TRANSPORTATION RESEARCH RECORD 751

Transportation System Analysis and Planning 1980

TRANSPORTATION RESEARCH BOARD

COMMISSION ON SOCIOTECHNICAL SYSTEMS
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C. 1980

Contents

Transportation Research Record 751
Price $5.20
Edited for TRB by Susan Singer-Bart

modes
1 highway transportation
2 public transit

subject areas
12 planning
17 energy and environment

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Transportation system analysis and planning 1980.

(Transportation research record; 751)
Reports prepared for the 59th annual meeting of the Transportation Research Board.
Includes bibliographical references.
1. Urban transportation–Planning–Addresses, essays, lectures. I. Title. II. Series.
TE7.H5 no. 751 [TA1205] 380.5s [388.4'068] 80-21248
ISBN 0-309-03061-7 ISSN 0361-1981

Sponsorship of the Papers in This Transportation Research Record

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

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

Committee on Transportation Systems Design
Gorman Gilbert, University of North Carolina, chairman
Virginia J. Ainslie, James H. Banks, William G. Barker, Edward A. Beimborn, Melvyn Cheslow, Lewis P. Clopton, Robert T. Dunphy, Jerry L. Edwards, Gary R. Erenrich, Joel Ettinger, 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, Robert L. Peskin, Richard H. Pratt, Sandra Rosenbloom, Larry H. Sams, William L. Smith, Darwin G. Stuart, Antti P. Talvitie

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

Committee on Statewide Multimodal Transportation Planning
Kenneth W. Shiatte, New York State Department of Transportation, chairman
John A. Bivens, Jr., Malcolm F. Brenan, Wm John Cameron, Roger L. Creighton, Robert E. David, Clarence W. Friesen, Arne L. Gausmann, Antonio W. Gonzales, Charles H. Graves, Philip I. Hazen, Lester A. Hoel, Thomas F. Humphrey, Colin Ian MacGillivray, Michael I. Schneider, Isaac Shafran, Carl N. Swerdloff, William C. Taylor, Thomas J. Thompson, D.L. Wieman, Norbert Y. Zucker

Committee on Transportation Planning Needs and Requirements of Small and Medium-Sized Communities
C. Michael Walton, University of Texas at Austin, chairman
Edward G. Bates, Jr., James A. Bautz, Jimmie D. Benson, Michael J. Demetsky, David A. Doctor, Antonio G. Hobeika, R. Keith Jones, Joseph C. McBride, Carl C. McChesney, Monty C. Murphy, Edward S. Neumann, Marion R. Poole, Donald B. Rhodes, Robert R. Roberts, George E. Schoener, Kenneth W. Shiatte, Vergil G. Stover, Montie G. Wade, John H. Waggoner, 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, 1979.

Authors of the Papers in This Record

Caldwell, Lawrence C. III, Virginia Department of Highways and Transportation, 1221 East Broad Street, Richmond, VA 23219

Cheslow, Melvyn D., Evaluation Research Corporation, 8330 Old Springhouse Road, Vienna, VA 22180

Demetsky, Michael J., University of Virginia and Virginia Highway and Transportation Research Council, Charlottesville, VA 22903

Dooley, Thomas, Research and Special Programs Administration, Transportation Systems Center, U.S. Department of Transportation, Kendall Square, Cambridge, MA 02142

Fernandez, J. Enrique, Department of Civil Engineering, Universidad Catolica de Chile, Santiago, Chile

Friesz, Terry L., Department of Civil and Urban Engineering, University of Pennsylvania, 113 Towne Building, D3, Philadelphia, PA 19174; formerly with the Department of Civil Engineering and the Operations Research Center of Massachusetts Institute of Technology

Han, Anthony Fu-Wha, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

Herald, William S., Alan M. Voorhees and Associates, 7798 Old Springhouse Road, McLean, VA 22102

Meyers, Catherine E., Planning and Research Bureau, New York State Department of Transportation, 1220 Washington Avenue, State Campus, Albany, NY 12232

Neveu, Alfred J., Planning and Research Bureau, New York State Department of Transportation, 1220 Washington Avenue, State Campus, Albany, NY 12232

Perl, Jossef, Department of Civil Engineering, Northwestern University, Evanston, IL 60201

Schoon, John G., Department of Civil Engineering, Northeastern University, 360 Huntington Avenue, Boston, MA 02115

Teply, Stan, Department of Civil Engineering, University of Alberta, Edmonton, Alberta, Canada

Tourreilles, Francisco A., Department of Civil Engineering and Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02139

Zakaria, Thabet, Delaware Valley Regional Planning Commission, 1819 J. F. Kennedy Boulevard, Philadelphia, PA 19103

Zerrilo, Robert J., Planning and Research Bureau, New York State Department of Transportation, 1220 Washington Avenue, State Campus, Albany, NY 12232

Issues in the Evaluation of Metropolitan Transportation Alternatives

Melvyn D. Cheslow

This paper was prepared as part of a study by the Urban Institute to improve the cost-estimation methods and evaluation procedures for the Urban Mass Transportation Administration alternatives analysis process. Several important issues related to evaluation of alternatives are discussed. They include problems and advantages of several analytical evaluation techniques such as cost-benefit analysis, cost-effectiveness analysis, and project scoring. Issues in evaluating user benefits are also discussed. Much of the complexity in the evaluation of alternatives comes from the different interests or groups that must be involved and the differences between the federal and local viewpoints. This paper is not a handbook on evaluation and assumes a basic knowledge of analytical evaluation techniques.

