges from 17.5 to nearly 30 percent. Policies that restrict automobile access and seek to preserve demand for downtown shopping by improving transit service will likely fail and precipitate instead a decline in the downtown retail center. Transit improvements that are politically and financially possible may be more effective offsets to reductions in automobile access in cities like Boston and New York, where transit service is already extensive enough to attract large fractions of riders to downtown destinations. The vast majority of American cities, however, resemble Denver more closely than they resemble Boston or New York.

The availability of two estimates of policy impact obtained by using completely dissimilar techniques illustrates an important benefit of the narrowly focused modeling strategy described in this paper. The ability to model the same policy simultaneously with alternative models is a result of the reduced cost inherent in less comprehensive models. The alternative results can be compared for consistency, and their quantitative forecasts can be used as bounds for policy impacts. If the forecasts produced by different methodologies are at least qualitatively consistent with one another, as in this case, confidence in their accuracy is enhanced.

The structure of the disaggregate model makes it possible to identify two separate sources for the decline in retail activity that accompanies a reduction in automobile access: an initial response to the change in level of service and subsequent adjustments that reinforce the first response as shoppers react to changes in the size of retail centers and the shopping opportunities they offer. The immediate impact of transportation policy may represent only a fraction of the ultimate change in activity. According to the disaggregate model, a 2.5-min increase in the time of an automobile trip leads directly to a 5.5 percent decline in retail activity. However, the chain of further adjustments by shoppers who find the smaller retail center less attractive induces a further decline of 24.3 percent.

For any reduction in automobile access, the aggregate forecast of overall decline in retail activity is substantially smaller than the disaggregate forecast. This disparity may well result from an excessive disaggregate estimate of consumer response to changes in the size of retail centers. Recall that the source of this estimate is the equilibration procedure described earlier in this paper and that both this procedure and its results are subject to the uncertainties mentioned there. Further insight into this issue must await further investigation into the convergence properties of the disaggregate system.

Finally, the impacts described in this paper are those that result solely from transportation policies. Transportation control plans may also include compensatory nontransportation measures that enhance the attractiveness of downtown amenities or the uniqueness of downtown retail opportunities. If reductions in automobile access to downtown retail areas are required to meet environmental standards, careful consideration should be given to such measures so that efforts to improve air quality do not unintentionally further erode the urban core. For example, total automobile restriction in certain areas accompanied by improved pedestrian amenities has, at least in Europe, succeeded in offsetting the negative influence of reduced accessibility. This is only common sense. If shopping in the CBD offers a unique or especially pleasant experience, because of the amenities or the products available there, shoppers should be more willing to bear marginal reductions in convenience to shop downtown.

## ACKNOWLEDGMENTS

This research was performed at Cambridge Systematics, Inc. It was funded in part by the Office of Environmental Affairs, U.S. Department of Transportation, and by the U.S. Environmental Protection Agency. We gratefully acknowledge the valuable contributions of Gregory K. Ingram to the specification of models used in this research and the able research assistance of Ellyn Eder.

## REFERENCES

1. Implementation and Administration of Air Quality Transportation Control Plans: An Analysis of Denver, Colorado. Cambridge Systematics and U.S. Department of Transportation, Aug. 1977.
2. T. R. Lakshmanan. An Approach to the Analysis of Intraurban Location Applied to the Baltimore Region. Economic Geography, Vol. 40, Oct. 1964.
3. T. R. Lakshmanan and W. G. Hansen. A Retail Market Potential Model. Journal of the American Institute of Planners, Vol. 31, May 1965.
4. T. R. Lakshmanan and W. G. Hansen. Market Potential and Its Application to a Regional Planning Problem. HRB, Highway Research Record 102, 1965, pp. 19-41.
5. K. J. Lancaster. A New Approach to Consumer

Theory. Journal of Political Economy, Vol. 74, April 1966.
6. C. F. Manski. The Analysis of Qualitative Choice. Department of Economics, MIT, Cambridge, PhD dissertation, June 1973.
7. D. McFadden. Conditional Logit Analysis of Qualitative Behavior. In Frontiers in Econometrics (P. Zarembka, ed.), Academic Press, New York, 1974.
8. M. E. Ben-Akiva and S. R. Lerman. Disaggregate Travel and Mobility Choice Models and Measures of Accessibility. In Behavioral Travel Modelling (D. A. Hensher and P. R. Stopher, eds.), Croom-helm, London, in preparation.
9. T. J. Atherton. Approaches for Transferring Disaggregate Travel Demand Models. Department of Civil Engineering, MIT, Cambridge, MS thesis, Feb. 1975.
10. T. J. Atherton and M. Ben-Akiva. Transferability and Updating of Disaggregate Travel Demand Models. TRB, Transportation Research Record 610, 1976, pp. 12-18.
11. W. Pecknold and J. Suhrbier. Tests of Transferability and Validation of Disaggregate Behavioral Demand Models for Evaluating the Energy Conservation Potential of Alternative Transportation Policies in Nine U.S. Cities. Cambridge Systematics and Office of Energy Conservation Policy, Federal Energy Administration, April 1977.
12. S. R. Lerman. A Disaggregate Behavioral Model of Urban Mobility Decisions. Department of Civil Engineering, MIT, Cambridge, PhD dissertation, June 1975.
13. Task Termination Report—Internal CBD Travel Demand Modeling. Barton-Aschman Associates and Cambridge Systematics, Aug. 1976.

*Publication of this paper sponsored by Committee on Urban Activity Systems.*

# Optimizing Urban Mass Transit Systems: A General Model

Alan Black, Graduate Program in Planning, University of Texas at Austin

**This paper describes a model for determining the general dimensions of an optimal mass transit system for an idealized urban area. The model is based on a circular city with a definite center and with density declining uniformly from the center in all directions according to the negative exponential function. The transit system consists of radial routes that emanate from the center and contain discrete stops. Only trips to or from the center are considered, and travel is assumed to occur only in radial and circumferential directions. The model represents total community costs of the system, defined to include travel time, operating costs, equipment, and construction. A recursive procedure was devised to find a simultaneous minimum with respect to the spacing of routes, number and spacing of stops on each route, and average headway. Numerical analyses were conducted for six hypothetical cities by using varying values for the parameters of the density function. In each case, three types of transit systems were compared: conventional bus service, buses on exclusive lanes, and rail rapid transit. The optimal system in the largest city examined was exclusive bus lanes; in the other five cases, the optimal system was conventional bus service. Other interesting relations that appeared in the results are summarized.**

The United States has entered a new era of massive investment in urban mass transit, prompted by the willingness of Congress to authorize billions of dollars in federal aid for local transit systems. However, there is yet no systematic procedure for allocating these resources and determining whether a transit proposal is worthwhile. Each proposal is evaluated on an ad hoc basis, and considerable weight is given to the zeal of the proponents, political pressures, and the current availability of funds. Choice of technology has become a major issue in many areas, and the question of whether medium-sized cities should proceed with huge investments in fixed-guideway transit systems is particularly controversial.

This paper summarizes a dissertation aimed at determining the dimensions of an optimal transit system for an idealized urban area (1). The approach was to hypothesize a circular city with a definite center and with density declining uniformly from the center in all directions. The transit system consists of routes that emanate from the center and contain discrete stops. By use of integral calculus, a model was derived that represented the total community costs of building and using such a system. By use of differential calculus, a procedure was developed to optimize the principal design variables in the system: the number of radial routes, their length, and the number and spacing of stops on each route. Numerical analyses compared three common forms of conventional transit: buses on city streets, buses on exclusive lanes, and rail rapid transit.

Such an abstract model cannot be mechanically applied to the complex, irregular pattern of a real city. Abstraction is an unavoidable compromise if a model is to be made mathematically tractable. Similar approaches have been followed in many previous studies of transit optimization. A few of these will be cited here; a fuller review can be found elsewhere (1).

Most previous studies can be divided into two geometrical approaches. One of these assumes a gridiron network of transit routes laid on a homogeneous infinite city, usually with a uniform density of trip ends. What may have been the first study of this type was done by Creighton and others (2) and involved both highway and transit grids; the object was to find the optimal combination of investment in the two modes. Holroyd (3) assumed a single grid of bus routes and derived a solution for the optimal spacing of routes and frequency of service. Two dissertations, one by Mattzie at Carnegie-Mellon (4) and the other by Woodhull at Rensselaer Polytechnic (5), also dealt with grid systems of transit routes.

The second approach is to examine a single transit line. Often one terminal is assumed to be in the central business district (CBD), and only trips to or from this terminal are considered. The object is usually to find the optimal spacing of stops. Some studies have assumed uniform spacing, but the more interesting have allowed for variable spacings. An early study of variable spacing of stops was made by Schneider (6). Vuchic (7) later did a fuller analysis of the problem. Both researchers assumed a constant density of boarding passengers along the line and came to the conclusion that interstation spacings should increase as one approaches the CBD. This is opposite to conditions normally found on radial rail routes in which the spacings become smaller near the CBD.

A third geometrical approach was taken by Byrne (8), who assumed a circular city of given radius and, for the case in which population density varies only with distance from the center, derived a solution for the optimal number of radial routes. Byrne presented results for four density functions including the negative exponential. The model described here has a similar geometry and also uses the negative exponential function, but it involves simultaneously solving for the optimal number of radials and optimal length of radials as well as other variables.

## DESCRIPTION OF THE MODEL

The hypothetical city is a complete circle, uniform throughout its 360 degrees, uninterrupted by barriers or irregularities, and extending to infinity. The city has a center that is taken to represent the CBD. The transit network consists of an unknown number of radial lines that emanate from the center and extend an unknown distance. Each line has discrete stops, and access is possible only at these fixed points, which must be determined. Because of the assumed symmetry of the city, the radial lines are equally spaced. Each has the same number of stops, spaced in the same way, and is of the same length.

The transit system serves a given amount and distribution of travel demand. Demand is highest at the center and declines with increasing distance from the center. Demand is assumed to be constant (i.e., land use, trip generation, and modal split are held constant). Every trip must be made, and the only alternative to using transit is walking.

To make the problem mathematically manageable, only trips to or from the CBD (assumed to be trips to or from the point at the center) were considered. In most cities, such CBD-oriented trips represent roughly half of all transit trips. This is the largest market for transit, and it dominates the design of transit networks.

It was assumed that travel can occur only in radial or circumferential directions but that otherwise it can occur anywhere on the city's surface. There are no circumferential transit routes; all circumferential travel is on foot. Each inbound traveler starts from his or her origin and walks in a circumferential arc to the nearest radial transit route. There the traveler has a choice between walking inward or outward until reaching a stop. He or she chooses the stop that minimizes the total time from origin to destination. Close to the center, some travelers find it faster to walk all the way and not use the transit service.

It was assumed that inbound and outbound trips are equal in number and form mirror images of each other. On each radial transit route, vehicles shuttle back and forth between the center and outer terminal, stopping at all designated stops.

### Objective

The objective selected was to minimize total community costs, measured in dollars. Total costs were defined to include capital investment for guideway and vehicles, operating costs, and the door-to-door travel time of travelers. It would be desirable to include other types of community costs, such as externalities and intangibles, but they were omitted because of the difficulty of measuring them or converting them to dollar values.

The approach to measuring time costs was to calculate the distance traveled from door to door and to estimate average speed on the portions of the route traversed. There are several components in door-to-door travel time. One of these is the time spent in the transit vehicle, which must be divided into two parts. The first is the time that would accrue if the vehicle moved at its cruising speed from the rider's point of embarkation to the point of debarkation. The second component consists of the additional time penalties incurred when the vehicle accelerates, decelerates, and waits at stops to load and discharge other passengers.

Two other components of door-to-door travel time were included. One is walking time—the time it takes a traveler to walk from origin to boarding stop or from where he or she gets off to his or her final destination. The other is waiting time, which depends on the scheduling of vehicles. It was assumed that average waiting time is half the scheduled headway (the time between successive buses or trains).

Operating costs depend on a number of factors, but here they were based solely on vehicle kilometers. This appears to be the most significant relation and also the simplest.

There are two major types of capital costs: fixed facilities (such as roadbed, structures, and stations) and running equipment (buses and trains). In analyzing a specific proposal, detailed estimates of capital costs are based on engineering drawings. This cannot be done for a hypothetical city; hence, the cost of fixed facilities was based on kilometers of guideway and number of stations. Equipment costs were estimated on the basis of the number of buses or train cars required to serve peak-period demand plus an allowance for vehicles out of service.

The daily time and operating costs and the one-time investment costs were put on a comparable basis through the annual cost method although it is actually average weekday costs that are represented in the model. For capital costs, an expected life span and interest rate were assumed, and the equivalent annual cost was calculated. This was converted to average weekday cost by assuming a number of weekday equivalents for a year.

### Decision Variables

The most important decision variables in designing such a system are

1. The number of radial routes (N),
2. The number of stops on each radial route (z),
3. The length of each radial route ($x_z$), and
4. The spacings between stops on each route.

The spacings between stops are implied in the set of variables $x_1$, $x_2$, $x_3$, ..., $x_z$, where $x_i$ is the distance from the city center to the ith stop. Thus, the spacing between the second and third stops is given by $x_3 - x_2$.

Another variable—scheduling of service—is also within the control of the transit authority. There is an important relation between frequency of service and route spacing, and they should be optimized simultaneously. Frequency of service was represented by average headway over the full day (h).

## Role of the Density Function

To develop the model, one must know the locations of the trip ends on the surface of the city. Each trip has one end at the center, but the other end is elsewhere. The approach was to adopt a function that relates the density of these outer trip ends to distance from the city center.

There remained the question of what function to use. A considerable body of literature, starting with the landmark article by Clark (9), suggests that the negative exponential function represents the relation between population density and distance from the city center. However, few studies have dealt with the density of trip ends. Furthermore, a review of the literature—described elsewhere (10)—revealed that there is now much debate over the proper function for population density. Researchers have used a variety of equations and obtained close fits to empirical data with many of them.

Therefore, some empirical research was conducted by using population and travel data from transportation studies of 12 metropolitan areas of the United States that range in size from New York to Syracuse. Regression analysis was used to fit data on six population and trip-end variables to four alternative equations: linear, exponential, power curve, and normal curve. The findings are described elsewhere (11).

Of particular concern was the density of CBD-oriented transit trip ends. Data for this variable were available only for six medium-sized cities. The analysis showed that the exponential function yielded the highest correlation coefficient for four of the cities. For the other two cities, the normal curve gave the highest correlation but the exponential function was almost as good. In all cases, the exponential function had a high correlation that ranged from 0.880 to 0.993.

As a consequence, the negative exponential function was selected for inclusion in the model. The specific equation is

$$D(r) = Ae^{-br} \quad (1)$$

where

$D(r)$ = density of trip ends at distance r,
r = distance from the center,
e = base of natural logarithms, and
A and b = parameters.

## DERIVATION OF THE MODEL

The model consists of a single equation, derived by integral calculus, that represents the total community costs of the transit system. Deriving the equation involves summing up the kilometers traveled, minutes spent in travel, and costs for the entire city. Because of the assumed circular nature of the city, the problem lends itself to the polar coordinate system rather than the Cartesion coordinate system. In the polar system, any point $r, \theta$ is identified by its distance from the origin r and the angle $\theta$ between a ray from the origin and a given axis.

Each stop on a route draws travelers from a tributary area that can be drawn on a map. There is also a circular area around the center from which people walk to the center and do not use transit. Therefore a route with z stops serves a sector that is divided into z + 1 tributary areas. This is shown in Figure 1 for a route with only three stops (plus the CBD terminal).

The boundary between the tributary areas of adjacent stops must be determined by finding the point at which a traveler is indifferent; that is, the total travel time to the CBD is equal whichever stop is used. This calculation is incorporated into the model. The variable $g_i$ represents the distance from the center to the boundary between those who walk in to stop i - 1 and those who walk out to stop i.

Describing the derivation of the entire equation would be impossible in this space. Therefore, the approach is illustrated by deriving the person kilometers traveled on transit vehicles. This is done for inbound travelers for a system with three stops, as shown in Figure 1. This requires determining the number of passengers who board at each stop and multiplying by the distance from the stop to the center. For the first stop, it is necessary to integrate the density function between the inner and outer boundaries of the tributary area ($g_1$ and $g_2$) and to multiply this by $x_1$. The resulting derivation is

$$\int_0^{2\pi} \int_{g_1}^{g_2} Ae^{-br} x_1 r\,dr\,d\theta = (2\pi A x_1/b^2)\,[e^{-bg_1}(1 + bg_1) - e^{-bg_2}(1 + bg_2)] \quad (2)$$

Notice that this is integrated over the full circle; it represents the travel by all persons who board at the first stop on all radials. Since the city is radially symmetrical, the behavior on each radial is identical, and there is no point in calculating them separately.

For passengers who board at the second stop, the only differences are that the limits of the tributary area are $g_2$ and $g_3$ and each passenger rides $x_2$. Hence, the expression for person kilometers has the same form as Equation 2, or

$$\int_0^{2\pi} \int_{g_2}^{g_3} Ae^{-br} x_2\, r\,dr\,d\theta = (2\pi A x_2/b^2)\,[e^{-bg_2}(1 + bg_2) - e^{-bg_3}(1 + bg_3)] \quad (3)$$

For persons who board at the third stop, the expression is

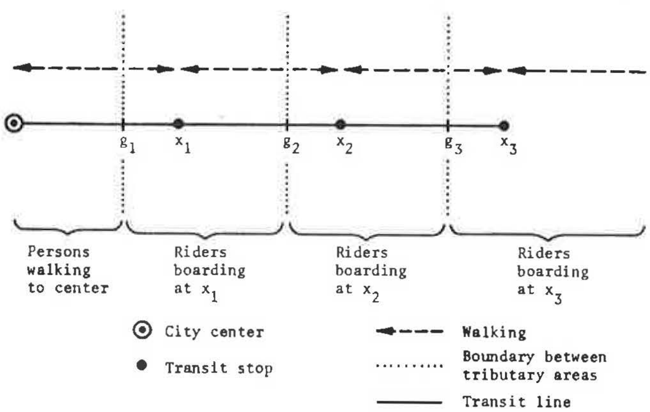
$$\int_0^{2\pi} \int_{g_3}^{\infty} Ae^{-br} x_3 r\,dr\,d\theta = (2\pi A x_3/b^2)\, e^{-bg_3}(1 + bg_3) \quad (4)$$

Adding these together, we get the total person kilometers traveled by transit:

$$(2\pi A/b^2)\,[e^{-bg_1}(1 + bg_1)(x_1) + e^{-bg_2}(1 + bg_2)(x_2 - x_1) + e^{-bg_3}(1 + bg_3)(x_3 - x_2)] \quad (5)$$

This is for a system with just three stops on each radial. For a larger system, there would be more terms since it is necessary to make a separate derivation for each

Figure 1. Relation between tributary areas and locations of transit stops.



stop. However, all the expressions are identical in mathematical form; only the designation of the variables must be altered. Thus, the equation for a three-stop system can be extrapolated to a system with more stops.

The derivation proceeded in a step-by-step (or stop-by-stop) fashion. The most difficult part was determining the kilometers walked in a radial direction. This had to be done separately for those who walk inward and those who walk outward. For a system with z stops, it was necessary to derive expressions for 2z + 1 separate areas. But again, all the expressions turned out to be identical in mathematical form.

Kilometers traveled was converted to person minutes of travel time by applying appropriate values of average speed. Waiting time and time for delays from stops were added to obtain total person minutes of travel time. This was converted to dollars by applying the assumed monetary value of travel time. Then expressions for operating, construction, and equipment costs were derived and added to get total community costs.

When the expressions for all cost components are added together, a considerable amount of cancellation and simplification is possible. The total cost equation for a system with three stops is as follows (certain portions of the calculations given in this paper were done in U.S. customary units, and in these instances no SI units are given):

$$\begin{aligned} y = \{ & (4\pi A c_2 t/b^3)\,[1 + e^{-bx_1}(2 + bx_1) + e^{-bx_2}(2 + bx_2) \\ & + e^{-bx_3}(2 + bx_3) - e^{-bg_1}(2 + bg_1) \\ & - e^{-bg_2}(2 + bg_2) - e^{-bg_3}(2 + bg_3)]\} \\ & + (2\pi^2 A c_2 t/N b^3) + (2NKqx_3/h) \\ & + [(2mNV/ph)(c_1 x_3 + 3d + L)] + (INx_3 + 3JN) \end{aligned} \quad (6)$$

where

- $y$ = total costs,
- $x_i$ = distance from center to i th stop (km),
- $g_i$ = distance from center to boundary between tributary areas for stop i - 1 and stop i (km),
- $c_2$ = walking speed (min/km),
- $t$ = value of travel time (dollars/min),
- $N$ = number of radial routes,
- $K$ = length of daily transit service period (min),
- $q$ = operating cost per vehicle mile (dollars),
- $h$ = headway between buses or trains (min),
- $m$ = spare vehicle factor,
- $V$ = equivalent daily cost of a vehicle (dollars),
- $p$ = ratio of peak-period headway to average all-day headway,
- $c_1$ = cruising speed of transit vehicle (min/km),
- $d$ = delay for a stop (min),
- $L$ = layover time (min),
- $I$ = equivalent daily cost of a mile of guideway (dollars), and
- $J$ = equivalent daily cost of a station (dollars).

The right-hand side of Equation 6 is divided into five parts, each of which has a recognizable significance, so that the equation can be rewritten in verbal form as follows: Total costs = radial travel, delay, and waiting time + circumferential travel time + operating cost + equipment cost + construction cost.

## FINDING THE OPTIMAL SOLUTION

The total cost equation for a system with z stops on each radial has z + 2 decision variables. These are N, h, and the set of z variables that represent the distance from the center to each of the stops ($x_1$, $x_2$, $x_3$, . . . , $x_z$). $x_z$ is also the length of the route.

To find a global minimum for all the variables, one starts by taking the partial derivative of the equation with respect to each variable and setting each result equal to zero. This yields a set of z + 2 equations that contain z + 2 unknowns. The next step is to find a simultaneous solution for the z + 2 equations that will specify the optimal solution. But all of the equations are nonlinear, and no general analytical method exists for finding the simultaneous solution to a set of nonlinear equations. Consequently, an approximating procedure was developed to find the optimal solution.

The set of equations does have a special structure that can be exploited to develop a recursive procedure. Most of the equations contain three unknowns, but there is one that contains only two. The equations can be sequenced so that, by assuming a value for one variable, values can be calculated for all the other variables. When this has been done, one equation remains. Inserting the previously calculated values in this equation provides a check on whether a simultaneous solution has been obtained.

Usually a simultaneous solution will not occur at first because the process began by guessing the value for one variable. Therefore, the recursive procedure must be embedded in a search procedure to find the right starting value. A classical one-dimensional search technique known as the regula falsi method was adopted for this purpose. This involves making two trials with arbitrarily selected values, comparing the results, and calculating a "best guess" for a third trial. The method continues with successive trials, always comparing the results of the two previous trials, until convergence is reached. The recursive procedure and the regula falsi method were implemented in a computer program that proved efficient in approximating a simultaneous solution for the set of equations.

There is a further, unusual dimension to the problem. To carry out the procedure just described, one must specify the number of stops because this determines the number of unknowns and equations. However, the number of stops is itself an unknown of some importance. Thus, at the start, the number of equations to be solved is unknown! This problem was handled by embedding the above procedure in an overall search for the optimal number of stops. This was done by using the Fibonacci search method, which successively eliminates groups of integers until it locates the integer that gives the optimum.

## NUMERICAL ANALYSIS

The computer program was used to calculate optimal transit systems for a number of hypothetical cases. Two interests were of primary importance in the selection of the tests:

1. Choice of technology—What is the best transit mode for a particular city? Tests were run for the three best known types of conventional transit service: (a) ordinary bus service in which buses run in mixed traffic on surface streets, (b) exclusive bus lanes, and (c) rail rapid transit. These are referred to as local bus, busway, and rail.
2. Impact of the density profile—What effect does the density of trip ends have on the optimal transit system? Tests were made for six hypothetical cities with different values for parameters of the density function. The three transit modes were compared for each city so that optimal transit systems were calculated for 18 cases.

The table below gives the regression and correlation coefficients obtained for the exponential function for six metropolitan areas:

| City | A | -b | -r |
|---|---|---|---|
| Detroit | 3427 | 0.286 | 0.966 |
| Cleveland | 4043 | 0.275 | 0.992 |
| Pittsburgh | 2345 | 0.355 | 0.880 |
| Buffalo | 2145 | 0.384 | 0.974 |
| Rochester | 2705 | 0.724 | 0.993 |
| Syracuse | 1285 | 0.632 | 0.935 |

The value of A is the density of trip ends per square mile at the city center (these numbers represent only the outer ends of trips, which is why they seem low). The value of b indicates the rate of decline in the density of trip ends per square mile with each mile of increased distance from the city center. The higher the absolute value of b, the more compact the city.

It was decided to make tests with combinations of two A values (2000 and 4000) and three b values (0.25, 0.50, and 0.75). Each city is given a "name" that signifies the two parameter values: for example, city 4/25 is the case where A = 4000 and b = 0.25. The total number of person trips made in each city can be determined from the following result:

$$\int_0^{2\pi}\int_0^{\infty} Ae^{-br}\, r\,dr\,d\theta = (2\pi A/b^2) \qquad (7)$$

Table 1. Dimensions of optimal local bus systems.

| City | Number of Stops | Length of Each Route (km) | Average Spacing (km) | Number of Radials | Total Route (km) | Headway (min) |
|---|---|---|---|---|---|---|
| 4/25 | 19 | 25.42 | 1.34 | 60 | 1526.2 | 8.60 |
| 2/25 | 19 | 25.43 | 1.34 | 43 | 1082.8 | 12.20 |
| 4/50 | 13 | 12.54 | 0.96 | 22 | 273.6 | 12.50 |
| 2/50 | 13 | 12.55 | 0.97 | 16 | 195.0 | 17.80 |
| 4/75 | 10 | 8.27 | 0.83 | 12 | 101.3 | 15.80 |
| 2/75 | 9 | 7.72 | 0.86 | 10 | 75.6 | 20.60 |

Note: 1 km = 0.62 mile.

The person-trip totals for the six hypothetical cities are given below:

| City | Person Trips |
|---|---|
| 4/25 | 402 124 |
| 2/25 | 201 062 |
| 4/50 | 100 531 |
| 2/50 | 50 265 |
| 4/75 | 44 680 |
| 2/75 | 22 340 |

Numerical analyses also require specifying values for a large number of parameters that mostly represent cost and performance characteristics of the transit system. The literature was surveyed to ascertain reasonable values for all parameters. Many of these depend on the transit mode being analyzed, but some (such as the interest rate—assumed to be 10 percent—and the value of travel time—set at \$2.40/h) are common to all modes.

## Local Bus System

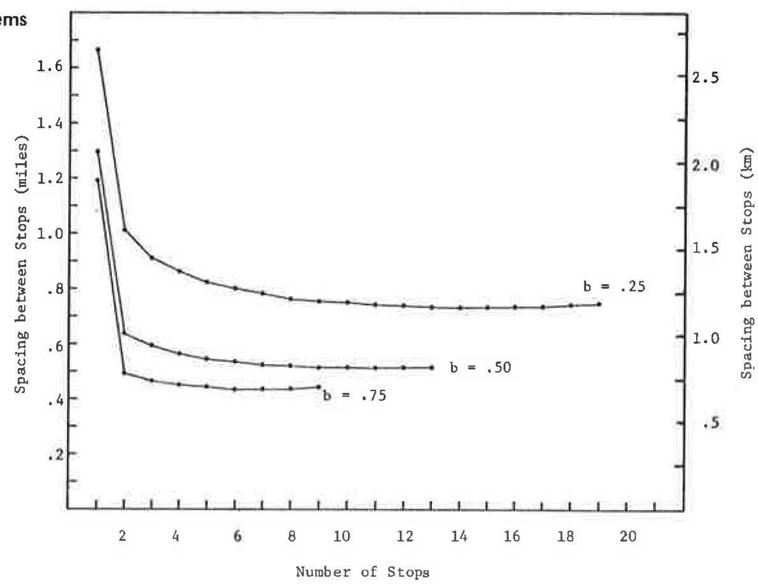
Dimensions of the optimal systems for the six cities are given in Table 1. As one would expect, the larger the city is, the larger is the optimal transit system. The optimal number of stops ranges from 9 to 19; the optimal length of each route from 7.72 to 25.43 km (4.8 to 15.8 miles); and the optimal number of radials from 10 to 60. Thus, there is considerable variation in the optimal values, which clearly depend on the density profile.

The value of A has little effect on the optimal number and spacing of stops or the optimal route length. The value of b has much more influence: The more compact the city is, the shorter are the routes, the fewer are the stops, and the closer is the spacing between stops.

Where A does have an impact is on the number of radials (although the b parameter remains dominant). City 4/25 has twice as many trips as city 2/25, and the optimal number of radials increases from 43 to 60. This suggests that the response to a uniformly distributed increase in demand should be to increase the number of routes rather than the frequency of stops on existing routes.

The optimal value for average headway ranges from 8.6 to 20.6 min; it is influenced by both the A and b values. When the value of A is doubled, optimal headway

Figure 2. Spacing between stops for optimal local bus systems for cities with A = 2000.



is reduced by less than half. This suggests that, when bus trips are added to handle increased demand, they should be distributed between existing routes and newly created routes (which will reduce both walking and waiting).

The pattern of spacing between stops is also of interest. The same pattern was found in all cases and persisted in the busway and rail alternatives. To illustrate this pattern, the interstop spacings for the three cities with A = 2000 are shown graphically in Figure 2. The pattern has the following features:

1. Starting from the center, the spacing decreases outward to a point about four-fifths the length of the route.
2. From this point to the outer terminal, the spacing gradually increases.
3. The first stop has a much larger spacing than any other, but variation among the others is very slight. One could generalize by saying that, except for the first stop, the spacing should be approximately uniform. This seems more realistic than the optimal spacing pattern derived by Schneider (6) and Vuchic (7).

## Busway System

The principal difference in the busway system is that it involves construction costs for guideway and stations in return for which the buses achieve higher speeds and lower operating costs. Table 2 gives the dimensions of the optimal busway systems. These are substantially smaller than the local bus systems for the corresponding cities in terms of number of stops, number of radials, and total kilometers of route.

The average spacing between stops is greater in all cases than for local bus. This is a concomitant of faster bus speeds: the faster the speed of transit service is, the farther people will walk to use it. The delay for a stop also increases, which further increases the spacing.

The value of A has more impact on the number of stops and on route length than it does in the case of local bus. This is undoubtedly because each station and kilometer of route entails a construction cost. Doubling the number of trips distributes this capital cost more and justifies a higher level of investment.

Optimal headways are much lower for busway than for local bus. This is apparently because the busway involves high costs for stations and route kilometers (versus zero cost for local bus) but the operating cost per vehicle kilometer is much lower. The outcome is a smaller route structure with better service.

**Table 2. Dimensions of optimal busway systems.**

| City | Number of Stops | Length of Each Route (km) | Average Spacing (km) | Number of Radials | Total Route (km) | Headway (min) |
|---|---|---|---|---|---|---|
| 4/25 | 13 | 24.48 | 1.88 | 16 | 401.6 | 2.35 |
| 2/25 | 11 | 22.64 | 2.06 | 12 | 272.7 | 3.45 |
| 4/50 | 7 | 10.19 | 1.46 | 9 | 87.1 | 4.90 |
| 2/50 | 5 | 8.75 | 1.75 | 6 | 56.5 | 7.40 |
| 4/75 | 4 | 5.64 | 1.41 | 6 | 33.3 | 7.60 |
| 2/75 | 3 | 4.37 | 1.46 | 5 | 21.0 | 8.50 |

Note: 1 km = 0.62 mile.

## Rail System

The principal difference between the busway and rail alternatives is that rail is more capital intensive. For rail, the costs of building a station and a route kilometer were assumed to be twice as great. The cost of a railcar was assumed to be more than four times greater than the cost of a bus.

Table 3 shows the dimensions of the optimal rail systems. Each is smaller than the optimal busway system for the same city. The number of stops and the length and the number of radials are all reduced. The slightly greater average spacing between stations results from an assumed higher cruising speed and larger delay for a stop with the rail alternative.

The rail model was slightly different from the one used for the bus alternatives in that it included a feature that also optimized the average length of trains. The results are given in the last column of Table 3, which shows that, in cities with greater total demand, trains should be longer. This in itself is not surprising, but note that optimal headway is not greatly reduced in the larger cities. This suggests that greater demand should be handled by running longer, rather than more frequent, trains.

## Comparison of the Three Modes

The busway had the least total cost for the largest city examined—city 4/25. The local bus system had optimal cost for the other five cities. The rail system had the highest cost in all cases.

The computer program calculates many other characteristics of the optimal systems. It is of interest that in two cases the busway system had the highest average travel speed from origin to destination, whereas the local bus system was highest in the other four. The rail alternative turned out poorly in this regard because walking was assumed to be the only mode of access, and walking distances were quite high. This suggests the importance of supplementing rail lines with feeder bus service.

The optimal headways are of interest because increasing the frequency of service is a common policy objective. The busway system was best in all cities, and rail was uniformly second. This again demonstrates that when there is a construction cost the optimum produces a small route structure with frequent service. When there is no construction cost, the route structure is much larger, and service on each route is less frequent.

**Table 3. Dimensions of optimal rail systems.**

| City | Number of Stops | Length of Each Route (km) | Average Spacing (km) | Number of Radials | Total Route (km) | Headway (min) | Average Number of Cars per Train |
|---|---|---|---|---|---|---|---|
| 4/25 | 11 | 23.85 | 2.17 | 11 | 265.0 | 8.55 | 3.58 |
| 2/25 | 9 | 21.17 | 2.35 | 8 | 179.3 | 10.00 | 2.75 |
| 4/50 | 5 | 8.64 | 1.73 | 6 | 55.7 | 8.00 | 1.45 |
| 2/50 | 4 | 6.97 | 1.74 | 5 | 35.8 | 9.05 | 1.03 |
| 4/75 | 3 | 4.40 | 1.47 | 5 | 20.9 | 9.20 | 1.00 |
| 2/75 | 1 | 1.93 | 1.93 | 5 | 9.7 | 9.70 | 1.00 |

Note: 1 km = 0.62 mile.

## CONCLUSIONS

The model appears to give a reasonable representation of total costs for different types of transit systems in cities with different density profiles. The results indicate considerable sensitivity to the form of transit service and the parameters of the density function. There are weaknesses in the current formulation; improvements and extensions are certainly possible. It would be desirable to make transit demand (the number of trips) and land-use configuration (implied by the density profile) sensitive to the provision of transit service, to include trips not going to or from the CBD, and to add some type of feeder routes to the radial lines.

The study indicated that an areawide rail transit system, without supplementary conventional bus service, is less economical than an areawide busway system with the same limitation within the range of density parameters examined (roughly those of medium-sized American cities). This finding depends on the values assumed for the cost and performance parameters, especially construction cost. Some medium-sized cities may contain sectors that have atypically high densities that would justify a rail line. There also may be situations in which alignments can be obtained at unusually low cost—perhaps underused railroad rights-of-way or the median strips of freeways. Any situation that involves atypically low costs for land acquisition and construction is more likely to warrant a rail line. This also applies to the busway system, which proved more expensive than conventional bus service in five of the six hypothetical cities.

It is surprising that ordinary bus service did so well in the comparison. This was largely because a dense network and close spacing of stops produced substantially shorter walking distances than did the alternatives. This underlines the importance of complementing high-speed main-line facilities with a pervasive feeder system or parking facilities at stations or both.

Both the A and b parameters of the density function affect the optimal transit system, but the b parameter has much more influence. Its major impact is on the length of radial routes and the number of stops. The average spacing between stops did not vary much from city to city.

Factors that vary in response to the A parameter can be interpreted as sensitive to scale. The results indicated that some economies of scale exist, but they are not overly significant. They seem to be largest when construction costs are involved.

Historical evidence is conclusive that the values of A and b have been declining in cities all over the world, which means that cities are becoming more dispersed and less centralized. It is noteworthy that the only city in which the busway alternative was optimal had the lowest value of b (city 4/25). Therefore, the historical decline of b values—although certainly related to increasing use of the automobile—does not necessarily spell doom for fixed-guideway transit systems. What happens is that trip lengths become longer, which makes it more worthwhile to introduce high-speed capital facilities.

This view conforms with the understanding of transportation planners. In most cities, the total number of trips to and from the CBD has remained fairly constant for years. However, homes are moving outward, and people are coming to the CBD from farther and farther away. This means that some radial transit improvements that could not be justified in the past may be warranted in the future.

## REFERENCES


1. A. Black. Optimizing the Design of Urban Mass Transit Systems: A General Model. Cornell Univ., PhD dissertation, 1975.
2. R. L. Creighton and others. Optimum Investment in Two-Mode Transportation Systems. HRB, Highway Research Record 47, 1964, pp. 23-45.
3. E. M. Holroyd. The Optimum Bus Service: A Theoretical Model for a Large Uniform Urban Area. In Vehicular Traffic Science (L. C. Edie, R. Herman, and R. Rothery, eds.), American Elsevier, New York, 1967, pp. 308-328.
4. D. E. Mattzie. Ubiquitous Rapid Transit Service in an Urban Area. Carnegie-Mellon Univ., PhD dissertation, 1968.
5. J. Woodhull. Urban Public Transportation: Specification of Service Parameters for a Two-Mode Cooperative Hierarchy. Rensselaer Polytechnic Institute, PhD dissertation, 1970.
6. M. Schneider. Travel Time Minimization– Transit. Chicago Area Transportation Study Research News, Vol. 4, June 2, 1961, pp. 10-16.
7. V. R. Vuchic. Interstation Spacings for Line-Haul Passenger Transportation. Univ. of California, Berkeley, PhD dissertation, 1966.
8. B. F. Byrne. Public Transportation Line Position and Headways for Minimum User and System Cost in a Radial Case. Transportation Research, Vol. 9, 1975, pp. 97-102.
9. C. Clark. Urban Population Densities. Journal of Royal Statistical Society, Vol. 114, 1951, pp. 490-496.
10. A. Black. Urban Density Profiles: A Guide to the Literature. Council of Planning Librarians, Monticello, IL, Exchange Bibliography 1160, 1976.
11. A. Black. Urban Density Profiles in Mass Transit Planning. Paper presented at 22nd North American Meeting of Regional Science Association, Nov. 1975.


*Publication of this paper sponsored by Committee on Urban Activity Systems.*

# Making the Concept of Equity Operational

Douglass B. Lee, Jr., Institute of Urban and Regional Research, University of Iowa

**In an effort to improve communication among transportation planning professionals and with the public, definitions of "equity"—a term commonly used by professionals, politicians, and citizens in discussion of planning issues—and related concepts are proposed. Two examples are offered to show how horizontal and vertical equity can be made operational. Supposed trade-offs in transportation between efficiency and equity are also explored, and it is concluded that, contrary to conventional wisdom, they are more often complementary than conflicting.**

Rather than invent entirely new words to round out their jargon, technical professions often borrow common words that have meanings somewhat related to the technical concept in need of a name. One of the problems in doing this is that a word such as "equity" is used by many people in discussing the same issue but with little overlap in meaning and, hence, limited communication. For many people, equity refers to their own (often private) definition of fairness, whereas for others equity may mean equal treatment or the distribution of income. Because equity has become such a popular word in transportation planning, some efforts at presenting an operational form of the concept seem justified.

In this paper, some definitions are first proposed, and then two examples are presented and discussed. The first example compares alternative policies for allocating gasoline during a shortage, and the second estimates empirically the vertical impact of financing rail rapid transit construction out of property and sales taxes, as was done in San Francisco. The conclusion is drawn that inequities in the transportation sector are the result of inefficiencies rather than a consequence of the conflict between efficiency and equity.

## DEFINITIONS OF HORIZONTAL AND VERTICAL EQUITY

Equity generally refers to the distribution of something that has value—i.e., costs or benefits—among entities—i.e., people, regions, or factors—and whether that distribution is good or bad or better or worse. Part of the problem in the use of the term is that equity is both descriptive (what the distribution is) and normative (whether it is good or bad).

Many of the standard works on public finance (2, 9, 10) include brief sections on definitions and alternative concepts of equity. Current literature is sometimes helpful (7), but more often it is directed at remote theoretical points. Although an occasional extended empirical work (6) will include transportation as one component, applications in the transportation field are limited (1, 3, 4, 5, 11). An extensive literature treats the shifting and incidence of the property tax. The definitions offered below are generally consistent with this literature although there is considerable disagreement and ambiguity within it. A first step toward a definition of equity is to group applications of the equity concept under two main headings:

1. Horizontal equity—In formal terms, this is the equivalent treatment of individuals in equal circumstances and relates most directly to popular notions of fairness. Suppose, for example, a large transportation investment creates benefits to landowners according to the schedule shown in Figure 1 but taxes are levied uniformly within the two jurisdictions benefited. It can be seen that (a) some persons in each jurisdiction pay for benefits they do not receive while others receive more than they pay and (b) one jurisdiction is paying more than its share of the total bill.

2. Vertical equity—The other side of equity refers to the distribution of income between different classes of incomes. Views on this subject tend to reflect one of two lines of thought: (a) The existing income distribution is unacceptable and another is preferred, usually one that is more egalitarian, or (b) the present distribution is tolerable, but the effects of proposed programs and policies should be evaluated to be sure they at least do not worsen the situation. The second approach is the one taken here, but roughly the same analytic skills are required in either case. This means that we are primarily interested in equity impacts, i.e., the incremental change in the aggregate distribution of income that results from a project.

Two types of criteria are used to evaluate impacts of vertical equity. One assumes that the size of the pot is fixed (there are no efficiency impacts) and the result of the policy is labeled either favorable (low incomes gain at the expense of high), unfavorable (high incomes gain at the expense of low), or neutral (there are no net redistributive effects). The other type of criterion is more general, applies to the distribution of costs, benefits, and net benefits, and is measured in proportion to income: Costs (taxes) that increase faster than income as a proportion of income are progressive as are benefits that increase less than proportionately; costs that increase less than proportionately or benefits that increase faster than income are regressive; and costs or benefits that are a constant proportion of income are neutral.

Three examples are shown graphically in Figure 2. Empirical estimates of these distributions will be less smooth (because of grouping of data by income class) and less monotonic than the diagrams shown.

## EFFICIENCY CRITERION

Maximum social welfare is obtained when, for all outputs, the marginal social benefit of the last unit is equal to the marginal social cost in terms of what society must give up in order to obtain that unit of output. In a perfectly functioning market, benefits are reflected by willingness to pay and can be represented diagrammatically by a demand curve; social opportunity costs are similarly represented by the supply curve, and the intersection—the optimum level of output and price—results automatically from the market processes. In economic theory, this is what is known as efficiency. Horizontal equity is satisfied because equal payment is made for equal use, and vertical equity is neutral as long as the initial distribution of income is acceptable.

No actual markets function perfectly, so the policy question becomes that of determining what kinds of market failure exist and what public intervention is warranted. Despite the pervasive presence of the public sector in transportation, the types of market failure that justify public intervention (notably, the natural monopoly characteristic of a large capital investment in a network) are few in number. In particular, there is no reason why users of the systems should not pay the full social costs of constructing and operating those systems. Although there are those who argue otherwise, it is assumed here that transportation does not create external benefits. External costs such as pollution and noise are ignored.

INEQUITY OF EFFICIENT PRICING

Some persons object to efficient pricing because it is

Figure 1. Horizontal equity between jurisdictions.

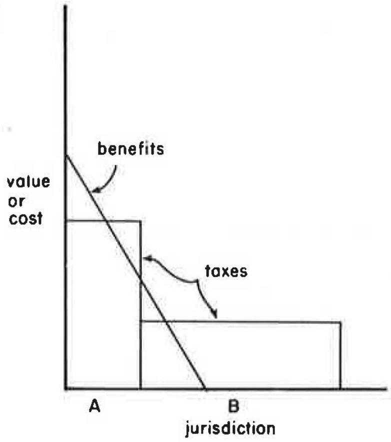
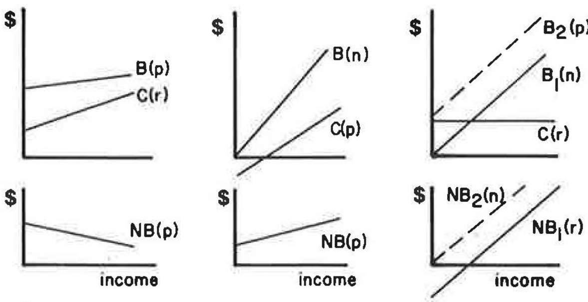


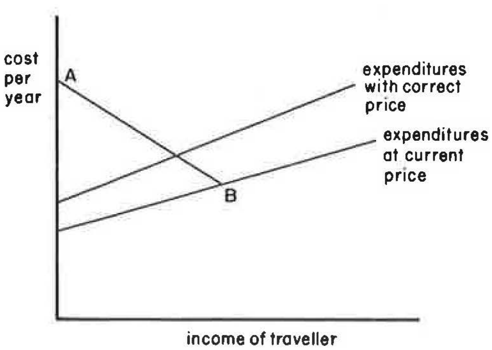

Figure 2. Possible cost and benefit distributions by income.




B = benefit
C = cost
NB = net benefits

n = neutral
r = regressive
p = progressive

Figure 3. Equity effects of correct pricing.




claimed to be inequitable. Of particular concern are those potential system users who are dissuaded by the level of the user charge. Clearly, many persons who would like to use a transportation facility and who choose not to in the face of the high price may be of lower than average income. The claim may be made that it is the poor who are "tolled off" the facility. The reasoning for this argument can be stated as follows: For any good or service that is in the broad category of being a necessity or is simply generally consumed, the amount spent for this item by each household will rise with the income of the household (overall, wealthier households will spend more for the item), but the proportion of income spent will decline. An increase in price, then, will operate like an excise tax, falling more heavily on lower income households as a proportion of income.

An example of the effects of correct pricing is shown in Figure 3. Indeed, the increase in price, by itself, is regressive. This observation should, however, be placed in context:

1. Equity impacts cannot be estimated without specifying the null alternative. When the price to users is below cost, then the deficit must be made up by a transfer from some group of taxpayers to the group of consumers. If, for example, the burden of the subsidy falls as shown by line AB in Figure 3, shifting the full cost burden onto consumers would result in a net improvement in equity. In the no-free-lunch real world, equity is determined not by whether the user pays or not but by how things are paid for by users and nonusers together.

2. If the higher price is the correct (i.e., efficient) one, then the welfare gains exceed the costs to consumers, and it is possible to make everyone better off as a result. This can be done through the generation of income, if private markets are functioning properly, or through direct government action. If the correct price is achieved by a tax, then the revenues can be used to provide a rebate to low-income households, to improve service to persons most adversely affected by the higher price (e.g., better commuter bus service), to construct new facilities where the demand warrants, or all of the above. In most policy contexts there are several feasible ways to at least approximate neutral or favorable equity and at the same time improve efficiency, and these actions should be taken in conjunction with each other.