The alternatives analysis process has been developed by the Urban Mass Transportation Administration (UMTA) for the evaluation of metropolitan transportation alternatives. In this process choices of preferred alternatives are made at the local government level, and decisions about providing financial grants are made at the federal level. Evaluation of alternatives must take place at several decision points during the staff analyses and the public and governmental deliberations. Because of the many groups involved and the variety of issues that must be addressed during an evaluation of alternatives, no single method of evaluation can be used. Instead, the evaluation process must involve a variety of procedures, each selected for the appropriate decision.

This paper discusses several important general issues concerning the evaluation process. It first contrasts the two uses of evaluation: selection of the most preferred alternative and determination of the worthiness of the chosen alternative. This second use concerns whether an investment is worth making. The paper next discusses the various stages in the evaluation process and compares the information needs of each stage. The amount of information that must be considered should be reduced as the process moves toward a final decision. This allows attention to be focused on the better alternatives and more important attributes. A section then discusses several common problems of analytical evaluation procedures. For example, the choice of an alternative when members of a community have different goals and preferences imposes limitations on a completely analytical evaluation process.

PURPOSES OF EVALUATION

The alternatives analysis process is required so that UMTA can allocate a limited amount of capital grant monies among a large number of metropolitan areas that desire federal aid. UMTA must decide how to apportion the aid so that the national welfare is most improved. In UMTA's evaluation, investments in some urban areas will appear to produce fewer benefits to the nation than will investments in other areas. Because of the finite federal funds available, all areas that desire aid will not receive it. Hence, UMTA must determine the relative worthiness of each aid request. Each metropolitan area will help UMTA by determining its most preferred investments as part of the alternatives analysis process so that UMTA can make appropriate comparisons among areas that desire aid.

In a similar manner, the local governments in each urban area must determine whether a transportation investment is worthy of local financial support. Because of other demands for public funds as well as limitations on the amount of tax revenues that can be raised, even the best local transportation alternative may not prove worthy of local public support. Hence, each urban area should use the alternatives analysis process to determine the worthiness of a selected transportation improvement in terms of its cost and its ability to contribute to the attainment of the goals of the local area.

In order to find a worthy investment, the alternatives analysis process requires an examination of several alternatives and determination of the most desirable. This process may not be clear cut because the alternative selected for funding must be judged worthy both by the local government and by UMTA. One alternative may not be best from the perspective of both of these funding sources. This situation may occur because either (a) alternatives may differ in the production of national benefits that are not local benefits or (b) each urban area will have differing local goals with respect to which alternatives are selected, and all these local goals will not be commensurate at the national level.

When a local area proceeds through the alternatives analysis process by comparing alternatives and finding the most preferred one, it has completed only part of its responsibilities. Even the most-preferred alternative may not be worthy of investment. Doing nothing, or making no new investments, may result in more local or national net benefits than any of the investment alternatives. For example, some alternatives may produce large user benefits compared to their direct costs but may also produce such extensive damage to the community through pollution or neighborhood disruption that they are not desired by the whole urban area. Other alternatives may produce a small amount of community disruption (by going underground, for example) but produce direct user costs greater than user benefits. These alternatives, too, will not be desired by the entire urban area. Community values in some urban areas may be such that no investment in expanding transportation capacity will be worthy.

Although determination of the worthiness of a project may not be easy, the local area must treat the issue and should compare the best investment with no investment.

In summary, the evaluation of alternatives in the alternatives analysis process has two purposes:

1. To select the most-preferred investment alternative with respect to local goals and
2. To determine the worthiness of the most-preferred investment alternative for the local governments, for the nation, and with respect to no new investment.

STAGES OF THE ALTERNATIVES ANALYSIS EVALUATION PROCESS

Evaluation within the alternatives analysis process proceeds from the analysis of the impacts and costs of alternatives toward the final choice of a desired alterna-

tive (which may be the no-new-investment alternative). The alternatives analysis process has several steps. First, the local area carries out an analysis to demonstrate to itself and to UMTA that there is a reasonable chance that it will have a worthy investment. UMTA then evaluates this analysis. Next, the local area selects its most desirable alternative. Within this step, several modifications to the alternatives may be considered as different community groups take part in the evaluation. Following this stage, UMTA evaluates the submissions from each local area and perhaps suggests modifications to the procedures or the alternatives. The local area may then revise a portion of the analysis.

Within this complex alternatives analysis process, several evaluations take place. They are carried out by local governments, local citizens, the urban area as a whole, and UMTA. Each of these participants makes choices in the alternatives analysis process. Each of these evaluations occurs in three stages. In the first stage, the information generated in the analyses is brought together and organized so that it can be used for decision making. From the multitude of data produced in the analyses, the measures important for a given evaluation must be determined, based on such criteria as relevance and accuracy. The measures must provide useful information, necessary for UMTA comparisons among areas, about the extent to which local goals are met or about benefits and costs.

The second stage of evaluation reduces the number of measures considered so that the information can be processed in the human mind. (Psychological studies have shown that there are limits to the number of attributes that can be considered at one time in a decision-making process.) This dimensionality reduction can involve any of several analytical techniques to aid the decision makers. Benefit-cost analysis and cost-effectiveness analysis are two such techniques. None of the techniques is foolproof in ensuring that correct choices will be made. Since they all reduce the number of measures under consideration in the final choice, their use results in the loss of some of the information originally presented. But, if no analytic dimensionality-reduction procedures were used, then each decision maker would rely on informal, mostly subjective, processes that may be less orderly and hence more biased than the analytical procedures.

The third stage of evaluation is the actual choice process. In some cases, this process continues the analytical procedures from the second stage, which leads to selection of a single measure that is used to choose among alternatives. More often, choice will involve a judgment among a small number of alternatives based on a few measures that cannot be combined in any agreed-on analytic manner.

Organizing and Presenting Information: Stage 1

The information to be evaluated should be selected on the basis of relevance and accuracy. The chosen measures should indicate to each segment of the population the impacts of each alternative. Unfortunately, if many community groups have different values and concerns, the quantity of data required would be enormous. More importantly, the data may not be very accurate. Most forecasting models available to transportation planners are not sufficiently accurate to produce the detailed forecasts each group desires. Therefore, there must be trade-offs made between the detail and accuracy of the evaluation information. Furthermore, the number of measures selected should be limited so that the various participants are not overwhelmed by data that may be redundant, of marginal use, or difficult to understand.

The accuracy of the evaluation measures is different from their precision. Current computer models can estimate a value to 10 or more significant places but, if the models or the input data are not absolutely correct, the value would have little accuracy. In fact, the accuracy can be as low as 1 or 2 significant digits (errors of 100 percent or 10 percent, respectively). It is important to assess carefully the accuracy of the estimates for each measure used in an evaluation to determine whether the measure will discriminate among the alternatives.

Some analysts have suggested that the accuracy of a model may be much better in comparing the differences among several alternatives than in making absolute forecasts. However, the validity of this assertion depends on the relative errors in the model parameters—parameters for attributes that vary must be more accurately estimated than those for attributes that do not vary.

In presenting the evaluation measures, the information should be understandable; that is, the measures themselves should be meaningful to all participants. Furthermore, at this stage all relevant, accurate data should be presented (along with notes about which important measures do not discriminate among alternatives). If a particular group (geographically or socioeconomically different) has distinct concerns or values, then an attempt should be made to present evaluation measures important to that group. Federal officials should generate their own list of evaluation measures that local areas must respond to.

Reducing Dimensionality of the Evaluation Measures: Stage 2

In order to make decisions from the large amount of data provided in stage 1, most participants will reduce the amount of information they consider, by use of a variety of mechanisms. Many analytical processes are available for reducing the number of measures (or dimensions) for consideration; some, such as benefit-cost analysis, allow the number of dimensions to be reduced to one, thus making the choice process straightforward. Other procedures only partially reduce the dimensionality.

Each urban area must decide the extent to which it will rely on analytical procedures to lead to its decisions. Some may use analytical approaches entirely, without having different local interests trade off their choices politically. Others may have each interest group independently use analytical procedures to reach its choice but allow bargaining among interests. Finally, some areas may use these methods only to reduce dimensionality and then use a subjective procedure for the choice process. Even if analytical procedures are not part of a formal evaluation process, they may be useful for educational purposes; they will improve understanding by the different groups of their attribute valuations and trade-offs.

Several valuable approaches exist for reducing dimensionality; however, each has defects if it is used single-mindedly. Each can introduce bias into the evaluation process by omitting or misevaluating some measures. Analysts and decision makers must all be aware of the benefits and deficiencies of each approach. A brief listing of various approaches is presented here. Four major analytical approaches are currently used extensively in program evaluation:

1. Cost-benefit analysis,
2. Cost-effectiveness analysis,

3. Scoring or rating methods, and
4. Reduced table of measures.

These approaches can be used not only for reducing dimensionality but also, in principle, for the complete evaluation process. Except for the reduced table of measures, each approach can reduce the choice problem to comparison of a single measure for each alternative. Most participants in the evaluation process, however, will probably choose to use nonanalytical approaches to complete the decision process.

In addition to the four major analytical approaches, several others provide ways to reduce the dimensionality of the list of measures or to reduce the number of alternatives being considered. These include the following:

1. Trade-off analysis of attributes,
2. Threshold analysis (required minimum level of attributes),
3. Comparison of pairs of alternatives,
4. Dominance analysis for attributes,
5. Stochastic dominance, and
6. Marginal analysis (setting two alternatives at same level of selected attribute).

Decision and Choice: Stage 3

The actual decision by each participant in an evaluation process involves the weighing of analytical and subjective information. Even when all information used is generated by analytical tools, implicitly a subjective decision has been made to omit nonanalytical input.

There are arguments for using only analytical measures to determine choice: consistency of approach can be ascertained and comparability among many decisions can be ensured. These arguments require that each decision maker represent his or her values analytically with respect to the various measures. This includes his or her approach to uncertainty (e.g., whether risks should be minimized, whether the expected value of uncertain futures should be optimized, or whether some other approach should be used). Most evaluators probably find a totally analytical approach too restrictive and use subjective methods in at least part of the evaluation process.

An important problem in carrying out an evaluation is that of social choice. The problem exists when the values of the members differ, so that different preferences for alternatives occur. In this situation, it is necessary either for an arbitrary decision or for bargaining to occur among the members. No analytical procedure can represent both the preferences and the intensity of preferences of the various interests.

Analytical procedures can be used advantageously in this stage when choices are made by small groups who have homogeneous interests. In some urban areas the decisions may be made by this type of group, but the requirements for citizen participation in the alternatives analysis process will generally preclude this situation.

Analytical procedures are more important in stage 2 to reduce the list of evaluation measures to a small number, which then must be appraised in the decision stage. They are also useful in providing education and insights to the participants about their evaluation of the various measures. The extent to which analytical techniques can be used will vary with the urban area and the types of goals and alternatives considered.