3. If the user is to be undercharged because it is more equitable, then the question becomes, How much subsidy? Once a major component of the system is subsidized, it becomes harder to deny subsidies to others; city bus companies and railroads have joined the ranks at the trough in the last decade or so, and the taxicabs and intercity bus companies are now starting to get hungry.

It is preferable, then, to separate—analytically—equity and efficiency and not attempt to achieve equity by sacrificing efficiency. Typically, the gains will be overwhelmed by the losses when, with a little care, it is quite possible to achieve both.

EXAMPLES

Response to a Gasoline Shortage

Three alternative policies for dealing with the situation in which there is excess demand for gasoline at prevailing (controlled) prices have been selected from among those discussed, proposed, or placed in practice. They have been simplified somewhat for discussion purposes, and no attempt has been made to test empirical assumptions used in evaluating the three alternatives.

1. Plan A imposes a tax on the price of gasoline that is large enough to reduce demand to the level of supply and uses the revenues to provide a tax rebate on the basis of income (no other test, such as automobile ownership, is considered).
2. Plan B allocates available supplies to regions according to previous consumption levels. Within those regions, the stock of gasoline is allocated to those willing to pay the controlled price plus wait in line for the gasoline.
3. Plan C issues rationing stamps to all licensed drivers according to need, the total number of stamps being equal to the total supply of gasoline. Need is hard to define precisely, but it appears to include such notions as the lesser need for gasoline among persons living in areas served by transit, greater need among persons who live far from where they work, and need based on automobile ownership and previous consumption.

The three plans are listed in decreasing order of efficiency (net social benefits). Plan A directs supplies to those who benefit most as expressed by willingness to pay; plan B includes a time price, which is a less efficient rationing device; and plan C is least efficient because it both creates heavy transaction costs and tends to encourage at least some inefficient consumers to maintain their previous levels of consumption.

### Horizontal Equity

If persons who consume equal amounts of gasoline make equal sacrifices, then horizontal equity is served; in other words, persons should pay in accordance with the amount consumed. Plan A would be the most equitable, then, because it would require each consumer to sacrifice in accordance with the amount of gasoline consumed. Plan B is less equitable because consumers in equal circumstances (i.e., who consume equal amounts of gasoline) will sacrifice varying amounts in terms of time and inconvenience depending on such factors as region, location within region, time schedule, and availability of stand-ins such as wives and children. But at least the costs are fully borne by consumers of gasoline. Plan C has the effect of creating income (the stamps have a value approximately equal to the optimal tax in plan A) for a particular group of consumers (those with automobiles, high gasoline consumption, and without access to transit) in a way that is arbitrary from the standpoint of horizontal equity; plan C is, in fact, perverse because it rewards those who are least deserving from the standpoint of horizontal equity (not necessarily the same as vertical).

### Vertical Equity

Plan B has the most favorable impact on vertical equity, but the reasons are somewhat unattractive. If it is assumed that persons with higher incomes also generally place a higher value on their time, then the time component of the price of gasoline extracts a greater sacrifice from them than from those with lower incomes; in other words, vertical equity is achieved by making everyone worse off but those with higher incomes more worse off than those with lower incomes. Both efficiency and equity can be improved somewhat by allowing persons with higher than average values of time to hire persons with lower than average values to stand in line for them. This becomes, in effect, a transfer payment from higher income to lower income people as a function of how much time those with lower incomes are willing to waste waiting in line.

Plan A also has a favorable vertical equity impact because high- as well as low-income people pay the higher price but only those with lower incomes receive the rebate. Depending on how the surplus revenues (above the amount of the rebate) are used, the vertical equity impact could be improved or worsened.

Plan C again has the least favorable impacts. The extent to which the distribution of income would be worsened by this plan depends on the distribution of income of needy persons (those with automobiles, a driver's license, or high previous consumption) versus the distribution of income of nonneedy persons. Whether the result would be favorable or unfavorable requires matching empirical information with a precise definition of need, but it appears plausible that most of the needy would be affluent suburban commuters. In addition, persons who do not have a driver's license (the poor and the elderly) are more likely to come from low-income than high-income households.

The efficiency of plan C could be improved slightly by allowing recipients to sell their stamps, which would permit a household with high consumption to decide whether to maintain previous levels of consumption or sell the stamps and consume less, but the equity impact would be unaffected by this transaction. Selling the stamps simply means that the income in kind (gasoline) can be exchanged for money income, and the distribution of income is unchanged.

A summary comparison of how the three plans rank in dealing with gasoline shortages is given below:

| Ranking | Efficiency | Horizontal Equity | Vertical Equity |
|---|---|---|---|
| Best | A | A | B |
| Second best | B | B | A |
| Worst | C | C | C |

In comparing the three plans for dealing with gasoline shortages, a conflict or trade-off between efficiency and equity appeared only once, and that was where vertical equity could be enhanced by making everyone worse off. In general, the efficient plan was the most equitable or could be made the most equitable by imposing modest side constraints. Planners should be looking for ways to impose these constraints on efficient solutions rather than attempting to redistribute income through transportation policy.

## Vertical Impact of Bay Area Rapid Transit Financing

The cost of constructing the Bay Area Rapid Transit (BART) system was paid for primarily from two local general revenue sources: a property tax with an effective rate of about 0.13 percent and a sales tax of $0.005 that exempts groceries. Given information about the income-related characteristics of taxpayers and users of the system, estimating the magnitudes of flows of costs and benefits between income groups requires four steps [the empirical information used in discussing this example is derived from Hoachlander (8)].

### Direct Incidence

Ideally, the property tax paid by each property owner in each income class would be calculated by applying the tax rate to the value of the owner's property, and sales taxes would be calculated by applying the sales tax rate to annual local expenditures. A number of difficulties make the reality considerably more crude, but only the more important ones will be described. First, data are grouped into large classes by income, and average values of income and property must be used. Second, the original source of the information was the 1970 U.S. Census, and property and income data are only provided for residential property so that the distribution of the impact of the property tax on commercial and industrial property is assumed to be the same as the distribution on residential property. Third, spill-ins and spill-outs (e.g., sales taxes paid by tourists) are assumed to be negligible or no different from the estimated distribution of impact based on local residents.

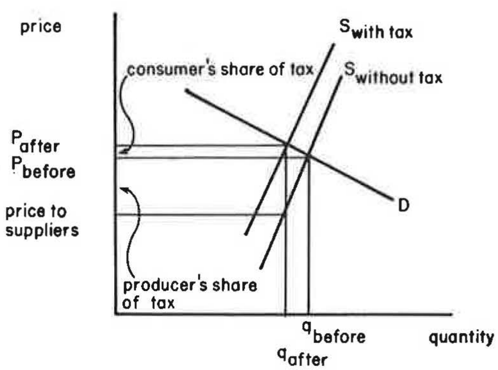
Market Adjustments

In many situations, the imposition of the tax will cause a change in behavior on the part of those on whom the tax is levied. If the policy change were a price increase, then consumers could be expected to adapt in various ways so as to lessen the impact of the higher price or take advantage of a lower price. Sales and property taxes can cause consumers to shop in jurisdictions where the tax is lower, and property taxes can encourage households and firms to locate in other jurisdictions. Although this would be possible in the Bay Area case, the tax rates are low enough that substantial attempts to escape them were probably not made. Of course, if the estimates are being made retrospectively, then the actual distributions of households after the tax was levied can be used.

Tax Shifting

The extent to which the burden of a tax falls on consumers versus producers (or landlords versus renters) depends on the relative elasticities of supply and demand. One example in which supply is fairly inelastic and demand is elastic is shown in Figure 4. In this situation, the tax falls more heavily on producers because consumers drop out of the market with even small increases

Figure 4. Incidence of tax on consumers and producers.




in price but producers cannot so easily adjust supply. If things were the other way around—inelastic demand and elastic supply—then consumers would bear most of the burden of the tax. It is important to note that it is not on whom the tax is levied but market conditions that determine incidence. For the conditions shown in the diagram in Figure 4, the tax could have been charged to consumers (a sales tax) instead of producers (an excise tax quoted in the price), and the results would still be the same.

Hoachlander assumed that homeowners absorb the full burden of the tax on owner-occupied property, that the tax on rental property is fully shifted forward onto tenants, and that sales taxes are fully borne by the consumer (he used the Internal Revenue Service estimates tabled for purposes of itemizing income-tax deductions), and the estimates for the cost burden given in Table 1 (8) reflect these assumptions. On the benefits side, because only users (or their households) are assumed to benefit, the passenger kilometer was chosen as a measure of benefit. By using recent ridership surveys and thus distribution of patronage and average trip length, an index of aggregate passenger kilometers of travel by income class can be constructed. For convenience, total benefits (net of fares) are assumed to be equal to the total taxes contributed, so the passenger-kilometer index was scaled to give the same total as that for costs. This allows the costs and benefits for each class to be compared on the basis of relative gain or loss (zero sum). The results of estimating benefits and also of subtracting costs are given in Table 2 (8).

Interpretation

A good way to represent the distributional results in graphic form is to measure costs or benefits per household on the vertical scale and let the width of each band be proportional to the size of the income class. In Figure 5, it is clear that high-income groups have gained at the expense of low-income groups, but the magnitudes are placed in perspective because the area of each segment indicates the amount of the transfer into or out of each income class.

Several points should be kept in mind in interpreting these results:

1. Benefits calculated per trip (instead of per kilometer) would appear to be less redistributive but nonetheless unfavorable.

2. Property taxes are not generally fully shifted and, to the extent that this is true, estimates of the cost burden are biased downward; i.e., higher income groups actually pay more tax than that shown.

3. As noted, the estimates of costs are based on residential property taxes only, which make up about 53 percent of BART property taxes. To the extent that the distribution of costs initially levied on commercial and industrial property differs from the distribution of costs levied on residential property, the cost estimates are inaccurate.

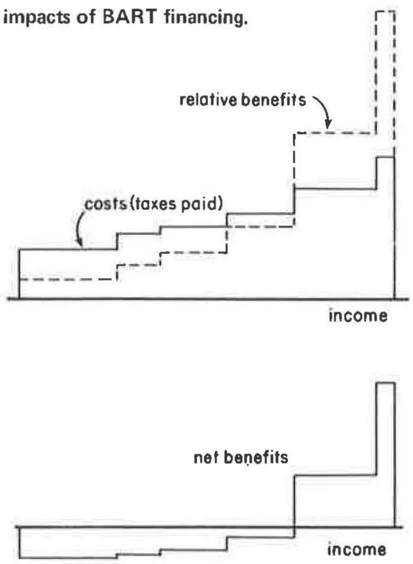
Table 1. BART taxes by income class.

| Income Class ($) | Tax on Homeowners ($) | | Tax on Renters ($) | | Total ($) | Households | | Tax per Household ($) |
|---|---|---|---|---|---|---|---|---|
| | Property | Sales | Property | Sales | | Number | Percent | |
| 0-5 000 | 1310 | 498 | 2367 | 1555 | 5 730 | 206 915 | 27 | 28 |
| 5 000-7 000 | 609 | 290 | 1068 | 795 | 2 762 | 83 502 | 11 | 33 |
| 7 000-10 000 | 1432 | 770 | 1508 | 1216 | 4 926 | 132 348 | 17 | 37 |
| 10 000-15 000 | 3351 | 1938 | 1525 | 1313 | 8 127 | 180 632 | 23 | 45 |
| 15 000-25 000 | 3501 | 2373 | 828 | 856 | 7 558 | 129 139 | 17 | 58 |
| ≥25 000 | 1381 | 916 | 213 | 220 | 2 730 | 37 139 | 5 | 73 |
| Total | | | | | 31 833 | 769 675 | 100 | |

Table 2. BART benefits by income class.

| Income Class ($) | BART Ridership (%) | Average Trip (km) | Passenger-Kilometer Benefits ($000) | Benefits per Household ($) | Net Benefits per Household ($) |
|---|---|---|---|---|---|
| 0-5 000 | 10.5 | 2.75 | 2 267 | 11 | -17 |
| 5 000-7 000 | 6.8 | 2.82 | 1 502 | 18 | -15 |
| 7 000-10 000 | 12.6 | 3.38 | 3 338 | 25 | -12 |
| 10 000-15 000 | 21.6 | 4.32 | 7 321 | 40 | -4 |
| 15 000-25 000 | 30.6 | 4.74 | 11 370 | 88 | 29 |
| ≥25 000 | 17.8 | 4.32 | 6 033 | 162 | 89 |
| Total | 100 | | 31 833 | | |

Figure 5. Vertical equity impacts of BART financing.




4. The conclusion that BART created an unfavorable income redistribution in the Bay Area cannot be accepted without establishing what would have occurred otherwise. Previous investments in highway capacity had drawn from similar sources in similar proportions and to the benefit of similar groups. It is quite likely that the unfavorable equity impacts of BART are not much, if at all, worse than the impacts of a corresponding investment in highways; BART only looks bad in comparison with an ideal sector that is equitably priced and financed.

5. In principle, fares for the high-quality service used by higher income travelers could be set at a level somewhat above costs as a bias toward progressivity. In the case of BART, however, these same travelers have available to them the most heavily subsidized alternative—commuting to and from the suburbs by automobile. BART could probably raise its fares by a modest amount, but it is severely constrained by prices set on competing modes.

## COMPLEMENTARITY OF EQUITY AND EFFICIENCY

If the characteristics of a good or service are such that (a) benefits of consumption are entirely captured by users (and perhaps passed on along with costs) and (b) the existing distribution of income is generally acceptable, then equity and efficiency can both be served most easily by charging full costs to users in accordance with use. Accomplishing this with complete accuracy in transportation would require that user fees at least vary by network segment, time of day, and type of vehicle. The system would be entirely self-supporting (covering all opportunity and administrative costs) and would contribute to sales and property tax revenues.

Such, of course, is not the case. In general, the transportation user underpays, and the underpayment is erratic but tends to be greater the higher the cost of the service is. Facilities for which demand is either very high or very low are especially underpriced. Moreover, inefficiencies in resource allocation and utilization also lead to undesirable equity impacts, such as the following:

1. The shortfall must be made up from some other source—normally a general revenue instrument, usually the property tax. This violates horizontal equity (nonusers pay for services that do not benefit them) and may have unfavorable vertical impacts as well.
2. The nature of the service offered—such as the balance between modes—is biased toward higher income users. Suburban commuters receive large subsidies per trip, whereas transit-dependent travelers receive far less service than they would get if all subsidies were eliminated.
3. Minor cross subsidies that might be desirable (e.g., for the elderly or school children) are impossible because all users underpay.
4. Attempts to correct inequities on one mode are frustrated by the ease with which the relatively affluent can escape higher user charges by shifting to another mode.

Certainly, it is not a simple task to evaluate the various kinds of equity impacts, but the methods and concepts are available and they are no harder to use than those related to efficiency. Much improvement in the state of the art needs to be made, but in the effort it might be discovered that, far from having sacrificed equity to efficiency, we have achieved neither.

## ACKNOWLEDGMENTS

Many helpful suggestions and comments on a previous draft were provided by Kevin Laverty and several anonymous reviewers, and their efforts are greatly appreciated.

## REFERENCES

1. N. S. Dorfman. Who Will Pay for Pollution Control? National Tax Journal, Vol. 28, No. 1, March 1975, pp. 101-115.
2. J. F. Due and A. F. Friedlaender. Government Finance: Economics of the Public Sector. Richard D. Irwin, Homewood, IL, 5th Ed., 1973.
3. C. D. Foster. The Regressiveness of Road Pricing. International Journal of Transport Economics, Vol. 1, No. 2, Aug. 1974, pp. 133-141.
4. M. Frankena. Income Distributional Effects of Urban Transit Subsidies. Journal of Transportation Economics and Policy, Vol. 7, No. 3, Sept. 1973, pp. 215-230.

5. A. M. Freeman. On the Incidence of Auto Pollution Control Costs. In Benefit-Cost and Policy Analysis 1974 (Richard Zeckhauser and others, eds.), Aldine, Chicago, 1975, pp. 269-282.
6. K. V. Green, W. B. Neenan, and C. D. Scott. Fiscal Interactions in a Metropolitan Area. Heath, Lexington, MA, 1974.
7. W. L. Hansen and B. A. Weisbrod. Distributional Effects of Public Expenditure Programs. Public Finance, Vol. 27, No. 4, 1972, pp. 414-420.
8. E. G. Hoachlander. Bay Area Rapid Transit: Who Pays and Who Benefits. Institute of Urban and Regional Development, Univ. of California, Berkeley, Working Paper 267, July 1976.
9. R. A. Musgrave and P. B. Musgrave. Public Finance in Theory and Practice. McGraw-Hill, New York, 2nd Ed., 1976.
10. N. M. Singer. Public Microeconomics. Little, Brown, Boston, 2nd Ed., 1976.
11. J. W. Zupnick. The Short-Run Incidence of a Tax Induced Rise in the Price of Gasoline. Journal of Economic Issues, Vol. 9, No. 2, June 1975, pp. 409-414.

*Publication of this paper sponsored by Committee on Transportation Systems Design.*

# Who Favors Work-Schedule Changes and Why

Anis A. Tannir,* Saudi Arabian Parsons Limited, Jeddah
David T. Hartgen, New York State Department of Transportation

**Factors that influence attitudes of white-collar employees toward alternative work-schedule changes are examined to determine whether the desire to avoid traffic congestion is a primary determinant of such attitudes. A random sample of 110 employees from the main office of the New York State Department of Transportation in Albany, New York, were given a short questionnaire on travel patterns, attitudes toward components of work schedules, and perceptions of impacts of work-schedule changes on family life, travel patterns, and working environment. An attitude scaling technique known as trade-off analysis was used to determine the most preferred programs and the characteristics of those in favor of and those opposed to schedule changes. Results showed the basic motivation behind favoring work-schedule changes is the employee's desire to introduce flexibility into family, leisure, and work activities; the desire to avoid traffic congestion is a contributing, but not a major, factor. The most preferred arrangements are 5-d variable hours, 4-d variable hours, and 5-d individual-specific hours, all with over 65 percent support. Support was strongest among younger employees who had children in school and weakest among single and older employees and car poolers. The policy implications for transportation planning are discussed.**

Considerable research has been published on the application of staggered work hours as a device to relieve commuter congestion in public transit facilities (1, 2, 3, 4). The conclusion of these studies is that peak demands in transit facilities can be reduced by 10 to 30 percent through widespread use of such policies. Studies of the impacts of the 4-d workweek on highway congestion (5, 6, 7, and a paper elsewhere in this Record by Tannir and Hartgen) and other studies (8, 9) support variable work hours and 4-d workweek policies as a possible policy for low-cost shifting of travel to reduce traffic congestion. All of these studies, however, have concentrated on large metropolitan areas.

The impacts of staggered work hours and 4-d workweek schedules on firms and their employees have been studied and generally found to be positive (10, 11, 12, 13). General benefits include improvement in employee morale and productivity, reduction in absenteeism and overtime, better use of capital assets, extended hours of service to clients, improved driving conditions during the trip to work, and, under certain conditions, reductions in energy consumption.

## DATA AND METHOD

The New York State government offices located at the State Campus in Albany, New York, were selected to be surveyed in this inquiry. The site is located approximately 6.4 km (4 miles) west of downtown Albany in a predominantly residential area, and there are approximately 10 042 employees. The campus is accessible by way of a highway network of local streets, major arterials, and expressways. New York State is the only employer on the campus, employment density is high, and public transportation does not play a major role in the daily movement of employees to and from their jobs. White-collar workers constitute the majority of these employees.

Employees on the State Campus were surveyed to determine employee characteristics, attitudes toward changes in work schedules, and perceived impacts. For several reasons, the main office of the New York State Department of Transportation (NYSDOT) was selected as the focal point for the employee survey. First, it is located on the State Campus. Second, the department population is generally representative of the entire campus population. Third, it was convenient because the researchers were familiar with the organizational structure and functional units of the department. And, finally, permission to conduct such a survey was obtainable from management and employee representatives of NYSDOT.

A random sample of 140 employees from the NYSDOT main office staff of 1771 were selected and contacted. Of these, 110 completed returns were used in the analysis. The returned sample was representative of the main office population (Table 1). Respondents were administered a questionnaire that covered travel and demographic characteristics, general attitudes toward work-schedule changes, perceived impacts of these

Table 1. Comparison of sample population and population of NYSDOT main office.

| | Sample Population | | Main Office Population | |
|---|---|---|---|---|
| Category | Number | Percent | Number | Percent |
| Sex | | | | |
| Male | 86 | 78 | 1314 | 73.7 |
| Female | 24 | 22 | 469 | 26.3 |
| Total | 110 | 100 | 1783 | 100.0 |
| State grade level | | | | |
| 1-9 | 29 | 28.7 | 615 | 34.7 |
| 10-19 | 42 | 41.6 | 676 | 38.2 |
| 20-29 | 22 | 21.8 | 387 | 21.9 |
| 30-38 | 7 | 6.9 | 68 | 3.8 |
| Unclassified | 1 | 1.0 | 24 | 1.4 |
| Total | 101 | 100 | 1770 | 100.0 |
| Bargaining unit | | | | |
| Administration | 33 | 30.5 | 517 | 29.0 |
| Operational | 2 | 1.8 | 32 | 1.8 |
| Professional, scientific, and technical | 60 | 55.1 | 1031 | 57.8 |
| Management and confidential | 14 | 12.8 | 203 | 11.4 |
| Total | 109 | 100 | 1784 | 100.0 |

Figure 1. Partitioning of the sample.

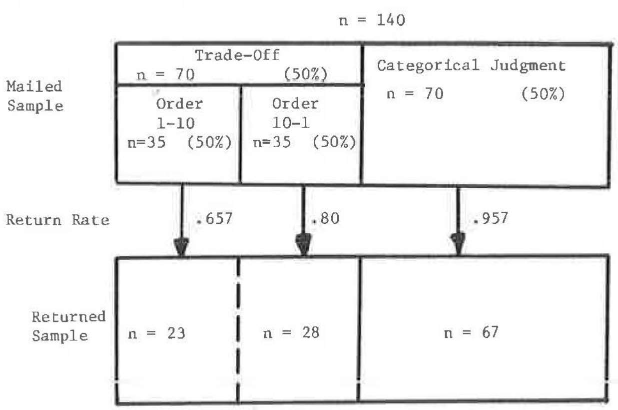
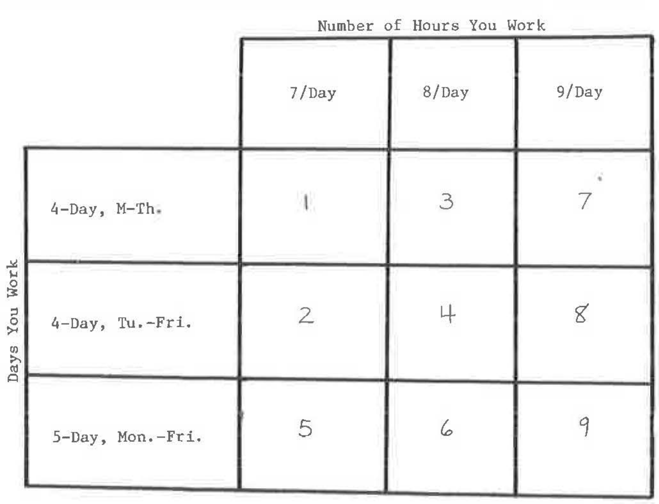



Figure 2. Example of trade-off matrix used in the survey.




changes, and attitudes toward various attributes of work schedules. In the case of the last item, 50 percent of the sample received a questionnaire in the categorical judgment format (short form), and 50 percent of the sample received a questionnaire that contained attitude questions in the trade-off analysis format (long form). To control bias resulting from respondent fatigue, this group was further divided into two equal subgroups (each 25 percent of the total sample), and the sequence of questions was reversed (Figure 1). The differences in these approaches are analyzed elsewhere (14, 15, 16).

A technique known as trade-off analysis (14, 17, 18, 19) was used to develop alternative feasible policies on changes in work schedules. This technique requires the respondent to rank order the cells of a two-variable matrix from most preferred to least preferred, as shown in Figure 2. A set of (n) (n - 1)/2 such matrixes is administered for the n attributes. These attributes and their levels are given below (current schedules call for 7.5-h workdays):

| Attribute Level | Mean Utility | Standard Deviation |
|---|---|---|
| Days worked | | |
| Four, Monday-Thursday | 0.38 | 0.08 |
| Four, Tuesday-Friday | 0.37 | 0.07 |
| Five, Monday-Friday (current schedule) | 0.25 | 0.11 |
| Number of hours worked per day | | |
| Seven (current schedule) | 0.43 | 0.11 |
| Eight | 0.33 | 0.05 |
| Nine | 0.24 | 0.09 |
| Times worked | | |
| Fixed (current schedule) | 0.29 | 0.08 |
| Specific for individual | 0.32 | 0.07 |
| Variable | 0.39 | 0.11 |
| Parking location | | |
| Wherever desired (current schedule) | 0.39 | 0.09 |
| Special place if car pool | 0.27 | 0.08 |
| Same space every day | 0.34 | 0.01 |
| Cost of parking | | |
| Zero (current schedule) | 0.58 | 0.09 |
| $1/month | 0.40 | 0.03 |
| $1/week | 0.12 | 0.08 |

The trade-off algorithm (17) uses these rank-order preferences to produce estimated utilities for each respondent by minimizing the differences between the observed rankings and the rankings of the cell-product utilities. Mean utilities are given in Table 2. This information is then inputted to a simulation routine that estimates preferences for alternative work-hour programs based on the preferences of all individuals. Shares are computed as follows: Let $u_{ik}$ = the utility that respondent i places on attribute level k, $U_{ip}$ = the utility that respondent i places on future p, and $P_{ip}$ = the preference (percentage favorability) given future p by respondent i. A linear utility function is assumed:

$$U_{ip} = \sum_k u_{ik} \quad (1)$$

A Luce share model is assumed for preference calculation:

$$P_{ip} = U_{ip}/(U_{ip} + U_{iq} + \ldots) \quad (2)$$

Aggregations of $P_{ip}$ over all respondents reveal total market preference, and detailed breakdowns of support by demographic and other characteristics reveal which groups stand to gain or lose under different work-schedule policies.

Table 2. General attitudes toward alternative work schedules.

| Attitude | Congestion Experienced on Work Trip | Response (%) Unfavorable | Neutral | Favorable | No Response | Percentage of Total (n = 110) |
|---|---|---|---|---|---|---|
| Toward variable work hours | None (A)[a] | 50 | 6 | 44 | 0 | 16.4 |
| | Somewhat (B)[a] | 21 | 15 | 64 | 0 | 48.2 |
| | Considerable-severe (C, D, E)[a] | 15 | 8 | 72 | 5 | 35.4 |
| | Average | 24 | 11 | 64 | 1 | |
| Toward 4-d workweek | None (A)[a] | 28 | 17 | 56 | 0 | 16.4 |
| | Somewhat (B)[a] | 30 | 8 | 62 | 0 | 48.2 |
| | Considerable-severe (C, D, E)[a] | 13 | 10 | 72 | 5 | 35.4 |
| | Average | 24 | 10 | 64 | 2 | |

[a]Letters in parentheses denote level of service as described in the Highway Capacity Manual (21).

Table 3. Summary of attitudes toward personal impacts of work-schedule changes.

| Area of Impact | Response (%) Very Unfavorable Impact | Somewhat Unfavorable Impact | Neutral | Somewhat Positive Impact | Very Positive Impact | Total ($N_t$) | Weighted Mean |
|---|---|---|---|---|---|---|---|
| Second job | 25 | 5 | 52 | 14 | 5 | 107 | 2.70 |
| Fatigue | 9 | 17 | 52 | 8 | 14 | 107 | 3.03 |
| Communication | 7 | 13 | 56 | 14 | 10 | 106 | 3.08 |
| Rush-hour congestion | 11 | 9 | 40 | 20 | 20 | 107 | 3.27 |
| Leave time | 5 | 5 | 49 | 20 | 21 | 107 | 3.49 |
| Gasoline savings | 6 | 4 | 45 | 24 | 21 | 107 | 3.50 |
| Productivity | 0 | 7 | 47 | 24 | 22 | 107 | 3.60 |
| Job satisfaction | 1 | 5 | 42 | 27 | 25 | 107 | 3.71 |
| Family time | 3 | 4 | 21 | 20 | 52 | 108 | 4.16 |
| Leisure time | 2 | 2 | 17 | 29 | 50 | 108 | 4.25 |

## RESULTS

General results of the survey are given in the table below:

| Characteristic | Percent |
|---|---|
| Family size, number in household | |
| 1 | 13.6 |
| 2 | 28.2 |
| 3-4 | 36.5 |
| 5-6 | 15.4 |
| ≥7 | 4.5 |
| Other (blank) | 1.8 |
| Automobile ownership, number of automobiles | |
| 0 | 3.6 |
| 1 | 45.5 |
| ≥2 | 49.1 |
| Other | 1.8 |
| Mode to work | |
| Drive | 80 |
| Automobile passenger | 18 |
| Other | 2 |
| Car pooling | |
| None | 70.6 |
| Occasional | 13.8 |
| Car poolers | 15.6 |

The key findings are that

1. The level of automobile ownership for the sample is very high,
2. The sample is average in relation to family size,
3. The automobile predominates in travel to work, and
4. Frequent and occasional car pooling is quite common.

The survey also revealed that, for car poolers and non-car-pool users, average work-trip length was 24 and 18 km (15 and 11 miles) respectively.

Data given in Table 2 show that the sample is generally very much in favor of variable work hours or 4-d weeks (64 percent overall). Favorability appears to be sensitive to the perceived level of traffic congestion: The higher the congestion is, the greater is the inclination to favor variable work hours. The impact of congestion level on attitudes toward a 4-d workweek is less pronounced. These findings suggest that the desire to avoid traffic congestion has at least a moderately important influence on favorability toward work-schedule changes.

Table 3, however, suggests that other factors may be more important. The respondents felt that work-schedule changes would have a very positive impact on family and leisure time but only a marginally positive impact on ability to avoid rush-hour congestion. These results suggest that the desire for flexibility in personal activities and the desire for more leisure time are the primary determinants of attitudes toward work-schedule changes.

### Attitudes Toward Work-Schedule Arrangements

As described earlier, the trade-off procedure allows the analyst to compare alternative work-schedule arrangements with the current fixed-hours schedule. The attributes and levels given previously are the basis for formulating various work-schedule programs.

Theoretically, it is possible to structure 243 possible programs or five attributes of three levels each. However, based on a literature search and previous surveys, it became apparent that most of the programs that can be structured are either unrealistic or not implementable from management, employee, or legal viewpoints. Therefore, only selected programs, called "futures," were structured for testing. Figure 3 shows eight such tests along with demographic breakdowns of support. It is apparent that only futures 5, 6, and 1 are preferred to the current policy.

**Figure 3. Support for proposed policies (or futures).**

| Attribute | | Ind. Specific Times | | | | Variable Hours | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Ind. Specific | 4-Day Workweek | 4-Day Pkg.Pref. | 4D Pkg Pref $1.0 Charge | Basic Var. Hrs. 5D | Var. Hrs. 4 Days | Var. Hrs. Pkg.Pref. | VarHrs.Pkg. Pref $1./Mo. |
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Days Worked | | 5D M-Fri. | 4D M-Thur | 4D M-Thur | 4D M-Thur | 5D M-Fri. | 4D M-Thur | 4D M-Thur | 4D M-Thur |
| Number of Hours | | 7.5 | 9.4 | 9.4 | 9.4 | 7.5 | 9.4 | 9.4 | 9.4 |
| Times Worked | | Specific | Specific | Specific | Specific | Variable | Variable | Variable | Variable |
| Parking Location | | Anywhere Assign.Lot | Anywhere Assign Lot | Special Carpool | Special Carpool | Anywhere Assign.Lot | Anywhere AssignLot | Special Carpool | Special Carpool |
| Parking Costs | | Free | Free | Free | $1.0/mo. | Free | Free | Free | $1.0/Mo. |
| Intensity of Preference (Present/Future) | | 48%/52% | 50%/50% | 58%/42% | 69%/31% | 43%/57% | 46%/54% | 53%/47% | 65%/35% |
| First Preference (Present/Future) | | 35%/65% | 57%/43% | 67%/33% | 78%/22% | 27%/73% | 35%/65% | 59%/41% | 75%/25% |
| Age | 25-34 | 78 | 57 | 43 | 36 | 86 | 79 | 64 | 36 |
| | 35-44 | 67 | 53 | 47 | 27 | 80 | 67 | 47 | 33 |
| | 45-54 | 67 | 33 | 25 | 8 | 67 | 58 | 33 | 17 |
| | 55-64 | 57 | 29 | 14 | 14 | 43 | 57 | 14 | 14 |
| Family Size | 1 | 29 | 43 | 43 | 29 | 43 | 43 | 57 | 43 |
| | 2 | 47 | 21 | 11 | 0 | 74 | 68 | 11 | 0 |
| | 3 or 4 | 89 | 61 | 44 | 39 | 72 | 61 | 56 | 39 |
| | 5 or 6 | 83 | 50 | 50 | 17 | 100 | 83 | 67 | 33 |
| Bar-gain. Unit | AdmSer | 54 | 15 | 7 | 8 | 46 | 38 | 23 | 17 |
| | PS & T | 72 | 59 | 41 | 31 | 86 | 72 | 48 | 31 |
| | M/C | 66 | 33 | 59 | 17 | 83 | 100 | 50 | 33 |
| Car-pool ing | No | 66 | 34 | 20 | 11 | 77 | 63 | 29 | 14 |
| | Sometimes | 75 | 62 | 50 | 25 | 62 | 62 | 62 | 38 |
| | Yes | 43 | 57 | 71 | 57 | 57 | 71 | 71 | 57 |
| Traffic Conges tion | None | 44 | 11 | 0 | 0 | 44 | 44 | 22 | 0 |
| | Somewhat | 74 | 61 | 43 | 30 | 78 | 78 | 52 | 35 |
| | Consid. | 64 | 29 | 36 | 21 | 86 | 64 | 36 | 21 |
| | Severe | 60 | 60 | 40 | 20 | 60 | 40 | 40 | 40 |
| | | 50% | 50% | 50% | 50% | 50% | 50% | 50% | 50% |

**Table 4. Profiles of groups that most strongly support or oppose the three preferred programs.**

| Program | Overall Preference (%) | Strongest Support Group | Percent | Strongest Opposition Group | Percent |
|---|---|---|---|---|---|
| Future 5 (5-d workweek, variable work hours) | 75 | Age 25 to 44 | 80-86 | Age 55 and over | 43 |
| | | Professional, scientific, and technical | 86 | Administrative services | 46 |
| | | Non-car-pool users | 77 | | |
| | | Family size ≥2 | 72-100 | One-person family | 43 |
| | | Travel time ≥30 min | 75-100 | | |
| | | Some and considerable congestion | 78-86 | No congestion | 44 |
| Future 6 (4-d workweek, variable work hours) | 65 | Ages 25 to 44 | 67-79 | Age 19 to 24 | 50 |
| | | Professional, scientific, and technical; management and confidential | 72-100 | Administrative services | 38 |
| | | Car poolers | 71 | | |
| | | Family size ≥2 | 61-83 | One-person family | 43 |
| | | Travel time 20 min | 74 | Travel time 10 min | 44 |
| | | Some congestion | 78 | Severe congestion | 40 |
| Future 1 (5-d workweek, individual-specific work hours) | 65 | Age 25 to 34 | 78 | | |
| | | Professional, scientific, and technical; management and confidential | 66-72 | | |
| | | Occasional car poolers | 75 | Car poolers | 43 |
| | | Family size ≥3 | 83-89 | One-person family | 29 |
| | | Travel time ≥30 min | 75 | | |
| | | Some congestion | 74 | No congestion | 44 |

## Future 5

Future 5—the 5-d workweek with variable hours—received the broadest support: 73 percent versus 27 percent support for the current schedule. This preference is especially strong among the 25 to 34 age group (85 percent) and 35 to 44 age group (80 percent). This strong preference can be explained by the fact that employees in these age groups tend to be in the child-raising stage, and variability in work start times provides the flexibility needed to reconcile job and child-care activities. Employees from families of three to four members also favor such arrangements (72 percent support); those who have more than two children tend to favor it even more, which further supports the hypothesis.

The only group that does not show enthusiasm is the 55 to 64 age group (only 43 percent support). This can be explained by observing that employees in that age bracket are very much used to the current schedule, and any change from it may cause hardships. As expected, car poolers only moderately support this program (57 percent) since car pools might be dissolved. Those who occasionally car pool are more enthusiastic about it (62 percent) since they most likely drive an automobile to work. Traffic congestion also plays a part in making this program the most preferred: Persons who experience more traffic congestion tend to support this program more. When the traffic problem is nonexistent, the support is 44 percent; when there is some congestion, the support becomes 78 percent; when congestion is considerable, the support for variable work hours is 86 percent.

### Future 6

Future 6—the 4-d workweek with variable hours—also generates strong overall support (65 percent). A close analysis of the support estimates shows that the preference trend among the various groups follows lines similar to those for future 5. Preferences by age groups indicate that the 25 to 34 group favor this schedule the most (79 percent) compared with 86 percent favorability for future 5. The 35 to 44 group gives 67 percent support to future 6 compared with 80 percent to future 5. The 45 to 54 age group gives 58 percent support to future 6 as compared with 67 percent to future 5. This degree of favorability underlines the inference that the attitudes of employees in these age groups are influenced by their desire to reconcile their work schedules with their family obligations. This point is further emphasized by the percentage preference based on family size. It is evident that households with three or four persons and those with five or six persons are strongly in favor of future 6—by 61 and 83 percent respectively. The 19 to 24 age group is evenly split in its support for future 6. Members of this group are most likely members of one- or two-person households. The support from this group is only 43 percent.

Future 6 does not appeal to drivers who do not experience any traffic congestion during their morning trip to work. Only 44 percent support this schedule. However, those who occasionally encounter traffic delays would greatly support this policy (78 percent). This may be caused by their desire to improve driving conditions through earlier work start times. On the other hand, those who experience considerable traffic delays support this policy by 64 percent. Their support may be based on the assumption that a 4-d workweek would spread the peak-hour demand, which would result in improved traffic conditions.

A third feasible schedule is a 5-d workweek with individual-specific hours (future 1 in Figure 3). The pattern of support follows similar general lines as those for futures 5 and 6. This underlines the desire of employees to reconcile their work and personal schedules. The remaining policies are not analyzed here; they are left to the reader to contemplate.

Support and opposition profiles are summarized in Table 4.

## CONCLUSION

In this study, broad support was found among white-collar state-government employees for changes in work-schedule arrangements. The most preferred programs are those that feature variable work hours. A program of this kind with five 7.5-h days is most favored and is followed by a variable-hour program with four 9.4-h days.

The desire for flexibility in work and family schedules is the basic motivating factor behind attitudes toward work-schedule changes. This is reflected through age and family size, which are the prime demographic factors in favoring a given program over the current schedule. An analysis of perceived impacts shows that leisure time and family activities would be the primary aspects of personal life that would benefit from such changes. A desire to avoid traffic congestion does not seem to be a dominant factor in attitudes toward alternative work schedules. However, those who experience some or considerable traffic congestion tend to be in favor of variable work hours.

These findings have broad policy implications for transportation planners and decision-makers who are concerned with ways to reduce traffic congestion. Since traffic congestion is not the primary factor influencing attitudes toward work-schedule changes, attempts to sell such programs on the basis of potential travel benefits are likely to be ineffective. A better approach would be to emphasize positive, achievable impacts on family life and leisure time and treat avoidance of traffic as an ancillary benefit. Even so, planners should recognize that all employees will not be equally affected: Flexible work hours will primarily benefit young households that have children at home to the detriment of single-person households, older employees, and car poolers. Thus, actions taken to relieve rush-hour congestion by introducing flexible work hours may be partially offset by the dissolution of current car pools and greater difficulty in car-pool formation. To deal with such trade-offs, the transportation planner must increasingly understand structural relations in family and work environments so that actions in one sphere will not be offset by unexpected detrimental effects in another.

## ACKNOWLEDGMENTS

We wish to acknowledge the assistance of the Office of Manpower and Employee Relations of NYSDOT, Paul Zuber of Rensselaer Polytechnic Institute, and Elene Donnelly and Richard Albertin of NYSDOT for auspices and assistance provided during this study. The assistance of Wilma C. Marhafer in preparing the manuscript is gratefully appreciated.

## REFERENCES

1. L. B. Cohen. Work Staggering for Traffic Relief: An Analysis of Manhattan CBD. Praeger, New York, 1968.
2. B. O'Malley. Staggered Work Hours Project in Lower Manhattan. TRB, Transportation Research Record 348, 1971, pp. 152-165.
3. J. C. Plewes and M. H. Yeates. The Urban Rush Hour: An Analysis of the Yonge Street, Toronto Subway System. Traffic Quarterly, Vol. 26, 1972, pp. 209-229.
4. C. Selinger. Managing Transportation Demand by Alternative Work Schedule Techniques. TRB, Special Rept. 172, 1977, pp. 67-74.
5. V. R. Desimone. Four-Day Work Week and Transportation. Transportation Engineering Journal, ASCE, Vol. 98, 1972, pp. 705-714.
6. S. E. Rowe and S. S. Taylor. The Effect of the Four Day Work-Week on Transportation in the Civic Center Area. Los Angeles Department of Traffic, 1972.
7. J. Guttman. Predicting the Effect of Variable Work Hours on Peak Period Congestion. Transportation, in preparation.

8. J. H. Dupree and R. H. Pratt. Low-Cost Urban Transportation Alternatives: A Study of Ways to Increase the Effectiveness of Existing Transportation Facilities. U.S. Department of Transportation, 1973.
9. R. Remak and S. Rosenbloom. Peak Period Traffic Congestion. NCHRP, Rept. 169, 1976.
10. H. L. Krieger. Statement Before the Subcommittee of the Senate Labor and Public Welfare Committee on Employment, Poverty, and Migratory Labor. U.S. General Accounting Office, April 7, 1976.
11. R. Poor, ed. Four Days, Forty Hours. Mentor Books, New York, 1973.
12. K. E. Wheeler, R. Gurman, and D. Tarnowieski. The Four-Day Work Week. American Management Association, New York, 1972.
13. A. Tannir. The Impacts of Feasible Staggered Work Hours and Compressed Workweek Policies on Highway Networks, Transportation Economics, Organizations and Employees. Planning Research Unit, New York State Department of Transportation, Albany, Preliminary Research Rept. 129, 1977.
14. P. M. Eberts. Trade-Off Versus Categorical Judgment: A Comparative Analysis of Two Attitude Scaling Methods for Transportation Planning. Planning Research Unit, New York State Department of Transportation, Albany, Preliminary Research Rept. 114, 1977.
15. P. M. Eberts and P. Koeppel. Trade-Off Analysis: Recent Empirical and Structural Findings. TRB, Transportation Research Record 673, in preparation.
16. P. Koeppel and P. M. Eberts. The Trade-Off Model: Empirical and Structural Findings. Planning Research Unit, New York State Department of Transportation, Albany, Preliminary Research Rept. 123, 1977.
17. S. M. Howe, E. P. Donnelly, and J. A. DesChamps. Trade-Off Analysis: Theory and Applications to Transportation Policy Planning. High Speed Ground Transportation Journal, Vol. 11, 1977, pp. 93-110.
18. Carpooling Impact Study: Trade-Off Model and Policy Simulation. Peat, Marwick, Mitchell and Company and U.S. Department of Transportation, 1976.
19. Statewide Public Opinion Poll on Transit Operating Assistance. Market Facts, Inc., and New York State Department of Transportation, Albany, 1975.
20. Highway Capacity Manual. HRB, Special Rept. 87, 1965.

*Publication of this paper sponsored by Committee on Transportation Planning Needs and Requirements of Small and Medium-Sized Communities.*

**Mr. Tannir was with the New York State Department of Transportation when this research was performed.*

# Traffic Impacts of Work-Schedule Changes in Medium-Sized Urban Areas

Anis A. Tannir,* Saudi Arabian Parsons Limited, Jeddah
David T. Hartgen, New York State Department of Transportation

**A test is made of the hypothesis that changes in work schedules can significantly reduce traffic congestion in medium-sized automobile-oriented cities. By using an extreme case—a single high-density employer in a residential area—estimates are made of the change in peak trips that would result from three alternative work-schedule changes. The impact on the surrounding street system is then evaluated by using traffic-assignment techniques. Results show that even a maximum-impact policy (4-d workweek) would have only a marginal effect on local traffic, reducing regional travel costs by 0.4 percent and costs in the immediate surrounding area by 2.2 percent. Of all the traffic benefits accrued, over 90 percent flow to actual participants, primarily through the reduced number of required work trips. Because of the institutional problems associated with implementing such policies on a large scale, it is concluded that efforts to reduce highway congestion in medium-sized automobile-oriented cities by use of alternative work schedules may not be cost-effective.**

The congestion-reducing approach of shifting travel in time and space so as to fit it within existing system capacity is receiving increasing attention. Numerous recent studies (1, 2, 3) describe the potential savings in traffic congestion achievable through such methods, and recent federal guidelines on transportation systems management require the analysis of such methods on a continuing basis. Some of the most attractive demand-shifting approaches involve the shifting of work schedules to permit greater use of limited facility capacity over a longer peak period. Work-shift policies have been given considerable attention in relation to transit service, and it has been concluded that such policies are capable of reducing peak-period congestion in transit facilities (particularly terminals and stations) by as much as 10 to 30 percent. However, considerably less is known about the effect of such policies on highway operations, particularly in small or medium-sized urban areas. Although several studies (2, 4, 5) have identified potential reductions in congestion as one of the primary benefits of such proposals, it is clear that cities in which a large portion of peak-hour trips do not currently use transit services will find the implementation of work-schedule changes a less feasible method of reducing congestion than larger urban areas might find it to be.

This paper is one of several (6, 7) that have investigated employee attitudes toward work-schedule changes and the impacts of such schedule changes on highway networks in medium-sized cities. The emphasis of this study is on cities in which the automobile is the primary transportation mode and public transportation services and their use are at a minimum because it is believed that such findings are more applicable to cities in the United States and elsewhere than are the findings of studies of transit services. There are only a few cities large enough to warrant transit service studies.

## STUDY LOCATION

The site selected for analysis was the State Campus in Albany, New York. The site is ideally suited for the analysis: It is a "spike" of employment (10 000 employees) in an otherwise residential area, New York State is the only employer on the site, patterns of employee work travel are primarily automobile-oriented, and transit plays a minor role in access to the site.