PROBLEMS OF EVALUATION

In order to choose between alternatives, each participant must determine which one is best. Where several relevant measures are needed to describe an alternative's outcome, they must be combined in some fashion for a choice to be made. In most decisions of limited complexity this combination occurs subjectively. When the problem is complex, analytical approaches are often used as aids in combining measures and making a choice. Every analytical evaluation procedure attempts to accomplish this combining process in one way or another. Some of their common problems are discussed here.

Choosing a Common Valuation Basis

In order to compare alternatives with respect to several different measures, an evaluator must somehow find a common unit for all measures so that their impacts can be combined. In day-to-day decisions, people do this without realizing that they are doing so. For complex situations, where several alternatives are involved, subjective approaches are still often used. Analytical approaches provide a useful alternative.

Some analytical evaluation procedures convert all evaluation measures to a common unit. For example, cost-benefit analysis uses a monetary unit. Common scoring approaches give each measure a weight based on its importance and give each alternative's outcome for that measure a score. Then, all weighted scores are added to arrive at a combined score for each alternative. In these scoring approaches the common unit is the unit weight. Cost-effectiveness analysis selects a single measure (or small group of measures) as being of primary importance and compares alternatives with respect to this measure while costs are held constant (or vice versa). In this approach all other measures are determined to be of lesser importance.

Time streams of impacts are often simplified by aggregating the information from several years. The approach normally used for monetary measures is to discount each year's values to an equivalent base-year value or present value. Discounting could also be used for nonmonetary, quantitative measures, such as environmental impacts. The present value of emissions or pollution-caused health effects could be calculated.

Nonlinear Valuation Functions

All approaches that reduce the dimensionality of the evaluation problem must determine valuations for the full range of each measure. In other words they must determine, for example, if one alternative has an impact five times larger than another alternative for some measure, whether the value of that impact is five times larger for all levels of the measure. Most of the valuation functions, including the monetary ones, used in cost-benefit analysis do have the simple assumption of a linear relationship between measure and value. However, the possibility of more complex functions should always be considered. For example, very large pollution impacts might eventually cause extreme societal effects whose value would not be simple extrapolations of the effects of low pollution levels. Hence, the value of pollution impacts would increase faster than would the level of the impacts.

Nonlinear valuation functions can be used in any of the analytical evaluation procedures. In evaluation approaches that do not determine a single valuation unit, the analytical measures are in physical units and any introduction of nonlinearities in the valuations is done subjectively.

Independence of Valuations

Most analytical evaluation processes omit any consideration that the value of a measure can be influenced by the level of a second measure. Hence the value of user benefits would be unchanged by the level of environmental impacts even to travelers who were affected by the negative impacts. In other words, the valuations are simply additive.

Both cost-benefit analyses and scoring functions can incorporate interdependence of evaluation measures, in principle. The valuation of one measure would be a function of the levels of other measures. Although it increases complexity, the analytical treatment of interdependence should be considered if any of the analytical approaches is used. Assumptions of independence of evaluation measures is probably one major reason why many simple analytical evaluation procedures yield results different from subjective evaluations.

Community Valuation of Alternatives

A problem with all dimensionality-reduction approaches is that they may be used in a situation where many interests are involved and each interest may have differing relative values for the evaluation measures. Cost-benefit analysis essentially weights each interest by its market power since the monetary value of measures is used in the evaluation. Alternatives to these weightings are obvious and, in fact, can even be used with cost-benefit analysis. The net monetary benefits of each group can be weighted by some alternative factor. Nash, Pearce, and Stanley describe approaches to alter the market-power weightings (1).

Weighting the importance of various groups is also a part of any scoring process. If all interests are aggregated so that a single set of scores and weights is used, then some implicit weighting of the interests is assumed. If technical staff members of the planning agency assign the values, then their assessment of the importance or power of each interest is the basis of the weightings. Alternatively, a more heterogeneous group of community members can collectively determine the weights and scores in an aggregate scoring approach. If this group is representative of the total community, then its agreed-on values will produce decisions that may be more politically acceptable than those based on the staff members' evaluations. Even in this case, however, there is no easy way to ensure that the community participants are truly representative of all interests.

An alternative to analytical combination of different interests is to provide several calculations for the chosen dimensionality-reduction approach by using weights for the various measures that different community groups have suggested or that the planning staff estimates would represent the range of different groups. These calculations would then be used by the planning staff to determine preferred alternatives for each actual or hypothetical group. The various community interests would then negotiate a most-preferred alternative, perhaps by suggesting new compromise alternatives.

Determination of Project Worthiness

Through use of any of the various analytical techniques for dimensionality reduction, coupled if necessary with negotiation by community groups, each urban area can determine its most-desired alternative. Whether this alternative is deemed worthy to the community (i.e., preferred to the no-new-investment alternative) can be ascertained if the level and distribution of financial impacts are accurately estimated.

A more difficult problem is determining the worthiness of each area's best alternative to the nation as a whole. UMTA will compare the evaluations prepared by each urban area to make this decision. It will be easier for UMTA to compare the proposed investments if each area performs an evaluation in such a way that it can be used to determine national worthiness. No evaluation technique is trouble-free for this determination, but cost-benefit analysis is the only one that attempts to value total benefits and total costs in units that are comparable across urban areas (i.e., dollars). Since UMTA has a limited amount of grant money for each year, the highest net benefit for that level of funding can be determined with cost-benefit analysis by comparing the ratio of discounted benefits to discounted costs for each alternative and selecting the projects that have the highest ratios until the budget limit is reached.

MAJOR APPROACHES TO DIMENSIONALITY REDUCTION

All of the primary evaluation approaches can be used for dimensionality reduction and for the choice of the preferred alternative. The techniques have been discussed in great detail in other sources. The purpose of this section is to present some of the problems entailed in the use of each of the analytical techniques. However, as mentioned above, the decision to use none of these techniques also involves problems, since judgments are not error free.