## EMPLOYEE ATTITUDES TOWARD WORK SCHEDULES

The main office of the New York State Department of Transportation (NYSDOT), which has approximately 1771 employees, was selected as the agency for the employee survey. A representative sample of 110 employees returned a questionnaire on travel patterns, attitudes toward work-schedule arrangements, and relative importance of the attributes of such programs. The following three policies (7) generated the greatest support against the current 5-d fixed-time policy:

1. A 5-d workweek with variable work hours (73 percent support),
2. A 4-d workweek with variable work hours (65 percent support), and
3. A 5-d workweek with individual-specific work hours (65 percent support).

A full analysis of the results is available in a report by Tannir (6) and in a paper by Tannir and Hartgen elsewhere in this Record. The general conclusion of the survey, however, was that attitudes toward alternative work schedules are influenced primarily by the desire of employees to increase flexibility in their personal and family lives.

## SHIFTS IN PEAK-PERIOD TRAVEL

The following analysis shows how trip ends were calculated for a 4-d workweek policy. Data on the percentage of trips that currently arrive at the campus during the peak (morning) hours are available from traffic counts. Of the trips that arrive each hour, it is assumed that 65 percent (from the employee survey) would shift to a 4-d workweek, which results in the data given below (total trips are based on 10 042 daily employees):

| | | Trips | | |
|---|---|---|---|---|
| Hour of Day (a.m.) | Percentage of 24-h Trip Ends | 65 Percent Shifting to 4-d Schedule | 35 Percent Not Shifting to 4-d Schedule | Total |
| 5:00-6:00 | 1.5 | 98 | 53 | 151 |
| 6:00-7:00 | 10.0 | 653 | 352 | 1005 |
| 7:00-8:00 | 30.3 | 1978 | 1065 | 3043 |
| 8:00-9:00 | 28.3 | 1847 | 995 | 2842 |

The new distribution of arrival times under the new policy can be calculated as follows. Trips not shifting to the new policy are assumed to stay at their present arrival times. For trips shifting to the new policy, the survey data show that between about 25 and 50 percent will move backward or forward in time to accommodate the new schedule (Figure 1). Thus, for example, the final arrival-time distribution for 5:00 to 6:00 a.m. can be calculated as follows:

| Category | | Number of Trips |
|---|---|---|
| Trips opting not to shift to new policy | | 53 |
| Trips shifting to new policy | | |
| Staying at 5:00-6:00 a.m. arrival time (0.25) (98) | = | 25 |
| Moving to 5:00-6:00 a.m. from 6:00-7:00 a.m. (0.25) (653) | = | 62 |
| Moving to 5:00-6:00 a.m. from 7:00-8:00 a.m. (0.25) (1978) | = | 495 |
| Moving to 5:00-6:00 a.m. from 8:00-9:00 a.m. (0) (1847) | = | 0 |
| Total | | 735 |

Similarly, trip ends for other time periods can be calculated as follows:

| Arrival Period (a.m.) | Number of Trips: Remaining on Original Schedule | On 4-d-Workweek Schedule | Total | Original Distribution | Difference (%) |
|---|---|---|---|---|---|
| 5:00-6:00 | 53 | 682 | 735 | 151 | |
| 6:00-7:00 | 352 | 1500 | 1852 | 1005 | |
| 7:00-8:00 | 1065 | 1112 | 2177 | 3043 | -28.4 |
| 8:00-9:00 | 995 | 0 | 995 | 2842 | |

These data show that the 7:00 to 8:00 a.m. period is still the peak hour but the peak is much flatter. The reduction in peak-hour traffic is 28.4 percent. In similar fashion, percentage reductions for the other two feasible policies are estimated at 4.0 and 6.0 percent respectively. Since the 4-d-workweek policy results in the greatest reduction by far in peak-hour demands for the campus area, the analysis of the impact on the local street system was continued for this policy only. It is assumed that other policies that result in a smaller reduction in peak-hour demand would have less impact on local street congestion.

## TRAFFIC ASSIGNMENT

The NYSDOT traffic-assignment model was then used to estimate the impact of the reduction in peak-hour demand on the local street system. The NYSDOT package is similar to the well-known Urban Mass Transportation

Figure 1. Trip shifts by hour of day (4-d workweek).

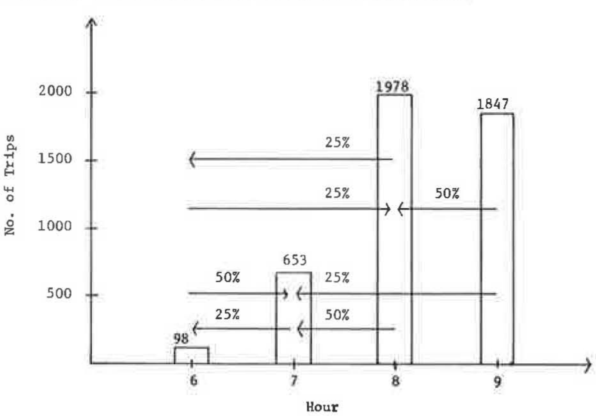


Table 1. Highway system operating conditions: null versus 4-d workweek.

| Item | Total System | | | District 38 | | |
|---|---|---|---|---|---|---|
| | Vehicle Kilometers of Capacity | Vehicle Kilometers of Travel | Speed (km/h) | Vehicle Kilometers of Capacity | Vehicle Kilometers of Travel | Speed (km/h) |
| Condition | | | | | | |
| Null | 6 285 300 | 1 539 200 | 29.84 | 339 700 | 69 000 | 25.00 |
| Test | 6 277 700 | 1 536 500 | 30.00 | 334 500 | 67 700 | 25.32 |
| Change (Δ) | -7 600 | -2 700 | +0.16 | -5 200 | -1 300 | +1.3 |
| Percentage change | -0.1 | -0.2 | +0.5 | -1.5 | -1.9 | +1.3 |

Note: 1 km = 0.62 mile.

Table 2. Impact of 4-d workweek on vehicle kilometers of travel.

| Location | Vehicle Kilometers of Travel on Arterials | | | Vehicle Kilometers of Travel on Expressways | | |
|---|---|---|---|---|---|---|
| | Null | Test | Difference (%) | Null | Test | Difference (%) |
| Campus area | 82 | 79 | -3.9 | 413 | 374 | -9.4 |
| Ring 1 | 1669 | 1669 | 0 | 3434 | 3405 | -0.8 |
| Ring 2 | 4184 | 4152 | 0.8 | 3050 | 3019 | -1.0 |
| Ring 3 | 1287 | 1289 | +0.1 | 1195 | 1185 | -0.7 |

Note: 1 km = 0.62 mile.

Administration package and operates by assigning trips over minimum time trees computed through the highway network. Full capacity-restraint options and distribution of trip ends by means of the opportunity model were used in the tests. The testing sequence to estimate the impact on the local street network was as follows:

1. Determine the flow of present (null case) traffic by conducting a full traffic assignment with the campus zone (49) at the current number of vehicle trips (N = 6396);
2. Determine the spatial extent of the traffic impact caused by the 4-d-workweek policy by running an assignment with the campus-zone trip ends reduced to 4542 (28 percent reduction as determined above);
3. Quantify the impact of this reduction on local travel by comparing vehicle kilometers of travel by zone and link type in the above two tests; and
4. Determine the actual travel impact of the policy by rerunning the full assignment with adjusted peak-hour factors to represent the smoother flow conditions in the peak hours generated by the removal of some of the campus-bound trips.

## RESULTS

A summary of the existing null conditions is given in Table 1. Of interest is the average speed on the network of 29.84 km/h (18.5 mph). This speed is the average of all speeds on all the network segments over a 24-h period. It takes into consideration speeds under free-flow conditions (usually high) and congested conditions (usually low). Note that district 38, which contains the campus zone, has an average operating speed of 25 km/h (15.5 mph).

A second traffic assignment was run to project traffic volumes under the conditions of the 4-d workweek. This run was identical to the first assignment except that the number of vehicle trip ends in the campus zone was reduced from 6396 to 4542. This run simulates the situation in which all State Campus employees are given an opportunity to be on a 4-d workweek. The results of this test are given in Table 2. Generally, the decrease in traffic volumes is slight. The drop in traffic of 3.9 percent on arterials and 9.4 percent on expressways in the campus area is clearly a much smaller drop than the 28 percent reduction in trip ends that caused it.

Table 2 also indicates that the effect of the 4-d workweek is highly localized around the campus area and dissipates quickly through the surrounding zone structure. Since this effect is entirely peak-hour travel, new peak-hour factors may now be computed for each affected zone that showed a reduction in vehicle kilometers of travel. The new peak-hour factors are computed by applying the percentage difference between the vehicle kilometers of travel to the base peak-hour factors in the null condition. The new peak-hour factors thus computed are used to update the distribution of trips among the various zones and network lines.

The analysis above assumes that the trips "removed" from the peak hour actually disappear—i.e., are not made. This is not the case: All of these trips are being made but at periods other than the peak hour. This means that only a small degree of relief in the traffic volume is expected to result from such reduction in the peak-hour demand.

To simulate this effect, the second assignment took into consideration the new peak-hour factors. Results are summarized in Table 1 ("null") and the following table [the economic analysis assumes 1970 dollars and a value of time of about $2.75/h (7)]:

| Item | Benefits ($) From Smoother Traffic | From Trips Not Made | Total |
|---|---|---|---|
| Travel time | 710 | 6664 | 7 374 |
| Operation | 73 | 2215 | 2 288 |
| Accidents | 12 | 352 | 364 |
| Total | 795 | 9231 | 10 026 |

Clearly, these results show only minimal impact as a result of even this extreme policy. A comparison of the two tests reveals that the daily travel cost savings to the region from the institution of a 4-d workweek on the State Campus would be only $795. These savings are concentrated in the immediate area surrounding the campus (district 38). Significantly greater travel savings also accrue directly to State Campus employees who participate in the program as a result of the 20 percent reduction in weekly work trips (table above).

## CONCLUSIONS

The central conclusions of this paper are the following:

1. The congestion-reducing impact of alternative work-schedule policies in highway-oriented cities is small. Even in surrounding areas of high employment, the impact dissipates quickly into the surrounding traffic.
2. Transportation benefits will accrue primarily to those who participate in such programs as a result of (a) the reduced number of required work trips and (b) avoidance of peak-period congestion. The general (non-participating) peak-hour commuting public will benefit only marginally.

These findings are particularly disturbing because they are based on an ideal test: A single, high-density, white-collar employer is assumed to adopt the policy that has maximum potential traffic impact. Most cities, however, have a wide mixture of types of employers and jobs dispersed over a wider area, and these employers are not all likely to choose the same policy. In central business districts, where work concentrations are high, the cooperation of numerous small employers may be difficult to achieve. Hence, in the general case, the congestion-reducing potential of work-schedule changes in automobile-oriented cities is probably small. Given the inherent problems of implementing such policies on a broad scale, even in small communities, the results here suggest that the traffic-reduction payoff may not be significant. These findings further suggest that, in the vast majority of American cities, transportation planners should not view alternative work schedules as a panacea to effectively reduce traffic congestion. Attempts to implement such policies, therefore, should not be motivated solely by potential reductions of traffic congestion but also by the other real personal benefits in job and family activities that they can provide.

ACKNOWLEDGMENTS

We wish to acknowledge the assistance of the Office of Manpower and Employee Relations of NYSDOT, Paul Zuber of Rensselaer Polytechnic Institute, and Elene Donnelly and Richard Albertin of NYSDOT for auspices and assistance provided during this study. The assistance of Wilma C. Marhafer in preparing the manuscript is gratefully appreciated.

REFERENCES

1. J. H. Dupree and J. K. Pratt. Low Cost Urban Transportation Alternatives: A Study of Ways to Increase the Effectiveness of Existing Transportation Facilities. U.S. Department of Transportation, Jan. 1973.
2. R. Remak and S. Rosenbloom. Peak Period Traffic Congestion. NCHRP, Rept. 169, 1974.
3. B. O'Malley. Staggered Work-Hour Project in Lower Manhattan. HRB, Highway Research Record 348, 1971, pp. 152-165.
4. J. Guttman. Predicting the Effect of Variable Work Hours on Peak Period Congestion. Paper submitted to Transportation, Feb. 1977.
5. C. S. Selinger. Managing Transportation Demand by Alternative Work Schedule Techniques. TRB, Special Rept. 172, 1977, pp. 67-74.
6. A. A. Tannir. The Impacts of Feasible Staggered Work Hours and Compressed Workweek Policies on Highway Networks, Transportation Economics, Organizations and Employees. Planning Research Unit, New York State Department of Transportation, Albany, Preliminary Research Rept. 129, 1977.
7. P. A. Albin and J. K. Lussi. Documentation of the Highway User Cost Accounting Program. New York State Department of Transportation, Albany, Rept. PUPR 4, Nov. 1971.

*Publication of this paper sponsored by Committee on Transportation Planning Needs and Requirements of Small and Medium-Sized Communities.*

**Mr. Tannir was with the New York State Department of Transportation when this research was performed.*

# Development of the California Transportation Plan: 1973-1977

William S. Weber,* Southern Pacific Transportation Company, San Francisco

**The California Transportation Board adopted Recommended Statewide Transportation Goals, Policies, and Objectives in March 1977, marking completion of the first of six elements of a comprehensive transportation plan that 1972 legislation originally had mandated must be completed by January 1, 1976. Because of administration and public criticism of the first draft and legislative failure to adopt the necessary transportation goals and policies, that mandate could not be fulfilled. The controversy that arose over the initial plan element suggests that remaining elements may not be completed. However, the California Transportation Board feels that several recommendations in the completed (or policy) element may eventually be adopted, if only piecemeal. This paper describes the evolution of the California Transportation Plan from development of the initial draft by the California Department of Transportation through various iterations and examines the difficulties that surround creation of an objective document in the face of interests that benefit from maintaining the status quo.**

In March 1977, the California Transportation Board adopted Recommended Statewide Transportation Goals, Policies, and Objectives. This action was the culmination of 4 years of work that, although it did not result in a completed plan, represented an intensive cooperative effort by the board, its staff, the California Department of Transportation (Caltrans), and an interdisciplinary task force.

The goals, policies, and objectives (or policy) element was the first and most important segment of the plan because it was to guide development of the remaining four elements. That the plan did not advance beyond this element was attributable to the controversy that enveloped this first stage. Legislation that mandated the drafting of the plan (Assembly Bill 69 of California Legislature, Chapter 1253, Statutes of 1972) also stipulated that the legislature must adopt the goals, policies, and objectives before subsequent elements could be completed. Instead, controversy negated legislative approval and brought the planning process to a halt. One result was the introduction and passage of legislation to restructure both the transportation organization and the planning process.

The California Transportation Plan had its genesis in the legislation that created Caltrans in 1973. This legislation also mandated a specific transportation plan process including the preparation of regional and state transportation plans. The regional plans were to be an integral part of the state plan.

## CONTENT OF CALIFORNIA TRANSPORTATION PLAN

The original legislation mandated specific elements in addition to the regional transportation plans such as statewide transportation goals, objectives, and policies; statewide forecasts of transportation needs and deficiencies; and an implementation program.

Also, two progress reports were to be submitted to the legislature—the first by July 1, 1974 (1), and the second by January 1, 1975 (2).

The legislation specified that each progress report should contain information such as a definition of statewide transportation goals, policies, and objectives; recommendations concerning regional goals and policies; and the manner in which economic, land use, taxation, and other specific criteria would be incorporated in the plan.

Caltrans was to prepare the plan and progress reports under board supervision. In addition, the board was to hold public hearings before adopting the plan and transmitting it to the legislature by January 1, 1976. However, a critical caveat included in the legislation stated that the board could not adopt the plan until the legislature had approved (or modified) the statewide goals, objectives, and policies that the board was to submit in the second progress report.

## SPECIAL REPORT TO THE LEGISLATURE

In conjunction with legislative deliberation and ultimate passage of the requirement for a transportation plan was a request by the legislature that the board prepare a special report (3), in advance of development of the plan, that would discuss pertinent transportation issues. The issues to be studied were provisions for local control over future transportation development, discussion of the need for creation of regional planning agencies with authority to implement their plans, provisions designating authority and responsibilities for the control of resource allocations for transportation, and proposed changes to be made in planning practices. These issues proved to be very controversial. However, preparation of this report provided a good background for the board's subsequent work on the plan itself.

## FIRST WORK ON THE PLAN

After the official establishment of Caltrans in July 1973, much of the first year was devoted to organization of the new department; some preliminary work on the plan was initiated. In April 1974, the first progress report was submitted to the board and was in turn transmitted to the legislature with the board's independent evaluation of progress.

In its transmittal to the legislature of this first progress report, the board identified issues that it thought required early legislative consideration, including approval of a statewide transportation goal, efforts to eliminate federal bypass grants to local agencies, and authorization of the board to serve as arbitrator in resolving conflicts between regional and local levels (4). In addition, the board noted areas in which, in its view, the progress report was insufficiently detailed. The board stressed the view of the "ultimate transportation system as multimodal . . . resisting domination by any particular mode."

The second progress report, an update of the first progress report that was to include the first three of seven mandated studies, reported on work activities completed or nearing completion. This report included a caveat that "the first version of the plan could fall short of the expectations of a limited few" (2). This caveat was added because in meetings between Caltrans and the board several members had expressed the view that some of the more pertinent issues were not being included in the plan. Although these issues were not included in the Caltrans second progress report, Caltrans assured the board that many would be included in the final draft plan.

In transmitting the second progress report to the legislature, the board stated that the issues that should be included in the plan and that the department had assured the board would be included were issues dealing with energy conservation, air quality, transportation deregulation, leverage of private capital in providing transportation, transit operating subsidies, the role of innovative modes and new technology, and noncapital alternatives to transportation improvements.

In response to a formal statement by the board of its concern about these issues in its transmittal to the legislature, Caltrans reported that it would contract for a series of issue papers on these topics. The papers were to be completed in time for inclusion in the final plan; however, Caltrans cautioned the board that, although it would have the special consultant reports to augment its staff work, not all issues could be addressed in the first version of the plan. Caltrans stated, for example, that there would be some recommendations but no final studies on energy conservation, air quality, transit operating subsidies, and noncapital alternatives.

The board agreed that Caltrans could not fully address each issue, but it would expect some work to be undertaken. In addition, the board stated that it would not expect Caltrans to prepare original research on each issue since research on all these issues had been carried out by others. Caltrans was urged to make use of such research.

## TRANSPORTATION PANEL

In February 1975, the California Transportation Board convened a panel of well-known experts in transportation planning to obtain their views of the plan (5). The panel cautioned the board not to be overly optimistic on what a plan could accomplish. Panel members stated that a planning document should be an aid in decision making and should be specific. Many plans were too general for fear of being offensive. The plan should also be comprehensive and recognize the relation between lifestyles, land use, and transportation.

In seeking various views, the board also asked its staff to prepare a paper on what the plan should include based on staff views and review of research in the field. Essentially, the staff paper focused on the need for a policy orientation in the state-level plan.

In response to the board transmittal of the second progress report to the legislature, the views of the panel, and research work by the board staff, Caltrans attempted to shift plan development closer to what the board wanted. However, Caltrans cautioned that, if the plan was to have more of a policy orientation, much of the work would have to be redone and the July 1 schedule to begin public hearings probably could not be met. The board stated that it did not want Caltrans to make a major shift in the work but, to the extent possible, to develop the issues that the board had requested.

## ROLE OF BROWN ADMINISTRATION

When the administration of Governor Edmund G. Brown, Jr., took office in 1973, the developing California Transportation Plan was one of the first major transportation issues facing his appointees. By early spring, these appointees had reviewed Caltrans' work on the plan and expressed definite concerns. In his first public statement about the plan, the new secretary of business and transportation referred to it as a "veritable wind tunnel of rhetoric." Administration appointees involved with transportation policy agreed with the board that the plan should include a discussion of policy alternatives before proceeding to programs.

In the summer of 1975, a draft (6) was circulated for public hearings. Although this draft included a series of alternatives that attempted to respond to the board and the new Business and Transportation Agency representatives, these alternatives were considered incomplete and limited. Although most highway support groups and local governments generally supported the plan, most public comment was critical. After the hearings, the board met to discuss the plan with a representative of the secretary of business and transportation.

In summarizing public hearing comments, Caltrans viewed the hearings as very supportive and recommended steps it proposed to take to complete the plan. The board and the administration representative countered that in their view there was little support for the plan and the direction in which Caltrans was proposing to complete the plan was leading to conclusions and recommendations not substantiated in the plan.

Board staff summarized the various points that had been raised during the public hearings and also noted numerous policy deficiencies in the draft plan that the board had previously stated should be covered before specific programs, project recommendations, or conclusions were developed. In addition, in October 1975 the board staff prepared recommendations concerning subsequent action on the plan (Issue Memorandum 35). The primary recommendation was not to proceed with adoption of the plan because the document did not provide an adequate base for providing recommendations in conformance with either legislative requirements or currently accepted planning practices and because of the lack of response to concerns previously expressed by the board on policy issues. The staff recommended deferral of the adoption mandate and presented several alternatives for board consideration that ranged from adoption of the recommended draft with comments to rejection. The staff also suggested a new approach for developing a satisfactory plan, which included creation of an interagency task force that would draw on Caltrans, other state agencies, and outside consultants.

The workshop discussion crystallized the opinions of board members, and they found themselves in almost unanimous agreement that the plan in its current draft form was unacceptable. The board sent a letter to Governor Brown and the legislature that in effect represented a rejection of the document; although acknowledging that the draft represented tremendous dedication and hard work by Caltrans, it pointed out that the members had been expressing their concerns for many months. The board recommended that an ad hoc multidisciplinary task force be appointed by the secretary of business and transportation to analyze and redirect development of the plan.

## CALTRANS VIEW OF PLAN REJECTION

Management and most of the planning staff of Caltrans believed the board was very wrong in its views, and the Caltrans director wrote the board expressing his disappointment in the board's action. He also cited what he viewed as ambiguity, overambitiousness, and fuzzy direction on the part of the legislature, the administration, the board, and the board staff. In his view, a 20-year master plan had no chance for broad-based endorsement, and he urged the board to work with him to prepare a 5-year plan to assist in state-level transportation decisions.

In contrast, the secretary of business and transportation thanked the board for its thoughtful action, said he did not feel the draft plan was usable as a policy or program guide for transportation, and agreed to organize a new planning effort along the lines suggested by the board.

## NEW PLAN

In October 1975, the assistant to the secretary of business and transportation presented a work program (7) for preparing a new policy-based plan. The basic approach was the development of various analytical studies or issue papers from which elements would be drawn as directed by the board to form the plan. The issue papers were categorized loosely in two major groupings: background studies and inventories (set 1) and issue analyses (set 2).

To be discussed under background studies and inventories were issues such as the definition of statewide interest or significance in transportation, characteristics of passenger travel and commodity movement, and statutory requirements that affect transportation. These studies and inventories paralleled legislative requirements that Caltrans either had not produced or had only partially developed.

To be discussed under issue analyses were issues such as air quality, energy, land use and transportation, the transportation disadvantaged, involvement of the private sector in transportation, new technology, and alternatives to public investment in and operations associated with transportation.

The work program proposed a small interdisciplinary task force organized under the Business and Transportation Agency. This task force was to consist of nine persons from various disciplines such as economics, with specialization in public investment and cost-benefit analysis, regulatory and market analysis, and welfare economics, political science, law, environmental planning (land use and urban planning), and transportation planning. Caltrans would provide backup staff and support assistance. Three advisory groups were proposed to "review products and provide comments on analysis, findings, and recommendations." The first would be an interagency advisory group composed of representatives of other state agencies involved with transportation. The second would include representatives from the private sector—both those directly engaged in transportation, such as modal operators, and those concerned with transportation, such as business people, environmentalists, and academics. The third group would be Caltrans management.

The schedule proposed in the work program envisioned starting work in November 1975 and completion by July 1976. As a result of delays in recruiting and organizing the task force and later delays involved with public hearings and redrafts, completion dates were rescheduled several times. Ultimately, completion was set for March 1977.

The skills of task force members actually recruited differed from the skills originally sought. The project manager was an engineer, and the staff included an engineer, a lawyer, two economists (one specializing in transportation economics and regulation), a financial expert, an environmentalist, a city planner, and a transportation planner.

Statewide Significance

Two issue papers that never completely satisfied the board were those on statewide significance and financing, and the board therefore considered the plan weak in these areas.

The concept of statewide significance derived from the costs and benefits of transportation activity, particularly "spillover" costs and benefits. To justify state action or involvement in any transportation activity, the activity should involve significant spillover of costs and benefits beyond local or regional boundaries. Similar spillover criteria could be used to judge whether an activity was of regional significance or only of local significance.

Criteria on statewide significance were intended to provide the legislature and the administration with the opportunity to identify those transportation activities that are of significance to the people of the state generally and for which the state should be prepared to pay. The board did not intend to prejudge but to provide the decision maker with a framework for evaluating issues and making policy judgements. The board recognized that spillover criteria had to be tempered in situations where sudden imposition of spillover criteria could cause severe dislocation or hardship. However, it was the board's desire that future decisions of the legislature and the administration, as well as those of regional and local officials, would gradually come to be based on spillover.

The board also sought a broader base for determining statewide significance than economic criteria alone. But the issue paper on this subject did not, in the board's view, adequately define the issue so that a determination could be made easily based on other than economic criteria or even assist in the many borderline cases.

As work on the plan progressed, the board began to refer to the document as the policy element of the plan. Believing that the plan could not do justice to all the requirements of the statute, especially the implementation program, the board retitled the document California Transportation Plan–Recommended Statewide Transportation Goals, Policies, and Objectives.

After completion of the policy element, an implementation element that met other requirements of the legislation could be prepared by Caltrans. The implementation element, of course, was to be based on the policy element. It was the lack of a policy base that had caused the board to be so critical of the 1975 draft and that the board felt had led to the unsubstantiated conclusions in that document.

Public Hearings

Six public hearings on the new draft were scheduled for November at various locations in the state. In advance of the hearings, the Business and Transportation Agency coordinated a full-time public information program to inform the public about the plan and the scheduled hearings.

Because the task force had been created in the agency, the director and the agency viewed the draft plan as an administration document prepared for the board under the board's direction. Only when the board adopted the plan, after any redrafting based on the public hearings, would it be a board document. The board, supported by the agency, purposely attempted to avoid any advocacy position in advance of the public hearings to ensure the board's objectivity throughout the hearings. This point is of interest because later, when the plan became very controversial, the administration and the agency quietly withdrew mention of their involvement in preparing the document.

The information program conducted by the Business and Transportation Agency included over 150 public information meetings throughout the state, newspaper advertisements announcing both the information meetings and the public hearings, and numerous radio and television appearances by representatives of the agency, the task force, and the board staff to discuss the plan and generate interest in the public hearings. To avoid the 1975 experience when too few copies of that plan were printed and distributed, 4500 copies of the new draft plan were printed and circulated early in October, approximately a month before the first hearing.

Misunderstandings Created by Newspaper Reports

Shortly before the draft plan was completed and circulated for public comment, newspaper articles appeared that caused considerable misunderstanding about the plan. These news stories, which drew from the content of both the issue papers and the draft plan, were the primary cause of controversy that enveloped the plan from that time on. In drawing from the issue papers, the news articles in some instances reported on alternatives or recommendations that the board had rejected. In referring to the draft plan, some articles gave misleading interpretations of the intent of the plan. Unfortunately, the general public, reacting only to the newspaper stories without ever seeing the plan, wrote the governor and their legislators to express alarm and indignation.

The public reacted to stories that reported the state was proposing such programs as tolls for use of freeways, increases in the gasoline tax of as much as \$0.13/L (\$0.50/gal), and other actions to "force" people out of their automobiles. These stories were based on issue paper discussion of the extent to which individuals were not paying the full cost of their use of transportation facilities and alternative methods by which a greater share of the total costs of transportation services (including environmental costs) might be assigned to users. The issue papers cited corroborating examples: e.g., the rush-hour freeway use of a vehicle occupied only by a driver, an obvious instance of a user's receiving more service than that for which he or she pays.

Although the issue paper acknowledged the impracticability of assessing or collecting for such peak-hour use or collecting a full-cost charge for use of the highway at any time of day, newspaper articles interpreted this discussion as a proposal that the state was planning to increase gasoline taxes by as much as \$0.13/L and to place tolls on the freeways.

Concerns Expressed

As the public information meetings and public hearings progressed, other misunderstandings became evident, and these the board eliminated or modified in the final document. This, of course, was the purpose of the public hearings. However, the controversy generated by the initial news stories and the public reaction was never completely overcome. Groups concerned about highway programs in their areas were especially critical of the plan. They believed that if the policies were implemented there would be drastic reductions in the funds available for highway construction and maintenance. Interestingly, individuals were more favorably disposed to the plan than were representatives of organized interest groups such as chambers of commerce, local government, the trucking industry, and automobile clubs.

The information meetings and extensive media coverage generated considerable interest in the public hearings. Over 1000 written comments were submitted to the board by individuals who could not personally attend the public hearings.

### Revisions Made After Public Hearings

Extensive textual changes in the plan were undertaken as the result of the public hearings. The board attempted to clarify policies and to define the intent of principles. Although documentation of each change here would be superfluous, a few significant examples illustrate the tenor of the amendments.

#### Chapter 1—Alternative Directions for California Transportation

This chapter of the plan was principally background discussion devoted to an exposition of transportation problems—funding, environmental, and operational—and assuming the many advantages of the existing system. However, heavy criticism was directed at the negative tone. The board accordingly restructured the chapter to provide an expanded description of existing system benefits, which, in the opinion of many at the public hearings, were extensive.

#### Chapter 2—Basic Principles

This chapter describes eight basic principles essential to the transportation decision-making process. Two principles caused special consternation at the public hearings.

"Full social accounting" was the term used to describe the importance of considering environmental, social, and economic advantages and disadvantages in making any transportation decision—a concept that goes well beyond the typical evaluation of simple financial cost-benefit relations. The term was confused with "social engineering," and thus the principle was retitled "full consideration of effects" and the supporting narrative clarified.

Of even greater importance was the concern that arose from the principle that called for full assignment of costs to system users. The public, in general, apparently believed that, with the exception of mass transit, users do in fact pay for services received. The gasoline tax in California had long been cited as an example of a direct charge on highway users that could be equated with the cost of highway construction and maintenance. The plan draft proposed extension of this principle to encompass full cost assignment, including the social, environmental, and economic costs that arise from use of a transportation service or system. Accomplishing this would, in users' eyes, have required a considerably heavier tax burden (i.e., a stiff boost in gasoline taxes), a proposition that proved controversial. The principle was therefore redefined to clarify that the concept could only be implemented over a long period of time and even then only with adequate assistance for low-income groups and safeguards to minimize dislocations. The plan narrative was amended to point out that any increase in user charges must be softened by reductions in other taxes and elimination of cross subsidies.

Finally, as the result of intensive pressure from the trucking industry (both management and labor), policies that called for eliminating regulation of intercity trucking were softened. Although the public hearing draft recommended elimination of economic rate and entry regulation, in rewriting the plan the board stated that changes were needed in regulation and further studies should be carried out to determine the best approach for bringing this about. The basic view of the board, though, was that the public would be better served by immediate steps toward deregulation in all areas of transportation.

### The Eight Basic Principles

The policies in the plan are based on the eight basic principles in Chapter 2. The policies deal with specific institutional issues, resource and environmental issues, and transportation issues by mode: highways; public transit; bicycles; freight transportation by highway, rail, air, water, and pipeline; and intermodal freight transportation:

1. Government role—The state should allow decision making by private enterprise to prevail in as many areas of transportation as possible. When government does involve itself in transportation decision making, it should strive to provide services that are as effective and efficient as possible.
2. Transportation management—Actions to make the use of the existing transportation system more efficient and effective should be considered before decisions are made to add to the system. These actions should be adopted when they can be expected to increase transportation efficiency and effectiveness or to improve the social, economic, and natural environment or both.
3. Alternatives—State, regional, and local decisions should consider a wide range of reasonable alternatives. Analysis of these alternatives should take into account different value systems or points of view held by various elements of the public that may be involved.
4. Full consideration of effects—Government decisions of major significance should be informed by a full analysis and disclosure of the advantages and disadvantages of the decision, including environmental, social, and economic effects and identification of the different interests that are affected.
5. User charges—Whenever possible and equitable, user charges should be encouraged. Users should be required to pay a fair share of the costs that occur from their use. User charges should be adjusted gradually and only after a careful analysis of their impact.
6. Equity—Where user charges do not cover at least a fair share of the costs, taxpayers who receive services should be the ones that pay for them, and those who suffer burdens or damages should be compensated. The costs charged should be in proportion to the benefits received, and the compensation to the general public should be in proportion to the damages suffered.
7. Basic transportation—Some form of basic transportation should be available to people who need it. Transportation policy should not result in low-income and handicapped individuals having to pay a disproportionately large share of their resources for necessary transportation in comparison with other individuals.
8. Government regulation—The California Trans-

portation Board should cooperate with appropriate entities such as the legislature, the public utilities commission, the federal government, and affected groups in review of regulation of interregional movement of goods and people. The amount of regulation should be no greater than that required in the public interest. Necessary long-range regulatory reform of transportation should be sought. Short-term reform of transit and paratransit regulation within a region or local area should be sought with concurrence of the local agencies involved. The state should develop and maintain environmental protection regulations that are necessary to preserve a safe and acceptable quality of both community and natural environments. The state should maintain quality-of-service regulation where it is necessary to protect the public welfare. Regulation of safety, health, and financial liability should safeguard individuals from hazards they are not able to perceive or account for but should not interfere with normal risk-taking choices.

## CHANGES IN THE ADMINISTRATION VIEW

The California Transportation Board held two meetings to review public comment and give direction to the task force for redrafting the policies. Between the two meetings the board received a joint letter from the director of Caltrans and the secretary of business and transportation that expressed administration concerns about the plan as it was circulated for public hearing.

The current assistant to the agency secretary had participated in most board discussions of the plan throughout 1976. Previously, the Caltrans director, in her former position with the agency, had participated with the board during public hearings on the earlier plan the board rejected and had drafted the work program for preparing the new plan. Thus, the board had the impression that these administration representatives had been informing the governor about the plan. As events developed, the governor apparently was not that well informed but became actively aware of the plan when critical newspaper articles and letters began to arrive in his office.

The joint letter from the agency secretary and the Caltrans director was sent to the board at the insistence of the governor. By that time, the board had already directed that clarifying changes be made in the plan in most of the areas the administration discussed. Major points in the administration letter questioned the equity and practicability of the pricing strategy that called for vehicle stickers for air quality, "smog taxes" imposed by regional agencies, freeway tolls, transportation stamps, and other sophisticated pricing mechanisms; opposition to any general increase in the level of taxation; and concern that board policies on deregulation of the transportation industry could have profound economic consequences for the state. (The plan did not propose a general increase in the level of taxes. It did propose that, whenever possible, users pay a reasonable share of facility costs and that other taxes be decreased. In addition, opposition to deregulation had been strongly expressed by the trucking industry to the administration and to individual legislators who in turn communicated their views to the governor.)

The letter also stated that the plan failed (a) to acknowledge regional transportation plans and unique problems associated with rural California, (b) to separate immediate transportation problems from longer term issues, and (c) to define precisely the phrase "full social accounting" and take into consideration the full range of benefits, as well as the costs, in the decision about any particular transportation investment.

In closing, the letter expressed concern about what the administration believed was the generally negative approach of the plan in dealing with the various transportation issues and asked the board to consider greater use of incentives with correspondingly less use of disincentives. Concern was also expressed about the way in which the plan was developed and specifically expressed the administration view that there had been inadequate public participation and consultation with the legislature and local governments.

The letter from the administration was widely reported by the press and interpreted as a reversal by the administration of previous support for the plan. The press pointed out that the task force was under the Business and Transportation Agency and that the agency secretary and the Caltrans director had, until they wrote the letter, been strong supporters of the policies in the plan. After receipt of the administration letter, however, all reference to the secretary and the Business and Transportation Agency was removed by the agency in the document prepared for final adoption.

## CONFUSION OF PLAN WITH HIGHWAY PROGRAM

Concurrent with the public hearing process, the draft policy plan became confused with another transportation document that generated considerable controversy in California—a 6-year highway program prepared by the director of Caltrans. This program proposed drastic reductions in funding of new highway construction and placed greater emphasis on highway maintenance and safety projects. The California Highway Commission, the majority of which was composed at this time of appointees of the previous administration, was upset about the proposed reductions in funding new construction projects as were the legislature and various highway support groups. The administration was viewed by the members of the commission, the legislature, and highway groups as philosophically opposed to construction of new highways.

The conflict between the director and Caltrans on the one side and the California Highway Commission and the legislature on the other generated considerable publicity. Because both documents dealt with transportation, the 6-year highway program was often confused with the policy element of the California Transportation Plan. Much of the legislative criticism directed at the administration's highway policies therefore overflowed into criticism of the California Transportation Plan.

## FINAL DRAFT PLAN

While controversy and confusion were engulfing both the transportation plan and Caltrans, the board redrafted the plan in response to comments from the public hearings and the letter from the administration and scheduled a final public hearing to determine whether the public viewed the redraft as meeting all or most of its earlier concerns. Copies of this draft were distributed to the public a month in advance of the final public hearing on March 17, 1977. During this period, an interim secretary of the Business and Transportation Agency replaced the secretary who had represented the administration during 1975 and 1976. This change required time to acquaint the new interim secretary with the plan and the history of events involving the plan. (Shortly after the plan was adopted, the interim secretary was replaced by a new permanent secretary who also had to be acquainted with the plan and its history.) The interim secretary wrote the board that in his view major improvements had been made toward overcoming the concerns expressed by the public, the legislature, and the administration about the earlier draft, and he looked forward to the orderly completion of the board process and the transmittal of a completed policy element to both the administration and the legislature by mid-April 1977.

## ADOPTION OF POLICIES

Without changing the general philosophy of the policies, the plan was redrafted to clarify the areas that had been misunderstood by the public and in response to criticism. At the final public hearing, the board directed that a few final changes be made based on the public comment received at the hearing. Most of the changes were not substantive. Some suggested changes that the board put aside for future consideration. In adopting and transmitting the policies to the legislature, the board recognized that legislative action on the policies was not likely and that the remaining planning elements mandated by the legislation would not be undertaken.

Several members of the legislature, although still critical of many of the policies, admitted it was not an unreasonable plan and that the document was better written and more objective than the October draft. However, newspaper articles at the time of the first public hearing the previous fall had done irretrievable damage. The criticism continued, and too much controversy had developed around the document for it to be adopted as a whole.

## SUPPORT FOR THE POLICIES

It was the board's belief that many of the policies would eventually be enacted into legislation in California, in other states, or at the federal level. Although the Caltrans director had been a joint signatory of the letter to the board in December, she continued to express general satisfaction and support for the policies. Outside of state government, supporters of the policies have included environmental groups such as the Sierra Club, the League of Women Voters, the railroad industry, a number of conservative economists, and numerous shippers. As stated earlier, the critics have included automobile clubs, the trucking industry, chambers of commerce, and numerous local governments.

## DIFFICULTY OF PLAN DEVELOPMENT

Board staff, and to a lesser extent board members, were viewed by Caltrans as expecting too much to be accomplished in the first plan. If the board had not rejected the plan prepared by Caltrans, it is possible that, after several updates, it might have approximated the desires of the board. But it is questionable whether a plan based on policy issues would have been developed.

Development of a transportation plan or transportation policies is difficult, as the California experience indicates. If the policies are significant and if difficult issues are treated objectively, long-established practices will be affected. Interest groups will oppose those recommendations that propose changes in benefits they receive from entrenched procedures. The public is apt to misunderstand the purpose of recommended changes, and some portions of the public also benefit from the status quo.

The board had been cautioned by the panel of experts it had convened in February 1975 that many plans were too general for fear of being offensive. In addition, criticism is often voiced by the public as well as by interest groups that government plans are too general and accomplish very little. However, based on the experience of developing objective transportation policies, it may be that only general, noncontroversial plans are possible.

In California, the policies proposed by the board dealt very specifically and objectively with a wide range of transportation issues. As a result of the controversy that enveloped the plan, the legislature indicated that it would not act on the document. It used criticism both of the plan itself and of the time taken to develop it as one reason for enacting legislation to overhaul the transportation planning process. In doing so, emphasis was placed on the development of short-term (up to 5 years) plans and improvement programs. This legislation also created an entirely new state-level planning body, the California Transportation Commission. The new commission must report on long-range issues as it believes appropriate—but without any guidance from the legislature as was required in the statute that governed the California Transportation Board.

It has been argued that the plan would not have become so controversial if there had been more discussion with the public and interest groups. This is debatable. There were extensive public information meetings and public hearings. Also, in view of the considerable length of time required in preparing the plan, it is unlikely that the political climate would have permitted further public hearings since these would have extended the completion date even further.

It is unlikely that far-reaching transportation policies and plans can be acted on or implemented from a single document. The major benefit of the policies will be educational. However, if the policies are objective and if the issues have been conscientiously studied, many of them will in time be accepted and implemented. Political pressures and interest groups will stall some policies, and a few of them will become outdated because of changing conditions. But many will survive, and this is the challenge.

## REFERENCES

1. California Transportation Plan—First Progress Report. California Department of Transportation, April 1974.
2. California Transportation Plan—Second Progress Report. California Department of Transportation, Dec. 1974.
3. Office of Transportation Planning and Research, California Transportation Board. Summary of Findings and Conclusions on State, Regional and Local Roles and Responsibilities in Transportation—A Report to the Legislature as Required by Section 13991. California Legislature, Sacramento, Statutes of 1972, Chapter 1253, Jan. 1973.
4. Evaluation of Progress and Recommendations to the Legislature. California Transportation Board, July 1974.
5. What Should Be Included in a Transportation Plan for California. California Transportation Board, Feb. 1975.
6. California Transportation Plan—A Draft. California Department of Transportation, July 1975.
7. Work Program. California Business and Transportation Agency, Sacramento, Oct. 1975.

*Publication of this paper sponsored by Committee on Intergovernmental Relations.*

**Mr. Weber was with the California Transportation Board when this research was performed.*

# Data Requirements for an Analysis of Intercity Passenger Travel by Bus

Stanley L. Ring, R. L. Carstens, and James D. Grove, Iowa State University

**Transportation planners in Iowa have initiated efforts to enhance the attractiveness of passenger travel by intercity bus to exploit more fully the inherent advantages of that travel mode. However, insufficient data have been available to provide a firm basis for such planning. Data traditionally provided to regulatory agencies are aggregated so as to afford little information on patronage for specific cities or specific route segments. Data from a 100 percent sample of ticket sales for a summer month from 23 focal communities in Iowa—including travel volumes, trip lengths, and origin-destination information—provided a basis for the analysis of bus travel in the state. Additional data were made available from a metropolitan station to indicate seasonal and daily variations in travel demand and the proportion of on-time service.**

Patronage of intercity buses was at essentially the same level in 1975 as it was in the late 1940s in both passengers and passenger kilometers. However, intercity travel increased by approximately 200 percent during that period. The bus share of intercity passenger kilometers decreased from 5.6 percent in 1948 to 1.9 percent in 1975 (1).

There are several reasons for the relative decline in importance of buses as carriers of intercity passengers. Travel by automobile offers substantially more flexibility and is generally perceived as more comfortable and as having other advantages, and air travel saves time. Consequently, these modes have experienced substantial increases in use during the past 30 years.

A large proportion of current intercity bus users are captive to the bus mode. The captive group includes, for example, a disproportionate number of persons without access to an automobile and those who are elderly, handicapped, or economically disadvantaged.

Transportation planners frequently express a desire to increase bus use by elective riders—those who have a choice of modes and are attracted to buses in the interest of economy or efficiency. This desire arises in part from the need to slow down the rate of growth in vehicular use of highways in response to reductions in the amount of improvement that can be effected under current highway funding programs. Concern about the consumption of energy also favors the increased use of buses. Although many factors influence comparisons of modal energy efficiency, such as circuity of routing and passenger load factors, research efforts generally have demonstrated that buses have pronounced advantages in terms of energy consumption per delivered passenger kilometer. Conclusions from one study, for example, were summarized as follows (2): "Buses are the most fuel efficient mode for all city pairs." Figure 1 (2) shows a comparison of modal fuel efficiencies for passenger modes and indicates clearly the inherent advantage of bus travel on the basis of fuel consumption.

It is possible for changes in government policies to exert a significant influence on the attractiveness of intercity bus travel. The amount of subsidy provided to competitive modes—rail and air—may be decreased. In respect to highway fuels, government-induced price changes or limitations on their availability will tend to decrease highway travel. The traditional government role in regulation provides additional opportunities to affect the relative use of different travel modes.

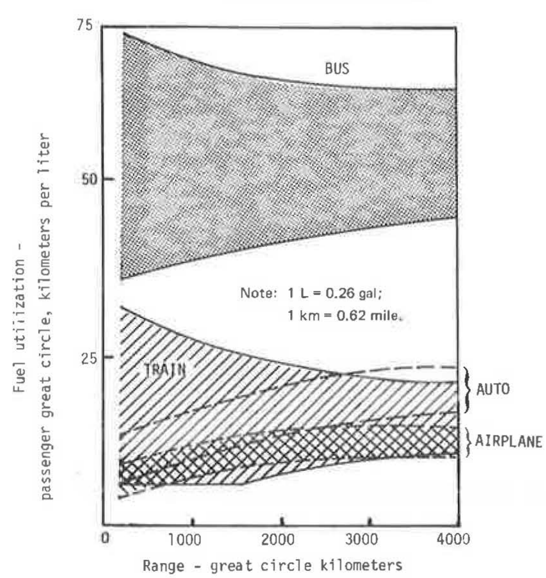
Unfortunately, transportation planners responsible for recommending policy changes are severely handicapped by limitations on the availability of data on the use of intercity buses. In contrast to the wealth of information available on most aspects of highway travel, little information is generally available on bus passenger movements on specific routes or from specific stations. Data comparable to highway origin-destination surveys, travel time and delay studies, or traffic volume counts simply do not exist in the context of intercity bus travel.

Data that do exist on bus passengers are usually aggregated to satisfy the needs of regulatory agencies and provide little information suitable for detailed planning. Some detail is available on vehicle movements (useful for highway planning) but not on person movements (essential for passenger travel planning). Nor are bus companies particularly concerned with gathering detailed passenger data. As profit-making enterprises, their concerns are directed toward the financial aspects of operations and result in the accumulation of data that may be adapted only with considerable difficulty to the needs of transportation planners. In addition, proprietary concerns about such data preclude their release in a form potentially useful to competitors. As a consequence, it was necessary to generate almost the entire data base concerning bus passenger movements from primary sources as part of the research effort reported here (3).

## DATA ON INTERCITY BUS PASSENGERS

A duplicate of each intercity bus passenger ticket is retained by the manager of the bus station. These receipts constitute a record of passenger revenue that is the usual basis for payment of commissions to station managers. Periodically (weekly, semimonthly, or monthly), these receipts are forwarded to a central office of the carrier. Larger carriers normally will not provide re-

**Figure 1. Modal comparisons of fuel efficiency.**



searchers with access to these records.