Cost-Benefit Analysis

Cost-benefit analysis provides a methodology for assessing both the most-preferred alternatives and the worthiness of investments. It transforms all impacts into monetary values by use of various techniques, some of which are complex, controversial, or both.

The major problems in using cost-benefit analysis stem from valuing impact measures that cannot be bought or sold in the private market. In some cases, calculations or experiments can be carried out to find these valuations. For example, monetary values of the effect of noise have been estimated from statistical analysis of land values for areas both near and far from noise sources. However, some researchers have argued that this economic approach does not really value the physiological or psychological impacts, while economists argue among themselves whether the economic valuation has been accomplished correctly.

There are also problems about valuing qualitative impacts like aesthetics. These impacts are generally omitted since they are often small (it is thought) compared with the more easily quantified measures.

A greater problem comes from valuing potentially large, abstract impacts, such as changes in urban form. In principle, the economic impacts of various forms could be estimated (e.g., employment, consumption, income, costs of services, and travel). But, in practice, current models cannot measure these impacts accurately for the marginal differences most transportation alternatives will make in urban form. The noneconomic differences of various urban forms (such as the psychological and social effects of living at high or low densities) are, of course, even more difficult to express in monetary terms. The difficulty in valuing these types of impacts for cost-benefit analysis exists for almost all the analytical approaches to dimensionality reduction. If any trade-offs can be made between the

values of the qualitative impacts and those of the easily quantified impacts, then some attempt can be made to calculate the values of the qualitative impacts in monetary terms.

The major advantage of cost-benefit analysis is that it allows the worthiness of alternatives to be determined because all the benefits can be directly compared with the costs in identical units. Hence, whether the benefits are worth more than the costs can be ascertained. Other approaches for converting benefits and costs into similar units exist (i.e., direct assessment of utilities) but do not allow comparability between evaluations in different areas as does cost-benefit analysis.

In some cases, an urban area may find it advantageous to use cost-benefit analysis to determine the worthiness of a selected alternative but rely on other approaches to choose the most-preferred investment. This comparative evaluation of alternatives can include relevant nonmonetary or qualitative measures, while the worthiness analysis would be limited to monetary values. It is useful to distinguish between this use of cost-benefit analysis and the more common use in economics, where it is the only evaluation tool. The limited use can be called net-monetarized-benefit analysis since benefits are expressed only in monetary terms.

Cost-Effectiveness Analysis

In some situations there are only one or two measures of impacts that are important in the evaluation of alternatives. For example, some engineering or defense problems use cost-effectiveness analysis in which the single measure of effectiveness of each alternative is compared with its costs. Cost-effectiveness ratios are calculations or points plotted on a graph of cost and effectiveness values. When more than one impact measure is of importance, they must be combined in some way or a multidimensional comparison must be made. In the former case, all the problems of other dimensionality-reduction approaches apply. In the latter case, the dimensionality of the evaluation process is not reduced. In some evaluations, cost-effectiveness ratios are calculated for several important measures, and these ratios are used as the main information source for decison making instead of the simple or unitary measures from which the ratios are derived. It is relatively easy to double-count impacts when ratios are presented. Hence, it is probably preferable to use only the unitary measures when several measures must be considered.

Scoring Methods

Scoring methods are used to rate or score different alternatives with respect to each of the important evaluation measures. Each measure is given a weight, and each alternative is scored with respect to the measure. Then, the weighted scores are summed to give a total score. These approaches are popular with transportation planners because of their ease of use and relative simplicity in concept. There are, however, several problems with their use that make them no less troublesome than cost-benefit analysis.

Since there are no rules for selecting the measures to use, several measures may be used that represent the same impacts. For example, the use of transit ridership, travel time savings, and reduced automobile use will double (or triple) count any benefits from switching automobile users to transit. When each of these measures is given some importance weight, there is the possibility that these benefits can be given greater total importance in the evaluation than is intended.

The evaluation can be biased toward a particular alternative by including several measures for which the alternative excels. These measures may be double counting the same benefit or just be cleverly selected for the desired outcome. This problem can exist with any evaluation technique, either analytical or subjective, but must be especially recognized for scoring methods because of their simplicity of use.

It has been found in some uses of scoring that the weights given various evaluation measures depend on the alternatives being considered. Hence, elimination or introduction of an alternative can change the weights and hence the scores of all alternatives. Countering this condition is not easy and probably should not be carried out, but recognition of the situation is important.

Nonlinear scoring functions are sometimes preferable to the simple linear ones that are commonly used. The score or valuation for a measure in this case is not simply proportional to the physical or measured level.

A complex, nonlinear scoring approach that has found some use in aiding business decisions is direct assessment of utilities. In this case the physical measures are converted to utility values by interviewing decision makers and assessing their utility valuations. They are aided in understanding their valuations by analysts who help them develop nonlinear transformations of measures to utilities. This approach is generally not recommended in the alternatives analysis process because of the large number of interests that must take part in the various decisions.

Reduced Table of Measures

This approach relies on unitary evaluation measures but selects a small set of them as being of major importance. This set of measures should number less than six or eight per alternative in order for each evaluator to assimilate the information. The problem with this approach is, of course, that someone has to select the most-important measures and can bias the final results in so doing. If this approach is used, it would probably be beneficial to select different reduced sets of measures for different interests. In particular, geography-specific impacts should be presented to the representatives of each jurisdiction or subarea.

Evaluators choose the preferred alternative from the reduced table of measures by using their judgment. Hence, this is not actually an analytical evaluation procedure. Analytical measures are presented, but no calculations are made for either reducing or combining the variety of measures. To aid in reducing the number of measures, techniqes that can be called partial approaches to dimensionality reduction can be used.

PARTIAL APPROACHES TO DIMENSIONALITY REDUCTION

It will often be possible to analytically combine or discard some evaluation measures even when analytical approaches to dimensionality reduction are not used on all measures. Trade-off analysis can aid in this effort. Other analytical approaches can allow some alternatives to be eliminated when compared with others.

Trade-Off Analysis of Attributes

Two attributes that are not commensurable can sometimes be made so by using trade-off analysis. Instead

of trying to value attributes in an abstract valuation unit such as utility, this approach essentially values one measure in terms of another. However, it is necessary either to elicit from evaluators their judgments about the trade-offs or to have staff members develop the trade-offs. The result is a graph or function, often nonlinear, that relates the value of one measure in terms of another. The process of trading off can sometimes allow the various interests to see that one or more evaluation measures is insignificant compared with others. Another possibility is that the approach will provide information that will be used in the subjective decision making of stage 3.

Threshold Analysis

One approach to reducing information assimilation requirements is to use thresholds. For each evaluation measure, a minimum (or maximum) acceptable level is set and then each alternative is given a plus, a zero, or a minus, depending on how it meets the minimum requirements. Often, no alternative meets all of the minimum requirements. Hence, the systematic use of this approach may eliminate all alternatives. In this case, a negotiation of new threshold levels must be carried out. A variation of the approach records the pluses and zeros but uses them only to aid the subjective decision making and not to eliminate alternatives. In some cases, thresholds may be used just for a subset of the evaluation measures, for example, environmental measures that have mandated threshold values.

Comparison of Pairs of Alternatives

The large amount of data provided to evaluators can sometimes be made less imposing by presenting information about only two alternatives at a time. It is much easier to assimilate this smaller information set because some evaluation measures will be considered unimportant in each pair of comparisons. The same measures generally will not be deleted in all comparisons and hence could not be omitted in a more comprehensive evaluation process. Therefore, by use of the comparisons of pairs of alternatives a simplification is allowed that could not be introduced if all alternatives were compared together.

To ensure that potential compromise alternatives are not discarded early in the evaluation process, evaluators may want to use each alternative in more than one pair of comparisons.

Dominance Analysis

In carrying out a comparison of pairs of alternatives, it is sometimes possible to find an alternative that is obviously better than another one for every evaluation measure. The first alternative is then said to dominate the second, and the latter can be eliminated from the set of alternatives that must be considered. This approach will seldom be useful in practice, especially when several interests have contributed their own evaluation measures.

Stochastic Dominance

A variation of dominance analysis can be used if the inaccuracy and uncertainty of the forecasted values for many evaluation measures are taken into account. The uncertainty about the levels for some impacts or costs may be so great that some or all of the alternatives are not significantly different for those measures. If one alternative dominates another for all measures except those where uncertainty leads to insignificant differences, it is said to have stochastic dominance. For example, the patronage and pollution impacts of two alternatives may not be statistically different but the second, for example, is significantly better than the first in costs and all other measures. Then the second alternative stochastically dominates the first. If an area can estimate the accuracy and uncertainty of its many cost and impact forecasts, it may find stochastic dominance to be a useful tool in reducing the number of alternatives.

Marginal Analysis

It is sometimes useful to modify a particular alternative so that one or more of its attributes is made similar to those of another alternative. This may make it easier to determine which is preferred. For example, if a bus alternative and a rail alternative are being evaluated, the service levels of the bus could be set equal to that of rail and then all the other impacts could be compared, or the rail alternative could be analyzed with the number of cars and trains reduced so that the service levels were similar to common bus service. These two new alternatives would allow evaluators to determine whether the changed measure (level of service) was a major factor in the different impacts of the two original alternatives or whether other differences between rail and bus were more important.

VALUING TRAVELER BENEFITS

Probably the most important element in evaluating transportation alternatives is determining the value of benefits to the alternatives' users. These benefits must be large enough to be considered worthy of the projects' costs since they are generally the main reason for making the transportation improvement.

There are various classes of travelers, all of whom must be considered in valuing total benefits. There are, first of all, the travelers who do not change their travel choices when a particular investment is made. These include those who use the mode being improved as well as those not directly affected by the investment. Then there are travelers who change modes or destinations due to the transportation improvement. Finally there are new trips that are made because of the changed transportation situation.

An original user of the improved mode generally receives a greater benefit than a traveler in any other group. Travelers who shift modes or destinations receive less benefits because at least some of the improvement produced no benefits to them. For example, if bus service were improved by a small amount, only a few (if any) automobile users might change modes. A larger improvement would bring about an additional number of mode shifts. Travelers in this second group of shifters receive no benefits from the smaller improvement since it does not convince them to change modes. Those who make a new trip due to an investment are similar to the shifters in that they do not all benefit from the full improvement. Finally, travelers on the unimproved modes may not benefit at all from investments in other modes or they may benefit from some congestion relief.

Some confusion and controversy have surrounded this issue of user benefits because several current methods of valuing the benefits can yield differing results for the same project. These various methods exist because of a variety of definitions of what the benefits actually are. An extensive literature on the subject of user benefits has produced no agreement on the correct definition of user benefits and hence on how to value

the benefits. Since there is no single accepted approach, this paper will not recommend one (although my preference may be clear). However, it will point out some of the different definitions of benefits that are being used. It is important that each alternatives analysis report document the approach it is using to value user benefits and use it consistently.