One aspect of importance to transportation planners is that these records also identify the origin and destination of each bus passenger. With the permission of the carriers involved, this data source was made available for research on bus passenger travel by the station agents at 23 selected cities in Iowa. The cities were selected on the basis that they provided a regional focus for much of the travel that uses common carriers within the state.

A 100 percent sample of all bus tickets sold at each of the 23 stations during a summer month was processed for this research. Tickets were reviewed immediately before being forwarded to the carrier, and a destination was recorded for each ticket. Thus, a complete record was available of all bus tickets sold at 23 of the primary

**Table 1. Bus tickets sold in selected Iowa cities in a typical summer month in 1976.**

| City | Destination: Iowa | Destination: Out of State | Destination: Total | City | Destination: Iowa | Destination: Out of State | Destination: Total |
|---|---|---|---|---|---|---|---|
| Ames | 686 | 431 | 1 117 | Iowa City | 1 490 | 1 003 | 2 493 |
| Atlantic | 112 | 167 | 279 | Marshalltown | 271 | 143 | 414 |
| Burlington | 504 | 284 | 788 | Mason City | 651 | 520 | 1 171 |
| Carroll | 97 | 78 | 175 | Muscatine | 212 | 153 | 365 |
| Cedar Rapids | 2 302 | 795 | 3 097 | Osceola | 100 | 68 | 168 |
| Clarinda | 33 | 70 | 103 | Ottumwa | 492 | 227 | 719 |
| Clinton | 174 | 271 | 445 | Sioux City | 800 | 1 510 | 2 310 |
| Council Bluffs | 370 | 96 | 466 | Spencer | 110 | 139 | 249 |
| Davenport | 754 | 569 | 1 323 | Waterloo | 1 038 | 740 | 1 778 |
| Decorah | 149 | 231 | 380 | West Union | 47 | 9 | 56 |
| Des Moines | 3 947 | 3 479 | 7 426 | Total | 15 608 | 12 266 | 27 874 |
| Dubuque | 877 | 1 050 | 1 927 | | | | |
| Fort Dodge | 392 | 233 | 625 | | | | |

**Table 2. Daily travel as a percentage of monthly travel for express and nonexpress bus service.**

| Month | Type of Service | Sunday | Monday | Tuesday | Wednesday | Thursday | Friday | Saturday |
|---|---|---|---|---|---|---|---|---|
| January | Express | 2.1 | 2.5 | 1.8 | 1.9 | 2.4 | 2.4 | 2.6 |
| | Nonexpress | 13.9 | 10.0 | 9.3 | 11.7 | 13.3 | 15.9 | 10.2 |
| | Total | 16.0 | 12.5 | 11.1 | 13.6 | 15.7 | 18.3 | 12.8 |
| February | Express | 3.3 | 2.3 | 1.9 | 1.9 | 2.4 | 2.9 | 3.0 |
| | Nonexpress | 17.5 | 11.1 | 8.2 | 9.4 | 11.0 | 16.1 | 9.0 |
| | Total | 20.8 | 13.4 | 10.1 | 11.3 | 13.4 | 19.0 | 12.0 |
| March | Express | 2.3 | 3.1 | 2.8 | 2.6 | 1.8 | 2.8 | 3.3 |
| | Nonexpress | 13.8 | 12.0 | 11.5 | 12.2 | 8.9 | 13.8 | 9.3 |
| | Total | 16.1 | 15.1 | 14.3 | 14.8 | 10.7 | 16.6 | 12.6 |
| April | Express | 2.3 | 2.7 | 2.2 | 2.0 | 2.2 | 3.5 | 2.8 |
| | Nonexpress | 11.7 | 11.5 | 8.5 | 9.1 | 14.4 | 17.6 | 9.7 |
| | Total | 14.0 | 14.2 | 10.7 | 11.1 | 16.6 | 21.1 | 12.5 |
| May | Express | 2.7 | 3.1 | 2.1 | 1.8 | 2.8 | 2.2 | 2.9 |
| | Nonexpress | 13.3 | 13.6 | 8.9 | 8.8 | 11.5 | 12.4 | 13.8 |
| | Total | 16.0 | 16.7 | 11.0 | 10.6 | 14.3 | 14.6 | 16.7 |
| June | Express | 2.7 | 2.7 | 3.0 | 3.4 | 3.0 | 3.2 | 3.5 |
| | Nonexpress | 11.2 | 11.1 | 11.7 | 11.3 | 10.1 | 13.0 | 10.1 |
| | Total | 13.9 | 13.8 | 14.7 | 14.7 | 13.1 | 16.2 | 13.6 |
| July | Express | 2.9 | 3.6 | 3.2 | 3.0 | 3.7 | 4.7 | 5.8 |
| | Nonexpress | 10.1 | 9.5 | 8.1 | 7.7 | 11.3 | 13.7 | 12.6 |
| | Total | 13.0 | 13.1 | 11.3 | 10.7 | 15.0 | 18.4 | 18.4 |
| August | Express | 4.6 | 4.5 | 3.9 | 3.1 | 3.5 | 2.9 | 4.1 |
| | Nonexpress | 11.5 | 10.9 | 11.1 | 8.7 | 9.9 | 11.9 | 9.5 |
| | Total | 16.1 | 15.4 | 15.0 | 11.8 | 13.4 | 14.8 | 13.6 |
| September | Express | 3.0 | 3.6 | 2.8 | 2.6 | 2.4 | 3.1 | 3.3 |
| | Nonexpress | 13.5 | 13.6 | 8.7 | 8.8 | 9.2 | 14.6 | 10.2 |
| | Total | 16.5 | 17.2 | 11.5 | 11.4 | 11.6 | 17.7 | 13.5 |
| October | Express | 2.2 | 2.4 | 2.1 | 2.3 | 2.0 | 2.5 | 2.4 |
| | Nonexpress | 12.1 | 10.0 | 9.3 | 11.0 | 12.5 | 19.5 | 9.5 |
| | Total | 14.3 | 12.4 | 11.4 | 13.3 | 14.5 | 22.0 | 11.9 |
| November | Express | 2.9 | 2.0 | 2.3 | 2.1 | 2.3 | 1.9 | 2.9 |
| | Nonexpress | 18.5 | 9.1 | 9.6 | 13.1 | 9.3 | 12.6 | 11.5 |
| | Total | 21.4 | 11.1 | 11.9 | 15.2 | 11.6 | 14.5 | 14.4 |
| December | Express | 2.7 | 3.9 | 3.2 | 2.6 | 1.9 | 2.5 | 3.2 |
| | Nonexpress | 10.1 | 14.2 | 12.1 | 12.3 | 8.1 | 12.5 | 10.8 |
| | Total | 12.8 | 18.1 | 15.3 | 14.9 | 10.0 | 15.0 | 14.0 |
| Average | Express | 2.8 | 3.0 | 2.6 | 2.4 | 2.5 | 2.9 | 3.3 |
| | Nonexpress | 13.1 | 11.4 | 9.8 | 10.3 | 12.5 | 14.5 | 10.5 |
| | Total | 15.9 | 14.4 | 12.4 | 12.7 | 15.0 | 17.4 | 13.8 |

Note: Data are taken from 1 year's bus passenger records (October 1975 through September 1976) obtained from Des Moines Union Bus Station.

intercity travel markets in Iowa in a typical summer month in 1976. These data are given in Table 1. Note that, although specific destinations have been recorded for each ticket sale, data in the table are summarized to indicate only the totals of interstate and intrastate sales.

Certain other data required to aid in understanding the characteristics of bus travel were obtained from the analysis of all arrivals and departures for a major carrier at a large terminal in Des Moines. A total of 166 126 passengers were recorded as boarding or disembarking as part of this survey. One of the parameters determined from this analysis was the seasonal and daily variation in bus travel. For example, bus loadings tend to peak in July, August, and December. During a typical week, passenger volume is highest on Friday. A summary of travel by day of the week and month and a further breakdown between express and nonexpress services are given in Table 2, and monthly travel as a percentage of total annual travel is given below:

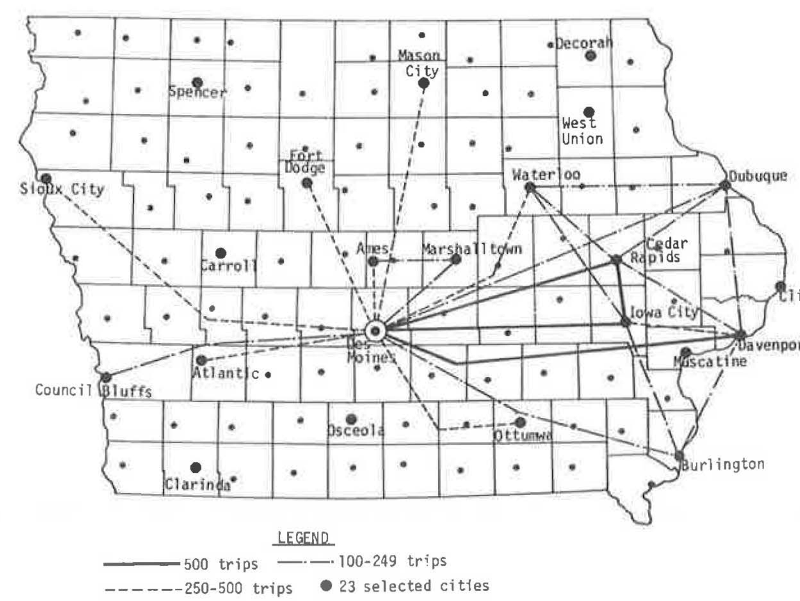
| Month | Travel (as percentage of total annual travel) |
|---|---|
| January | 7.26 |
| February | 6.42 |
| March | 7.72 |
| April | 7.39 |
| May | 7.43 |

**Figure 2. Matrix of monthly trip interchanges among 23 selected Iowa cities in 1976.**

| | Ames | Atlantic | Burlington | Carroll | Cedar Rapids | Clarinda | Clinton | Council Bluffs | Davenport | Decorah | Des Moines | Dubuque | Fort Dodge | Iowa City | Marshalltown | Mason City | Muscatine | Osceola | Ottumwa | Sioux City | Spencer | Waterloo | West Union |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ames | | 5 | 5 | 24 | 159 | 0 | 12 | 14 | 14 | 3 | 395 | 27 | 47 | 59 | 37 | 61 | 0 | 3 | 12 | 18 | 6 | 41 | 0 |
| Atlantic | | | 0 | 0 | 8 | 0 | 0 | 13 | 6 | 0 | 290 | 0 | 5 | 14 | 4 | 5 | 0 | 0 | 0 | 0 | 0 | 3 | 0 |
| Burlington | | | | 0 | 54 | 1 | 7 | 3 | 114 | 0 | 220 | 9 | 10 | 108 | 5 | 6 | 26 | 0 | 36 | 27 | 0 | 15 | 2 |
| Carroll | | | | | 32 | 0 | 0 | 8 | 2 | 1 | 40 | 6 | 4 | 7 | 4 | 3 | 0 | 0 | 4 | 2 | 2 | 6 | 0 |
| Cedar Rapids | | | | | | 0 | 63 | 16 | 147 | 51 | 523 | 187 | 37 | 504 | 90 | 80 | 23 | 3 | 65 | 5 | 9 | 171 | 21 |
| Clarinda | | | | | | | 0 | 35 | 1 | 0 | 12 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Clinton | | | | | | | | 5 | 24 | 1 | 59 | 5 | 3 | 30 | 7 | 1 | 2 | 0 | 2 | 1 | 0 | 5 | 0 |
| Council Bluffs | | | | | | | | | 3 | 2 | 159 | 0 | 1 | 18 | 8 | 1 | 2 | 2 | 5 | 19 | 13 | 11 | 0 |
| Davenport | | | | | | | | | | 2 | 504 | 127 | 8 | 327 | 7 | 23 | 90 | 0 | 35 | 4 | 0 | 53 | 0 |
| Decorah | | | | | | | | | | | 14 | 8 | 2 | 18 | 0 | 19 | 0 | 0 | 1 | 0 | 0 | 20 | 1 |
| Des Moines | | | | | | | | | | | | 119 | 350 | 752 | 155 | 336 | 58 | 95 | 319 | 263 | 26 | 491 | 5 |
| Dubuque | | | | | | | | | | | | | 21 | 77 | 7 | 26 | 3 | 0 | 13 | 50 | 4 | 241 | 2 |
| Fort Dodge | | | | | | | | | | | | | | 42 | 7 | 9 | 0 | 2 | 0 | 66 | 7 | 46 | 3 |
| Iowa City | | | | | | | | | | | | | | | 15 | 61 | 30 | 2 | 64 | 37 | 5 | 127 | 5 |
| Marshalltown | | | | | | | | | | | | | | | | 7 | 3 | 2 | 3 | 3 | 3 | 53 | 0 |
| Mason City | | | | | | | | | | | | | | | | | 8 | 13 | 5 | 17 | 30 | 70 | 0 |
| Muscatine | | | | | | | | | | | | | | | | | | 0 | 5 | 0 | 0 | 7 | 1 |
| Osceola | | | | | | | | | | | | | | | | | | | 0 | 0 | 0 | 1 | 0 |
| Ottumwa | | | | | | | | | | | | | | | | | | | | 6 | 1 | 8 | 0 |
| Sioux City | | | | | | | | | | | | | | | | | | | | | 3 | 25 | 0 |
| Spencer | | | | | | | | | | | | | | | | | | | | | | 4 | 0 |
| Waterloo | | | | | | | | | | | | | | | | | | | | | | | 17 |
| West Union | | | | | | | | | | | | | | | | | | | | | | | |
| Total | | 5 | 5 | 24 | 253 | 1 | 82 | 94 | 311 | 60 | 2216 | [illegible] | 489 | 1956 | 346 | 638 | 245 | 122 | 569 | 518 | 109 | 1398 | 57 |

Total Trips = 9986

**Figure 3. Monthly intercity bus trips in Iowa in 1976.**



| Month | Travel (as percentage of total annual travel) |
|---|---|
| June | 9.15 |
| July | 10.45 |
| August | 10.34 |
| September | 7.43 |
| October | 8.28 |
| November | 8.27 |
| December | 9.87 |

These factors were used to adjust the travel for a typical summer month to the same common base.

Records of bus arrivals and departures also provide a measure of on-time service. The results of this survey are given below:

| Month | On-Time Arrivals (%) | On-Time Departures (%) |
|---|---|---|
| January | 82.3 | 78.3 |
| February | 86.8 | 78.9 |
| March | 85.2 | 75.3 |
| April | 90.2 | 79.7 |
| May | 83.2 | 73.6 |
| June | 80.4 | 69.7 |
| July | 80.0 | 68.2 |
| August | 85.2 | 71.4 |
| September | 88.2 | 76.2 |
| October | 85.4 | 74.4 |
| November | 76.1 | 65.4 |
| December | 73.8 | 59.8 |
| Average | 83.1 | 72.9 |

| Day | On-Time Arrivals (%) | On-Time Departures (%) |
|---|---|---|
| Sunday | 87.7 | 76.9 |
| Monday | 86.4 | 78.1 |
| Tuesday | 83.8 | 72.6 |
| Wednesday | 82.5 | 73.5 |
| Thursday | 80.9 | 71.1 |
| Friday | 77.9 | 67.0 |
| Saturday | 80.0 | 67.9 |
| Average | 82.7 | 72.5 |

An on-time arrival or departure was defined for this purpose as one that operated not more than 10 min later than its scheduled time.

These surveys illustrate the point that carriers or station agents do maintain certain records that are helpful to transportation planners involved with intercity bus travel. However, such data are not widely disseminated and are not commonly available to planners. The surveys described above provided a foundation on which to build a planning effort directed to the improvement of intercity bus travel in Iowa.

## ANALYSES

An initial application of the data on intercity bus passenger destinations for the 23 study cities was to quantify the monthly trip interchanges among city pairs. The trip interchange matrix is shown in Figure 2. The next step was to plot these passenger demands as desire lines on a map of the state. Monthly volumes in excess of 100 trips between study cities are shown in Figure 3, which graphically identifies the principal corridors for intrastate bus travel demand. A distinct focus on Des

Figure 4. Monthly bus-passenger route volumes for intrastate travel between 23 Iowa cities.

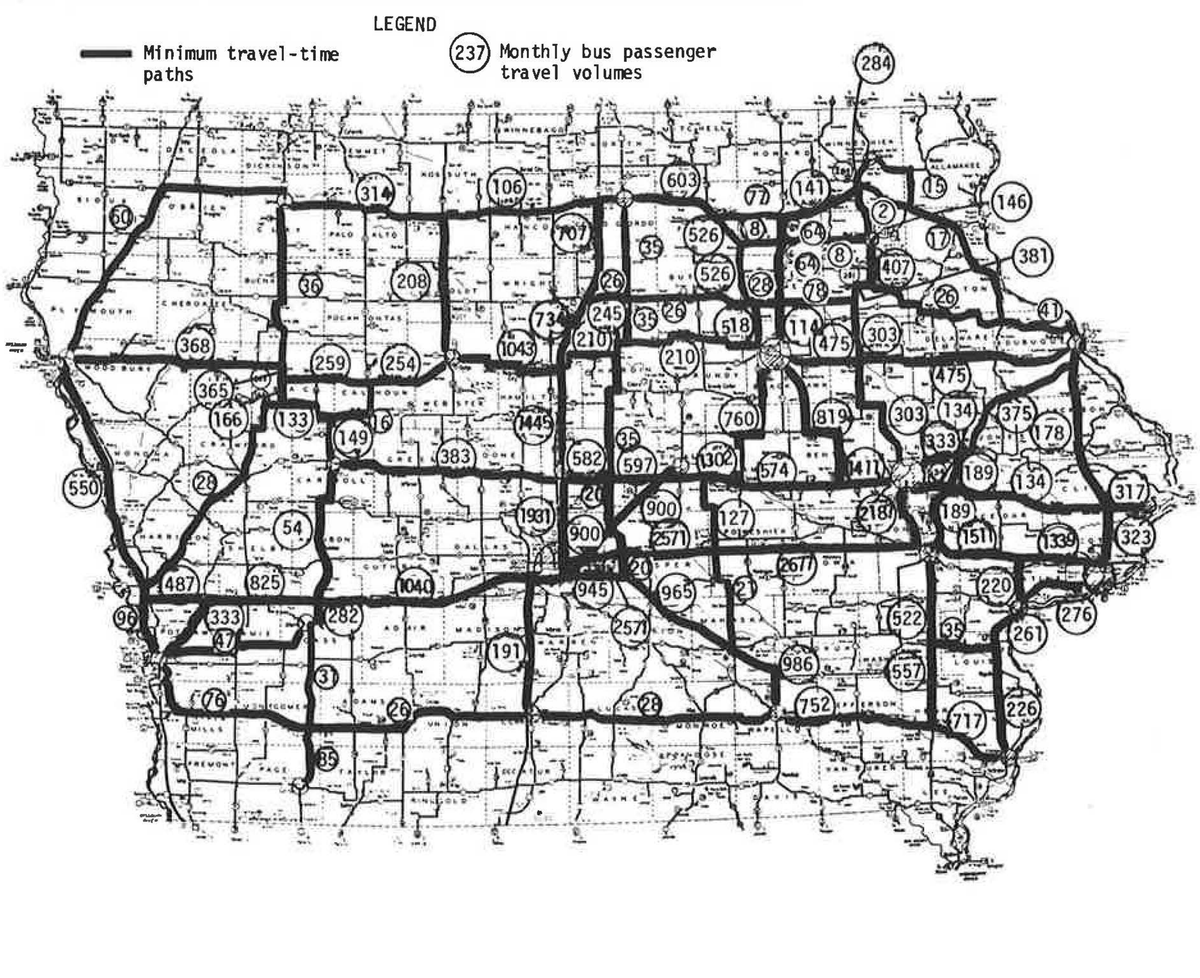
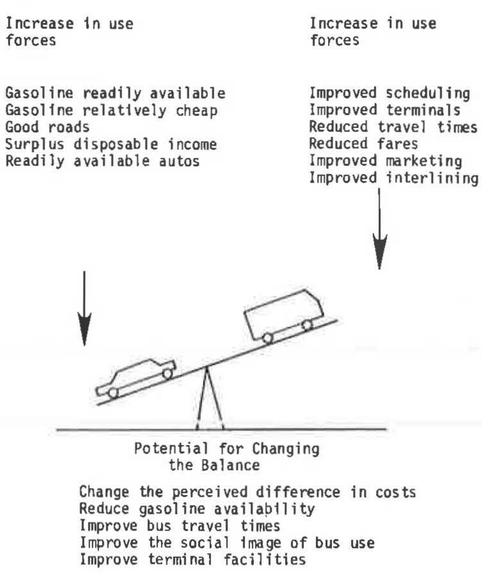


Table 3. Analysis of trip-length frequency.

| City | 50 Percent of Tickets Sold for Travel Distance <x km (km) | Percentage of Tickets Sold for Destination Distance Greater Than 241 km | 322 km | 644 km |
|---|---|---|---|---|
| Ames | 177 | 44 | 30 | 7 |
| Atlantic | 177 | 40 | 29 | 18 |
| Burlington | 177 | 44 | 24 | 10 |
| Carroll | 193 | 41 | 23 | 14 |
| Cedar Rapids | 177 | 29 | 23 | 6 |
| Clarinda | 209 | 45 | 44 | 18 |
| Clinton | 225 | 41 | 29 | 13 |
| Council Bluffs | 193 | 35 | 28 | 9 |
| Davenport | 241 | 46 | 21 | 11 |
| Decorah | 193 | 45 | 16 | 6 |
| Des Moines | 225 | 47 | 34 | 12 |
| Dubuque | 225 | 47 | 19 | 8 |
| Fort Dodge | 209 | 48 | 36 | 14 |
| Iowa City | 209 | 48 | 42 | 9 |
| Marshalltown | 145 | 38 | 27 | 10 |
| Mason City | 225 | 42 | 36 | 26 |
| Muscatine | 290 | 52 | 39 | 12 |
| Osceola | 257 | 51 | 38 | 11 |
| Ottumwa | 145 | 28 | 23 | 9 |
| Sioux City | 241 | 50 | 38 | 18 |
| Spencer | 306 | 67 | 23 | 10 |
| Waterloo | 177 | 42 | 33 | 9 |
| West Union | 161 | 25 | 9 | 5 |

Note: 1 km = 0.62 mile.

Figure 5. Potential for increasing use of motor buses.




Moines is also evident in Figure 3.

In a subsequent step, travel trees were developed for travel from each study city to all other 22 cities based on a minimum travel-time path along primary highways. By combining 23 minimum travel-time trees, composite volumes were assigned to specific highway links. The resulting composite tree and total volumes for each link are shown in Figure 4.

Table 3 gives certain summary information from trip-length frequency analyses of the monthly ticket sales for each study city. Most trips were shorter than 644 km (400 miles). Median trip lengths, with few exceptions, were 806 km (500 miles) or less.

The data generated in this study were also used to develop a forecasting model for intercity bus use. The monthly ticket sales in 1976, as given in Table 1 and adjusted to reflect monthly variations given in Table 2, were used as the dependent variable. Independent variables were representative of various social, economic, and demographic characteristics of the communities or of the level of bus service and quality of terminal facilities available to travelers. Several forms of regression equations were tested to evaluate the relation between the independent variables and the dependent variable.

## ROLE OF GOVERNMENT

Various options are available to a government that may elect to foster and support travel by intercity bus. An objective of government concern in such a case would be to increase the motor-bus share of an intercity passenger travel market. However, several factors tend to inhibit extensive use of buses today. Abundant automotive fuel at relatively low costs and an extensive highway network significantly enhance the relative attractiveness of the automobile as a travel mode. The current imbalance and the forces that influence the use of buses and automobiles are shown graphically in Figure 5.

A state government can exert relatively little influence on many of the factors that affect modal use. The amount of disposable income available for transportation and the cost and availability of fuel are factors that are largely beyond state control. However, a state can influence a choice to travel by bus among elective riders by inducing improvements in the level of service, through changes in routes and schedules, or by upgrading terminal facilities. This suggests a financial inducement to carriers by the state.

Even in the absence of a policy of subsidizing intercity bus carriers to reduce the current imbalance in factors of modal choice, it is possible that such a policy may arise as a result of limitations in the supply of motor fuel. In this case, an objective would be to reduce the adverse impact on the life-style of a highly mobile society that was developed on the basis of affluence and an abundance of energy.

In either case, an enlightened decision as to the appropriate role of government must be based on facts concerning bus use and travel characteristics. These facts are not currently available to transportation planners in suitable form. The types of data that can be obtained and used to analyze changes in policy were demonstrated by this study. Without data that indicate bus use in terms of volumes, trip lengths, and various measures of service level, there is little basis on which policy makers can justify significant expenditures to assist intercity bus carriers.

## ACKNOWLEDGMENTS

The research summarized here was sponsored by the Planning and Research Division of the Iowa Department of Transportation. It was carried out and partially supported by the Engineering Research Institute of Iowa State University. The points of view expressed are ours, however, and do not necesarily reflect the points of view of the Iowa Department of Transportation or the Engineering Research Institute.

## REFERENCES

1. Transportation Facts and Trends. Transportation Association of America, Washington, DC, 12th Ed., 1976.
2. Intercity Passenger Transportation Data. Boeing Commerical Airplane Co., Seattle, 1975.
3. Intercity Passenger Carrier Improvement Study. Engineering Research Institute, Iowa State Univ., Ames, 1977.

*Publication of this paper sponsored by Committee on Transportation Information Systems and Data Requirements.*

# Marginal Weighting of Transportation Survey Data

L. Taylor Ollmann,* John Hancock Mutual Life Insurance Company, Boston
Stephen M. Howe,* North Central Texas Council of Governments, Dallas
Kenneth W. Kloeber,* Calocerinos and Spina, Buffalo
Gerald S. Cohen, New York State Department of Transportation

**Two procedures are discussed for calculating cell weights to be applied to sample data to ensure that the expanded sample jointly represents the population on several background distributions. Deming's procedures minimize the squared deviation of cell weights from unity, whereas Johnson's procedure iteratively produces weights that match the marginal distributions, are all positive, and have low variance from unity. The theory and calculation procedure of these two approaches are developed and are illustrated by a simple example. Advantages and disadvantages of the procedures are described, and applications to various transportation problems are highlighted. The paper concludes that, although the Johnson technique is less well grounded in theory, it produces reasonable weights, prevents negative estimates, and converges quickly. Its use is recommended.**

A common problem in transportation planning is the selection and processing of a sample of observations so that they represent a larger population. Such activities are central in most analyses of transportation policies because planners are often concerned with the impacts of policies on the general population and its various subgroups. Methods of simple random sampling ensure representations within statistical limits, but may not be cost-effective or feasible. Other sampling procedures (stratification, clustering, multistage, or quota) can also be used, but each has its biases and limitations. These problems are further compounded by "micro" samples used in disaggregate modeling, and numerous sophisticated aggregation procedures have been evolved to ensure model application, even from highly nonrepresentative samples.

Many conventional analyses, however, still require that representativeness be demonstrable, if for no other reason than to increase the analyst's confidence in the validity of the data. In most surveys, representativeness is taken to mean that the distribution of sample data agrees (within statistical limits) with distribution of known demographic or geographic data for the population. The underlying assumption of such a test is, of course, that, if the demographics are representative, other data (such as trips and opinions) will also be representative of the population.

In many transportation surveys, however—particularly small-sample, telephone, and mail-out studies—it is difficult to achieve representations without very great cost. Hence, too many of one group may be sampled and not enough of another. To ensure that the data conform to census (or other population) distributions, one must generally weight the data by multiplying by appropriate weighting factors. For instance, 5 percent of a sample may consist of men aged 15 to 24 whereas census figures might show that 8 percent of the population is of that type. One would then weight the sample data for this group by a factor of 1.6.

## MARGINALS AND CELLS

Suppose that a survey sample is categorized according to sex, age group, household size, and number of automobiles owned per household. The data may thus be tabulated into a four-dimensional contingency table, one cell of which might consist of males aged 15 to 24 from two-person households that own one automobile. A marginal group may be defined as an aggregate of all cells that share a single characteristic—for example, the marginal group of persons aged 15 to 24, the marginal group of persons from two-automobile households, or the marginal group of females. [In tabulation, the marginal groups would appear in the table margins as row or column sums (Table 1).]

It would be very useful if population data were available for each cell because then the sample data for each cell could be weighted to match the proportion of that cell in the entire population. Unfortunately, population (e.g., census) data are rarely cross tabulated for more than two categories simultaneously. Thus, the best we can do is to compute expansion factors for each of the cells so that the resulting weighted samples conform to census proportions in each marginal group. Marginal weighting is any procedure for artificially computing those cell expansion factors. This paper describes such a method recently programmed by the New York State Department of Transportation (NYSDOT) (1) and used extensively in small-sample transportation studies.

## MARGINAL WEIGHTING

Several different procedures of marginal weighting are in use, and they generally produce different weights for the same data. An illustration is given below for two categories: age (in three marginal groupings with census proportions of 25, 50, and 25 percent) and sex (in two marginal groupings of 50 percent each). Table 1 gives sample proportions for each of the six cells and the five marginal-group target proportions mentioned above.

Table 2 gives four sets of weights that, when multiplied by the data in Table 1, yield proper marginals. For example, for the age 55 and over marginal group in Table 2, $0 \times 15 + 1.25 \times 20 = 25$ percent (procedure 1), $1 \times 15 + 0.5 \times 20 = 25$ percent (procedure 2), $0.82 \times 15 + 0.635 \times 20 = 25$ percent (procedure 3), and $0.79 \times 15 + 0.6575 \times 20 = 25$ percent (procedure 4). Since numerous sets of weights can be conducted, we must ask what criteria should be used to compare one possible set of weights with another and whether there is a best set of weights.

The following criteria seem reasonable:

1. The marginal groupings of weighted cells should match closely (if not exactly) the corresponding population (census) proportions.
2. The individual weights should tend to be near one. That is, the individual cells (particularly those that contain a large proportion of the sample) should be changed as little as possible.
3. Weights should all be positive. Factors of zero fail to use all available information, and negative factors make even less sense.

Procedure 1 in Table 2 shows a particularly bad set of weights that satisfy the first criterion but violate the

Table 1. Sample data.

| Age Group | Male (%) | Female (%) | Total (%) | Census Marginal (%) |
|---|---|---|---|---|
| 16 to 24 | 10 | 10 | 20 | 25 |
| 25 to 54 | 20 | 25 | 45 | 50 |
| 55 and over | 15 | 20 | 35 | 25 |
| Total | 45 | 55 | 100 | |
| Census marginal | 50 | 50 | | 100 |

Table 2. Examples of marginal weights.

| | Marginal Weighting | | | | | |
|---|---|---|---|---|---|---|
| | Age 16 to 24 | | Age 25 to 54 | | Age 55 and Over | |
| Procedure | Male | Female | Male | Female | Male | Female |
| 1 | 5 | -2.5 | 0 | 2 | 0 | 1.25 |
| 2 | 1 | 1.5 | 1.25 | 1 | 1 | 0.5 |
| 3 (Deming) | 1.342 | 1.1575 | 1.214 | 1.029 | 0.820 | 0.635 |
| 4 (Johnson) | 1.3644 | 1.1356 | 1.2253 | 1.0198 | 0.7900 | 0.6575 |

Table 3. Example cell weights for skewed samples.

| Item | | | Census Marginals (%) | Deming Weights | | Johnson Weights | |
|---|---|---|---|---|---|---|---|
| Cell percentages | 20 | 30 | 30 | 0 | 1 | 0.2863 | 0.8091 |
| | 30 | 20 | 70 | 1 | 2 | 0.8091 | 2.2863 |
| Census marginals, % | 30 | 70 | | | | | |
| Cell percentages | 20 | 30 | 20 | -0.5 | 1 | 0.1125 | 0.5917 |
| | 30 | 20 | 80 | 1 | 2.5 | 0.5917 | 3.1125 |
| Census marginals, % | 20 | 80 | | | | | |

next two. The weights in effect throw away all sample information about men over age 24 and reverse any sample input from young women! Procedure 2 gives a more reasonable set of weights, satisfying the first criterion and also doing a reasonable job on criteria 2 and 3. Generally, no set of weights is optimal with respect to all these criteria at once. Marginal weighting methods are known for optimizing some of the criteria but generally at the expense of others.

## Deming Method

One procedure is designed to optimize the second criterion subject to the constraint of satisfying the first exactly (2). Deming measures the second criterion by the formula $\Sigma_{i_1} \Sigma_{i_2} \Sigma_{i_n} a_{i_1 i_2 \ldots i_n} (W_{i_1 \ldots i_n} - 1)^2$ where, in an n-dimensional table of sample data, $a_{i_1 i_2 \ldots i_n}$ is the proportion of data in cell $i_1, i_2, \ldots, i_n$ and $W_{i_1 \ldots i_n}$ is the computed weight applied to that cell. This term can then be minimized subject to the constraint of exactly matching the census target proportions. If there are m marginal groups in all, the computation of cell weights applies the method of Lagrange multipliers to obtain m equations in m unknowns, which are then solved to provide one coefficient for each marginal group. The cell weights are then computed by adding to one the coefficients of all marginal groupings to which the cell belongs. Hence, in the example from Table 1, five simultaneous linear equations are solved to yield two coefficients for sex ($c_1 = 0.1849$ and $c_2 = 0$) and three coefficients for age ($d_1 = 0.1576$, $d_2 = 0.0289$, and $d_3 = -0.3650$). The cell weights (given in procedure 4 in Table 2) are then computed by $W_{ij} = 1 + c_i + d_j$.

The advantages of the Deming approach are that the weights can be directly calculated, the number of computations remains manageable, and the results are "optimal." The major disadvantages are that some weights can be very far from one or can be even zero or negative, especially when the sample marginal proportions are sharply skewed. Deming's weights do seem reasonable when the sample is not so severely skewed (Table 3). In addition, the computation, which involves linear algebra, calculus, and the method of Lagrange multipliers, is harder to understand.

## Johnson Method

An alternative weighting procedure proposed by Johnson (3) assigns a coefficient to each marginal grouping; the cell weights are then obtained by multiplying together the coefficients of all marginal groupings to which the cell belongs. Johnson describes a simple iterative procedure for producing such cell weights so as to match exactly the census marginal proportions (criterion 1).

Assume that sample proportions for n marginal distributions are given ($r_k$ marginal groups in the kth category). There are $(r_1)(r_2)\ldots(r_n)$ cell proportions $a_{i_1 \ldots i_n}$ and $r_1 + r_2 + \ldots + r_n$ marginal target proportions $M_{k,i_k}$ ($k = 1, n$, $i_k = 1, r_k$). First, sum the cell sample proportions for each marginal group in the first marginal:

$$S_j^1 = \Sigma_{i_2} \ldots \Sigma_{i_n} a_{ji_2 \ldots i_n} \qquad j = 1, r_1 \tag{1}$$

and then rescale the cells in each of these marginal groups to match target marginal sums in the first marginal:

$$a_{ji_2 \ldots i_n}^1 = a_{ji_2 \ldots i_n} \cdot M_{1,j}/S_j^1 \qquad j = 1, r_1 \tag{2}$$

The rescaled cell values $a^1$ then match the first marginal target sums exactly but probably do not match targets in other marginals. One then rescales a second time to match the second marginal targets:

$$S_j^2 = \Sigma_{i_1} \Sigma_{i_3} \ldots \Sigma_{i_n} a_{i_1 j, i_3 \ldots i_n}^1 \qquad j = 1, r_2 \tag{3}$$

$$a_{i_1 j, i_3 \ldots i_n}^2 = a_{i_1 j, i_3 \ldots i_n}^1 \cdot M_{2,j}/S_j^2 \tag{4}$$

and again for each marginal until the last:

$$S_j^n = \Sigma_{i_1} \ldots \Sigma_{i_{n-1}} a_{i_1 \ldots i_{n-1} j}^{n-1} \qquad j = 1, r_n$$

$$a_{i_1 \ldots i_{n-1} j}^n = a_{i_1 \ldots i_{n-1}}^{n-1} \cdot M_{n,j}/S_j^n \tag{5}$$

This completes the first iteration.

For the data in Table 1, these calculations are as follows:

$S_1^1 = 0.1 + 0.2 + 0.15 = 0.45$

$S_2^1 = 0.1 + 0.25 + 0.2 = 0.55$

$a_{1,1}^1 = 0.1(0.5/0.45) = 0.1111$

$a_{1,2}^1 = 0.2(0.5/0.45) = 0.2222$

$a_{1,3}^1 = 0.15(0.5/0.45) = 0.1667$

$a_{2,1}^1 = 0.1(0.5/0.55) = 0.0909$

$a_{2,2}^1 = 0.25(0.5/0.55) = 0.2273$

$a_{2,3}^1 = 0.2(0.5/0.55) = 0.1818$

$S_1^2 = 0.1111 + 0.0909 = 0.2020$

$S_2^2 = 0.2222 + 0.2273 = 0.4495$

$S_3^2 = 0.1667 + 0.1818 = 0.3485$

$a_{1,1}^2 = 0.1111(0.25/0.2020) = 0.1375$

$a_{1,2}^2 = 0.2222(0.5/0.4495) = 0.2472$

$a_{1,3}^2 = 0.1667(0.25/0.3485) = 0.1196$

$a_{2,1}^2 = 0.0909(0.25/0.2020) = 0.1125$

$a_{2,2}^2 = 0.2273(0.5/0.4495) = 0.2528$

$a_{2,3}^2 = 0.1818(0.25/0.3485) = 0.1304$

At the end of the first iteration, the rescaled cell values $a_{i_1 \ldots i_n}^n$ will not match exactly any target marginal sums except those of the last marginal. However, they will tend to match other marginal targets more closely than did the original cell values $a_{i_1 \ldots i_n}$. A second iteration can then be performed that repeats the above process for each category in turn and finally obtains cell values $a^{2n}_{i_1 \ldots i_n}$, which will tend to match each marginal target more closely than did the values $a^{n}_{i_1 \ldots i_n}$. One continues to iterate until all marginals are matched simultaneously to a desired degree of accuracy. In our example, three full iterations suffice to obtain values that match all five marginal targets with a tolerance of $10^{-7}$. Weights for each of the cells are then obtained by dividing the final scaled cell values by the original values, as follows:

$$W_{i_1 \ldots i_n} = a^N_{i_1 \ldots i_n}/a_{i_1 \ldots i_n} \quad \text{if } a_{i_1 \ldots i_n} \neq 0 \tag{6}$$

For our example, those weights appear in procedure 4 in Table 2.

The advantages of the Johnson technique are many. The iterative procedure is easy to understand and to program: One simply multiplies each marginal in one category (age, for example) by a factor that makes the weighted cell total match the census target for that marginal. One then repeats this process for each of the other categories and then again for all categories and so on until all marginals match their census proportions simultaneously. Further, the procedure converges rather quickly: Weights have been obtained to an accuracy of six decimal places in 5 to 15 iterations on small-sample data that have three categories of three, four, and six marginal groupings (72 cells).

Criteria 2 and 3 are also well satisfied: All cell weights are positive and tend to be near 1.0. Values of Deming's criterion tend to be low although not the minimum obtainable by Deming's method. The weights obtained look reasonable and seem to be uniquely determined by the requirement that a fixed weight be assigned to each marginal grouping and each cell weight be simply the product of the weights of those marginal groupings to which the cell belongs. Reasonable weights are obtained even in matrixes where 50 to 60 percent of the cells contain no samples as long as each marginal grouping is reasonably well sampled. The method is robust: The inputted sample can initially be highly skewed from desired marginals as in the case of daytime telephone samples.

## DISCUSSION OF RESULTS

One disadvantage of both the Deming and Johnson procedures is that the act of subdividing a marginal group into several smaller groups changes the weights assigned not only to cells of that group but also to cells of other groups that were left unchanged. For example, if the two lower age groupings in Table 1 are combined, the Deming and Johnson weights are as given below:

| Age Group | Male (%) | Female (%) | Total (%) | Census (%) |
|---|---|---|---|---|
| 16 to 54 | 30 | 35 | 65 | 75 |
| 55 and over | 15 | 20 | 35 | 25 |
| Total | 45 | 55 | 100 | |
| Census | 50 | 50 | | 100 |

| Deming Weights | | Johnson Weights | |
|---|---|---|---|
| 1.2556 | 1.0667 | 1.2709 | 1.0535 |
| 0.8222 | 0.6333 | 0.7916 | 0.6563 |

Notice that the weights assigned to the cells for the group aged 55 and over are not the same as the corresponding ones in procedures 3 and 4 in Table 2 even though the cell descriptions are the same. The differences are only a few percent, but the examples show that the numbers produced by either method have no natural validity and artificially depend on the way one chooses to define marginal groupings.

We attempted to overcome this difficulty by applying the Johnson technique to each cell individually. To weight one cell, we aggregated all marginals in each category except for the marginals that contained the particular cell in question. The result was a $2 \times 2 \times \ldots \times 2$ table of sample data (categorized only by the marginals that defined the cell and the complementary marginal groups), one cell of which was identical to the original cell. The Johnson technique was then applied to obtain a weight for that cell. Cell weights for each cell in the original data table were obtained one at a time in this way and are given below:

| Age Group | Male | Female |
|---|---|---|
| 16 to 24 | 1.3675 | 1.1325 |
| 25 to 54 | 1.2372 | 1.0102 |
| 55 and over | 0.7916 | 0.6563 |

The resulting weights depended only on the description of the particular cell, the sample proportions of that cell's own marginal groups, and the census targets for those groups. Hence, changing the definitions of some marginal groups would not change the weight of an unaltered cell. This approach and the pure Johnson procedure were then applied to small survey samples from four small upstate New York cities: Plattsburgh, Watertown, Glens Falls, and Elmira (4). The sample sizes ranged from 209 to 245 and were tabulated in three categories divided into three, four, and five marginal groups. Between 22 and 32 of the 72 cells in each survey were empty, and many of the remainder contained only one or two samples. The weights so produced tended to be closer to 1.0 than those produced by the pure Johnson algorithm; Deming criterion values were about 20 to 50 percent less. Unfortunately, this individual treatment of cells was unable to reproduce the target marginal totals (criterion 1). The marginal totals after weighting differed from the desired values by 1 to 10 percent and in a few cases by as much as 30 percent! This is unacceptably high. It may well be that such losses of marginal accuracy are necessary in any weighting procedure whose cell weight function depends only on the cell and directly related marginals and does not depend on the classification of all other data in the sample.

Returning to the two original weighting procedures that do yield correct marginals, we favor the Johnson approach for its previously mentioned advantages. Johnson points out that the procedure can fail to converge if a large proportion of the cells are empty. Our results on the small-city surveys showed that this proportion can be as much as 60 percent and the algorithm will still converge in less than 15 iterations.

The chief distinction of the Deming approach is the criterion it optimizes. We conjecture that the Johnson method also minimizes a function of the following form:

$$\Sigma_{i_1} \ldots \Sigma_{i_n} a_{i_1 \ldots i_n} \cdot W_{i_1 \ldots i_n} \cdot \log (W_{i_1 \ldots i_n}) \tag{7}$$

This is not unreasonable as a criterion. But we can give no "elegant" reason why this is an ideal function to minimize.

Marginal weighting techniques have many possible applications. The New York State Department of Transportation has found them useful in many of the following contexts:

| Analytical Problems | Studies |
|---|---|
| Nonrandom samples | Daytime telephone surveys |
| Surveys with missing data | Small area surveys |
| Limited budget or small samples | Transit on-board studies |
| Need for quick turnaround | Disaggregate data for mode choice |
| Limited population data | Community attitude studies |
| | Travel update studies |
| | Corridor studies |

NYSDOT has applied these techniques in many contexts, including a 1000-household survey, a 300-household survey on community attitudes toward transit, a survey of 1200 transit riders in Albany, and 200- to 300-household surveys of community attitudes in six towns. In all cases, the method has proved to be useful and time saving and to provide reasonable results. The procedure develops reasonable weights with manageable effort for real-world data even if those data are highly skewed with many empty cells. Examples may exist that cause the method to fail, but they must be concocted for the purpose, would probably not occur in practice, and would probably cause other methods of marginal weighting to fail also. We feel that both methods are valuable additions to the repertoire of survey analysis techniques used by transportation analysts, and we particularly recommend the Johnson procedure because of its ease of understanding and programming.

ACKNOWLEDGMENTS

We wish to recognize the important contribution of Richard Johnson of Market Facts in constructing and describing his procedure. The assistance of Wilma C. Marhafer in preparing this manuscript is gratefully appreciated.

REFERENCES

1. K. W. Kloeber and S. M. Howe. Marginal Weighting Procedures for Expanding Small Sample Surveys. Planning Research Unit, New York State Department of Transportation, PRR 97, Nov. 1975.
2. W. E. Deming. The Statistical Adjustment of Data. Dover, New York, 1964.
3. R. M. Johnson. Marginal Weighting. Market Facts, Chicago, Oct. 1972.
4. W. C. Holthoff, S. M. Howe, and E. P. Donnelly. Documentation of Expansion and Cross Tab Set Up for Small Urban Area Study, 1974-75. Planning Research Unit, New York State Department of Transportation, memorandum, Dec. 1975.

*Publication of this paper sponsored by Committee on Transportation Information Systems and Data Requirements.*

**Messrs. Ollmann, Howe, and Kloeber were with the New York State Department of Transportation when this research was performed.*

# Demonstration of a Simplified Traffic Model for Small Urban Areas

Donald Rhodes, New Hampshire Department of Public Works and Highways
Thomas Hillegass, Urban Mass Transportation Administration, U.S. Department of Transportation

**A simple and straightforward method for forecasting highway traffic that was applied in Plaistow, New Hampshire, is described. The method requires an external origin-destination survey, demographic forecasts, and good ground counts. No home interview survey or demand model calibration is required. The approach combines a Fratar expansion of external trip tables with a factoring of internal travel on a link-by-link basis to produce a forecast of total automobile travel by link. A unique feature is that the link-by-link factoring of internal travel is directly related to the growth of population and employment in internal zones. This is done by means of an artificial trip-generation and gravity model that requires no calibration. The method produced logical results for Plaistow and appears to have promise for other applications.**

There is widespread recognition today that traditional transportation planning techniques are somewhat more complex and costly than warranted by the types of problems to be addressed in smaller urban areas. This recognition and the desire to direct more of limited planning resources toward the solution of short-range problems have led to a search for simplified approaches to long-range travel forecasting that are tailored to the needs of smaller urban areas.

Planners for smaller urban areas are urged to perform an analysis of problems, growth, and related factors and to design study techniques that best suit these problems. Yet, lacking proven step-by-step alternatives to current forecasting methods, planners for these areas are justifiably reluctant to risk new approaches. The profession badly needs demonstrations of new methods, or new ways of using old methods, that are relatively straightforward and are suited to typical problems of small urban areas. Both the method and the outcome of these demonstrations as well as the types of situations to which they are adaptable need to be well documented. This paper attempts such a documentation of an approach taken by the New Hampshire Department of Public Works and Highways to forecasting automobile traffic for the town of Plaistow, New Hampshire.

## PLAISTOW

Plaistow is a 27-km$^2$ (10.5-mile$^2$) community with an estimated 1975 population of 5330 persons. Located on the Massachusetts border north of Boston and Haverhill, Plaistow is defined by the U.S. Bureau of the Census as part of the Lawrence-Haverhill (Massachusetts) urbanized area.

Plaistow is part of rapidly growing Rockingham County, New Hampshire, the population of which increased by 40 percent between 1960 and 1970. Plaistow itself grew by 60 percent during that decade. The county population is projected to increase by over 70 percent between 1975 and 1995. During the same period, Plaistow's growth is expected to be more moderate because of restraints on land use. The high rate of development in surrounding communities will contribute to the increase of commercial activities in Plaistow.

Approximately 50 percent of the land area in Plaistow is developed; of that, about 66 percent is residential, 21 percent is commercial, and the remainder is industrial and governmental. Nearly 90 percent of the undeveloped land is currently zoned for residential use.

As part of an urbanized area, Plaistow is subject to the requirements of the Federal-Aid Highway Act of 1962 for continuous, comprehensive, and cooperative transportation planning within urbanized areas. The Plaistow Highway Transportation Study is an element of the comprehensive transportation plan for the New Hampshire portion of the Lawrence-Haverhill urbanized area.

## METHODOLOGY

The traffic forecasting approach in Plaistow may briefly be described in the following steps:

1. Data preparation—(a) Functionally classify the road system and code the network; (b) conduct an external-cordon roadside origin-destination (O-D) survey and build trip tables (no O-D home interview is needed); (c) expand the highway counting program to cover all main facilities; and (d) assemble zonal population and employment estimates (retail and other) for the base year and make zonal forecasts of these to the future year.
2. External travel—(a) Assign the external-cordon trip tables to the highway network; (b) subtract the resulting assigned volumes from counted volumes on links where counts are available (the difference represents the internal volume on each link); (c) factor the external trip tables to the future year by using the Fratar method; and (d) assign the future external trip table to the highway network.
3. Internal travel—(a) By using base-year zonal population as a proxy for "productions" and employment similarly for "attractions," apply a pseudogravity model and, for F-factors, either use all equal to 1 or borrow factors from other sources (the gravity model output is in the form of a trip table but does not contain trips because the productions and attractions were not calibrated to equal trips); (b) by using future-year zonal population and employment, again solve the pseudogravity model by using the same method as in step 1 above; (c) assign both base-year and future-year internal tables to the highway network; and (d) compute a growth factor for internal travel as the link-by-link ratio of future load divided by base load.
4. Forecasts of link traffic—Factor base-year internal volume on the link (the difference between assigned external trip table and full link count as in 2b above) by the internal growth factor (3d above) and add to this the future external trips (2d above).

The approach described is basically a link count-factoring approach to travel forecasting that includes the following advantages:

1. External travel is forecasted separately and is actually assigned to the network. Thus, when diversion of such long trips to other routes (e.g., a new bypass) is a future possibility, this is accounted for in the process.
2. Internal travel is a link count-factoring process, but the factor itself is based on land-use changes (population and employment) and thus is more logical and policy sensitive than many other count-factoring approaches.

The approach requires no model calibration, but it does require a coded highway network and the ability to apply a Fratar expansion and a simple gravity model. Its disadvantages include the following:

1. Estimating internal travel on entirely new links is not straightforward and was not attempted in Plaistow, but it probably could be done.
2. The procedure of assigning external travel and subtracting this on a link-by-link basis from total link counts (to impute internal link counts) is very sensitive to errors in the assignment process. In Plaistow, the all-or-nothing assignment with some manual adjustment worked satisfactorily.

Generally, the method has the appealing feature of being nonsynthetic, i.e., tied directly into an external O-D trip table so that link counts and radical errors are unlikely and would probably be quite apparent if they did occur. The approach is somewhat more rigorous than that usually taken for an area the size of Plaistow; however, the small size of the prototype area helped make this demonstration feasible.

The outstanding weakness of the approach is in its inability to handle diversion of internal travel to entirely new links in the future year since there are no existing counts on such links to factor up. A method by which the new link is added to the base-year network first, a count is then inputted for the link based on reassignment of the base-year internal (pseudo) trip table, and then future factoring is done seems plausible but was not attempted. Overall, one might conclude that the method is best suited where the following conditions exist:

1. A small urban area of rather slow or moderate growth;
2. An external O-D survey, land-use data, and computer processes are available or feasible given staff and resources; and
3. Future alternative highways do not include major new facilities on entirely new alignments except when these are to serve predominately external travel (e.g., a bypass).

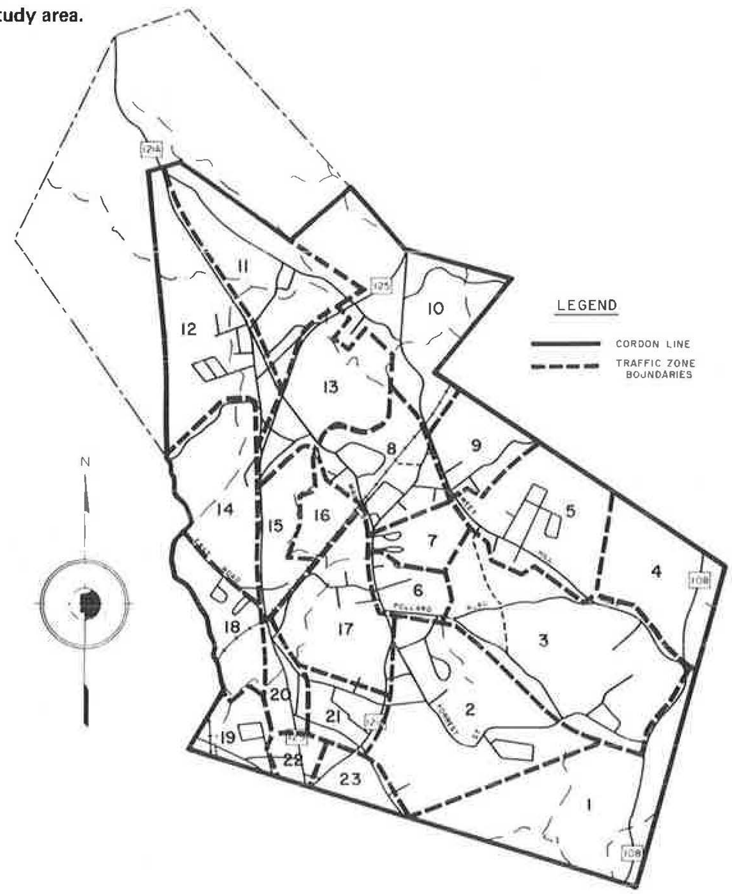
In regard to the third condition, it would be most difficult, for example, to estimate future internal travel on a new downtown river bridge if such a bridge were 3.2 km (2 miles) upriver from the existing bridge. The factoring of internal counts on the existing bridge would not produce relevant data for such a new bridge if it were located at such a distance from the existing bridge that it would attract trips between new internal O-D zones.

## APPLICATION OF METHOD TO PLAISTOW

### Data Preparation

The area included in the Plaistow Highway Transportation Study (Figure 1) includes the entire town of Plaistow except the lightly populated northwest portion along NH-

**Figure 1. Plaistow, New Hampshire, study area.**




121A. The study area was divided into 23 traffic zones. To the extent possible, the zones were partitioned so that the land use in each would be homogeneous. The land use in Plaistow, except along the NH-125 corridor, is predominately residential with a scattering of small businesses. It would have been impractical to establish a separate zone for each business; therefore, businesses were included in the residential zones. The transportation network was composed of all roads in Plaistow that are part of the federal-aid system. These are the streets and highways comprising the arterial and collector systems.

An external O-D study was conducted in Plaistow during August 1975 to survey travel patterns in the town. Twelve interview stations were established and were located wherever a collector or an arterial street or highway crossed the cordon line. From 6:00 a.m. until 8:00 p.m., as many vehicles as possible from each direction were stopped, and the drivers were interviewed. The sample rate varied from 32 to 70 percent depending on the traffic volume, the number of personnel doing the interviewing at a particular station, and the frequency of illegible and illogical interviews.

All interviews were coded and keyed for computer processing. All records were edited, and illogical trips were eliminated. The trip origins and destinations were coded in six-digit notation that conformed to the external zoning system of the New Hampshire Department of Public Works and Highways.

Tables developed for each station, by direction, summarized the trips between each O-D pair. These O-D tables were factored upward so that the total number of trips would equal the average summer weekday traffic (ASWT) at the station. The O-D tables were reduced to zone-to-zone tables by replacing the origin or destination or both (if outside the study area) by the station through which the trip entered or exited the study area.

During August 1975 and concurrent with the collection of driver interviews, short-term automatic traffic recorder counts were taken at selected locations in Plaistow. The traffic volumes were expressed in terms of ASWT. The ASWT was used rather than annual average daily traffic (AADT) because it is more representative of the traffic during the busiest period of the year—June through August. Eight-hour manual turning movement counts were made at all significant intersections in Plaistow during August 1975. These were factored to ASWT volumes by using traffic recorder data.

The population forecasts in this study are based on a cohort survival projection of the New Hampshire population by the New Hampshire Office of Comprehensive Planning. Planning region forecasts were made by using the same technique; the regional totals were controlled to the state totals. The regional projections were disaggregated to the community level on the basis of (a) the potential saturation population of the town, (b) its accessibility to nearby urban centers, and (c) its competitive advantage for attracting growth as compared with other towns of the region.

A dwelling unit count by traffic zone has been done for Plaistow. The current population was distributed among the zones proportionately to the zonal dwelling units. The forecast town population was allocated to traffic zones on the basis of the number of existing and anticipated housing units in each zone.

Existing employment was obtained from New Hampshire Department of Employment Security data. The employers' addresses were given so that the determination of employment by zone was straightforward. Future employment was forecast by analyzing the area available for commercial and industrial development and estimating the number of employees resulting from that development.

## External Travel

The next step in the process is to assign the base-year external trip table to the coded highway network. The PLANPAC program for doing this is LOADVN, which assigns the trips between a pair of zones to the minimum time path between the zones. This assignment is on an "all or nothing" basis; in other words, although two paths between a pair of zones may differ in time by only a fraction of a minute, all the trips are assigned to the shortest path. Travel times or speeds may be adjusted slightly to achieve an optimal assignment. However, unrealistic assignments still occur where there are two or more equally efficient paths between pairs of zones. In a large network, such "lumpiness" is often masked. However, the problem is more obvious when it occurs in a small network such as that of Plaistow.

In an attempt to model traffic assignments more accurately, a different procedure was tried. The UROAD program from the Urban Transportation Planning System (UTPS) has the capability for doing a probabilistic multipath assignment in which trips between each pair of zones are assigned to more than one path. Each "efficient" path receives a fraction of the trips that is proportional to exp $(-\theta^*DI)$, where $\theta$ is the user-specified diversion parameter and DI is the difference between the total impedance of the given path and the minimum path's impedance. As $\theta$ approaches zero, all efficient paths become equally likely and, as $\theta$ becomes larger (approaches 10.0), only the minimum path has a significant likelihood.

Although stochastic assignment programs have been available for many years, they have been used only infrequently because of their relative complexity, the user's lack of control over program operation, and the difficulty in determining how a change in the network (an increase or decrease of speed on a particular link) will affect the assignment. Furthermore, the user must specify the value of $\theta$. Trial and error—running the program and comparing the output to actual ground counts—constitutes the only method of reasonably estimating that parameter.

The suggested value of $\theta$—0.002—assigns traffic too evenly to alternative paths and results in obvious overassignment on numerous links and improved loadings on only a few links. Higher $\theta$s were tried until 0.200 gave results very close to the "all-or-nothing" assignment. The very high overassignment on several key links when the lower values of $\theta$ are used seems especially perplexing and is a good example of the unpredictable nature of stochastic assignments.

The volumes from the various assignments were compared with the actual ASWT values. The assigned volumes should be less than or equal to the actual traffic since only external trips are being assigned. The difference between the assigned traffic and the actual ASWT volumes provides an estimate of the internal trips on each link.

None of the several assignments proved to be optimal for all links. When the all-or-nothing assignment was used, certain links were overassigned whereas others appeared to have too little traffic. Further refinement of the network improved the assignments to certain links, but there were generally compensating disadvantages to other links. Similar problems arose with the probabilistic assignment. Certain links would be improved by the use of that method, but other links would have poorer assignments. As smaller values of $\theta$ were used, some links become severely overassigned.

After analysis of the results of the different assignments, it was apparent that the all-or-nothing and $\theta$ = 0.20 assignments were the most realistic. For this study, the conventional all-or-nothing procedure was finally selected. It provides the best assignment and has the added advantage of being more convenient to use.

The projection of the base-year external trip table was done by using the Fratar trip-distribution model, which is based on the assumption that the change in trips in an interchange is directly proportional to the change in trips in the O-D zones that contribute to the interchange. The required input data are a base-year trip table and growth factors for each origin zone. Actually, two separate, but merged, tables are input—the external-external trips and the external-internal trips.

The model for zonal growth factors was developed by means of regression analysis. The number of external-internal trips originating in and destined for each zone was obtained from the trip table. These trips were the dependent variable. The independent variables—population and trade and nontrade employment by traffic zone—were provided as discussed above. The initial run of program BPR02R—the PLANPAC regression program—included all 23 of the zones in the study area. The regression analysis can be summarized as follows:

$$\text{Trips} = 128.22 + 21.44\ (\text{trade employment}) + 0.89\ (\text{population}) + 1.72\ (\text{other employment}) \quad (1)$$

$$\text{Trips} = 259.49 + 12.42\ (\text{trade employment}) + 0.89\ (\text{population}) + 1.14\ (\text{other employment}) \quad (2)$$

$$\text{Trips} = 8.43\ (\text{trade employment}) + 1.61\ (\text{population}) + 1.99\ (\text{other employment}) \quad (3)$$

The F and $R^2$ statistics and standard errors of estimate for the three equations are given below:

| Equation | F | $R^2$ | Standard Error of Estimate |
|---|---|---|---|
| 1 | 41.91 | 0.869 | 593.54 |
| 2 | 14.62 | 0.745 | 171.32 |
| 3 | 60.95 | 0.920 | 208.97 |

The F and $R^2$ statistics of Equation 1 are very good, but the standard error of estimate of 593 is very high. The residuals were examined, and four cases that had the most extreme residuals (zones 6, 13, 15, and 22) were removed from the subsequent regression analysis. These zones have employment or population characteristics that result in atypical trip-generation characteristics. The second run of the BPR02R program produced Equation 2, which has lower F and $R^2$ values than the previous equation; both values remain high enough, however, to ensure a high level of statistical reliability. More favorably, the standard error has been reduced by 71 percent.

At this point, projected employment and population data were substituted into the model, and future trip ends by zone were computed. Future trips divided by computed existing trips produced growth factors for each zone. Trips to the four zones deleted from the regression analysis were estimated individually. Future trips from all zones were summed and divided by the total existing trips. The overall growth factor of 1.34 seemed unreasonably low since both population and employment are projected to increase at rates greater than 1.40. The reason for the relatively low growth rate is the constant term of 259.49 in the equation, which causes the model to be less sensitive than it would be if the constant was much smaller. To alleviate this problem, a third run of BPR02R was made in which the Y-intercept (the constant term) was forced to zero. This resulted in Equation 3, which is statistically valid and has a standard error of estimate of 209. Future-year trips for each zone were computed, and the zonal growth factors were obtained. The overall increase of external-internal trips is projected to be approximately 55 percent.

The growth factors at the external stations cannot be estimated by the preceding methodology. Where possible, analysis of the trend in traffic growth is an accepted procedure for developing growth factors at cordon-line stations. However, very few historical traffic data are available in the area of Plaistow. On NH-125 and NH-108, short-term counts have been taken fairly consistently over the past 6 or 7 years. The projections of these trends were used to temper the growth estimates on these two routes. The primary method of developing these growth factors was to determine in which town or towns the trips through particular stations were originating. The population projections made by the New Hampshire Office of Comprehensive Planning were examined, and growth rates from 1975 to 1995 were computed. These growth rates formed the basis for the external-station growth factors. For example, at station 30, 62 percent of the traffic originates in Atkinson and about 26 percent originates in Hampstead. The population of Atkinson is projected to increase by a factor of 3.30 and that of Hampstead by 2.31 between 1975 and 1995. A growth factor of 2.6, which moderates the traffic growth somewhat, was selected.

After inputting the growth factors and trip table, the FRAT program was run with three iterations. The ratio between the attracted trips computed by multiplying base-year attractions by the growth factors and the iterated trip attractions balanced within 11 percent for all zones and stations. The TRIPSO data set from FRAT was input to program TRPTAB, and the zone-to-zone trip table was output.

Future external trips were assigned to the highway network by using the same methodology as that used for base-year trips.

## Internal Travel

The forecasting of internal trip volumes is a more complex task than is the projection of internal-external and external-external trips. Traditionally, in larger studies, internal-trip-making data are gathered through home interview surveys. However, in smaller study areas and recently in many larger areas, home interviews have not been conducted because of the excessive time and costs involved. One cannot calibrate a mathematical model for forecasting internal trips when no home interview survey has been done. The Fratar model requires the base-year trip table, which is not available for internal trips. The gravity model input includes trip productions (Ps) and trip attractions (As) by zone and friction factors (F) for each impedance increment. Although the F-factors may be estimated from similar studies, Ps and As depend on socioeconomic characteristics of the study area, which must be surveyed.

The procedure adopted for estimating future internal trip volumes by link was a novel part of this study. A single-purpose pseudogravity model was the basis of this methodology. Two runs of PLANPAC program GM were made. In the first, zonal productions were set equal to base-year population, and zonal attractions were estimated as follows:

1. For zones whose trade employment is predominately retail shopping, A = 1.0 (nontrade employment) + 9.0 (trade employment); and
2. For zones whose trade employment is in other categories, such as sit-down restaurants, automobile sales, and other lower generating categories, A = 1.0 (nontrade employment) + 5.0 (trade employment).

These relations do not estimate the number of trips attracted. Rather, as in the case of productions, they estimate only a rough relative attractiveness of one zone versus another. These relative rates were based on trip-generation models calibrated elsewhere in New Hampshire.

In the second GM run, the same procedure was followed for the forecast year, and projected population and employment were substituted for the base-year values. The other two data sets input to program GM are friction factors and impedances, which were the same for both runs. The friction factors were all assumed to be equal to 1, which means that trips of all lengths are equally likely. This is not an unreasonable assumption given a small network and a study area with only a few travel-time increments. The impedances are output from the PLANPAC program BUILDVN and reflect characteristics of the network.

The resulting trip table from each GM run was loaded onto the network. These are artificial volumes that should be approximately proportional to actual internal traffic. The growth rates for internal trips by link were computed by dividing base-year assigned volumes into the assignment for the forecast year. These growth factors by link were multiplied by the counted base-year internal trip volumes to produce forecast-year internal volumes.

To test the effect of using F-factors all equal to 1, the gravity model program was run twice more by using the same 1975 and 1995 Ps and As but normalized friction factors that ranged from 60 to 2 for 10 time increments. The F-factors were estimated from Federal Highway Administration charts (1). The assigned 1975 and 1995 internal trips varied by as much as 30 percent when the revised F-factors were used. However, the growth factors for nearly all links changed by less than 10 percent when the more refined factors were applied.

## Forecasts of Link Traffic

The 1995 ASWT volumes are the sum of the assigned external trips plus the product of the 1975 internal counts times the link growth factor. Unadjusted computed values are not used indiscriminately because some individual links are unrealistic. Where two or more alternate routes exist, projected external and internal traffic on all links is summed and then divided between the alternate paths. If the base-year external trips had been severely overassigned or underassigned or if the artificial gravity model trips appeared very unrealistic, more reasonable estimates of link traffic volumes were developed.

It must be reemphasized that, when a small study area and network are being considered and the all-or-nothing assignment is being used, the resulting assignment is lumpy and many links have unrealistic traffic volumes. It is therefore inadvisable always to use the numbers directly as output from the assignment program.

Projected increases in traffic volume over the 20-

year forecast period on Plaistow's streets and highways vary from less than 50 percent on roadways that serve primarily local traffic to more than 150 percent on highways that carry traffic to and from more rapidly growing neighboring towns. The major through highway in Plaistow, NH-125, is the exception since it has a projected growth of 40 percent. Since these forecasts are based on certain anticipated growth trends, substantial changes in patterns of development will affect the projected traffic volumes.

Obviously, the anticipated traffic volumes on many of the links of the Plaistow network will result in congestion and less than ideal travel conditions. The next phase of the Plaistow Highway Transportation Study will compute the hourly capacity of the streets and highways in Plaistow and compare these with the forecast volumes. In this way, deficiencies in the existing network may be identified so that remedial measures can be planned.

REFERENCE

1. Urban Trip Distribution Friction Factors. Federal Highway Administration, U.S. Department of Transportation, 1974.

*Publication of this paper sponsored by Committee on Transportation Planning Needs and Requirements of Small and Medium-Sized Communities.*

# Evaluation of the Impact of Restricted Interchanges on Travel Demand

Ronald J. Fijalkowski and Mark Rosenberg, Delaware Valley Regional Planning Commission, Philadelphia

**A study was conducted to develop and apply a low-cost process for assessing the effect on travel demand of a less-than-optimal design for proposed freeway interchanges on I-476 in Delaware County, Pennsylvania. The principal strength of the procedure developed lies in its computational simplicity. It is primarily manual and, although it does not require the excessive computer costs associated with the usual set of transportation planning models, it produces the detailed information required for such an analysis. The process relies on travel demand, trip length, and service provided for each trip movement to isolate possible impacts. The diversions expected to result from the constrained interchange designs are then determined based on the conditions of the surrounding network. For the two interchanges considered in the study, the process provided detailed information not usually associated with typical planning models. Although little effect on demand was anticipated at either interchange, the information generated was used for incorporating design revisions at one of the locations analyzed to eliminate deficiencies in capacity.**

Provisions contained in the National Environmental Policy Act of 1969 allow for the transportation-related use of parkland or historical sites if and only if no feasible and prudent alternative can be found. In response to testimony delivered at public hearings on the draft environmental impact statement (EIS) for I-476, the Pennsylvania Department of Transportation (PennDOT) modified two interchange designs to avoid the taking of such lands. Inherent in these modifications was a lessening in the quality of service provided at each interchange, where traffic signals replaced previously free-flow ramp movements at specified locations.

Concerned over the possible diversion of travel from these locations and its effect on the surrounding communities, PennDOT requested the Delaware Valley Regional Planning Commission (DVRPC) to evaluate the impacts of the design changes on future travel. In reviewing the request, it was deemed impractical to perform the analysis by using conventional simulation techniques because of both scope and cost. Such an approach would likely conceal the interaction between the system and the users and consequently hinder proper evaluation. To avoid these shortcomings, it was decided to accomplish the analysis by relying on a primarily manual process and using accepted techniques wherever possible.

The procedure developed is divided into two principal phases. Phase 1 focuses on traffic demand and operating characteristics for the original design of the freeway and each interchange. Phase 2 assesses the constraints created by the redesigned interchange and determines the movements affected.

This report details the procedure developed by DVRPC and presents an evaluation of the modifications proposed for the Mid-County Expressway. Figure 1 shows an outline of the process developed for this analysis and a description of each of the phases discussed.

## PHASE 1—ORIGINAL DESIGN

### Define Immediate Area of Impact

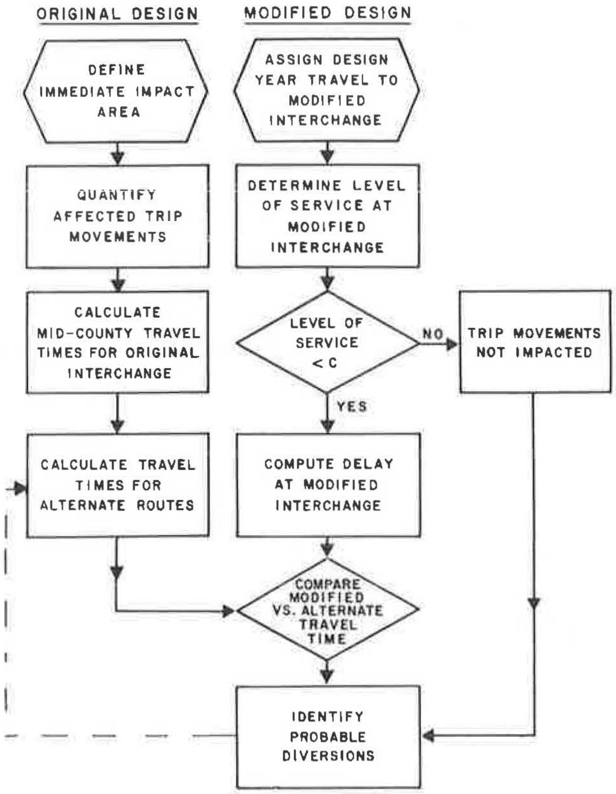
Trips that use the subject interchange will not be equally affected by the proposed modifications. Long-distance travel by the freeway will likely not divert to an alternative path regardless of the localized conditions at the interchange. The premise is that the time these trips save once they are on the freeway outweighs the time lost at the interchange.

This reasoning does not prevail for short-distance travel. Time saved on the freeway can readily be offset by time lost at the restricted movements and by the more circuitous routing caused by the freeway.

Trips that use the redesigned interchange and a short segment of the freeway should therefore be most severely affected, and the overall impact should be inversely related to distance traveled by way of the freeway. In conjunction with this hypothesis, trip movements by the modified interchange and a single segment of the freeway were first isolated for analysis as the most critical movements.

Since the facility under study is proposed, a "select link analysis" was performed to define trip movements that used the modified and adjacent interchanges. Ser-

Figure 1. Study methodology.




vice areas that surround each of these interchanges were then devised based on the select link trip matrixes and the logical trip patterns via the available highway network.

Quantify Affected Trip Movements

Next, the demand between the defined areas via the freeway is quantified by relying on select link data from an available traffic assignment. From this assignment, the number of trips that use I-476 was extracted for traffic analysis zones in each service area and aggregated to establish area-to-area traffic movements.

The selected link process used design-year (1996) traffic estimates forecast for the freeway. Design-year forecasts were used because they represent the adopted criterion for assessing proper design and denote the maximum impact caused by the revision.

Calculate Travel Times for Original Design

Travel times between service areas via the expressway are then calculated by using the original design scheme and accompanying travel forecasts. This calculation is performed by using peak-hour demand to reflect the critical period of maximum impact. The amount of travel is best defined during this period, and travel speeds can be estimated with reasonable accuracy.

For the I-476 analysis, evening peak-hour travel times between service areas were tabulated by using a relation for speed based on the volume, capacity, speed limit, and signal density of each link (1). Speeds were computed directionally because the proposed modifications will not provide balanced service. In addition, several modifications will create directional constraints.

Peak flow by direction was derived from the daily assignment by using adjustment factors created for the original design of the expressway and data provided by PennDOT on the proposed modifications (2).

Calculate Travel Times by Alternate Route

Peak-hour speeds are then similarly computed for competing alternative arterial paths between each service area. The predicted speeds that resulted were validated by using existing speed and delay data (3). To ensure compatibility, travel times for arterial and freeway paths are calculated from common points of divergence located at the approximate center of each service area.

PHASE 2—MODIFIED DESIGN

Assign Design-Year Travel to Modified Interchange

Concurrent with phase 1, design-year travel forecasts are also assigned to the appropriate movements at the redesigned interchanges. Again, evening peak-hour travel is allotted by direction.

Determine Level of Service at Modified Interchange

By using procedures described in the Highway Capacity Manual (4), a capacity analysis is conducted for all movements at each interchange, and the level of service is determined. For I-476, the analysis focused on the service provided at the signalized ramp termini introduced at each interchange. A signal phasing plan was developed to provide balanced operation wherever possible under geometric constraints imposed by the design. It was assumed in the analysis that the ramp designs were adequate to accommodate expected queues.

Level of service C, the accepted urban design standard, is the decision variable adopted in the process. At service levels A through C, drivers are not objectionably restricted and no more than a third of the cycles are fully loaded (4). Below level of service C (D through F), delays to approaching vehicles become substantial, queues develop, and vehicles are often detained for more than a single cycle.

Since drivers are not objectionably restricted until service falls below level C, it is reasonable that delays created by the modified designs will not be perceived by the public. If the capacity analysis then indicates future peak-hour service of C or better, the modified design is assumed to have no impact. If, conversely, level of service C is not achieved, the analysis must be continued to quantify the effect on travel.

Compute Delay at Modified Interchange

Where level of service C is not reached, the additional delay imposed by the modified design is needed to determine the new travel time between service areas via the freeway. For the analysis of I-476, this additional segment of travel time was estimated based on level of service and implied load factor. A relation developed by May and Pratt (5) that assumes that the surrounding signals will not be interconnected was applied. This relation is shown graphically in Figure 2.

Note that no estimate of delay exists for movements at level of service F at which demand exceeds capacity. The actual determination of delay in this instance is more complex and depends on (a) the period in which demand exceeds capacity, (b) the degree to which this occurs, and (c) the queue length that results from these

Figure 2. Average intersection delay versus load factor.

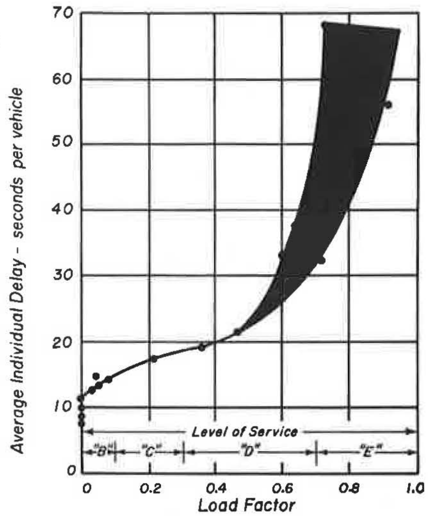
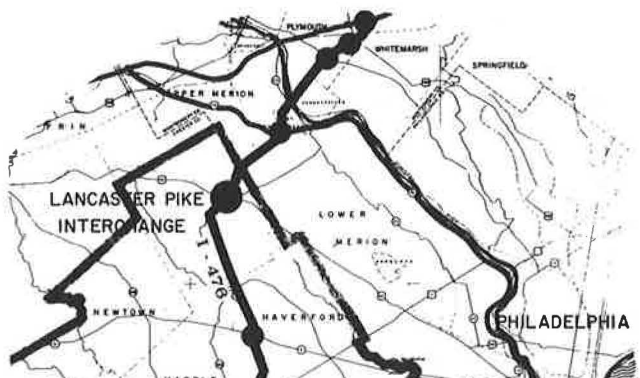



Figure 3. Map of I-476 area studied.




conditions. For the evaluation of I-476, a minimum value of 90 s was assumed in this instance by means of extrapolation of the graph shown in Figure 2. This simplified reapproximation rather than a more rigorous determination was applied because detailed flow data for the periods surrounding the peak hour were not available and subsequent design modifications were expected to correct these conditions.

Compare Modified and Alternate Travel Times

The appropriate signal delays calculated in the preceding step are then added to the travel times determined for each movement via the freeway for the original design. The estimated travel time for each trip when the interchanges are modified is then compared with the travel times previously determined for the alternative arterial paths between each service area.

Identify Probable Diversions

If travel times for the modified freeway designs exceed that for the alternative path, peak-hour trips for the identified movements are expected to divert from the freeway. Movements that provide faster service via the freeway are concluded to be unaffected by the modifications. As indicated, the alternative assignment assumes an all-or-nothing relation. Although a more complicated phenomenon assuredly exists, the small magnitude of diversions from I-476 renders a more sophisticated approach inappropriate.

Although peak-hour data are principally used in the analysis, the trends cited can be considered indicative of daily conditions and similar impacts can be expected throughout the day.

Depending on the results of this analysis, service areas may be extended to include the next adjacent interchanges along the freeway if the arterial paths indicate a marked improvement in travel time in comparison with the modified freeway. This extension would conceivably continue until an equilibrium point is reached where the freeway is obviously the faster alternative. Additionally, if a sizable magnitude of travel is diverted to the arterials, speeds should be recomputed to determine the impact and the process should be reiterated until a balance is achieved. For the I-476 analysis, neither of these feedback options was necessary.

MID-COUNTY EXPRESSWAY

I-476, the Mid-County Expressway, is a proposed six-lane facility extending 32.6 km (20.2 miles) through suburban Delaware and Montgomery counties in Pennsylvania. The expressway connects I-95 (the Delaware Expressway) with I-276 (the Pennsylvania Turnpike) and is designed to create a beltway facility around the city of Philadelphia. Along its route, 11 interchanges are proposed, including a third Interstate connection via I-76 (the Schuylkill Expressway).

The two interchanges for which modifications have been proposed are (a) US-30, Lancaster Pike; and (b) Baltimore Pike. The locations of I-476 and these interchanges are indicated on the map shown in Figure 3.

Both of these four-lane arterials that radiate from central Philadelphia currently accommodate approximately 25 000 vehicles/d with peak-hour flow near capacity. If the freeway is constructed as originally proposed, daily travel on these facilities is expected to climb in the design year to 40 000 vehicles/d at Lancaster Pike and 35 000 vehicles/d at Baltimore Pike.

Figures 4 and 5 show the original and modified design schemes and the location of the environmentally sensitive tracts that surround each interchange. As indicated, the directional interchange at Lancaster Pike has been revised so that all movements are accommodated in the southwest quadrant and these tracts are thus avoided. This design uses less right-of-way but requires signal control for all movements between Lancaster Pike and the Mid-County Expressway. Access to the freeway from the east is accomplished by way of dual left-turn lanes (movement B-D in Figure 4). Double left-turn lanes are also provided for traffic exiting the Mid-County Expressway from the north and destined west on Lancaster Pike (D-A).

At the Baltimore Pike interchange, the proposed modifications to the original cloverleaf scheme are less

Figure 4. Original and modified schemes for Lancaster Pike interchange.

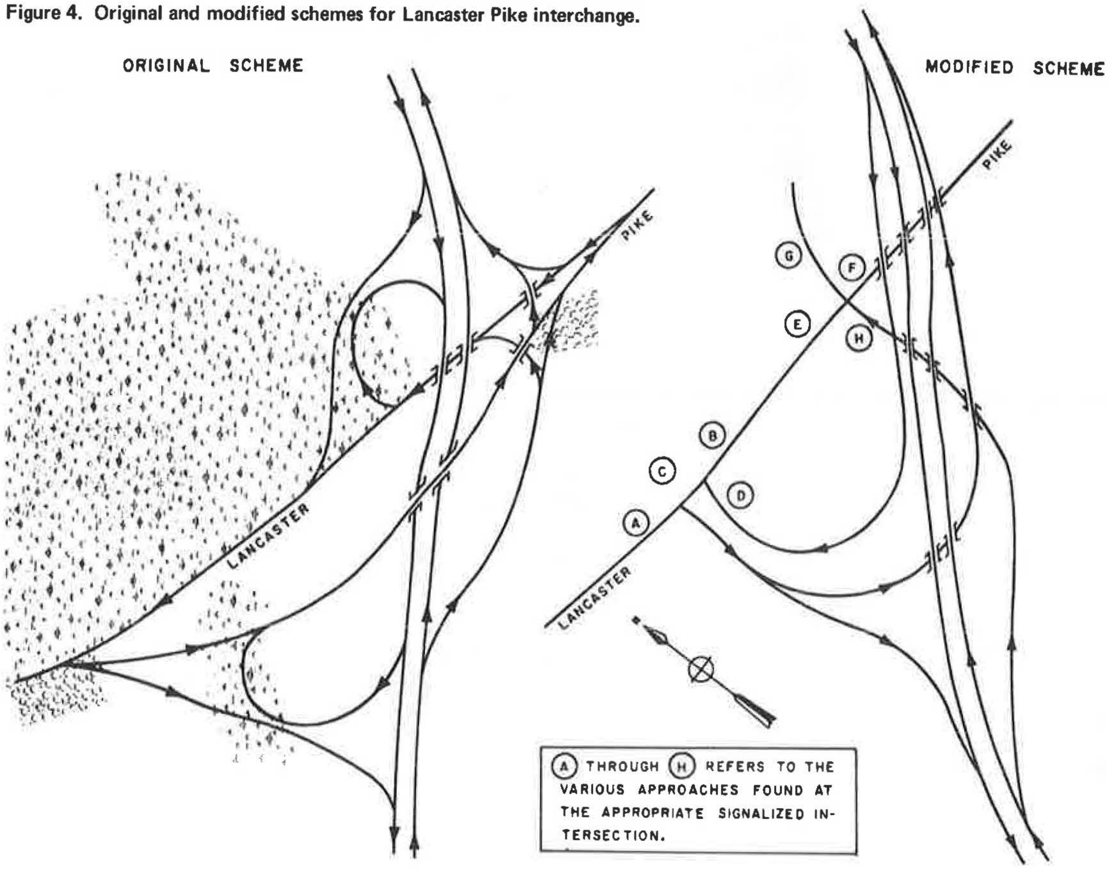
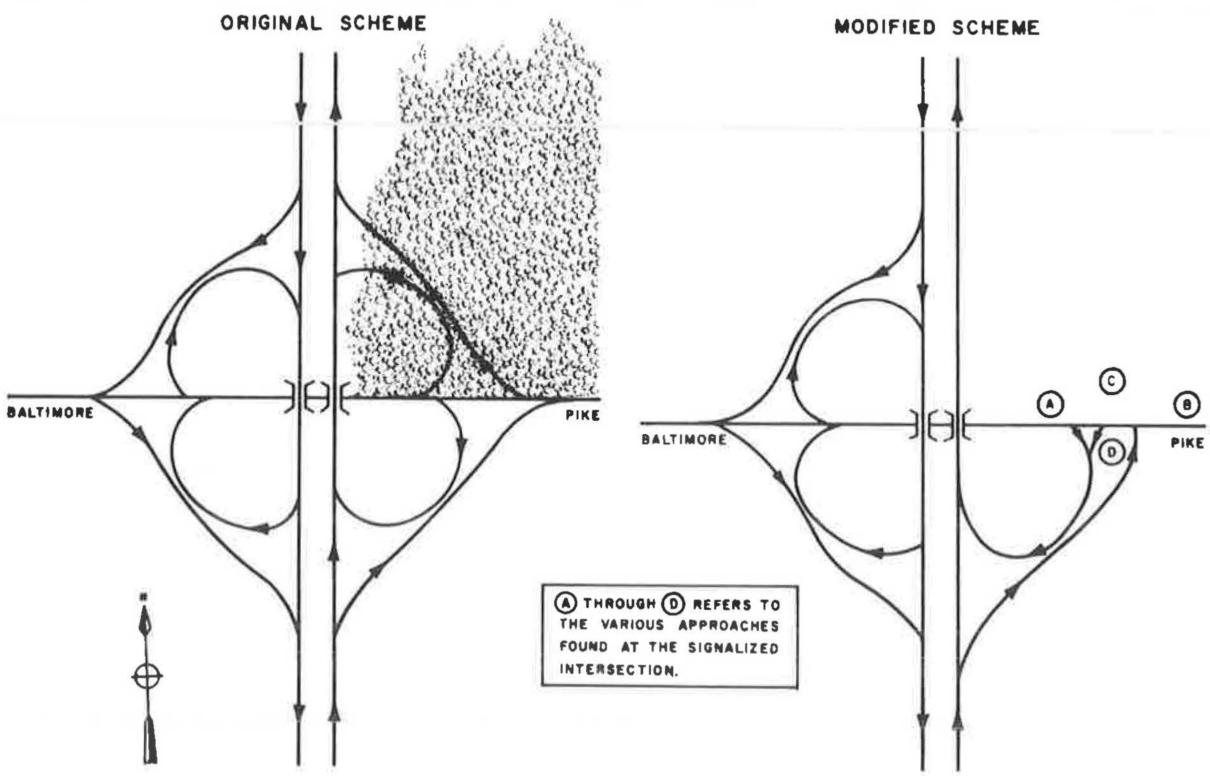



Figure 5. Original and modified schemes for Baltimore Pike interchange.



severe. Only those ramps in the northeast quadrant of the interchange have been relocated. This causes northbound travel both entering and exiting the freeway to proceed through a signalized intersection. Again, double left-turn lanes will be provided on (a) westbound Baltimore Pike for travel destined north on the Mid-County Expressway (B-D in Figure 5) and (b) the northbound exit ramp from the Mid-County Expressway for travel destined west on Baltimore Pike (D-A).

## Lancaster Pike

The study areas that surround the modified interchange at Lancaster Pike, including the adjoining interchanges at the Schuylkill Expressway and West Chester Pike, are shown in Figure 6. The east-west boundaries for the areas of analysis were delineated from the origins and destinations assigned to the expressway.

The major road network that services the delineated areas is composed primarily of two-lane undivided facilities and is also shown in Figure 6. Table 1 gives the calculated travel times for the original design between Lancaster Pike and the adjoining interchanges by way of the Mid-County Expressway and the surrounding parallel arterials. Note that the travel time between areas 2 and 6 is less via Matson Ford Road and Montgomery Avenue even under the original design plan for this interchange. Appropriately, no trips were assigned to the Mid-County Expressway between these areas.

A capacity analysis conducted for the southbound entrance and exit of the freeway showed that the signalized intersection would operate at capacity in the peak hour

Figure 6. Lancaster Pike interchange study area.

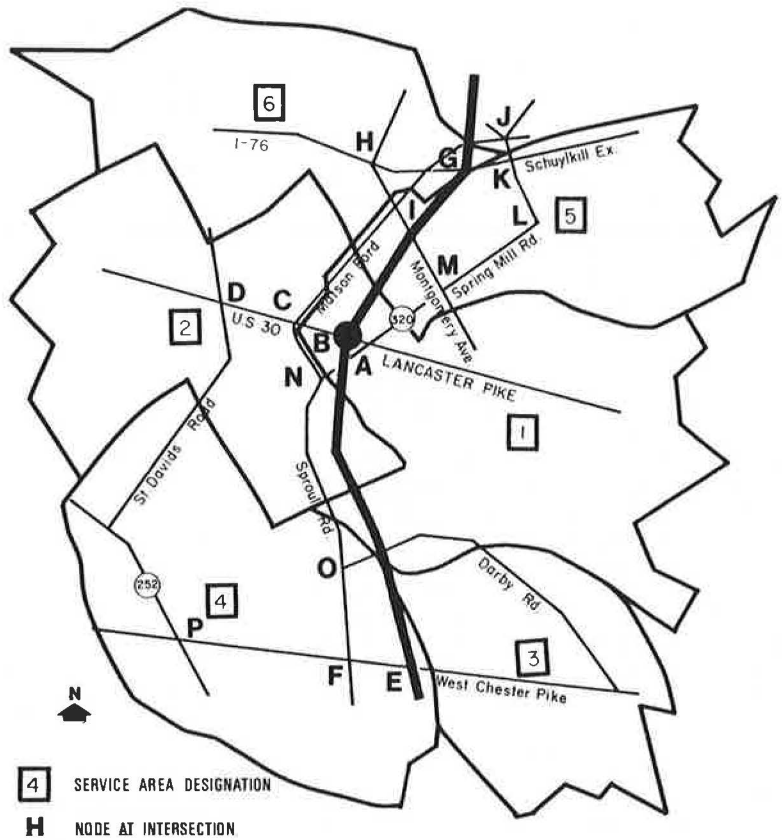



Table 1. Peak travel times between service areas for Lancaster Pike interchange (original design).

| Service Areas | Mid-County Expressway Path[a] | Minutes Between Areas Peak | Nonpeak | Arterial Path[a] | Minutes Between Areas Peak | Nonpeak |
|---|---|---|---|---|---|---|
| 1-3 | 1-A-B-E-3 | 23.5 | 16.3 | 1-A-N-O-3 | 31.9 | 28.0 |
| | | | | 1-A-N-O-F-E-3 | 37.7 | 31.8 |
| 1-4 | 1-A-B-E-F-4 | 21.1 | 19.3 | 1-A-N-O-F-4 | 22.3 | 21.5 |
| | | | | 1-A-B-C-D-P-4 | 28.5 | 21.5 |
| 1-5 | 1-A-B-G-K-5 | 7.7 | 6.3 | 1-A-M-L-K-5 | 9.8 | 9.0 |
| 1-6 | 1-A-B-G-H-6 | 7.1 | 10.1 | 1-A-M-I-H-6 | 13.6 | 11.5 |
| 2-3 | 2-D-C-B-E-3 | 32.9 | 25.2 | 2-D-C-N-O-3 | 38.3 | 32.7 |
| | | | | 2-D-P-F-E-3 | 42.3 | 37.9 |
| | | | | 2-D-C-N-O-F-E-3 | 44.2 | 38.5 |
| 2-4 | 2-D-C-B-E-F-4 | 29.0 | 26.0 | 2-D-P-4 | 32.3 | 25.1 |
| | | | | 2-D-C-N-O-F-4 | 32.6 | 31.1 |
| 2-5 | 2-D-C-B-G-K-5 | 16.8 | 15.7 | 2-D-C-I-J-K-5 | 21.6 | 19.4 |
| | | | | 2-D-C-I-H-G-K-5 | 25.0 | 18.8 |
| 2-6 | 2-D-C-B-G-H-6 | 19.1 | 17.5 | 2-D-C-I-H-6 | 18.1 | 15.3 |

[a]See Figure 6.

(level of service E). A similar analysis performed at the northbound off-ramp connection with Lancaster Pike indicated this to be the more critical juncture. Peak-hour volumes are expected to exceed capacity at this location and produce significant delays and intersection failure (level of service F). The table below gives the results of these analyses for specific movements at the interchange and also the delay incurred at the intersection because of the signalized control (no delay is added for movement E-F since the signal is currently in operation and level of service is not affected):

| Entrance or Exit Point for Mid-County Expressway | Movement | Level of Service | Additional Delay (min) |
|---|---|---|---|
| North Exit | F-E | F | 1.5 |
| North Exit | H-F | F | 1.5 |
| North Exit | E-F | A | None |
| North Exit | H-E | F | 1.5 |
| South Exit | D-B | B | 0.3 |
| South Exit | D-A | D | 0.7 |
| Entrance | A-D | E | 1.2 |
| Entrance | B-D | E | 0.7 |

The additional delays encountered for each movement produce the expected travel times when the modified interchange is in operation. For the Mid-County Expressway paths, Table 2 gives the travel times for each movement with the original and modified interchange, the number of peak-hour and daily trips between the respective service areas, and movements for which a faster alternative path exists.

In Table 2 the modified design will increase travel time through the interchange from 0.3 to 2.2 min depending on the specific movement. This range is applicable for all travel that passes through the interchange including the long-distance trips not directly analyzed.