Some of the several definitions of user benefits are the following:

1. Travel time savings,
2. Generalized cost savings,
3. Total net revenues,
4. Consumer surplus,
5. Trips by a preferred mode, and
6. Measure of service available as a surrogate for user benefit.

Each of these is discussed briefly below. One source of a more extensive discussion is the article by Haney (2). However, there is no good documentation for the planner or engineer who is not well grounded in economics.

Travel Time Savings

Probably the most common measure of user benefits for urban transportation investments in the United States is travel time savings. The time savings of all travelers are added together to produce an aggregate indicator of total user benefits. It is common in highway investment analysis to use this measure and to omit any consideration of induced travel (e.g., trips made only because of highway improvement). With transit investments, the benefits to travelers diverted from other modes are also included, even if induced travel is omitted.

Haney has shown that the use of travel time savings as a benefit measure may produce useless results when modal changes take place (2). For example, if a new system is introduced that is slower but cheaper than a competitive mode, modal shifts can occur, but the shifters will experience a travel time loss on the slower mode—hardly a benefit. The problem occurs because travel time is not a comprehensive enough measure of user benefits. Some benefits may occur from user cost savings or from greater comfort or convenience.

If all alternatives being examined would be attractive to users primarily because of their (door-to-door) speed, then evaluators may be able to get by with using only travel time savings. Even in this case, they must consider the fact that all time savings are not valued the same. Experiments have shown that very small savings (a few minutes or less per trip) are not valued as greatly as large time savings that can be put to another use. Hence, the use of a constant value of time may be too simplified an approach. I say may be because neither a single value nor a more complex valuation function is consistent with the demand model that produced the patronage estimates. This lack of consistency is what caused the problem Haney discussed, and the use of a more complex value of the time function will not solve it.

Some researchers have pointed out that the travel time savings do not actually occur but are converted into additional trips or longer trips to more preferred destinations. In other words, the experiments that control for trip destination and frequency overestimate the time savings from an improvement. The benefits from the changed travel patterns are more difficult to measure.

Generalized Cost Savings

When transportation alternatives that have both travel time and travel cost differences are being considered, an attractive approach for valuing user benefits is the use of a concept called generalized cost. A generalized cost is a weighted sum of all the costs, travel impedances, and discomforts of an alternative. Usually only travel time and monetary travel costs are included, with the travel time weighted by the value of time and added to the monetary costs. This measure of user benefit is more complete than the travel time savings alone.

It has been shown recently that even the use of generalized costs can produce results inconsistent with the demand forecasts. The demonstration involves complex algebra but shows that adding the benefits of several travelers must be done with a function that is consistent with and related to the demand model. For example, if a modal choice model is used that is a function only of the ratios of generalized costs (or travel impedances), then a linear summing of the generalized costs for each traveler will produce a consistent estimate of benefits; this approach was used frequently in the past. However, if a modal choice model of the logit form is used, a more complex function (called the log sum) must be used to consistently add up benefits.

The problem becomes more complex if a transportation investment produces changes in destination or trip frequency. Then the benefit valuation should take these into account also. Formulas can be derived, for example, that would be consistent with a logit modal choice model and a gravity model of destination choice or of any other model mixture. If these more complex forms are used, the generalized cost becomes a function of the travel impedance (or disutility) terms in all the relevant models.

The more complex forms of the generalized costs have not been used in any operational setting due to their conceptual and algebraic difficulties. However, the simpler forms may produce inconsistent results by indicating that travelers suffer disadvantages when they switch modes after a modal improvement.

Total Revenues

Another approach that is used to estimate user benefits is to suggest that a traveler's value of benefits can be estimated by his or her willingness to pay for the benefits. If he or she would pay very little for, say, a savings in travel time, then his or her valuation of the time must be low. A simplistic way of using this approach is to use the traveler's actual payment to represent his or her willingness to pay. This is wrong on two counts. First, there is the practical problem that improving travel time without raising fares would produce no benefits with this approach. More importantly, the revenues paid almost always underestimate the real benefits since most users of a facility would pay more than they are charged. (Those who would pay less than the fare are not users.)

Changes in Consumer Surplus

An approach exists for valuing the willingness to pay: it uses a concept called consumer surplus, which is the additional value a consumer receives over the price he or she pays for goods or services. It is derived from the demand model and, hence, is always consistent with the ridership predictions.

Consumer surplus calculations have not generally been used to analyze urban transportation investments

in this country for several reasons. First, since it is consistent with the demand model, it will value higher the benefits to more affluent travelers—they have a greater willingness to pay. (This occurs only if the demand model includes income as a variable. Some mode-choice models and most destination-choice models are independent of income.) Second, there are many technical arguments about the correct way to measure consumer surplus. For example, it will vary depending on the pricing policy used. A transit system that breaks even in the fare box will have different user benefits (as well as costs) from one that is partly subsidized. Further, some analysts have pointed out that the user of consumer surplus benefits will make subsidized public-sector investments appear more desirable than private-sector investments that are designed so that revenues cover costs. Hence, they suggest a more conservative measure of user benefit to be used.

The more complex generalized cost approach discussed above is similar to the consumer surplus measure since both are consistent with the demand model. However, the benefits calculated from the two approaches may not be the same.

Trips by a Preferred Mode

When a transit service is being proposed, it often may be thought that the more transit riders, the better. But, we cannot accept for the alternatives analysis that one mode or type of mode is inherently better than another. This is what the analysis has to demonstrate. Hence, simple ridership counts on some modes are not an acceptable measure of user benefits. A further problem with this approach is that it values all trips on the mode equally—those by the original users and those diverted and induced, which is not correct.