For short-distance trips, the average increase is 1.4 min/vehicle. With this change, two-way trips between the Villanova and Marple Township areas (1 to 4) and directional trips from the Villanova to Conshohocken areas (1 to 5) are provided a slightly faster routing by way of PA-320. Although the maximum time savings is limited to approximately 20 s, it is reasonable to expect these trips to divert when one considers the expected congestion near the interchange and its psychological impact on driver behavior.

The impact on the remaining trips that approach the interchange from the south and those that originate in the Villanova area (area 1) is less clear. The estimated travel times on the Mid-County Expressway are faster than on the alternative arterial paths available. However, all of these movements will encounter the level of service F conditions cited. Depending on the degree of congestion reached at this location, some of this traffic may choose to avoid the location.

Assuming that this condition is tolerable, the analysis indicates that approximately 600 trips will divert from the Mid-County Expressway to the parallel section of Sproul Road (PA-320) and an additional 100 trips will divert to the parallel section of PA-320 and Spring Mill Road if the modified interchange is constructed.

## Baltimore Pike

The service areas and major road network that surround Baltimore Pike are shown in Figure 7; Sproul Road (PA-320), Providence Road (PA-252), and Woodland Avenue (PA-420) are the principal alternative arterial facilities to the Mid-County Expressway. The alternative paths considered and their respective travel times are given in Table 3. Again, the area-to-area movements that have no trips are afforded faster travel by way of the arterials even with the original design plan for the Baltimore Pike interchange.

The capacity analysis conducted at the approaches is summarized below (see Figure 5 for movements):

| Entrance or Exit Point for Mid-County Expressway | Movement | Level of Service | Additional Delay (min) |
|---|---|---|---|
| North Entrance | B-D | D | 0.7 |
| North Entrance | A-D | Nonsignalized free-flow movement | — |
| North Exit | D-B | B | 0.3 |
| North Exit | D-A | D | 0.7 |
| Intersection | B-A | A | 0.1 |

Overall service at the signalized ramps is favorable: Level of service ranges between A and D for specific movements. The D conditions are recorded at the double left-turning movements both on and off the northbound section of the Mid-County Expressway.

A comparison of peak-hour travel times with the original and modified interchange design plan and the number of local trips affected are given in Table 4. Only 10 of the possible 16 movements are affected by the modified design, and the average additional delay is calculated at 0.4 min or 25 s/vehicle. Individual delays range from 0.0 to 0.7 min/vehicle movement.

This table, used in conjunction with the alternative route travel times given in Table 3, indicates that movements assigned to the Mid-County Expressway are still served better via the freeway. As a result, no diversion is expected with the modified design plan proposed at the Baltimore Pike interchange.

**Table 2. Travel times between service areas via Mid-County Expressway for 1996 Lancaster Pike interchange.**

| | Peak Direction | | | | | Nonpeak Direction | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Travel Time (min) | | Number of Trips[a] | | | Travel Time (min) | | Number of Trips | | |
| Service Areas | Original | Modified | Peak | Daily | Condition of Movements | Original | Modified | Peak | Daily | Condition of Movements |
| 1-3 | 23.5 | 25.0 | 80 | 600 | 1[b] | 16.3 | 18.5 | 60 | 600 | 1[b] |
| 1-4 | 21.1 | 22.6 | 40 | 300 | 2[c] | 19.3 | 21.5 | 30 | 300 | 2[c] |
| 1-5 | 7.7 | 9.9 | 10 | 100 | 2[c] | 6.3 | 6.6 | 10 | 100 | |
| 1-6 | 7.1 | 9.3 | 10 | 100 | 1[b] | 10.1 | 10.4 | 10 | 100 | |
| 2-3 | 32.9 | 34.4 | 380 | 2900 | | 25.2 | 26.4 | 250 | 2900 | 1[b] |
| 2-4 | 29.0 | 30.5 | 80 | 600 | | 26.0 | 27.2 | 50 | 600 | 1[b] |
| 2-5 | 16.8 | 18.0 | 70 | 500 | | 15.7 | 16.0 | 50 | 500 | |
| 2-6 | 19.1 | 20.3 | 0 | 0 | 2[c] | 16.5 | 16.8 | 0 | 0 | 2[c] |

[a] Data summarized from selected link.
[b] Movements that must use approach operating at level of service F.
[c] Movements with faster arterial path when Lancaster Pike interchange is modified.

Figure 7. Baltimore Pike interchange study area.

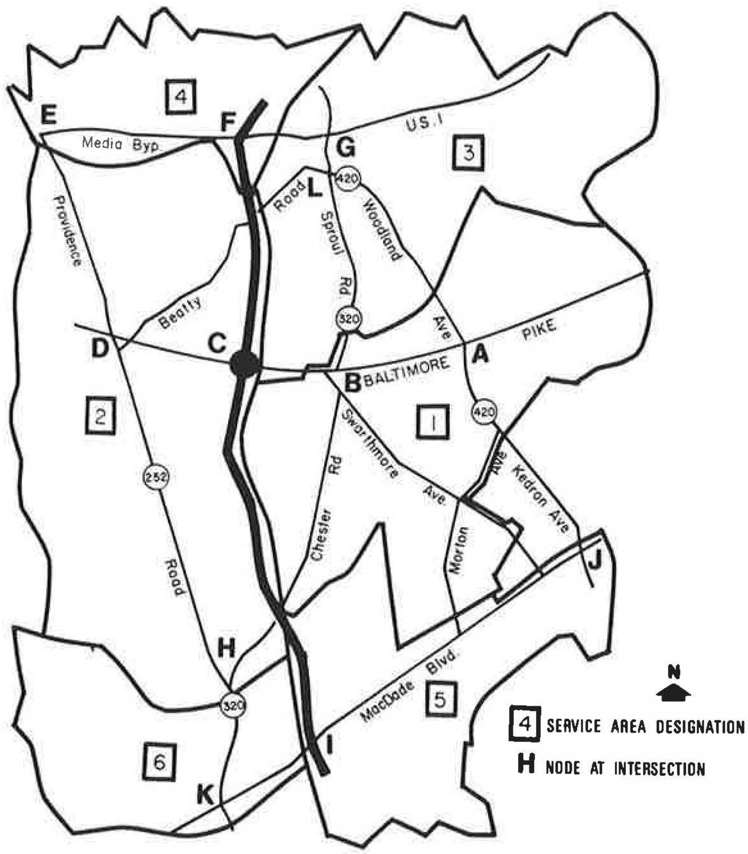



Table 3. Peak travel times between service areas for Baltimore Pike interchange (original design).

| | Mid-County Expressway | | | Arterial | | |
|---|---|---|---|---|---|---|
| | | Minutes Between Areas | | | Minutes Between Areas | |
| Service Areas | Path[a] | Peak Direction | Nonpeak Direction | Path | Peak Direction | Nonpeak Direction |
| 1-3 | 1-B-C-F-G-3 | 5.3 | 4.7 | 1-B-L-G-3 | 6.2 | 5.1 |
| | | | | 1-A-L-G-3 | 6.2 | 6.2 |
| 1-4 | 1-B-C-F-4 | 5.3 | 4.7 | 1-B-L-G-F-4 | 7.5 | 6.3 |
| | | | | 1-B-C-D-E-4 | 14.2 | 10.9 |
| 1-5 | 1-B-C-I-5 | 15.7 | 13.4 | 1-A-J-5 | 11.9 | 11.4 |
| 1-6 | 1-B-C-I-K-6 | 10.5 | 9.1 | 1-B-H-K-6 | 20.0 | 18.0 |
| 2-3 | 2-C-F-G-3 | 7.9 | 6.4 | 2-C-B-L-G-3 | 12.9 | 10.4 |
| 2-4 | 2-C-F-4 | 5.4 | 4.9 | 2-D-E-4 | 5.8 | 5.8 |
| 2-5 | 2-C-I-5 | 17.4 | 13.8 | 2-C-B-A-J-5 | 20.9 | 17.8 |
| 2-6 | 2-C-I-K-6 | 10.8 | 9.5 | 2-D-H-K-6 | 9.2 | 8.4 |

[a]See Figure 7.

Table 4. Travel times between service areas via Mid-County Expressway for 1996 Baltimore Pike interchange.

| | Peak Direction | | | | Nonpeak Direction | | | |
|---|---|---|---|---|---|---|---|---|
| | Travel Time (min) | | Number of Trips[a] | | Travel Time (min) | | Number of Trips | |
| Service Areas | Original | Modified | Peak | Daily | Original | Modified | Peak | Daily |
| 1-3 | 5.3 | 6.0 | 30 | 300 | 4.7 | 5.0 | 30 | 300 |
| 1-4 | 5.3 | 6.0 | 90 | 900 | 4.7 | 5.0 | 90 | 900 |
| 1-5[b] | 15.7 | 16.0 | 0 | 0 | 13.4 | 13.5 | 0 | 0 |
| 1-6 | 10.5 | 10.8 | 10 | 100 | 9.1 | 9.2 | 10 | 100 |
| 2-3 | 7.9 | 7.9 | 70 | 700 | 6.4 | 6.4 | 70 | 700 |
| 2-4 | 5.4 | 5.4 | 140 | 1300 | 4.9 | 4.9 | 140 | 1300 |
| 2-5 | 17.4 | 18.1 | 30 | 300 | 13.8 | 13.8 | 30 | 330 |
| 2-6[b] | 10.8 | 11.5 | 0 | 0 | 9.5 | 9.5 | 0 | 0 |

[a] Data summarized from selected link.
[b] Movements with faster arterial path when Baltimore Pike interchange is modified.

## CONCLUSIONS

The process developed achieved the objective of the analysis. By using negligible computer modeling and relying on accepted relations, specific information was produced that detailed the probable effects of each modification on travel movements.

The process appears most appropriate for analyzing the many alternatives that often result from the EIS process, as in the case of I-476. Its rational as well as computational simplicity allows for a clear presentation of impacts to the public, which facilitates development of an alternative that will achieve the desires of the community and its decision makers.

For I-476, the effect of the modified design plans at both interchanges will be predominately restricted to the localized area around each interchange. At Lancaster Pike, the probable diversion of 600 vehicles is relatively minute and should be considered negligible. The prime concern at this interchange should focus on alleviating the intersection problem expected at the terminus of the northbound off-ramp from the Mid-County Expressway. It is the magnitude of the associated congestion that will dictate the acceptability of the design and determine the extent to which travel through the interchange is altered.

For the Baltimore Pike, less severe modifications combined with smaller daily demand result in no travel diversion from the Mid-County Expressway. Here, capacity can adequately accommodate anticipated demand with only two movements below normal design standards.

In essence, the modifications do not measurably divert traffic, but a lower quality of service is provided to the users.

After the completion of the study, improvements were introduced by PennDOT at the Lancaster Pike interchange to eliminate the level F service cited in the analysis. Additional left-turning lanes have been introduced for movements G-F and H-E (Figure 4). In addition, widening of Lancaster Avenue to accommodate an additional lane of through travel from the east is assumed (6). These modifications improve the level of service to D in all cases.

## ACKNOWLEDGMENTS

This paper was financed in part by the Federal Highway Administration, U.S. Department of Transportation, and by the Pennsylvania Department of Transportation. The contents reflect our views and not necessarily the views or policies of the funding agencies.

## REFERENCES

1. Freeway-Surface Arterial VMT Splitter. Creighton, Hamburg and Federal Highway Administration, 1971, pp. 22-24.
2. Draft Environmental Impact Statement and Section 4(F) Statement. McCormick, Taylor and Associates, Philadelphia, 1976, plate IV-16.
3. Travel Time Survey. Delaware Valley Regional Planning Commission, Philadelphia, Stages 1, 2, and 3, 1972, 1973, 1974.
4. Highway Capacity Manual. HRB, Special Rept. 87, 1965, p. 130.
5. A. D. May and D. Pratt. A Simulation Study of Load Factor at Signalized Intersections. Traffic Engineering, Vol. 38, Feb. 1968, pp. 44-49.
6. Traffic and Transportation Basis Report. McCormick, Taylor and Associates, Philadelphia, Final Environmental Impact/Section 4(F) Statement, 1976, p. IV-88.

*Publication of this paper sponsored by Committee on Transportation Planning Needs and Requirements of Small and Medium-Sized Communities.*

# Interactive Computer Graphics for Station Simulation Models

Gregory P. Benz and Jerome M. Lutin, Princeton University

**A model developed for the Urban Mass Transportation Administration (UMTA) as an aid to the designers of transit stations—the UMTA Station Simulation (USS) model—was originally designed to give tabular output of numeric data on a batch-process computer. The potential for interactive running of the program with graphic as well as tabular output is studied. The 22 output reports of USS are examined for their ability to answer basic design questions. The following four types of computer graphics presentations are discussed as means of improving the ability of USS to answer such design questions: station animation and three types of static displays—histograms, station diagrams, and performance charts. Prototype graphic displays have been developed for each of the three static presentations and matched to present USS output.**

A computer program developed for the Urban Mass Transportation Administration (UMTA) as an aid to designers of transit stations—the UMTA Station Simulation (USS) model—is a discrete-event, Monte Carlo type of simulation model programmed in FORTRAN for use on IBM 360/370 computers. A semi-interactive version of USS was developed at Princeton University by using the IBM virtual machine (VM) operating system on an IBM 370/158 computer. A conversational monitoring system (CMS) EXEC program was written to create a user-oriented dialogue to handle language for file manipulation and job control. The semi-interactive version of USS reduced the cost of running the program by 95 percent in comparison with the conventional version (5).

The semi-interactive version of USS has made the program a more responsive part of the designer's creative process. Current research at Princeton and Aviation Simulations International is aimed at improving the program in the areas of data input and program capabilities to make USS more flexible and easier to use.

USS currently provides tabular output on times, volumes of people, and areas per person for the station being simulated. The development of graphic displays of data will improve the ability of USS to provide readily usable information and answer basic design questions. Data must be presented in a manner that not only answers these design questions but also allows the designers to perceive the performance of individual elements relative to the performance of the entire station network. Since designers communicate their ideas graphically, graphic display of information and ideas is the best way to communicate the "whole picture" and relate individual elements within their context. Thus, if a simulation model is to communicate effectively with designers, it should do so in a graphic form. This paper discusses several interactive computer graphic displays developed by the authors for incorporation in USS.

## BASIC DESIGN QUESTIONS

Designers of transit stations need answers to basic design questions such as the following:

1. How many turnstiles or fare-collection devices are required? In determining the proper number of devices, designers try to maximize passenger convenience while minimizing cost. This trade-off has become especially critical with the increasing use of automatic fare-collection devices, e.g., in the Washington, D.C., Metro system where magnetic card readers are required for both entry and exit.
2. How many escalators or stairs are needed? The provision of vertical circulation, particularly by mechanically assisted means such as escalators or elevators, also requires designers to balance passenger convenience against increased cost. Deep stations, although sometimes considered cheaper to build than cut-and-cover stations, increase the necessity for adequate vertical circulation capacity and also increase the cost per unit of vertical circulation.
3. Where might congestion occur and what will be the level of service? The area per person in a given space determines the level of service for pedestrian movement and comfort (3). Designers need to identify potential points of congestion that might cause low levels of service, and they also need to know the duration of the low levels.
4. How long does it take to walk through the station? Travel-time penalties imposed on the passenger by station design can influence the decision of an individual to use transit. The user should be able to walk as directly to his or her destination and as close to desired walk speed as possible.
5. How much area is needed for the platform? In sizing a platform, especially in a station that has a high volume of people, overdesign of the platform area can be avoided if the dynamics of passenger flow and crowding are examined over a time period.
6. What is the overall station performance? The proper arrangement of individual elements determines the overall performance of transit stations. This is true even though the design may contain a sufficient amount of space and number of devices for the design volumes.

Designers wish to know answers to the above questions under three station conditions: (a) high volume of inbound transit users (arriving at a constant rate, in groups, or both); (b) high volume of outbound transit users (departing in groups); and (c) high volume of inbound and outbound transit users (at platforms). The use of the term inbound has been chosen by convention to represent generally the direction taken by pedestrians who enter a transit station and proceed to a platform to board a transit vehicle or train. Outbound is the opposite direction (2).

## USS OUTPUT

USS produces up to 22 different reports of station operations. The first 8 are standard reports produced for every simulation run; the remainder are optional reports requested by the user. Each report is summarized here from the USS user's guide (2).

### Standard Output Reports

Report 1—link statistics in numeric order—shown in Figure 1 (2), provides summary statistics for each link, or pathway, of the network (in ascending numeric order), including the maximum number of people on the link at any one time, the minimum square feet per person (occupancy) in the link movement area (link length × link width) at any one time (because the model discussed here is calibrated for U.S. customary units of measurement, no SI equivalents are given), and the total volume of people who traversed the link.

Report 2—link statistics in ascending order by oc-

**Figure 1. USS report 1: link statistics in numeric order.**

USS REPORT 1

LINK STATISTICS IN NUMERIC ORDER

| LINK | MAXIMUM PEOPLE | OCCU-PANCY | TOTAL VOLUME |
|---|---|---|---|
| 1- 6 | 11 | 10.7 | 56 |
| 1- 31 | 5 | 15.4 | 22 |
| 2- 4 | 6 | 44.3 | 78 |
| 23- 3 | 10 | 17.5 | 30 |
| 24- 3 | 12 | 10.5 | 32 |
| 25- 3 | 13 | 9.1 | 39 |
| 26- 3 | 14 | 12.5 | 39 |
| 4- 8 | 6 | 29.0 | 53 |
| 4- 28 | 3 | 39.3 | 22 |
| 6- 9 | 10 | 27.4 | 45 |
| 8- 9 | 6 | 20.5 | 53 |
| 8- 28 | 3 | 33.3 | 22 |
| 8- 31 | 3 | 42.6 | 22 |
| 9- 10 | 4 | 23.7 | 36 |
| 9- 11 | 3 | 28.3 | 34 |
| 9- 12 | 4 | 23.7 | 27 |
| 10- 13 | 4 | 18.0 | 35 |

cupancy—is identical to report 1 except that the links are listed in ascending order of minimum square feet per person to aid in the quick identification of the worst spots of congestion in the station.

Report 3—node statistics in numeric order—provides a summary description of queuing and queuing-device activity at each node. Statistics are given for both inbound and outbound sides of the node including maximum number of persons in queue, the maximum recorded use of the queue area expressed as percentage of queue area capacity, the total volume of persons passing through the node, and the percentage of time that the device representing the node was in actual use (utilization).

Report 4—node statistics in descending order by use—is identical to report 3 except that the nodes are ordered in descending percentage of use of the queue area.

Report 5—total walk time for station—is a histogram that shows the distribution of walk time for passengers from the time they enter the station until they leave. Histograms are discussed in more detail later in this paper.

Report 6—total time in queue for station—is a histogram that shows the distribution of time in queue for passengers from the time they enter the station until they leave.

Report 7—total time in system for station—is a histogram that shows the distribution of the total time passengers spend in the station, walking and in queue, from the time they enter the station until they leave.

Report 8—overall station impedance by access-egress mode—shown in Figure 2 (2) has two parts. Part 1 gives for each station origin-destination pair the number of persons who arrived at the given origin and were assigned to leave at the given destination and the number of these who actually left the station. Part 2 gives the corresponding station impedance (minimum, mean, and maximum walk time, queue time, and total time in the system) in seconds as experienced by travelers who left the station.

## Detailed Output Reports (Optional)

Report 9—link occupancy report—lets the user see the dynamics of individual link activity at discrete intervals. The report is divided into two parts. Part 1 gives the number of persons who enter and leave the link movement area for both inbound and outbound directions, the number in the link movement area, and the proportion of persons on other links who shared part of their movement area with the subject link, all during the last time interval. Also included are the total number of persons and the area per person in square feet in the movement area. Part 2 gives the observations for the A-node and B-node queue areas inside the link (A-B) at the end of each time interval for the number of persons inside the queue area, the area in square feet required for those in the queue area, and the number of persons outside the specific queue area.

Reports 10 through 18—link occupancy histograms—show the distribution of individual link data in report 9:

1. Report 10—number of arrivals at link,
2. Report 11—number of departures from link,
3. Report 12—number in movement area on link,
4. Report 13—people from other links that compete on link,
5. Report 14—total people in the area associated with link,
6. Report 15—area per person in area associated with link,
7. Report 16—number in queue at node,
8. Report 17—required queue area for node, and
9. Report 18—people outside queue area at node.

Reports 19, 20, and 21—path impedance histograms—show the distribution of walk time, time in queue, and total time respectively along a path between a specified pair of nodes.

Report 22—individual path analysis—is a link-by-link trace of a single passenger as he or she moves from zone of station access to zone of station egress. This report is described in more detail later in this paper.

Table 1 gives the 22 reports provided by USS and indicates the ability of each report to answer specific design questions. Despite the tremendous amount of information available in these reports, in many cases the data are presented in a form designers cannot readily use. The time and uncertainty involved in transforming the vast amount of data into a readily usable form—especially for reports 9 through 21, which must be specified for each pair of nodes for which the designer wants information—may discourage the regular use of USS in the design process.

## POTENTIAL GRAPHIC OUTPUT

If data produced by USS could be graphically displayed

Figure 2. USS report 8: overall station impedance by access-egress mode.

USS REPORT 8

OVERALL STATION IMPEDANCE BY ACCESS/EGRESS MODE (PART 1)

| ORIG NODE | DEST NODE | ARRIVALS MODE/LINE | | ARRIVALS PEOPLE | DEPARTURES MODE/LINE | | DEPARTURES PEOPLE |
|---|---|---|---|---|---|---|---|
| 1 | 2 | BUS | 1 | 0 | | | 0 |
| 1 | 3 | BUS | 1 | 56 | RAIL | 1 | 36 |
| 2 | 1 | | | 0 | BUS | 1 | 0 |
| 2 | 3 | | | 56 | RAIL | 1 | 42 |
| 3 | 1 | RAIL | 1 | 31 | BUS | 1 | 22 |
| 3 | 2 | RAIL | 1 | 31 | | | 22 |

OVERALL STATION IMPEDANCE BY ACCESS/EGRESS MODE (PART 2)

| ORIG NODE | DEST NODE | WALK TIME MEAN | WALK TIME MIN | WALK TIME MAX | TIME IN QUEUE MEAN | TIME IN QUEUE MIN | TIME IN QUEUE MAX | TIME IN SYSTEM MEAN | TIME IN SYSTEM MIN | TIME IN SYSTEM MAX |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 1 | 3 | 58.1 | 39.2 | 95.8 | 42.7 | 8.5 | 85.2 | 100.8 | 63.5 | 159.4 |
| 2 | 1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 2 | 3 | 54.7 | 39.6 | 83.0 | 42.5 | 5.4 | 92.2 | 97.0 | 46.3 | 166.6 |
| 3 | 1 | 56.9 | 38.0 | 109.0 | 30.7 | 4.6 | 86.3 | 87.6 | 42.6 | 156.4 |
| 3 | 2 | 56.9 | 38.6 | 96.4 | 18.2 | 4.4 | 47.4 | 75.1 | 44.2 | 113.4 |

**Table 1. USS reports: ability to answer design questions and form of graphic output.**

| Type of Report | Report Number | Title | Design Question: Number of Devices | Vertical Circulation | Platform Area | Walk Time | Congestion (level of service) | Overall Station Performance | Form of Graphic Output |
|---|---|---|---|---|---|---|---|---|---|
| Standard output | 1 | Link statistics (numeric order) | X | X | X | | X | | D |
| | 2 | Link statistics (ascending order) | X | X | X | | X | | D |
| | 3 | Node statistics (numeric order) | X | X | X | | X | | D |
| | 4 | Node statistics (ascending order) | X | X | X | | X | | D |
| | 5 | Total walk time (station) | | | | X | | X | H |
| | 6 | Total queue time (station) | | | | | | X | H |
| | 7 | Total time in system (station) | | | | X | | X | H |
| | 8 | Station impedance | | | | X | | X | |
| Detailed occupancy | 9 | Link occupancy | X | X | X | | X | | 2 |
| Link occupancy histogram | 10 | Number of arrivals at link | X | X | X | | X | | H |
| | 11 | Number of departures at link | X | X | X | | X | | H |
| | 12 | Number of movements on link | X | X | X | | X | | H |
| | 13 | People from other links who compete on link | | | X | | X | | H |
| | 14 | Total people in link area | | | X | | X | | H |
| | 15 | Area per person in link area | X | X | X | | X | | H |
| | 16 | Number in queue at node | X | X | X | | X | | H |
| | 17 | Required queue area for node | X | X | X | | X | | H |
| | 18 | People outside queue area at node | X | X | X | | X | | H |
| Path impedance histogram | 19 | Walk time, node A to node B | X | X | X | X | X | | H |
| | 20 | Queue time, node A to node B | X | X | X | X | | | H |
| | 21 | Total time, node A to node B | X | X | X | X | X | | H |
| | 22 | Individual path analysis | X | X | X | X | X | X | 1 |
| Potential graphic aid | | | H | 1 | 2 | H | 1, 2 | 1, 2 | |

Note: D = station diagram, H = histogram, 1 = individual path analysis chart, and 2 = area performance chart.

**Figure 3. USS report 5: total walk time for station.**

```
                              USS        REPORT  5

TOTAL WALK TIME FOR STATION

   0    4    8   12   16   20   24   28   32   36   40
   +----+----+----+----+----+----+----+----+----+----+ P.C.   CUM. COUNT
  0.                                                   0.0    0.0     0
  5.                                                   0.0    0.0     0
 10.                                                   0.0    0.0     0
 15.                                                   0.0    0.0     0
 20.                                                   0.0    0.0     0
 25.                                                   0.0    0.0     0
 30.                                                   0.0    0.0     0
 35.                                                   0.0    0.0     0
 40........                                            5.7    5.7     7
 45..........                                          7.4   13.1     9
 50.........................                          19.7   32.8    24
 55..............................                     23.8   56.6    29
 60.......................                            18.0   74.6    22
 65...........                                         8.2   82.8    10
 70..........                                          7.4   90.2     9
 75...                                                 1.6   91.8     2
 80....                                                2.5   94.3     3
 85..                                                  0.8   95.1     1
 90..                                                  0.8   95.9     1
 95...                                                 1.6   97.5     2
100...                                                 1.6   99.2     2
105.                                                   0.0   99.2     0
110..                                                  0.8  100.0     1
REMAINING COUNTS ARE ALL ZERO
NUMBER OF OBSERVATIONS=      122.,  SUM=      7145.,  MEAN=  58.566,
VARIANCE=        165.361,  STANDARD DEVIATION=  12.859
```

in a form more easily used by designers, USS could be used more readily and be more valuable. The potential types of graphics discussed in this paper are (a) station animation, (b) histograms, (c) station diagrams, and (d) performance charts (individual path analysis chart and area performance charts).

Table 1 gives the graphic form that can summarize the information in each USS report. The graphic form that potentially can aid in answering the basic design questions is also indicated. One or two charts can do the job of many reports by combining the information contained in several reports and displaying it in a form that designers can readily perceive and use.

## Station Animation

Animation is a useful cognitive tool for visualizing and aiding in the comprehension of complex processes. Integrated with an interactive computer simulation model, computer animation can be created and viewed in real time directly at an interactive terminal. A demonstration project that applied computer animation to a station simulation model showed that, despite current high cost, time consumption, and program rigidity, animating a model enables designers to visualize intricate spatial and temporal relations of transit station-vehicle activities (1).

Aviation Simulations International is currently investigating the potential of USS results for animated computer graphics (4). Potentially, station animation can be very valuable. Currently, however, it requires sophisticated computer hardware and software beyond the capacity of most users for whom USS is intended. For the near future, therefore, any graphic presentation by USS should be within the capabilities of the most readily available output device—a line printer or simple cathode ray tube (CRT) terminal. Ultimately, however, station animation will provide valuable assistance in the visualization and comprehension of transit station designs.

## Histograms

Histograms that show the distribution of particular simulation data represent the most common output format of the original USS program (see the last column and the bottom line of Table 1). Figure 3 (2) shows an example of a USS report that uses a histogram as graphic output. In this histogram (report 5—total walk time for station), the distribution of simulated walk time for passengers from the time they enter the station until they leave is shown. The vertical scale on the left is total walk time in seconds divided into 5-s interval groups. The horizontal scale is the number of observations for each interval group.

## Station Diagrams

Diagrams that represent the station network could be produced by USS on a line printer or other output device to display the data from several USS reports (see the last column of Table 1). USS reports provide data on individual elements of the station but not on total systems or networks (except report 22). The performance of individual elements is not of much value to the designer unless it is easily related to the whole station network. One reason is that poor performance of one element may not be a problem of the element itself but may actually be the result of the poor performance of an adjacent element. Another reason is that an element may appear to be performing poorly when isolated from the station network and yet within the network may actually be performing properly. Finally, representation of individual elements is open to bias because of subjective judgments by the analyst in coding the station network. Station diagrams allow the designer to view the performance of an individual element relative to the station network as a whole.

Figure 4 (6) shows a computer-drawn plan of a typical transit station with a platform, concourse, and several corridors. In Figure 5 (6), this station layout is shown as a computer-drawn network representation with 22 nodes as required for analysis by USS. Node numbers are circled. The train is represented by node 1, the train doors by nodes 2 through 9, and the station exit by node 21.

Figure 6 (6) is a diagram of the station network that shows the percentage of the total passenger volume that used a particular link. Numerical results posted on the links are the percentage of total passenger volume that used each link. Rectangles are drawn to the right of link centerlines to indicate direction. The widths of the rectangles are proportional to link volumes (single lines represent trip volumes less than 1 percent of total origin-destination volumes).

Viewing this picture of the network annotated with numeric results provides an easier means of interpreting data than scanning through tables of numbers. Link volumes are shown here, but a similar diagram can show node volumes.

## Performance Charts

Two types of charts that will aid the designer in evaluating station performance have been developed: the individual path analysis chart (1 in Table 1) and area performance charts (2 in Table 1).

### Individual Path Analysis Chart

The individual path analysis chart is a graphic representation of the data produced in USS report 22—individual path analysis. Report 22 is a link-by-link trace of a single passenger as he or she moves from station entrance to exit or transit vehicle. When the passenger exits, the program begins tracing another individual beginning at a different entrance. A separate report is printed for each person traced. A sample report is shown in Figure 7 (2). The report shows the individual's exogenous attributes, i.e., an identification number, origin-destination nodes, desired walk speed, and mobility status (handicapped or nonhandicapped). The endogenous attributes, which are printed as the passenger completes travel along each link defined by the From Node and To Node columns, include a series of time measurements related to the simulation clock time printed under the Time column. All times are in seconds. The link time measurements include time in queue, cumulative time in queue, link walk time, cumulative walk time, time on link (sum of walk time and time in queue), and cumulative time in system. As shown in Figure 7, it took the passenger 76 s to travel from node 3 to egress node 1 (2).

To establish a graphic display of these tabular data, it is assumed that perceived travel time is more often used for evaluation of travel time by a transit user than actual travel time. Time spent walking is, within certain limits, perceived as making progress or time well spent. Therefore, it has a speed equal to the slope of a line defined by distance and time (speed = distance/time). Time spent in queue is not perceived as making progress or time well spent. Therefore, its slope is equal to zero.

The individual path analysis chart (Figure 8) is

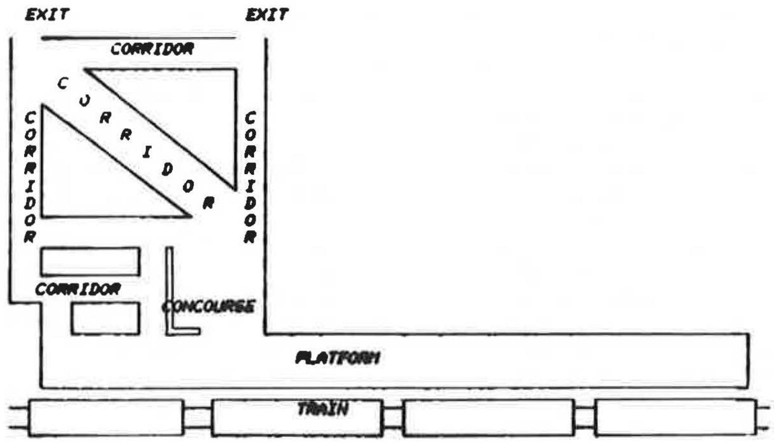



Figure 4. Transit station plan.

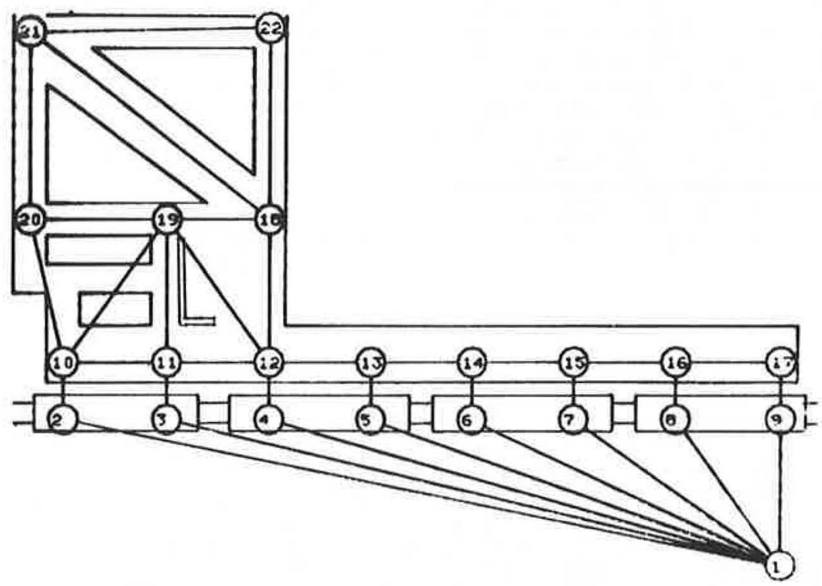



Figure 5. Network representation of transit station.

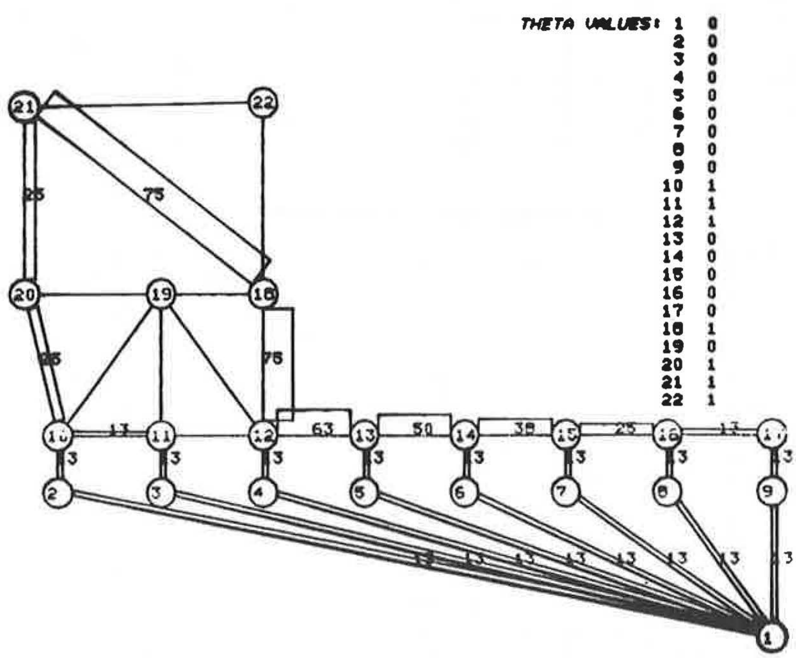



Figure 6. Percentage of total passenger volume using each link.

readily constructed from the information given in report 22 and is a graphic representation of the data shown in the report. Chart 1 consists of a planar graph in which the horizontal axis plots time in seconds from start (point 0) to finish and the vertical axis plots distance in feet from origin (0 feet) to destination (X feet away). The slope of any line is equal to the speed of the individual whose progress is being plotted: Slope = ($\Delta X \div \Delta Y$) = ($\Delta$FT/$\Delta$SEC) = speed (feet per second). Thus, the steeper the slope is, the faster is the speed. When the slope equals zero, the person is standing still (i.e., in queue). This chart plots the speed between the nodes along a desired path between two zones. The designer can now graphically view the progress of an individual from origin to destination, noting where queues occur and where speed is reduced because of congestion or some other factor. To facilitate the analysis, the node numbers are indicated adjacent to nodes along the path at their appropriate distance and time. On the right side of the chart, the link speeds (SP) of the individual are listed. When link time consists of walk time and queue time, walk time is plotted up to the node distance (slope = speed), and then the queue time is represented as a flat line (slope = 0) that graphically shows the amount of time spent in that queue (QT). This graphically reinforces the assumption that a person perceives time spent walking (progress) differently from queue time (no progress). The individual path analysis chart, accompanied by a printout of report 22, allows quick examination of the performance of a passenger's path through the station.

## Area Performance Charts

One measure of platform, or area, performance with a high volume of people is area per person. The measure determines the level of service for pedestrian movement and comfort (3). The less the area per person is, the poorer the level of service is. However, area-per-person criteria alone are not sufficient for evaluating the performance of a transit station platform (or area). Periods of platform crowding are tolerable if their duration is relatively short (perhaps less than 10 s). In the case of a transit station that has a high volume of passengers entering (inbound) and exiting (outbound) a transit vehicle during a given time period, extremely low values of area per person will exist while the inbound and outbound passengers are on the platform together.

The platform performance charts consist of two parts: The platform population chart plots the number of people on the platform (or other area) over time; platform area per person is the result of the total fixed area of the platform divided by the total number of people on the platform over time. Currently, USS does not have the ability to simulate the platform as an area, but future modifications should allow this.

### *Part A—Platform Population Chart*

The platform population chart (Figure 9) has along its Y-axis the number of people on the platform and along the X-axis time in seconds. The graph plots the number

**Figure 7. USS report 22: individual path analysis.**

USS REPORT 22

INDIVIDUAL PATH ANALYSIS

PERSON ID = 1 ORIGIN = 3 DESTINATION = 1
DESIRED WALKING SPEED = 215 HANDICAPPED =NO

| FROM NODE | TO NODE | TIME | TIME IN QUEUE | CUM TIME IN QUEUE | WALK TIME | CUM WALK TIME | TIME ON LINK | CUM TIME IN SYSTEM |
|---|---|---|---|---|---|---|---|---|
| 3 | 26 | 7 | 2 | 2 | 4 | 4 | 6 | 6 |
| 26 | 22 | 11 | 0 | 3 | 4 | 8 | 4 | 11 |
| 22 | 20 | 14 | 0 | 3 | 3 | 11 | 3 | 14 |
| 20 | 18 | 18 | 0 | 3 | 4 | 15 | 4 | 18 |
| 18 | 16 | 29 | 0 | 3 | 11 | 26 | 11 | 29 |
| 16 | 15 | 34 | 1 | 3 | 4 | 30 | 5 | 33 |
| 15 | 27 | 48 | 7 | 11 | 6 | 36 | 13 | 47 |
| 27 | 28 | 58 | 0 | 11 | 11 | 47 | 11 | 58 |
| 28 | 8 | 62 | 0 | 11 | 3 | 50 | 3 | 61 |
| 8 | 31 | 72 | 5 | 16 | 6 | 56 | 11 | 72 |
| 31 | 1 | 76 | 0 | 16 | 4 | 60 | 4 | 76 |

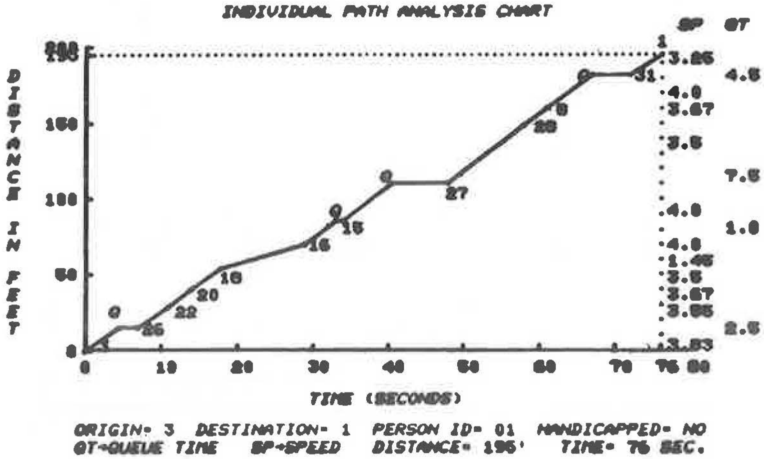


**Figure 8. Individual path analysis chart.**

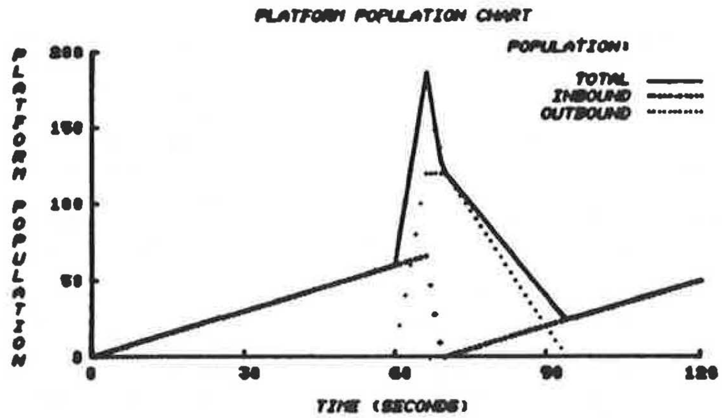



Figure 9. Area performance chart: part A—platform population.

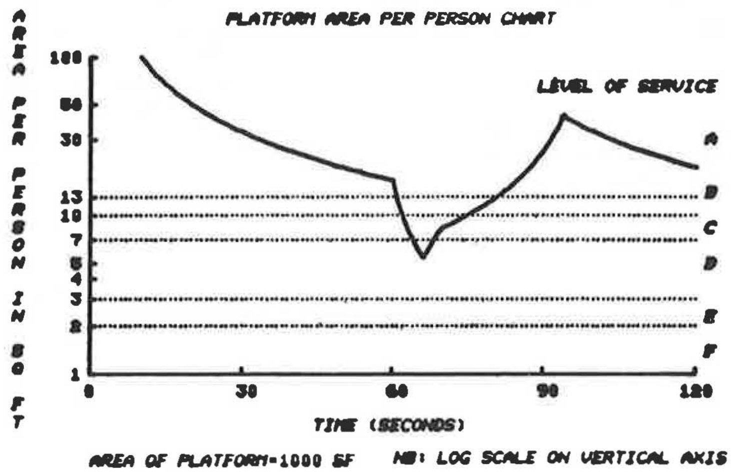



Figure 10. Area performance chart: part B—platform area per person.

of people on the platform at any one time—total, inbound, and outbound populations. Inbound passengers arrive either at some rate (people per second) or in groups at specific time periods ($Y_1$ people at $T_1$, $Y_2$ people at $T_2$, and so on). The train arrives at a specified time and discharges its passengers (outbound) at some rate (people per second). When the outbound passengers are off the train, those passengers who were on the platform (inbound) board the train at some rate (people per second). When all those who desired to board the train have done so, the train departs. Meanwhile, those who exited the train (outbound) are leaving the platform at some rate (people per second) after taking a specified number of seconds from time of exiting the train to reach the exits of the platform.

*Part B—Platform Area per Person Chart*

The platform area per person chart (Figure 10) has along its X-axis time in seconds and along its Y-axis area per person in square feet. This chart allows the designer to see the area per person at any time during the simulation. The total population of the platform at each second in part A is divided into the total fixed area of the station platform. The lowest value plotted on the chart is the minimum area per person during the simulation. This allows the designer to see the minimum area per person and when it occurred. The chart also allows the designer to examine the duration of low values of area per person. Low values are acceptable if their duration is for short intervals. Because the area of the platform is fixed, the change in area per person from X people to X + 1 people has a diminishing rate of significance: lim $\Delta$ area per person = 0 and $X \rightarrow \infty$. This problem is particularly apparent at the critical peak volumes. Therefore, a logarithmic scale is used along the Y-axis to improve the visual perception of the condition. The Fruin levels of service are listed on the right side of the chart. The chart summarizes the data given in USS reports 1, 3, and 9 through 20.

A variation of part B, which would also be derived from part A or USS report 9—link occupancy—would be a nomograph that showed the required area for a platform as a function of passenger population over time for the various levels of service.

## CONCLUSION

Interactive computer graphics has the potential to make the USS computer program an effective part of transit station designers' creative process. A review of the USS output reports showed that, although they can aid in answering basic design questions, the information in the reports often is not in a form that designers can readily perceive and use. Four types of graphics can improve the ability of USS to aid designers:

1. Station animation allows designers to visualize intricate spatial and temporal station-vehicle activities.
2. Histograms show the distribution of particular simulation data.
3. Station diagrams allow the designer to relate the performance of individual elements to the entire station network.

4. Two types of performance charts graphically summarize a great deal of information on (a) individual paths through the station and (b) the population and area per person in a particular space.

The last two types of graphics were developed by the authors from data already present in one form or another in USS reports.

Although the use of these graphics was described in the specific context of the USS computer program, their use is open to application in many other types of design problems that involve movement of people and vehicles. The use of computer graphics, and especially the development of the innovative display techniques demonstrated in this application, serve to highlight problem areas and focus attention on specific aspects of simulation modeling. Graphic tools provide a superior format for computer output that enables the designer and the analyst to interpret the simulation results more quickly and apply them to the design problem.

ACKNOWLEDGMENT

This research was supported by a grant from the Planning Methodology and Technical Support Division of the Urban Mass Transportation Administration, U.S. Department of Transportation.

REFERENCES

1. R. M. Baecker and T. R. Horsley. Computer Animated Simulation Models: A Tool for Transportation Planning. TRB, Transportation Research Record 557, 1975, pp. 33-44.
2. Software Development Program—Transit Station Simulation Users Guide. Barton-Aschman Associates and Office of Transit Planning, Urban Mass Transportation Administration, Draft Rept. UTP.PMM 75.1, 1975.
3. J. J. Fruin. Pedestrian Planning and Design. Metropolitan Association of Urban Designers and Environmental Planners, New York, 1971.
4. E. S. Joline and J. C. Hayward. Application of Computer-Drawn Motion Pictures for Demonstration of Urban Transportation Concepts. In Personal Rapid Transit III (D. A. Gary, ed.), Univ. of Minnesota, Minneapolis, 1976.
5. J. M. Lutin. Transit Station Simulation Model (USS) Testing and Evaluation. Urban Mass Transportation Administration, U.S. Department of Transportation, Final Rept. 76-TR-9, July 1976.
6. J. M. Lutin. A Perceptual Model of Path Choice in Urban Transportation Networks. Princeton Univ., Princeton, NJ, PhD dissertation, 1977.

*Publication of this paper sponsored by Committee on Computer Graphics and Interactive Graphics.*

# Use of Interactive Computer Model STREAK for Transportation Planning

Robert S. Scott and Bernard Mendes, De Leuw, Cather and Company, Los Angeles

**STREAK, an interactive set of computer programs for transportation planning, is described in both nontechnical and technical terms, and some experience in its use is discussed. The program uses portable desk-top terminals and conventional telephone lines. Outputs on the portable terminal include responsive dialogue (questions and answers), tabular arrays, and data plotted by coordinates for use with overlaid maps. Capabilities of the model include multimodal network development, pathfinding in multimodal networks, travel demand estimation for all modes, travel assignment, and outputs in map coordinates or tabular form under direct user control. The primary advantage of STREAK in planning studies is its ability to evaluate and report findings on an alternative in a matter of seconds. A secondary advantage is the ease with which the model performs data corrections and network modifications directly via the terminal. The model has proved its value in several studies.**

An interactive set of computer programs has been developed to assist planners with problems that require the testing of many candidate solutions. Although primarily designed for sketch-planning studies, the programs have proved to be useful for a variety of planning problems including many that involve only a few alternative solutions. The programs, called Strategic Transportation Evaluation and Analysis Kit (STREAK), are actuated by means of a portable desk-top terminal and a conventional telephone line. The purpose of this paper is to describe the role of STREAK in the planning process, provide a brief technical description of the models, and discuss experience to date with their use.

## STREAK AND THE PLANNING PROCESS

STREAK is well suited for network evaluation, travel demand forecasting, locational analysis, and accessibility calculations. Though primarily intended for sketch planning, it can also be used for detailed studies of small to medium-sized areas.

The major capabilities of the model include the following:

1. Multimodal network development and modification in which links can be rapidly added, deleted, or modified (specifically or by class codes) to activate alternative systems such as highway networks, bikeways, pedestrian paths, or bus lines on highway links;
2. Non-line-specific coding scheme for simplified transit network representation, which significantly reduces the time required to evaluate alternative transit networks;
3. Pathfinding in multimodal networks, subject to a variety of user-specified constraints;
4. Travel demand estimation by use of default trip distribution and mode-split models, default models with user-selected parameters, or user-specified trip distribution and mode-split models;
5. Assignment to optimal paths via minimum time, cost, or combined impedance; and
6. Inputs and outputs displayed in map (printer plot) or tabular form under direct user control.

STREAK essentially provides the same capabilities as standard transportation planning packages: network development, updating, pathfinding, travel demand projection, trip assignment, and so on. Assuming the same network and parameters, STREAK will produce results similar to those produced by other transportation planning packages. The major advantages of STREAK over standard batch processing are

1. Near-instantaneous response to many "what if" questions that a planner should ask in designing or evaluating a transportation system,
2. Minimal processing time, and
3. Quick and easy modification of network data, which permits many alternatives to be tested rapidly.

Although STREAK could be used for much of the standard work of travel estimation, its conversational capabilities are best suited to the following areas:

1. Initial development of transportation network alternatives, especially corridor evaluation, technology assessment, or alignment studies;
2. Comparative accessibility analyses of alternative networks to measure how well each system provides service to jobs, activity centers, population concentrations, transit dependents, or other relevant groups;
3. Location studies of public facilities such as transit stations, fire stations, schools, or activity centers to provide quantitative inputs in testing of alternative locations;
4. Immediate evaluation of proposals and suggestions in meetings of the public and planning committees.

STREAK operates in an interactive mode. The fundamental assumptions that underlie interactive planning are the following:

1. Overall network system performance can be evaluated on the basis of how a limited number of major or typical destinations are served—i.e., by evaluating an alternative network based on travel between all zones and, say, only the downtown area, the airport, a major shopping center, and a typical residential zone.
2. Locations of stations (or other facilities) can, in a first-cut approach, be analyzed on the basis of accessibility to population and employment sites—i.e., without regard to travel demand.

These two assumptions differ significantly from the standard practice of processing all zone-to-zone trips.

In the two cases above, influence areas, travel-time contours, service levels, patronage estimates, capacity evaluations, and link loadings can be provided in a matter of seconds by using a time-sharing environment and a limited number of representative destinations. An analyst would then keep modifying station locations, spacings, network connections, speeds, frequencies of service, and mode-choice parameters to improve the alternative. At the same time, he or she would be doing a sensitivity analysis of the alternative. The final plan developed in this fashion should be fully evaluated in the standard manner by processing all zones. This can be more economically produced off-line, but using the results of the STREAK analyses can make the batch runs far more cost-effective.

The interactive methodology, therefore, only complements and does not replace standard batch processing. Properly used, it should greatly increase the quality and scope of the comprehensive planning process and provide a variety of useful evaluation measures. It also lends itself to immediate answers at meetings and sessions, helping to channel discussions into profitable and concrete (as opposed to speculative) directions.

The planner communicates with STREAK in a brief but powerful command vocabulary. One subset of commands modifies the network components: nodes, links, and their attributes. Another subset is used to control the operation of STREAK pathfinding and other algorithms. A third subset modifies the models by imposing constraints, redefining parameters, or providing alternate sources of input. A final subset causes data to be displayed in report and printer-plot format.

## TECHNICAL DESCRIPTION OF STREAK

The STREAK interactive computer program contains several standard transportation planning functions within one logical shell. The program is amenable to sketch and detailed planning studies that involve composite-mode, abstract, or conventional transportation networks. STREAK is written in FORTRAN and is operational on the CDC NOS system.

### Planning Network

The input network is composed of links and nodes. The user selects a subset of nodes to serve as zone centroids (trip-end locations) or sources (roots of minimal impedance trees). Network encoding is simplified by a novel scheme that is capable of storing all connectivity information in (links + streets) cells.

Each network link or node may be tagged with four class codes. By using this feature, the program easily partitions the network (for example, by jurisdiction or by street type) for purposes of input, mode-split processing, or display of results. Other input data include coordinates, terminal capacities, impedances, demand data in various guises, scaling factors, program controls, and model parameters.

The first time a network is used, input is in the form of a card-image file. After structural changes have been made in the course of running the program, the required intermediate networks are saved in an equivalent binary format. The final modified network of one planning session then becomes the input to a later session.

### Tree Builder

The tree builder uses a variation of the Moore-Dijkstra minimal path algorithm (1), which makes the process suitable for allocating resources over networks subject to capacity constraints. The process builds on one impedance variable and accumulates three others, and in addition optionally applies link access-egress times, terminal times, behavioral factors, and transit waiting times. The expected wait time is computed from transit service data by means of the Dial-Loubal non-line-specific waiting-time algorithm originally developed for inclusion in the Urban Transportation Planning System of the Urban Mass Transportation Administration (UMTA) (2).

Trees are built simultaneously or iteratively. Simultaneous building from several root (sink) points produces nonoverlapping drainage areas. Iterative building processes root points in succession, reaching all network links with each tree. This logic is mandatory whenever trip distribution is performed.

## Demand Acquisition

Trip-related data may be input either as trip tables or as trip ends. When several purposes represented in the trip-table input are to be assigned, the user selects the order in which the tables are accessed and the trips are assigned.

Alternatively, trip productions and attractions are accepted by a procedure that proportions P's to A's or A's to P's as the user may desire according to the basic distribution formula

$$T_{ij} = (A_j f_{ij} P_i) \Big/ \left( \sum_k f_{kj} P_k \right) \tag{1}$$

where

$T_{ij}$ = trips from a network point i to point j,
$A_j$ = trip attractions associated with network point j,
$f_{ij}$ = friction factor computed as a function of $t_{ij}$, where $t_{ij}$ is the path impedance from point j to point i, and
$P_i$ = trip productions associated with network point i.

STREAK accepts friction factors defined by the user (fixed index interval) or assists the user by equating $f_{ij}$ to the expression $t_{ij}^{-a}$, the exponent of which (a) may be supplied by the user or by the model as a default value.

## Assignment

The trip-loading logic is nondestructive, backwards, all-or-nothing assignment. In a mode-split context, the planner can ask the program to assign the trips from either mode (or neither). The trips can be loaded conventionally onto nodes identified as zone centroids or directly onto every link within a zone instead of onto a zone centroid. For this latter option, the link loading is done in proportion to length of link in the zone divided by the total zonal link lengths. Specifically, the trip volume on a link is given by the following formula:

$$v_i = (V_j w_i d_i) \Big/ \left( \sum_i w_i d_i \right) \tag{2}$$

where

$v_i$ = trips on link i from zone j,
$V_j$ = trips in zone j,
$w_i$ = weight assigned to link i (1.0 unless the user supplies other values), and
$d_i$ = length of portion of link i contained in zone j.

## Mode Split

The program can separately save input and result data for two modes and exercise a mode-split process on the calculated total demand based on those data and external factors. STREAK has four built-in mode-split models: multivariate logit, simple logit, elementary step function, and a simple service ratio model. Specifications of model parameters are under the user's control in each case.

## Program Commands

By typing just one command on a computer terminal, the analyst sets in motion a chain of planning actions. For example, the command statement

R2ML

will build trees outward from previously designated nodes, distribute demand and perform mode-split calculations, and load demand onto the network.

The program has been kept flexible by triggering much of the operational logic through command statement options. Thus, in the example above, demand assignment would not occur unless the L were keyed in. The command cited, R, has more than 10 options, each with suboptions. Options are usually prespecified by the user, but the program strikes a balance by eliciting suboptions as part of the interactive dialogue.

Program commands fall into five categories. The first controls changes to the network, adding new links, deactivating old ones, creating new zones, and modifying their attributes' values. The second command group affects program parameters, such as the maximum permissible path length, or the index to the vector that contains the control impedances for tree building. The third classification controls planning model execution and is necessarily the most complex in contents and results. The fourth category implements the optional report capability. Input and generated data become displayable in tabular and print-plotted formats. The last set of instructions lets the user reinitialize the current terminal session or end it.

The experienced user has little need for computer prompting. STREAK permits this analyst to specify a series of commands that takes shortcuts through the logic (and dialogue) maze. However, other actions, such as adding new links, always require scrutiny because limits exist on the checks the program can perform for reasonableness. A user's manual has been prepared that describes network coding procedures, program logic, and program commands (3).

## Cost of Operation

Different network characteristics and control parameters made it difficult to provide the performance curves for the program. However, in the course of many runs in the Los Angeles South Bay study (4), it was noted (for the CDC Cyber-75 computer) that STREAK could build two 1000-link trees and distribute and assign trips over the total network in less than 3.5 s while testing a large set of constraints. Most interactive commands are processed in fractions of a second although normal time-sharing delay prolongs the response time. In a matter of a few seconds, then, STREAK provides the analyst with influence areas, travel-time contours, service levels, patronage estimates, capacity evaluations, and traffic volumes.

## USE OF STREAK

The initial input of data and network description is normally done in the conventional manner with cards or card images. Existing data bases and networks can generally be used as a starting point. A zone system must be overlaid on the network, link lengths must be measured node to node (and by segment within a zone for the optional assignment process), a destination node must be selected in each zone, and zonal trip origins must be determined. If one wishes to use the plot capabilities of STREAK, coordinates must be determined for each node and input to the data base.

After the initial input of data, all data corrections and updates and network modifications are made directly from the terminal keyboard. In addition, it is possible to run partial assignments of one zone or a related set of zones, to examine individual trees, or to perform various other quick tests. By operating in this manner directly from the terminal, the time required to correct the data files, clean up the networks, and calibrate the models is significantly reduced. Calibration of the distribution model generally involves the proper selection of the friction factor exponent to achieve the proper average trip length and total unit distance of vehicle travel for the study area. Calibration of a mode-split model involves the proper selection of time and cost coefficients for each mode.

After development of a clean data base and network and calibration of the models, the planning analysis can begin. In exercising the program, the user can select the entire data base (all zones) or a subset of data or zones of interest. Use of subsets of data and zones permits faster processing and quicker terminal response times and facilitates examination of a large number of alternatives. Examples of planning considerations that can effectively use data subsets are

1. Travel-time contours for site and network alternatives;
2. Population or employment within a given travel time (e.g., 5, 10, or 15 min) of activity centers or other key locations via alternative transportation networks;
3. Maximum time or distance needed to reach a given percentage (e.g., 50 percent of population or employment from various locations for each alternative);
4. Plots of fastest or shortest paths between points in a network; and
5. Determination of catchment areas for candidate sites—transit stations, schools, ambulance locations, and fire stations—and identification of the parts of a region best served by each site so that travel time is minimized.

In comparison with the processing of all zone-to-zone trips for such planning cases, in a time-sharing environment all the needed influence areas, travel-time contours, service levels, patronage estimates, capacity evaluations, and link loadings can be provided in a matter of seconds. An analyst would then keep modifying station locations or spacings, network connections, speeds, frequencies of service, and mode-choice parameters to find improved plan alternatives. By similar modifications, the analyst could perform a sensitivity analysis around any major plan alternative.

In studies that deal with determinations of modal choice or major facility analysis, all zones should be processed. This can also be done on-line, but for large networks it is more economical to print the results off-line on a batch computer. Examples of studies that use zonal subsets and those that use all zones are discussed in the following section.

## CURRENT EXPERIENCE WITH STREAK

Experience is growing in the use of STREAK as a planning tool. De Leuw, Cather and Company has used the models in several areas including Boulder, Colorado; Florida's east coast; Los Angeles County; Sacramento, California; Burke Mountain, British Columbia (a new town planned for the vicinity of Vancouver); Parramatta, Australia (a suburb of Sydney); and Melbourne, Australia.

### Burke Mountain, British Columbia

The Burke Mountain study illustrates the planning-speed capabilities of STREAK. The computer analysis for this study was accomplished in less than a week. In this study, transportation relations of the proposed new town to the Vancouver region were described in terms of travel-time contours, accessibility measures, and additional highway volumes projected to result from the new development.

### Boulder, Colorado

The Boulder study provides a good example of the use of STREAK in its primary role as a sketch-planning tool. Four transit systems were defined and analyzed for the study area: baseline bus, advanced bus, light rail with background bus, and elevated guideway with background bus. The following measures or maps were produced for each alternative:

1. Travel-time (isochron) contours, by highway and transit, for several activity centers and residential areas;
2. Population (or employment) totals within 10, 20, and 30 min of selected locations for various network configurations;
3. Travel time needed to reach 50 percent of the population (or employment) for different locations;
4. Maps of transit/automobile travel-time ratios for several locations;
5. Expected transit waiting times for various destinations from important origin locations; and
6. Forecast patronage of the transit system.

Figures 1 and 2 are taken from the Boulder report (5). The link values were produced on-line, a map was overlaid, and the caption was added to produce the final figures. Note in Figure 2 that the transit/automobile time ratios are high for short trips close to the node being examined. This is a result of the fact that walking plus waiting times for short trips are much larger than automobile times whereas for longer trips walking and waiting times become less significant and the travel-time ratio approaches the inverse operating-speed ratio of the two modes.

In the Boulder study, interactive capabilities of the model were used to examine a multitude of different configurations in the design of the three proposed transit alternatives. The interactive features were also used at many meetings between the consultant and community staff personnel, elected representatives, and citizen groups.

### South Bay, Los Angeles County

The South Bay STREAK analysis in Los Angeles County (4) focused on short- and long-term highway proposals. This study represented a departure from previous STREAK applications in several important aspects:

1. The study area was contained in a large metro-

**Figure 1. Use of STREAK in the Boulder transit study: transit access times for node 48 in tenths of a minute.**

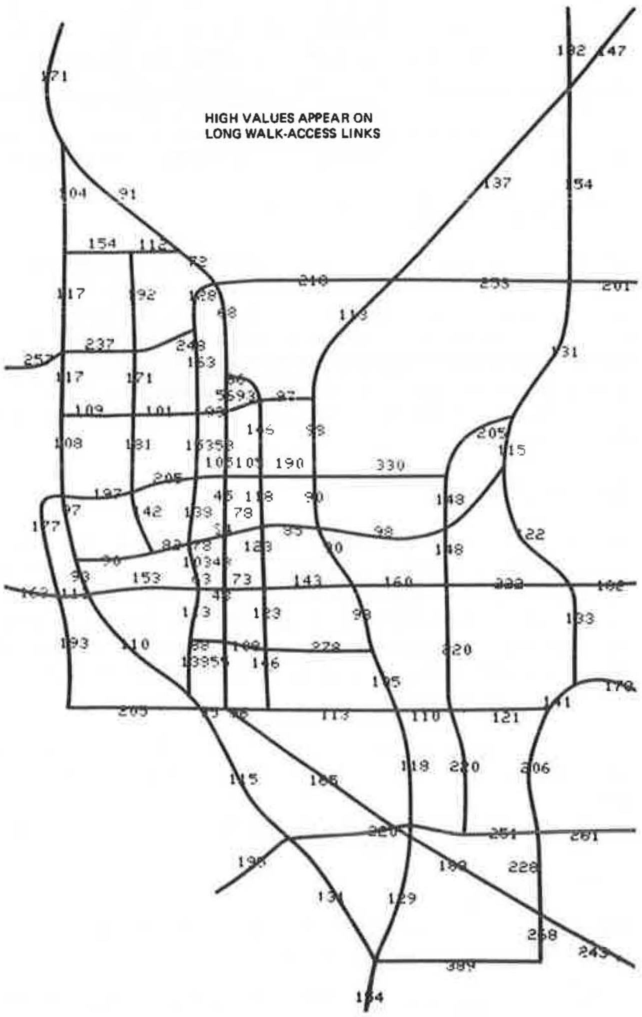



**Figure 2. Ratio of transit/automobile times (x 100) for node 48 in the Boulder transit study.**


politan area and consequently experienced a high percentage of external trips.

2. The highway network was much larger than had previously been analyzed with the model.
3. A high degree of detail for the highway assignment was desired for the purpose of analyzing highway improvements as well as network modifications.

These problems and model demands were successfully met by a combination of model modifications (to handle large networks and more zones) and an innovative methodology for isolating the study area from the rest of the metropolitan region (6). To cope with the highway complexities of the South Bay subregion, the original program was modified to permit the input of detailed networks that contained as many as 1500 links and 150 zones.

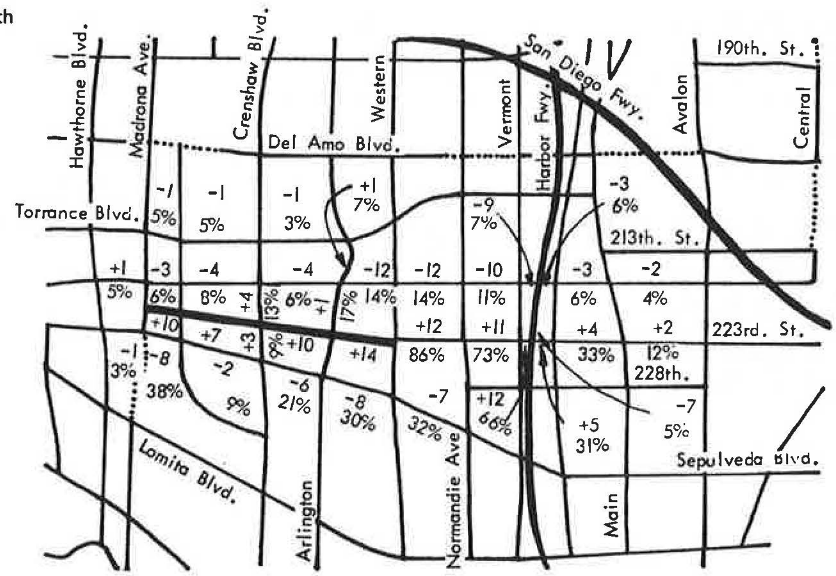
To isolate the study area from the rest of the region, the following actions were taken:

1. Future cordon crossings were obtained from a subregional traffic assignment previously prepared by the state of California. These volumes were modified in accordance with more recent demographic and travel forecasts.
2. A buffer area was established between the study area boundary and the cordon line for network analysis. The purpose of the buffer area is to provide a transition zone from the more detailed STREAK network simulation in the study area to the coarser subregional representation at the cordon line.
3. Penalties were developed for external trip attractors to represent the composite times and costs of travel outside the study area based on average internal-external trip length.
4. Different procedures were developed for assigning internal trips, internal-to-external trips, and external-to-external trips. The three processes were used in a single assignment by accumulating the link volumes for the three trip types.

Since detailed and accurate highway network volumes were desired, the complete set of zones was assigned during each of the model runs for the highway analysis. However, the interactive capabilities of the model were extensively used in data preparation, network development, model calibration, development of the study area isolation methodology, and development and evaluation of the highway alternatives. In fact, in every step except the final output runs, the interactive capabilities of the model—such as partial assignment, path building, data correction, and network change—were extensively used.

Outputs for the South Bay analysis included simulation of current traffic volumes, forecast 1995 volumes on the existing network, and forecast traffic volumes under various highway improvement alternatives. For ease of comparison, maps were prepared for each of the simu-

Figure 3. Use of STREAK in the Los Angeles South Bay study: traffic impacts of extending 223rd Street to Madrona Avenue.




lated highway alternatives to show the projected volume changes in absolute numbers and percentages. The maps for a given alternative show every link that is projected to experience a change in volume of ≥1000 vehicles/d. Hence, the maps show projected use of the proposed projects, changes in use on nearby facilities, and the total extent of the network impact of the proposal. Figure 3 shows an example of one of the maps from the South Bay study (this figure is redrawn since the original report used a colored basemap). Approximately 100 highway projects were analyzed during this study. Many of these projects were evaluated by means of screenline and volume-capacity analysis, whereas others were subjected to the more extensive analysis shown in the figure. Since the study area was large, an on-line printing of link values would have required five parallel strips of terminal printout paper. Hence, in this study the link values were printed in tabular form and then plotted on maps by hand.

## Parramatta, Australia

The study in Parramatta, Australia, involved the analysis of three public transport corridors, each of which has three possible modes and several possible route alignments. STREAK was used to analyze flows toward the regional center—Parramatta. The STREAK network was built to represent all possible combinations of route, mode, station location, and speed. Alternative systems were rapidly analyzed by means of the class code manipulation ability of STREAK. Patronage, mode split, diversion from existing rail lines, average travel times and trip lengths, and trip distribution within the subregion were some of the key characteristics examined by using STREAK. The corridor analysis team used STREAK to evaluate alternative station locations and the effect on patronage of various operational and route constraints. The STREAK exercises were performed over a period of 1 month, and a month was needed for network and data preparation.

## Melbourne, Australia

In Melbourne, Australia, the Ministry of Transport has obtained STREAK sketch-planning capabilities. The model has already been used on their strategic zone system, and future applications include investigation of the Melbourne underground, pedestrian movement in the city, and road staging studies.

Another application in Melbourne involved analyzing long-term impacts and construction staging for a complete outer ring freeway system. STREAK was used in conjunction with the De Leuw, Cather transport and land-use interaction model TRANSTEP to evaluate the effects on accessibility, travel time, and road loadings for different construction sequences of the outer ring. Eight selected areas were analyzed in detail, and STREAK provided an indication of how the benefits from each section of the ring were distributed throughout the total metropolitan area. The road network was coded by members of the Melbourne Joint Road Planning Group, who also participated in the STREAK on-line analysis sessions. The total STREAK exercise was completed within 2 months; the analysis was completed in 2 d.

## SUMMARY

An interactive computer planning program called STREAK has been used and debugged through application to several studies. When used in its primary role of sketch planning, the planning package has demonstrated advantages of cost and time savings in comparison with conventional batch-loaded planning programs. STREAK also provides ease of data-file correction and network modification, an ability to quickly examine many alternatives, and an ability to respond interactively to a wide variety of planning questions by means of a portable terminal. For the planner who wishes to examine several alternatives on a limited budget or in a short time, STREAK may be a more appropriate tool than the conventional batch-loaded computer program.

STREAK is not a planner's panacea. Although the model currently can handle networks as large as 1500 links, STREAK is not recommended for applications of large highway networks with capacity restraint and iterative assignments. Even in such studies, the model would still provide some advantages in data and network preparation, but most of its other special features, including much of the savings in cost and time, would be rendered useless. However, in sketch planning and in many other planning applications such as those discussed in this paper, STREAK offers several advantages.

In addition to economic and time considerations, many planners will welcome the ability to work directly with the data, the networks, and the alternatives without the intermediary steps of coding, keypunching, computer running, and printout. Instead of taking problems to the computer, perhaps planners should try bringing the computer to their problems.

REFERENCES

1. E. W. Dijkstra. A Note on Two Problems in Connection With Graphs. Numerische Mathematik, Vol. 1, 1959, pp. 269-271.
2. R. B. Dial, P. S. Loubal, E. Luther, and G. J. H. Brown. A Procedure for Long Range Transportation (Sketch) Planning. PRC Systems Sciences Co., McLean, VA, 1973.
3. P. S. Loubal and B. Mendes. STREAK Interactive Transportation Modelling Program User's Guide. De Leuw, Cather and Co., San Francisco, 1975.
4. South Bay Corridor Study—Phase II. De Leuw, Cather and Co., San Francisco, June 1977.
5. Boulder Transit Study—Analysis of Alternatives for the Boulder Valley. Denver Regional Transportation District, June 1977.
6. South Bay Corridor Study—Phase II. De Leuw, Cather and Co., San Francisco, Appendix B, Vol. 2, June 1977.

*Publication of this paper sponsored by Committee on Computer Graphics and Interactive Graphics.*

# Creation of Urban Transportation Network Models From DIME Files

Jerome M. Lutin, Princeton University

**The development of a set of computer programs that use U.S. Bureau of the Census geographic data to create urban transportation network models is reported. The process uses the Dual Independent Map Encoding (DIME) files created by the U.S. Bureau of the Census for each of the major standard metropolitan statistical areas of the United States. Manual coding of networks is reduced to a minimum, and highly detailed network models can be created. Functioning and use of the program are documented, and the way in which computer graphic displays are readily generated from DIME file data is demonstrated. Examples of computer graphic output from the program are presented.**

Transportation planners have long been faced with the problem of analyzing vast amounts of data. Although the use of computers has made data handling easier and advanced mathematical techniques have provided powerful tools to test hypotheses and find underlying structures, the planner still has difficulty in using all the numerical output received through the process. As a result, the planner, working under time constraints and monetary budgets, has less opportunity to analyze alternative plans.

This paper documents the development of a methodology to create models of urban transportation networks from available U.S. census data. The methodology revolves around the use of new computer software to minimize network data collection. In addition, it provides the basis for graphic displays of transportation system planning information.

Much of the transportation planner's task involves analysis of transportation systems such as streets, bus routes, and rail lines. Dealing with these systems numerically often requires creating an abstraction of the system by using graph theory in which street or line segments are represented as links or "edges" and intersections or stations are represented by nodes. Numerical values that represent speed, capacity, distance, and other system characteristics are assigned to each link and node, and this results in a network simulation that can be analyzed by using computerized mathematical models.

## MODELING TRANSPORTATION NETWORKS

Modeling a transportation network involves three stages of work. First, one must specify the network to be modeled. This involves making a number of important decisions including (a) the zonal structure or geographic subdivisions of the area being modeled, (b) the scale or level of detail of the model, and (c) the number of elements—links and nodes—to be included in the model. Second, one must prepare the data for machine processing. This includes (a) specifying link characteristics, such as length (distance), time, and capacity; (b) coding the network, which includes numbering nodes and determining nodal x-y coordinates; and (c) preparing the input data cards or records. At this point, the network model is ready for machine processing.

Machine processing of the network model includes several tasks. First, one must create a computer file, or historical record, of the network. Next, zone-to-zone routes through the system, or minimum paths, must be calculated. These files of sequential links and nodes, known as trees or vines, must be stored and verified. Often, manual searching through tables of numerical output is required to see if the computer-generated routes are reasonable and error-free. From these minimum path files, individual link travel times and costs are summed to derive zone-to-zone travel times and costs. Finally, zone-to-zone travel demand is matched to the network to determine the amount of traffic flowing over each route. Other types of network analysis are possible as well, but the procedure outlined above is the one most commonly used.

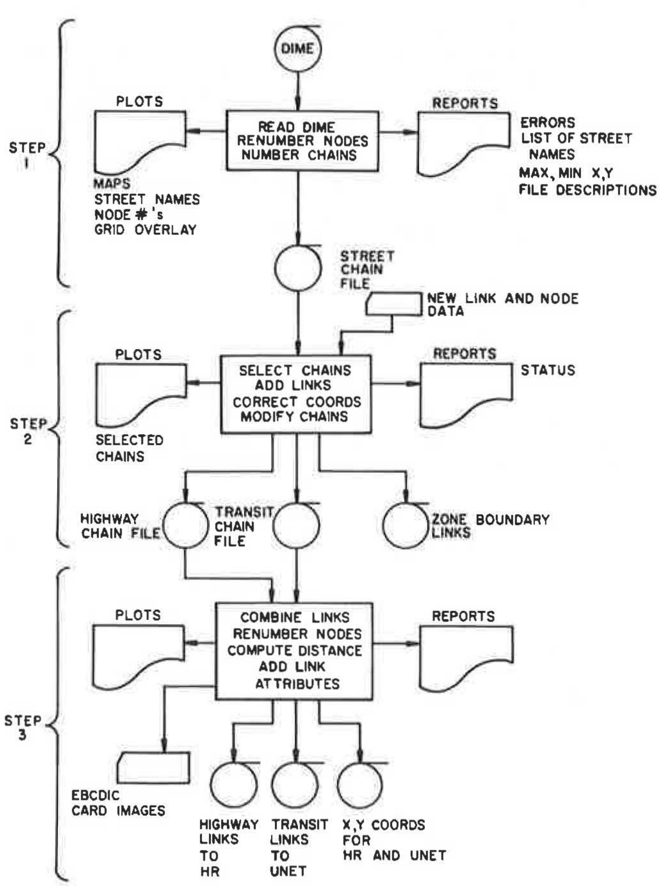
Two common difficulties have been found in creating and analyzing network models. First, a great deal of information must be collected to supply an adequate representation of a network. Urban areas typically have thousands of streets and hundreds of kilometers of bus lines. A computerized network model may have several thousand links and nodes that require thousands of data items punched on computer cards. This process, known as coding, becomes a time-consuming task that must be repeated for each network to be analyzed. Second, for each network, the analyst must examine vast amounts of computer output material that contains volumes of numbers. This complicates the task of evaluating the results of each network simulation.

Planners recognize that graphic tools, such as charts and maps, are often the best means of presenting numerical data for transit systems. For example, one may wish to use a graphic representation of flow volumes on a street network in which the width of the band indicates the number of trips that use each link of the system. This graphic presentation conveys far more information than the same data presented in tabular form.

Federally sponsored libraries of computer programs have been developed for the analysis of urban transportation systems and are currently widely used throughout the country. Some of these programs have been written or modified to produce graphic output of data. In particular, the Urban Transportation Planning System (UTPS) of the Urban Mass Transportation Administration (UMTA) and the computer programs of the Federal Highway Administration (FHWA) have been modified to produce graphic displays of networks on plotting devices attached to computers (and line printers as well).

However, the production of graphic output for transportation networks requires even more data collection. It is necessary to establish a coordinate system or grid to locate each point to be displayed. Two numerical coordinate values, one for the horizontal or x-direction and one for the vertical or y-direction, must be calculated for each point. These values must then be punched on computer cards or added to the data bank representing

Figure 1. Flow chart of network process using DIME files.




the coded links and nodes of the network. Typically, this process has been done largely by hand. The development of electronic digitizers now allows the analyst to record the network by tracing lines from a map onto a sensitized tablet that electronically records numerical x-y coordinates for each point. Even with digitizing equipment, however, the process of network creation is a laborious task.

## DUAL INDEPENDENT MAP ENCODING (DIME) FILES

For the 1970 Census, the U.S. Bureau of the Census developed the Geographic Base File (GBF) and a computer program—Dual Independent Map Encoding (DIME)—that contains x-y coordinate files for most streets and many natural boundaries. The resulting areas are roughly equivalent to city blocks. DIME files are currently available for most of the more than 200 major U.S. cities that the bureau describes as standard metropolitan statistical areas (SMSAs), and the system is being expanded to include all SMSAs. If these files were properly linked to transportation planning computer programs, the result would be near elimination of the manual effort to code and describe even the most complex network. It would provide the desired graphic display information as well.

The Princeton University Transportation Program has used the DIME files to produce x-y coordinates for each bus route in Trenton, New Jersey. The files were also used in conjunction with the Bureau of the Census ADMATCH program to allocate employment records to traffic zones.

In developing the DIME files for urban areas, the Bureau of the Census adopted a format for data storage that best suited its purpose—primarily, the location and coding of street addresses for automated processing of data and census data collection by mail. Thus, the DIME file records are grouped mainly by census block and are not directly usable for describing a transportation network. Each record in the DIME file is 300 characters long and represents a block face or geographic boundary segment. Since most streets and natural boundaries are located in the DIME files, it is possible to extract required data and reconstitute a street network that is appropriate for most urban transportation planning. The main problem involves the manipulation of DIME records into "chains" of links that represent entire streets.

## SOFTWARE DEVELOPMENT

This project has developed a method whereby census data can be more effectively used by planners within the framework of the Urban Transportation Planning System (UTPS). A set of three computer programs has been developed that can read the DIME files and, on user request, output a selected file of nodes and node coordinates in a variety of formats (such as tape files or punched cards). It will be possible to use these, in turn, as inputs to the UTPS network-building programs UNET and HR. The end product will be a computer package that provides the interface between census data and the UTPS package.

Through use of these programs, it is expected that the effort required to code data for many transportation planning projects can be greatly reduced. It will enable transportation planners to model and analyze more alternative transportation networks within a given time and cost budget. Basically, the provision of a software link between the DIME files and transportation network simulation models has three objectives: (a) to provide

Figure 2. Bureau of the Census block map of Trenton urbanized area.

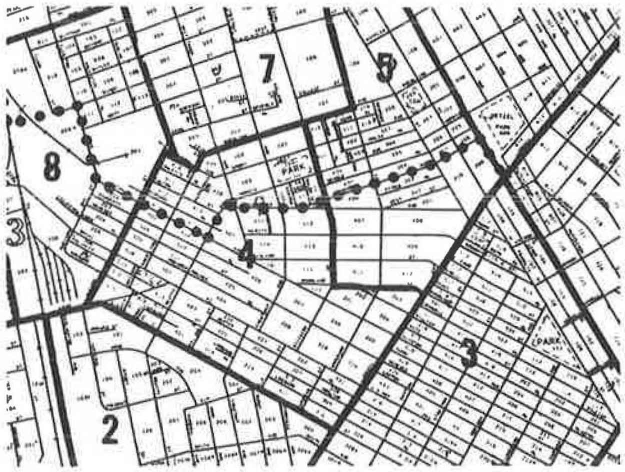


Figure 3. Computer-drawn map of Trenton DIME file.

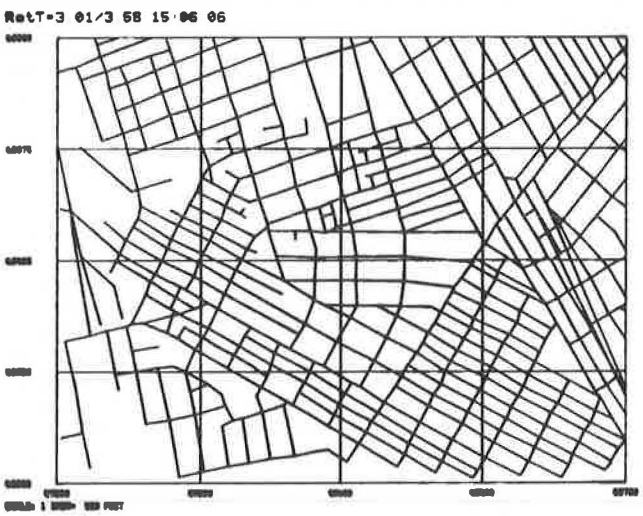
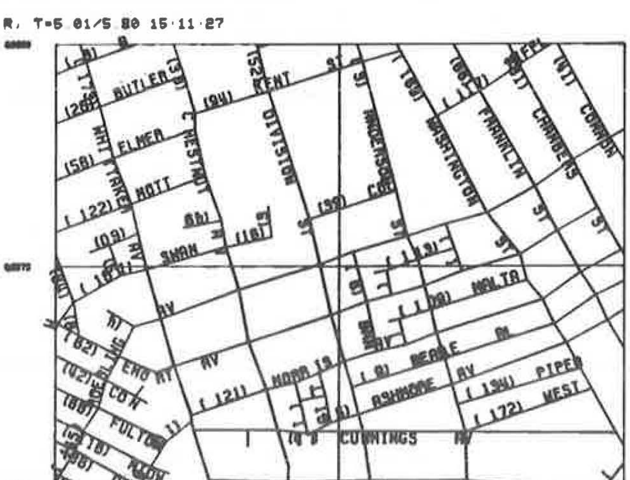


Figure 4. Enlargement of Trenton DIME file map.




a data base suitable for graphic display of analysis results, (b) to reduce data-handling requirements for transportation planning, and (c) to provide the capability for developing very detailed network models for specialized projects.

## PROGRAM OPERATION

Figure 1 shows a general flow chart of the operating sequence of programs currently under development. Functioning of the programs is described below.

### Step 1—DIME Editing

The first step involves reading the raw DIME files. Here, the user is provided with measures that indicate the completeness of the DIME files. Errors (tape errors, missing data, and out-of-range values) are given, counts of records are taken, street names and boundary names are listed, and maximum dimensions of the network are determined.

The program outputs printed reports plus, if desired, plots of the network, with street names, node numbers, and a planning grid overlay, for determining new coordinates. At the same time, the program creates an output file of links chained together according to street or boundary name. Each chain is assigned a number for ease of reference. All nodes in the DIME file (each record contains both a "from" node and a "to" node and four-digit node numbers by tract) are assigned unique, new node numbers. Each DIME file is grouped according to map number to correspond to the Metropolitan Mapping Series (MMS) of the Bureau of the Census. Thus, the user can select various segments of the urban area for inclusion in the network.

### Step 2—Editing of Chain Files

The second step in the process involves reading the chain files produced by step 1. In this step, the user can add new links to the chain file, modify and restructure street chains, and correct node coordinates. Step 2 produces several output files of links chained together at the option of the user. For example, one can obtain a zone boundary file, a transit network file, and a highway network file. This stage also produces reports of changes and plots of the networks.

### Step 3—Creation of Link Files

This stage produces output files of links in forms suitable for analysis by network-building programs. The program reads selected chain files, combines links (by deleting unused nodes), and renumbers nodes to correspond with conventions used in network-building programs. Link lengths or distances are calculated, and various attributes, such as type of facility and speed limit, can be added to each record. In many cases, default values for link characteristics are supplied to simplify network coding. Files of x-y coordinates are produced so that the output link files can be graphically displayed. The user-selected output files produced at this stage are correctly formatted for use by UMTA network processing programs HR and UNET for highway and transit analysis respectively.

## PROGRAM OUTPUT

A number of output reports are produced at each stage of processing. For example, in stage 1, a comprehensive street chain directory is produced. For each chain, the report lists chain number, name, number of links, number of nodes, and the node number and x-y coordinates for each node in sequence. Other reports, including reports of errors found and a dictionary of corresponding new and old node numbers, are also produced.

However, the main value of the program lies in its ability to produce graphic output and to produce networks capable of graphic display. Figure 2 (1) shows a segment of the Bureau of the Census MMS map of a portion of Trenton, New Jersey. One can see that each street is included as well as block and tract identification, some major open areas (parks and cemeteries), and major rail lines. This map is the basis for the geocoded DIME file for the Trenton area. Figure 3 shows a computer-drawn map of the same approximate area that uses chain files extracted from the Trenton DIME file. A coordinate grid (of the same scale as that used in the DIME file) is overlaid on the map. In Figure 4, the scale has been enlarged four times, and street chain numbers and street names are both included.

These maps were drawn on the face of a Tektronix 4013 cathode ray tube (CRT) terminal by using a preliminary version of the software described above. Standard CALCOMP plotting routines were used for legends and alphanumeric characters. These figures show the great level of detail available in the DIME files.

## CONCLUSIONS

Although some work remains to be done in testing and improving the programs, the work reported here has attained two major objectives:

1. The feasibility of using the DIME geocoding system as a basis for transportation network modeling has been established.
2. A methodology for processing the DIME data files has been developed.

Increasingly, transportation planners are being asked to provide more detailed and more specific analysis of transportation systems in urban areas. This requires detailed network models. Use of the software described here and the available Bureau of the Census DIME files will enable the planner to create detailed network models with a minimum of manual intervention. Although many errors and gaps have been found in DIME files, the Bureau of the Census is striving to obtain complete and correct DIME files for all SMSAs. The software described here provides for human intervention, primarily to detect and correct errors. Although the process cannot be completely automated, the amount of effort required to create a network model will diminish as the quality of DIME file data improves. In the future, it is expected that DIME-based networks, which can be readily matched to other demographic census data, will be in widespread use among transportation planning agencies.

## ACKNOWLEDGMENTS

I would like to express my gratitude to Granville Paules, Linda Moore, and Robert B. Dial of the Urban Mass Transportation Administration for their support of this work. I would also like to thank Warren B. Powell, William D. Collins, and Brian P. VonHersen for their work in programming the software described above. This research was performed under a university research and training grant and a research and development grant from the Urban Mass Transportation Administration.

## REFERENCES

1. Block Statistics Trenton, NJ-PA Urbanized Area. U.S. Bureau of the Census, Metropolitan Map Series, U.S. Government Printing Office, HC(3)-151, Map Sheet P-21, 1971.
2. M. A. Meyer and J. Silver. Present Status and Future Prospects of the Census Bureau's GBF/DIME CUE Program. U.S. Bureau of the Census, presented at Conference on Geographic Base File Systems—A Forward Look, Boston, April 16-17, 1974.

*Publication of this paper sponsored by Committee on Computer Graphics and Interactive Graphics.*

# Computer Geocoding of Travel Surveys

Priscilla Sawyer Cornelio and Joseph M. Manning, Urban Transportation Systems Associates, Inc., Newton, Massachusetts

**Computer geocoding of travel survey data, which involves use of a geographic base file and a series of user-oriented computer programs, is a workable and preferable alternative to manual geocoding, which is tedious and time-consuming. The basis for the computer geocoding system is the Dual Independent Map Encoding/Geographic Base File (DIME/GBF) developed by the U.S. Bureau of the Census. The DIME/GBF, which exists for all standard metropolitan statistical areas, contains detailed information on street segments, including street name, direction, range of house numbers, and census tract and block. Two programs developed by the U.S. Bureau of the Census, ZIPSTAN and UNIMATCH, are used to perform the actual geocoding. ZIPSTAN is a preprocessor program that arranges the addresses before being linked by address to the DIME/GBF by use of the UNIMATCH program. Geocoding on-board and travel surveys in the Boston area indicate that 70 to 80 percent of all addresses can be geocoded to a detailed zone level at a processing cost of $2.12/1000 addresses. Addresses not geocoded automatically are generally incomplete or contain invalid information. The basis for this system, the methods and procedures used in the Boston area, the results of the matching operation, and the costs involved in the effort are presented and discussed.**

Planners and engineers frequently use travel and attitude surveys to determine existing travel patterns and latent demand for new or improved transportation facilities.

These surveys are an important tool to officials who use the data generated to determine the feasibility of improvements.

Trip tables are necessary to analyze origin and destination survey data effectively and are generally constructed to indicate traffic or person movements between analysis zones. These zones can be of any size, ranging in area from a block to a town. To build a trip table, the origins and destinations indicated on each survey form must be assigned to analysis zones, a process known as geocoding.

In many surveys, geocoding is performed manually. This involves locating the address either on a map or in an address coding guide and coding the tract or zone on the survey form for data transcription. The complexity of manual geocoding is proportional to the level of analysis zone used. If a town code is sufficient, geocoding is a simple operation. However, if there are different analysis zones for odd and even house numbers on each street and each block or group of blocks represents a zone, the time involved to code each survey can be substantial.

Obviously, when a large number of addresses must be geocoded, the manual process can be expensive because of the staff time required to code and quality-check the forms. In addition, it is inevitable that results may be inconsistent when the judgment of a coder is necessary—for example, in determining if a building is on the odd- or even-numbered side of the street. A substantial amount of human error also can be introduced during analysis zone coding and data transcription phases when digits can be transposed or incorrectly transcribed.

Computer geocoding is a workable and preferable alternative to manual geocoding. The process uses a geographic base file and a series of user-oriented computer programs. It can result in a better product for less money because the computer can perform sophisticated functions rapidly, economically, and consistently.

## DEFINITIONS

Before describing the automatic geocoding process, it is necessary to understand the three major inputs to the process. These are

1. The Dual Independent Map Encoding/Geographic Base File (DIME/GBF),
2. The ZIP Code Standardizer (ZIPSTAN) program, and
3. The Universal Matching (UNIMATCH) program.

### DIME/GBF

The DIME/GBF, developed by the U.S. Bureau of the Census, forms the basis of the reference file used for geocoding. The DIME/GBF contains records that describe street segments, legal boundaries, and topographic features including rivers, railroad tracks, and canals (3, 4, 6). Each segment contains a variety of information. Included for street addresses are the street name, street prefix and suffix, highest and lowest house number ranges for both sides of the street, city code, county code, state code, ZIP code, and census tract and block designations. In addition, the file contains X-Y coordinates and to and from node numbers that can be used for mapping.

Figure 1 shows a sample of the map used to construct a DIME/GBF. Each street segment between node points has its own data record, which contains the information given in Table 1 (4).

The U.S. Bureau of the Census, with substantial assistance from local areas, has constructed a DIME/GBF for the urbanized areas of each of the 233 U.S. standard metropolitan statistical areas (SMSAs). The file for any SMSA may be purchased from the bureau for about $80. However, the files for some cities may be more complete than those for others. The bureau encourages planning agencies to correct and update their files to eliminate errors such as incorrect ZIP codes or missing streets. Some DIME/GBFs have been carefully edited, corrected, and updated whereas others have not.

### ZIPSTAN

The ZIPSTAN program is a preprocessor that arranges the addresses before being linked by address to the DIME/GBF by using the UNIMATCH program. ZIPSTAN standardizes addresses by ensuring that house numbers and city names are in specific locations on the data records. The program uses equivalency tables that contain possible abbreviations and common misspellings of street prefixes, suffixes, names, and street types and provide the appropriate conversion of each. Because the DIME/GBF uses numeric city, county, and state codes, an alphanumeric table of equivalents is required input for ZIPSTAN. For example, a city table converts the city of Boston to a county-city code of 017005. ZIPSTAN appends to each address record a 100-character matchkey that includes the standardized address (2).

### UNIMATCH

The UNIMATCH program is used to match the addresses to be geocoded with the DIME/GBF and to attach zone information to the data record (1). The UNIMATCH program can be set up to attempt an exact match of house number, street name, and city code. If an exact match is not possible, the program applies weights and penalties set by the programmer to misspelled street names and blank or nonmatching street types, directions, and house numbers. Thus, data records are linked with reference file entries that have a high probability of being the same but have some missing elements in their address constructions.

To illustrate how the UNIMATCH program operates, if the address to be geocoded is missing a street-type suffix (for example, 100 Summer ___), and the reference file has a Summer Street with a house-number range of 1 to 50 and a Summer Road with a house-number range of 75 to 300, the program can be set to select the most probable reference file entry, 100 Summer Road. The programmer also has the option of instructing the computer to assume Street for all addresses with a blank suffix. Thus, if the reference file contained entries for 1 to 200 Summer Street and 75 to 300 Summer Road, the program would select the Summer Street entry.

The UNIMATCH and ZIPSTAN programs have certain computer requirements. They are written in IBM System/360 assembler language and are designed for computers that use the IBM System 370/360 operating system. The programs will operate under the following configurations: (a) primary control program, (b) multiprogramming fixed task, (c) multiprogramming variable task, and (d) VS1 or VS2, the virtual systems.

## GEOCODING PROCESS

### Initial Steps

To take advantage of automatic geocoding, certain data elements are necessary. The origin and destination addresses must contain a house number or street name including prefix and type. If more than one city or state

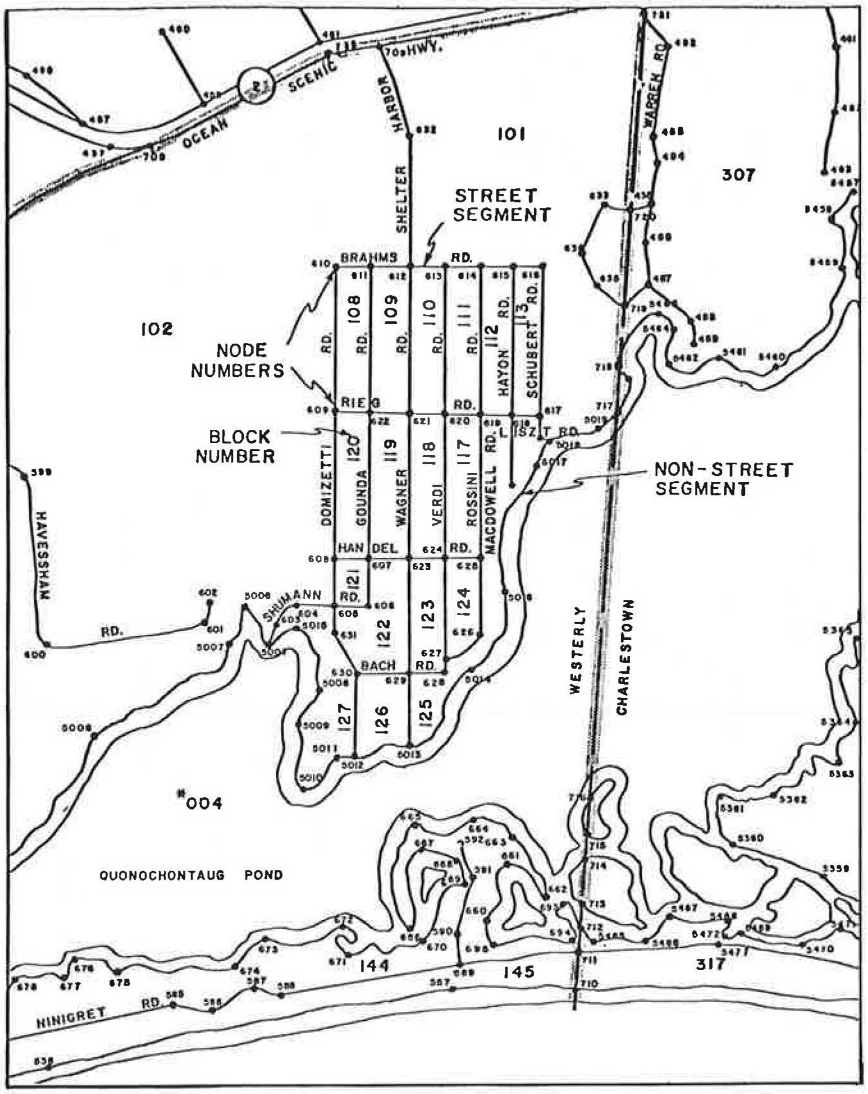



**Figure 1.** DIME/GBF map.

**Table 1.** Geographic elements contained in DIME/GBF.

| Item | Characters |
|---|---|
| Street prefix direction | 1-2 |
| Street or nonstreet feature name | 3-22 |
| Street type | 23-26 |
| Street suffix direction | 27-28 |
| Nonstreet feature code | 29 |
| 1970 enumeration district left (1970 nonmail census areas only) | 30-34 |
| Blank (census use only) | 35-40 |
| 1970 enumeration district right (1970 nonmail census areas only) | 41-45 |
| From map (basic number) | 46-48 |
| From map (suffix) | 49-50 |
| To map (basic number) | 51-53 |
| To map (suffix) | 54-55 |
| Coding limit flag | 56 |
| Left low address | 57-62 |
| Left high address | 63-68 |
| Right low address | 69-74 |
| Right high address | 75-80 |
| File code | 81-84 |
| Record number | 85-89 |
| Check digit | 91 |
| Census tract left (basic) | 92-95 |
| Census tract left (suffix) | 96-97 |
| Census tract right (basic) | 98-101 |
| Census tract right (suffix) | 102-103 |
| ZIP code left | 104-108 |
| ZIP code right | 109-113 |
| Standard metropolitan statistical area | 114-117 |
| Street code (1970 mail census areas only) | 118-122 |
| From node | 123-126 |
| To node | 127-130 |
| Place code left | 131-134 |
| Place code right | 135-138 |
| State code left | 139-140 |
| County code left | 141-143 |
| Minor civil division code/ census county division code left | 144-146 |
| 1970 congressional district left | 147-148 |
| 1970 area code left | 149-151 |
| Block left (basic) | 152-154 |
| Block left (suffix) | 155-156 |
| 1960-1970 annexation code left (1970 mail census areas only) | 157 |
| State code right | 158-159 |
| County code right | 160-162 |
| Minor civil division code/ census county division code right | 163-165 |
| 1970 congressional district right | 166-167 |
| 1970 area code right | 168-170 |
| Block right (basic) | 171-173 |
| Block right (suffix) | 174-175 |
| 1960-1970 annexation code right (1970 mail census areas only) | 176 |
| From state plane code | 177-178 |
| To state plane code | 179-180 |
| From map set mile (X coordinate) | 181-186 |
| From map set mile (Y coordinate) | 187-192 |
| To map set mile (X coordinate) | 193-198 |
| To map set mile (Y coordinate) | 199-204 |
| From latitude (Y coordinate) | 205-210 |
| From longitude (X coordinate) | 211-217 |
| To latitude (Y coordinate) | 218-223 |
| To longitude (X coordinate) | 224-230 |
| From state plane (Y coordinate) | 231-237 |
| From state plane (X coordinate) | 238-244 |
| To state plane (Y coordinate) | 245-251 |
| To state plane (X coordinate) | 252-258 |
| Blank (census use only) | 259-300 |

is involved, municipality and state names are required. City names often can be replaced with ZIP codes because ZIP codes generally follow municipal boundaries. Street intersections are not valid input unless an intersection reference file has been created.

Before keypunching is initiated, all questionnaires should be edited to ensure completeness and consistency. The editing procedure consists of checking for blank fields and transferring erroneous field entries to the correct location. For example, a house number that was written in the street-name field should be moved to its proper position on the form. Although many of these deficiencies can be overcome by ZIPSTAN, this editing procedure consistently leads to better matching in subsequent steps.

### Computer Geocoding

Figure 2 shows the process involved in a simple geocoding project. In a more complex task, manipulation of the DIME/GBF or additional matches against reference files may be necessary.

Before geocoding, input data record addresses are reformatted to one address per record. If a record contains three addresses (origin, destination, and parking location), the record is reformatted to three separate records, each containing an address and a unique serial number and record position number. These elements permit rejoining the addresses in subsequent processing. Because the ZIPSTAN and UNIMATCH programs reference only one location per record for matching, reformatting the data allows processing of all addresses in one pass.

The steps involved in the geocoding process are the following:

1. The DIME/GBF is arranged in alphabetical order. The records are sorted by city or town and alphabetized by the street name records in each town. They are further sorted within house number, and the file is then ready for accessing. A printout of the file could be used as a guide for manual geocoding.
2. Organization of the data records is initiated, and ZIPSTAN is used to arrange them in a consistent format.
3. After ZIPSTAN has standardized the addresses, they are sorted in the same fashion as was the DIME/GBF. Following that sorting, a coder could manually code most of the data by visually inspecting each list and identifying the matches.
4. The UNIMATCH program is used to link the addresses to be geocoded with the DIME/GBF and to attach the appropriate census tract, block, and traffic zone to the travel data record.

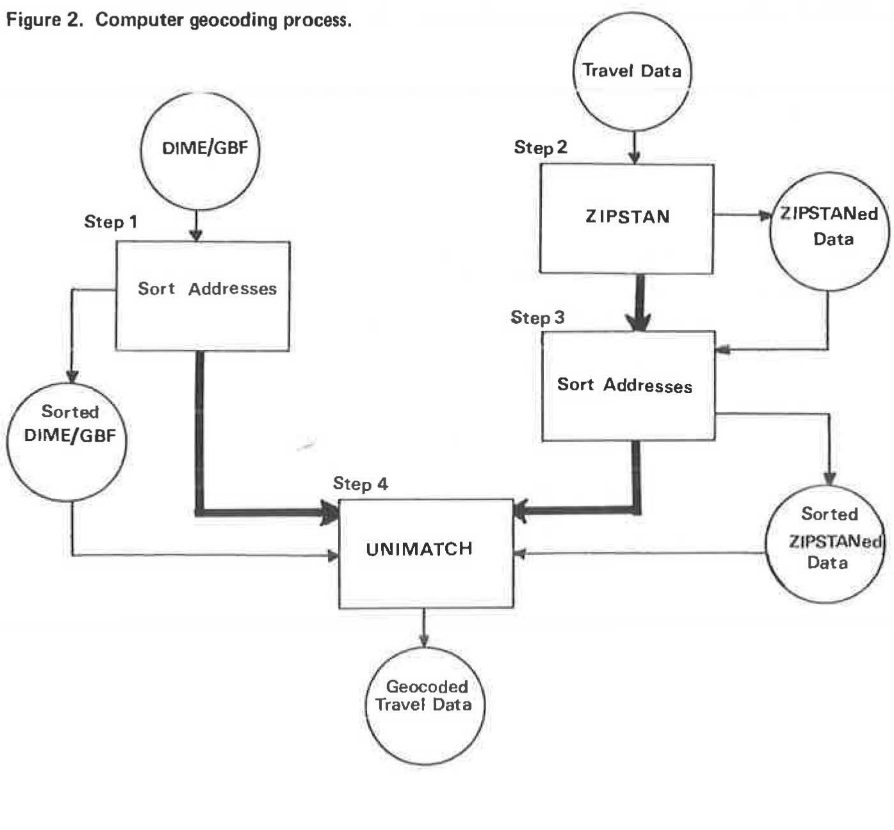
## RESULTS

The success rate for computer geocoding varies according to the quality of the respondents' answers, the quality of editing, the accuracy of data transcription, and the allowable weights and penalties specified by the user of the UNIMATCH program.

The editing process is more important in computer geocoding than in manual coding because, to a point, the machine cannot interpret what the respondent meant. For instance, unless specifically informed in the ZIPSTAN preprocessor, the computer does not regard BOS to be the same as Boston. In addition, the ZIPSTAN and UNIMATCH programs do not have the capability of determining that Beaconstreet is actually Beacon Street. Beaconstreet would be treated as a street name, missing a street type. These problems can be solved easily by using the ZIPSTAN equivalency tables; however, it is difficult to anticipate what misspellings and other errors will occur, and to correct such errors requires a second pass.

Inaccurate data transcription can negate the effects of precise editing. A space or letter punched before the beginning of a name, such as  Beacon or BBeacon, or a

**Figure 2. Computer geocoding process.**



reversal of letters, such as Ebacon, precludes a successful match. Experience has shown that the computer time and storage requirements for a match that allows all letters to be misspelled are excessive because almost every reference file entry becomes a possible candidate for a match. A solution to this difficulty is to specify program weights to require an exact match on the first two or three letters in a name, such as Bea___, and allow misspellings for matching the remaining letters. Thus, Beacno Street and Beanco Street can be matched with Beacon Street in the reference file.

The final variables that affect the match rate are the programmer-specified weights and penalties. It is possible to require an exact match in which house number, street name, prefix, suffix, and city must correspond exactly to the reference file entry, but exact matches have a low success rate. If street misspellings and omissions of prefix, suffix, and house numbers are allowed, the match rate increases proportionately.

Experience with on-board surveys conducted for the Massachusetts Bay Transportation Authority (MBTA) and with travel surveys performed as part of the Central Artery Study in Boston has shown that a 70 to 80 percent match rate can be expected with computer geocoding if flexibility is provided within the UNIMATCH program (5). If exact matches had been required in the MBTA and Central Artery Study surveys, the match rates would have been 6 and 4 percent respectively. Note that the UNIMATCH program is, in effect, making the same probabilistic choices that would be made in manual coding. The difference is that the computer is always consistent and unbiased when making these choices whereas a coder is not.

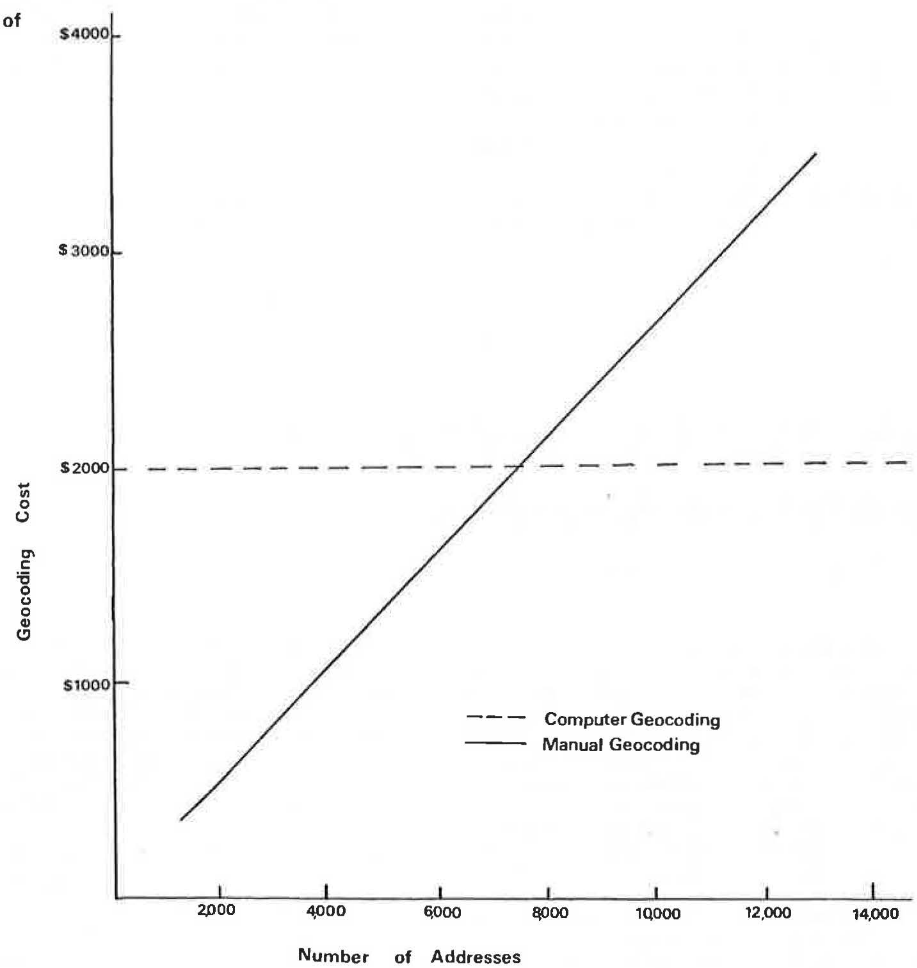
In the MBTA survey, a major generator file, later expanded for the Central Artery Study, was constructed. This file contained the names of buildings such as City Hall and the State House and places such as Harvard Square and Copley Plaza, which were not in the DIME/GBF. The locations of these generators were manually coded, transcribed on tape, and processed against addresses that did not match with the DIME/GBF. In smaller surveys, this additional step is often unnecessary because the addresses of certain major generators can be entered on the survey form during the editing step. However, in large surveys, this secondary match is required.

The table below gives a summary of the results of geocoding in the Central Artery Study survey:

| Condition | Number of Addresses | Match Rate (%) |
|---|---|---|
| Exact match | 5 026 | 4 |
| Probable match | 84 375 | 69 |
| Generator file match | 22 939 | 19 |
| Nonmatch | 10 524 | 8 |
| Total | 122 864 | 100 |

The supplemental match with the generator file increased the match rate from 73 to 92 percent. Altogether, 112 340 records were geocoded to a detailed zone level. The remaining 10 524 addresses were subsequently coded

Figure 3. Cost comparison of computer and manual geocoding.



to a town level. One hour and 17 min of computer processing unit (CPU) time and 576 K of virtual storage on an IBM 370 computer were required to geocode 89 401 addresses by using the DIME/GBF exact and probable match capabilities of the UNIMATCH program. The match with the generator file, which had 1900 entries, took about 13.5 min and used 64 K of virtual storage. As these figures indicate, the CPU time and storage requirements are a function of the number of fields being matched (city, street name, prefix, suffix, and house number versus city and building name) and the number of reference file and data file entries.

The MBTA and Central Artery Study addresses were geocoded on state computer facilities. On other geocoding projects, the computer costs associated with ZIPSTAN and UNIMATCH processing on a private computer facility average approximately $2.12/1000 addresses. Depending on the complexity of the DIME/GBF, the entire geocoding process, including development of ZIPSTAN tables and UNIMATCH specifications, can be developed, set up, and pretested at a cost of from $1000 to $4000. The costs for a major DIME/GBF such as Boston's, which has over 90 000 records, is at the higher end of the scale, and the costs for smaller DIME/GBFs are correspondingly lower.

As previously mentioned, some manual coding was necessary to construct the major generator file. Of the 1891 entries in the file, about 700 were unique locations; the remainder were variations of a place or building name. The 700 unique locations required about 48 person hours, or about 4 min/address, to geocode. At a cost of $4/h, excluding overhead and supervision of coders, this results in a cost of $0.267/address or $267/1000 addresses.

Figure 3 shows a graph that illustrates the cost relation between manual and computer geocoding. In constructing this figure, the setup cost for the computer geocoding was assumed to be $2000 and the processing cost $2.12/1000 addresses. The manual cost is $267/1000 addresses. As Figure 3 indicates, at these costs it is more economical to geocode by computer when there are about 7500 addresses or more. However, if smaller surveys, such as on-board bus surveys, are to be conducted on a continuing basis, it is more cost-effective to develop the computer geocoding capability.

Once they are set up, the DIME/GBF, major generator file, and ZIPSTAN and UNIMATCH programs can be used almost interchangeably on any on-board or other type of travel survey an agency may wish to conduct. Although a somewhat substantial start-up cost is required to implement a computer geocoding system, it can be more than recovered through the multiple uses of the system.

ACKNOWLEDGMENTS

The computer geocoding system described in this paper was prepared for the Community Affairs Division of the Massachusetts Bay Transportation Authority, the Central Artery Study of the Massachusetts Department of Public Works, and the Federal Highway Administration, U.S. Department of Transportation.

REFERENCES

1. Census Use Study: UNIMATCH 1 Users Manual. U.S. Bureau of the Census, 1974.
2. Census Use Study: ZIPSTAN Address Standardizer. U.S. Bureau of the Census, 1974.
3. CREATE: Coders' Manual. U.S. Bureau of the Census, 1976.
4. Documentation for the GBF/DIME-File. U.S. Bureau of the Census, 1976.
5. Documentation of On-Board Surveys—Volume 1. Urban Transportation Systems Associates and Massachusetts Bay Transportation Authority, 1976.
6. GBF/DIME System: A Geographic Dimension for Decision-Making. U.S. Bureau of the Census, Conference Proceedings, April 1976.

*Publication of this paper sponsored by Committee on Computer Graphics and Interactive Graphics.*

# Network Base File System for Transportation Analysis

Claus D. Gehner, University of Washington

**Development of the Network Base File System (NETBASIS), which has been under way at the University of Washington since 1974, is described. NETBASIS, which is implemented in an interactive graphic computer environment, explicitly builds on geographic base file data to provide a general-purpose transportation network data base together with the required data manipulation and display software. In addition to the standard capabilities of data base input, editing, and retrieval provided by the host data base management system, software has been implemented to allow the user to extract geographic or functional subsets of the network (e.g., arterials only). This extracted network can then be abstracted to remove nonintersection nodes and thus produce an efficient network model that can be used within most existing transportation planning tools. An interactive graphic network editor is also provided that allows the user to modify the extracted and abstracted network to reflect planning options to be analyzed. Segment-specific transportation data, such as transit lines, speed, and capacity, are added to the data base by a FROM-ON-TO coding scheme, which significantly reduces costly and error-prone intermediate data coding. Use of a commercial data base management system with a FORTRAN interface has resulted in significant cost savings in the development of the data input, editing, and retrieval software required to allow users easy access to the network data and to provide maximum flexibility of base file content and structure for adaptation to changing user requirements.**

The existence of U.S. Bureau of the Census geographic base files (GBFs) for most larger urban areas offers a significant resource for the network models required as part of many transportation studies. The thrust of the development of the Network Base File System

(NETBASIS) at the University of Washington is to build explicitly on this existing GBF data resource to provide a general-purpose transportation network data base together with the required data manipulation and display software. The purpose of this paper is to present a status report on the development of NETBASIS as of April 1978.

Initial development of NETBASIS was begun in 1974 and was documented by Gehner in 1975 (1). This work was found to be similar in a number of ways to a project conducted by Lutin at Princeton University (2).

A network base file system such as that described here is intended to provide a useful data resource for the entire spectrum of transportation planning activities, from long-range sketch planning to short-range implementation planning. Based on these intended uses and users, NETBASIS is characterized by requirements in two major dimensions: information content and functional capabilities.

## INFORMATION CONTENT

The backbone of the network data base is a complete description of the street network in the urbanized portion of the Seattle-Everett standard metropolitan statistical area (SMSA). The geographic and topological description of the street network is derived from the local Dual Independent Map Encoding/Geographic Base File (DIME/GBF) (3, Section 2.1) and is shown in summary in Figure 1. Appended to this network description are a number of data items relevant to various transportation analyses. A summary of the current set of such transportation attributes is shown in Figure 2.

The functional codes (items 27 to 35) are used to indicate primary and secondary functions of a given street segment. Classification into freeways, arterials, collectors, and local streets is based on the Highway Capacity Manual (4). Municipal engineering departments tend to use these or Urban Arterial Board route and section codes (items 46 and 47) or both in their own functional classification schemes. The additional codes for transit, bicycle, and walk are used to flag segments used for those modes either exclusively or in conjunction with a primary mode. Note that a given segment can have more than one functional code. Thus, for example, a segment may be a collector street with a bikeway, usable by pedestrians, or it may be an exclusive bikeway.

The walk-access and drive-access functional codes are derived codes to flag segments that, in an extracted and abstracted network, provide walk or drive access from zone centroids to the network. This follows the network conventions used by transportation analysis software of the Federal Highway Administration (FHWA) and the Urban Mass Transportation Administration (UMTA) (5).

Currently, the base file is capable of storing three separate speeds per direction. This follows the convention used in the UMTA Urban Transportation Planning System (UTPS), which differentiates network speed characteristics for the morning peak, the evening peak, and off-peak periods. The sources for this data are municipal and county engineering departments. The Seattle Engineering Department, for example, conducts speed-delay studies on a regular basis on all arterials within its jurisdiction.

Segment length and travel time are derived attributes. Length is calculated from coordinates (items 21 to 24 in Figure 1) based on the state plane coordinate system. Each coordinate unit corresponds to 3 m (10 ft). The travel times are in turn calculated from the segment length and the particular speed of interest depending on the proposed use of the network.

By definition, each segment can accommodate only one transit stop per direction. The transit stop number (item 40) refers to a record in a transit stop file that describes such characteristics of stops as transit lines served and location relative to the nearest intersection. This information is available from transit operators who collect such stop inventories for maintenance purposes. The other dimension of the transit system definition in NETBASIS is given by a transit route file that, for each transit line (or route) in the system, stores the ordered sequence of segments traversed by that line.

## FUNCTIONAL REQUIREMENTS

To be useful as an information base for transportation analysis, NETBASIS must have the capability to input and update the data base, edit the street and transit network, extract areal and functional subsets of the network, and abstract the extracted network down to the level of detail required for a given analysis. Some simple network analysis capabilities, such as minimum path calculations, have also been found to be useful.

Assuming, as we do here, that the backbone of the

**Figure 1. Contents of geographic base file.**

| SEGMENT IDENTIFICATION | | | | | | NODE ID | | | | ADDRESS RANGE | | | | SEGMENT GEOGRAPHY | | | | | | COORDINATES | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | FROM | | TO | | LEFT | | RIGHT | | LEFT | | | RIGHT | | | FROM | | TO | |
| MAP NO. | SEQ.NO. | PRE DIR | STREET/NON-STREET FEATURE NAME | TYPE | SUF DIR | MAP | NODE | MAP | NODE | LOW | HIGH | LOW | HIGH | TRACT | BLOCK | AREA | TRACT | BLOCK | AREA | X | Y | X | Y |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |

**Figure 2. Contents of network base file.**

| NETWORK CHARACTERISTICS | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | FUNCTION CODE | | | | | | | | | LEFT | | | | | RIGHT | | | | | | | |
| | | | | | | | | | | | | SPEED | | | | | SPEED | | | | | | |
| NETWORK CODE | SEGMENT LENGTH | TRANSIT | FREEWAY | ARTERIAL | COLLECTOR | LOCAL | BICYCLE | WALK | WALK ACC. | DRIVE ACC. | TRAVEL TIME | AM | PM | OFF | TRANSIT STOP NO. | TRAVEL TIME | AM | PM | OFF | TRANSIT STOP NO. | UAB ROUTE | UAB SECTION | ADDITIONAL NETWORK DATA |
| | 26 | | 28 | | 30 | | 32 | | 34 | | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | |

data base is given by the GBF, the maintenance of which, for the time being, is outside the scope of NETBASIS, the input and update function refers primarily to transportation-related data such as link types (e.g., arterial or freeway), speeds, and capacities, which must be added to the GBF data. Specification of the data for input or update from the user's point of view should be independent of the internal organization and structure of the data base (e.g., it should not require knowledge of the node and segment numbering scheme). Furthermore, such data specifications should be in a format that is as close as possible to the format in which the data are currently maintained and collected to avoid error-prone intermediate coding.

Network editing consists of selective addition or deletion of network links (or street segments) or shifting of the location of a node (or street intersection) or both. Past experience has shown that this function is significantly aided by graphic input-output capabilities.

Network extraction is defined here to encompass the selection of a subset of the network data base. The selection criterion can be geographic (e.g., windowing for use in a detailed corridor analysis) or functional (e.g., selection of an expressway and arterial network for use in a regional highway-oriented study).

Network abstraction, finally, is the process of aggregating links to the point where only the minimum set of nodes is retained. Thus, if one had previously extracted an arterial-expressway network, only those nodes would be retained that are intersections between arterials and expressways. This function ensures that the resulting network model is the most efficient possible in terms of applying to it the many existing network analysis algorithms (e.g., minimum path), the computational cost of which is often proportional to the number of nodes and links. This abstracted and extracted network can become a new, derived data base to which the above input-update and network editing functions can be applied, allowing the user to "design" the specific network model needed for a given analysis.

## IMPLEMENTATION

### Objectives

The implementation of NETBASIS was guided by four general objectives:

1. Minimize the time and cost required for software development,
2. Minimize data collection requirements,
3. Maximize system flexibility in terms of types of data files to be accessed and modification and expansion of functions, and
4. Maximize accessibility of the system to the user.

### Strategies

In pursuing the four stated objectives, four implementation strategies were identified.

#### GBF as a Network Resource

A large amount of time and money has been expended nationally to develop, validate, and maintain geographic base files and the associated geoprocessing software. These existing GBFs represent a valuable data resource that the NETBASIS development has exploited to obtain detailed transportation network base files at a low cost and within a short start-up time. This is possible because the GBF is essentially a network model that contains node and link information for all street segments (and nonstreet features such as natural boundaries) in an urban area.

#### Use of Data Base Management System

An important aspect of information system development is the provision of easy, quick, and inexpensive capabilities of data storage, retrieval, modification, and report generation. Developing, testing, and putting in operation these low-level data-handling capabilities in most instances consumes a large portion of the person hours required in development of a data base system.

This project has circumvented much of this development cost by using the capabilities of commercially available data base management systems (DBMSs). Although they were developed primarily for business applications, modern DBMSs are sufficiently flexible and general in purpose to be applied to other types of projects, such as the implementation of NETBASIS. Specifically, aside from providing powerful data management capabilities, many DBMSs provide for a "host language interface" (FORTRAN or COBOL), either in terms of a precompiler or through a set of subroutines, which allows high-level application programs to access the data bases and make use of the efficient data retrieval capabilities. It is primarily through this host language interface that this project has made use of DBMS capabilities (6).

#### Use of Interactive User-Oriented Command Language

In an effort to maximize accessibility of the system to a wide range of users, including those with little or no background in computer applications, a highly user-oriented command language is used to input to the programs the data necessary for proper execution. Thus, the user interacts with the software through a series of commands that consist of keywords and data to direct program execution.

Use of an existing software package for design and interpretation of command syntax—LANGPAK (7)—contributed to achieving the objective of minimizing the costs of software development. LANGPAK includes a feature for the interactive design and testing of syntax for new or modified commands. This feature significantly adds to the ability to modify and expand the NETBASIS package to meet new and evolving user needs.

#### Use of Graphic Input-Output Devices

A final important implementation strategy for NETBASIS is that it uses an interactive graphic computing environment. The selection of an appropriate network model for a given analysis or planning effort through the use of editing, extraction, and abstraction functions is basically a design process in which real-time interaction with the computer can be very important. The implementation of NETBASIS thus makes use of a graphic cathode ray tube (CRT) terminal as the primary interface between the computer and the user. Specifically, the Tektronix 4014 terminal allows the display of maps, networks, and other spatial data and allows the user to input points and vectors in two-dimensional space by means of user-controlled cross hairs.

The use of graphics significantly enhances the flexibility of the system in that it allows the user much more control over the types of data displays that best suit the particular needs of the analysis.

In addition to the CRT terminal, displays can also be routed to a CALCOMP plotter, which produces high-quality graphics suitable for presentations and reproductions.

## APPLICATIONS OF NETBASIS FUNCTIONS

In describing the NETBASIS functions implemented to date, a window of the Seattle-King County network is used. We start with the basic network file and describe the addition of a freeway that runs through the selected area but was not part of the original GBF. This demonstrates the network editing functions. Some network attributes are then added to the base file—specifically, a definition of the subset of arterial streets and their speeds to demonstrate input-update capabilities. The extraction and abstraction functions are demonstrated by deriving a network that consists only of the freeway and arterial streets. Finally, the utility of the minimum path algorithm is used to calculate and display travel-time contours.

The user-computer dialogue occurs in the form of a command language, as discussed above. Commands basically consist of key words and (where needed) data items. The specific commands described in this report are only intended as examples and do not show the full range of options available within each function. A * is used to prompt the user that the system is ready to accept a user command.

Figure 3. Base network window display.

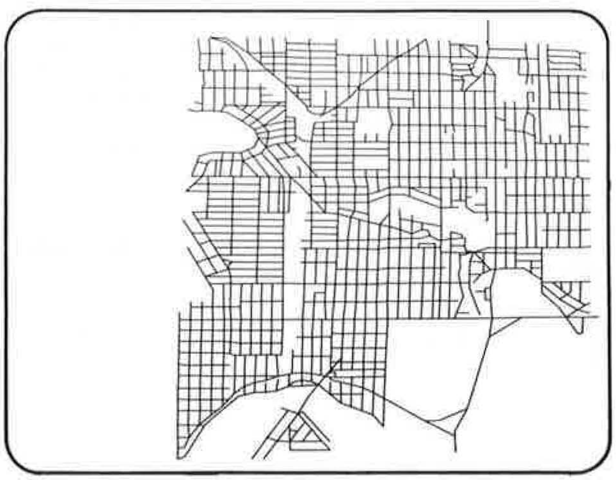
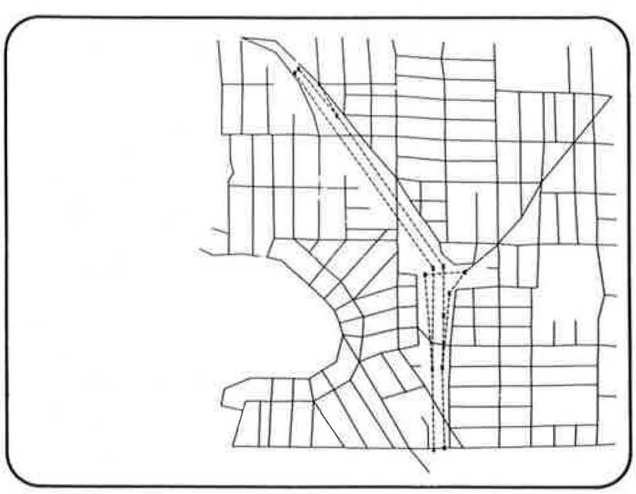


Figure 4. Window display after addition of several new links.



In addition to the command input described above, a number of functions call for the graphic input of coordinates. This implies that a (portion of a) network is displayed on the CRT screen with reference to which coordinates are located. In general, the transmission of a set of coordinates will cause the node nearest to that point to be selected. If, however, the coordinates are transmitted with the N key, the programs will interpret this as a new node for insertion into the base file. Transmission of coordinates with 0 (zero) will terminate graphic input.

### Network Windowing

To select that portion of the Seattle network on which the description of the example applications will be based, the network windowing function, which is equivalent to the selection of an areal subset, is invoked. The following command is used:

```
*WINDOW MAPS 5,6,14,15,23,24
*
```

In this case, we have chosen to specify the window on the basis of map numbers that are used to organize the area into square-mile sections (the system is calibrated in U.S. customary units of measurement). The other alternative is to explicitly specify coordinate extrema, either numerically or graphically, by using the cross hairs. The second option implies that a subset of the currently displayed window is to be selected ("zooming in"). The final * is displayed after the current function is completed, which indicates that the programs are ready for the next command.

### Displaying the Currently Selected Network Subset

At any time, the currently selected window of the network can be displayed on the CRT screen. This is accomplished by the following command:

```
*PLOT NETWORK
*
```

This command results in a display such as that shown in Figure 3.

### Network Editing

Network editing is accomplished by four separate subfunctions:

1. Divide a segment by inserting a new node anywhere between its current end points.
2. Insert new segments between existing and new nodes.
3. Move an existing node.
4. Delete an existing segment.

The function of dividing a segment is currently used primarily to prepare for the addition of new segments. Thus, for example, if a freeway exit ramp is to be inserted that merges with an existing street at its midpoint, that intersection must first be created. This function is invoked by the command

```
*DIVIDE SEGMENT
```

The user then graphically specifies the end points of the segment to be split and subsequently points to the location of the new node. This sequence of three graphic inputs for each segment to be split can be repeated until a 0 (zero) is transmitted.

The process of inserting new segments is the most complex in terms of the input required from the user. A typical dialogue, which results in the new segments shown in Figure 4, is as follows:

```
*NETWIN MAP 5
*PLOTN
*ADD SEGMENTS
≥INTERSTATE 5 FY:FW=!,AMR=55  (graphic input of segments)
≥82 EXIT RP N:FW=1,AMR=30  (graphic input)
≥LAKE CITY EXIT RAMP NORTH:FW=1,AMR=35  (graphic input)
≥INTERSTATE 5 FREEWAY:FW=1,AMR=55  (graphic input)
≥LAKE CITY ENTER RP S:FW=1,AMR=35  (graphic input)
.
.
.
≥
*
```

After specifying and displaying a smaller window (only map 5) in an effort to obtain better resolution during the editing process, the add-segment function is invoked by using the appropriate command. The user is then prompted with > to input the name and any desired attributes of the segments to be added. Segment attributes are specified by a name or abbreviation, as defined in the DBMS (FW = freeway function code and AMR = a.m. peak speed on right segment side), and the values they are to assume. In this example, we are specifying that the new segments are part of the freeway system and have varying a.m. peak speeds.

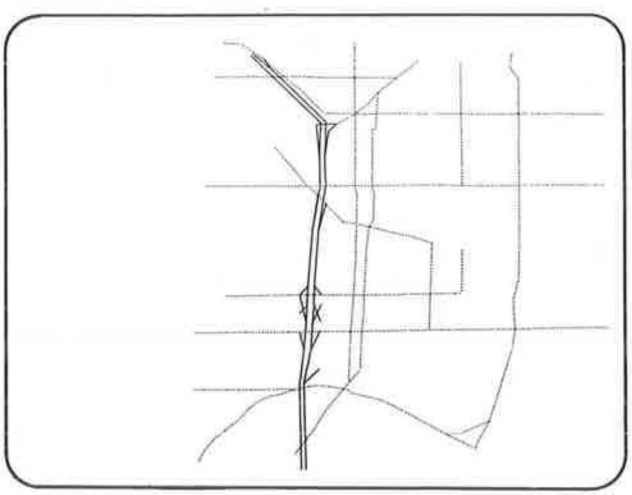
After each such input, the cross hairs with which the user can input one or more (continuous) segments appear on the CRT screen. End-point nodes for each new segment that is inserted are displayed with a *, and the new segment appears as a dashed line (Figure 4). A new sequence of segments is terminated with the 0 key. At this point, a new street name and attribute input is prompted for. The user can terminate the add-segment function by inputting a carriage return.

The ability to move an existing node is often useful for "cleaning up" a network. Thus, if, for example, inspection of a network display in comparison with a base map shows one or more nodes out of place, the user first points to the node to be moved (graphic input) and then to its new location.

Finally, to complete the network editing capabilities, a function to delete segments has been implemented. The user again uses the cross hairs to point to the end points of segments to be deleted.

Figure 5. Highlighting of arterials and freeways in network display.

Figure 6. Display of extracted network.



## Input-Update Network Attributes

Based on experience with a number of local agencies, it was found that large amounts of transportation-related data are collected and maintained in a format keyed to street names and street intersections. This has led to the development of a general-purpose FROM-ON-TO data input scheme. This specifies the segments to be affected as having the ON street name between the intersections with the FROM and TO streets respectively.

A typical dialogue that uses this input scheme to specify arterial streets and their speeds would be as follows:

```
*LINK ATTR TO BALARD.NET
≥CORLISS WAY N & BANNER WAY NE & NE 75TH:MA=1,AMR=30,
   AML=25
≥NE 75TH & 34TH AVE NE:
≥BAGLEY AVE N & N 80TH ST & 1 AVE N:
≥NE 70TH & ROOSEVELT WAY NE & NE 50TH:MA=1,AMR=25,
   AML=15
.
.
.
≥
*
```

The input-update function is invoked by use of the command, as shown. TO BALARD.NET specifies that the attributes are to affect the network data base stored under that name. The implicit input mode in the example shown is alphanumeric from the terminal keyboard. There are two other options available. The key word and data in the command FROM<filename> would indicate that a batch input file named filename has been prepared that contains all information normally typed in from the keyboard. This is useful for large volumes of data. Alternatively, the key word GRAPHIC would indicate that the cross hairs are to be used to point to intersections (nodes), replacing the FROM-ON-TO specifications.

After invoking the function, the user is prompted to input the FROM-ON-TO street names, specifying the sequence of segments to be affected and the attributes to be changed. The attribute input is identical to the one described in the previous section. In the example shown, we are specifying that the segments are to be classed as major arterials (MA=1) and that their right and left a.m.

peak speeds are to be 30 and 25 mph respectively.

Figure 5 shows the set of arterials in the test area that were specified in this way. Note that in this display the arterials have been differentiated from other segments.

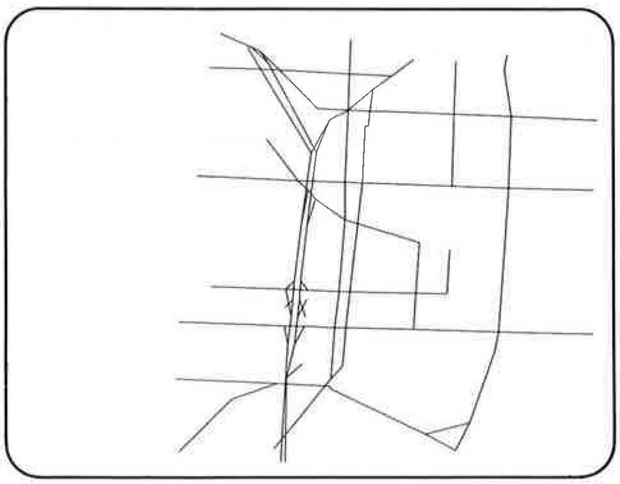
Network Extraction

Network extraction, defined earlier, is defined here as the selection of an areal or functional subset of a network. The command to invoke this function is

```
*SUBSET TO <filename> SEARCH MA.EQ.1.OR.FW.EQ.1
```

The example shown indicates a SEARCH for all segments that have an arterial or freeway function code equal to 1. The command also specifies that the extracted network data base is to be stored under <filename>. A display of the extracted network is shown in Figure 6.

**Figure 7. Display of abstracted network.**



Network Abstraction

The network abstraction function operates on previously extracted networks, as defined above. The function is invoked by the command ABSTRACT, and no additional specifications are required. As defined here, network abstraction consists of aggregating segments of a given street name between its end points or remaining bona-fide intersections or both.

The resulting network is shown in Figure 7. The number of segments has been reduced from 362 in the extracted network to 113 in the abstracted version. As mentioned previously, this derived network can now become the subject of editing and input-update functions, providing the user with great flexibility in designing a network model to suit any given transportation analysis.

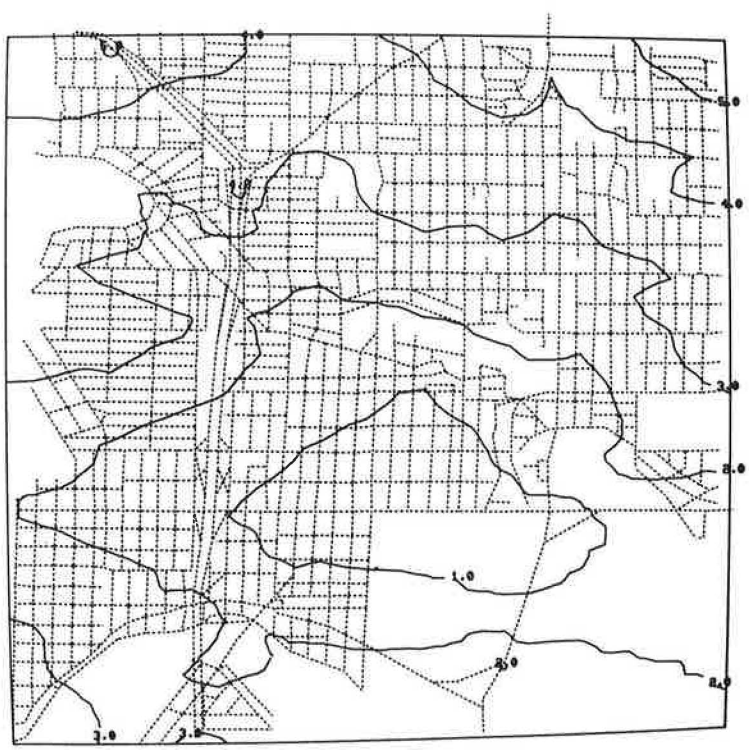
Calculation and Display of Travel-Time Contours

The calculation and display of travel-time contours involves a number of separate functions that are not described in detail here. The following set of commands was used to generate the contour overlay shown in Figure 8:

```
*DERIVE TTR=LL/AMR/1.47,TTL=LL/AML/1.47
*MINPATH TIME ROOT NE 45TH & 17TH AVE NE
*CONTOUR TO CONTR.DAT, BET 0 600 INT 60
*PLOTC FROM CONTR.DAT
*
```

The first step is to derive segment traversal times based on link length (LL) and a.m. peak speeds (with conversion to seconds). The second step involves the execution of the minimum path tree algorithm based on time (as opposed to distance) from a single root at NE 45TH and 17TH AVE NE. The third step consists of

**Figure 8. Display of travel-time contours.**



using this minimum path data to calculate contours between 0 and 600 s at intervals of 1 min. The final command requests that these contour data be displayed on the CRT screen.

## FUTURE DIRECTIONS FOR NETBASIS

During the current academic year, the Urban Transportation Program at the University of Washington, sponsored by a continuing research grant from UMTA and a contribution from Seattle Metro Transit, is conducting a detailed analysis of an on-board origin-destination survey recently completed by Metro. For this purpose, development of the NETBASIS data base is continuing with the addition of transportation attributes for additional subareas within Seattle-King County and the coding of a substantial portion of the transit stops and routes. Furthermore, an existing interactive graphic analysis and display system is being coordinated with NETBASIS.

Using these tools, the research project will emphasize interactive graphic analysis of origin-destination data at a very detailed scale in an effort to evaluate the usefulness of the tools and the data to the ongoing planning and analysis functions of Metro Transit and other cooperating agencies. The experience gained during the course of this project is expected to result in significant changes and additions to the functional requirements for NETBASIS. Future projects will have to be defined to implement these changes.

## REFERENCES

1. C. D. Gehner. Development and Testing of the Network Basefile System (NETBASIS). Urban Transportation Program, Univ. of Washington, Seattle, Working Paper 38, 1975.
2. J. M. Lutin. Creation of Urban Transportation Network Models From U.S. Census Geocoded DIME Files. Paper presented at CENDOT Workshop-Conference, Long Beach, CA, Nov. 8-12, 1976.
3. Documentation for the Geographic Base (DIME) File. Bureau of the Census, U.S. Department of Commerce, Aug. 1974.
4. Highway Capacity Manual. HRB, Special Rept. 87, 1965.
5. Urban Transportation Planning—General Information and Introduction to System 360. Federal Highway Administration, U.S. Department of Transportation, March 1972.
6. System 1022 Reference Manual—Version 113-2. Software House, Inc., Waltham, MA, Sept. 1976.
7. L. E. Heindel and J. T. Roberto. LANG-PAK—An Interactive Language Design System. Elsevier, New York, 1975.

*Publication of this paper sponsored by Committee on Computer Graphics and Interactive Graphics.*