Measure of Service as a Surrogate

Due to the confusion and complexity some evaluators find with the various user-benefit measures, they occasionally decide not to use any of them and to rely, instead, on a measure of transportation service, such as highway lane kilometers, bus kilometers, or seat kilometers. The greater the level of service, the greater is the implied user benefit. This approach is not satisfactory because it cannot discriminate between alternatives that travelers would use and ones that they would not. A system operated in areas of low population density or where good competitive service already exists will not attract as many riders nor provide as many user benefits as a better-placed alternative. The evaluation must assess which alternative actually benefits users more.

CONCLUSIONS

The evaluation of urban transportation alternatives is a complex undertaking that involves several local and federal interests. The local evaluations must include a consideration of the worthiness of the preferred alternative so that UMTA can choose among the several projects that are competing for federal funds. No analytical procedure is foolproof as a tool for the local evaluation, but the total reliance on subjectivity (judgment) precludes UMTA from comparing the various evaluations prepared by the several competing local areas.

A recommended evaluation process would combine both analytic and subjective decision-making techniques. One way to carry this out would be to use the following procedure:

Step 1—staff and community selection of alternatives,
Step 2—staff development of information (analytical),
Step 3—community and local government reduction of dimensionality (analytical and nonanalytical),
Step 4—repetition of step 2 with more precision,
Step 5—repetition of step 3,
Step 6—staff determination of worthiness of remaining alternatives (analytical), and
Step 7—choice of preferred worthy alternative (nonanalytical, requires bargaining among community members).

ACKNOWLEDGMENT

This paper was prepared while I was at the Urban Institute under a contract to the Transportation Systems Center of the U. S. Department of Transportation. None of the views expressed here are necessarily those of the sponsor. Most of the thoughs presented in this paper crystallized during extensive discussions with my colleague, Kiran Bhatt. Many useful conversations were also held with Leonard Bronetsky of the Transportation Systems Center.

REFERENCES

1. C. Nash, D. Pearce, and J. Stanley. Criteria for Evaluating Project Evaluation Techniques. American Institute of Planners Journal, 41, March 1975, pp. 83-89.
2. D.G. Haney. Problems, Misconceptions, and Errors in Benefit-Cost Analyses of Transit Systems. HRB, Highway Research Record 314, 1970, pp. 98-113.

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

Analysis of Downtown People Mover Systems by Using the DPM Simulation Model

Thomas Dooley

Downtown people movers (DPMs), a class of automated transit system that operates on exclusive guideways, are being considered by many cities as a possible solution to their circulation and distribution problems. This paper describes how a discrete event-simulation model developed by the Transportation Systems Center can be used in the planning and design of DPM systems. The paper identifies the variables that can be studied and that affect system ridership, cost, and performance by the model. The key inputs, the modeling functions, and outputs of the model are discussed in the context of an example, the 1990 Los Angeles DPM system. Use of the model to determine the feasible combinations of fleet size, vehicle capacity, and operating headway to meet the Los Angeles DPM system performance goals for the year 2000 is discussed. Finally, the use of the model to examine the effects of a vehicle failure on passenger service and system operation and to evaluate three algorithms for system recovery is illustrated.

Downtown people mover (DPM) systems are a subset of automated guideway transit (AGT) systems, a class of transportation system that is characterized by unmanned vehicles operated on fixed exclusive guideways. The first generation of DPMs will consist largely of elevated systems in which proven technologies are deployed in central business districts (CBDs) and adjacent areas of larger U.S. cities.

As part of its program for transportation planning support to urban areas, the Urban Mass Transportation Administration (UMTA) has sponsored the development of special demand and supply analysis techniques. The Downtown People Mover Simulation (DPMS) model was developed to provide a tool for planners and designers:

1. To determine the sensitivity of system performance to the range of AGT design parameters (such as capacity, speed, and safe headway) and to variations in the magnitude and spatial distribution of demand;
2. To determine potential system bottlenecks created by the dynamic interaction of demand and service;
3. To examine the impact on service of infrastructure decisions that affect system operation, such as station size or guideway curvature;
4. To determine the effect of anomalies such as demand surges or equipment failures on passenger service; and
5. To examine a variety of system operating strategies.

THE DPM PLANNING AND DESIGN PROCESS

The planning process examines the feasibility of DPM in comparison with other modes by determining the ridership of alternative route alignments, station locations, and intermodal concepts together with the trip-making characteristics of the deployment area.

The DPMS model provides a tool for the linking of the planning process and the design process. Figure 1 illustrates this concept. The design process defines the detailed system characteristics that will provide the level of service that was assumed in the demand estimation and planning process. Hence, the system designer takes the guideway layout, which includes station locations and the distances between stations, and the baseline station-to-station demand matrix as given. System service characteristics used in the planning process, such as headways and travel times, represent constraints.

The system designer incorporates network constraints and demand profile assumptions into the scenario representation. Sets of system operating and hardware characteristics that affect the baseline service characteristics are defined. Table 1 lists the scenario and design variables that can be manipulated by the DPMS and the corresponding service characteristic variables that are model outputs or can be derived from model outputs. Sensitivity analysis should be performed in the areas shown in Figure 1. These include network constraints, demand projections, alternative system characteristics, and anomaly analysis. The final products of the design process include the sensitivity of service characteristics to system variables, the cost impacts of system variables, and the performance specifications. The variables addressed by the DPMS that affect cost and performance specifications are also shown in Table 1. The design variables used as model inputs will be discussed later in the context of the Los Angeles DPM example.

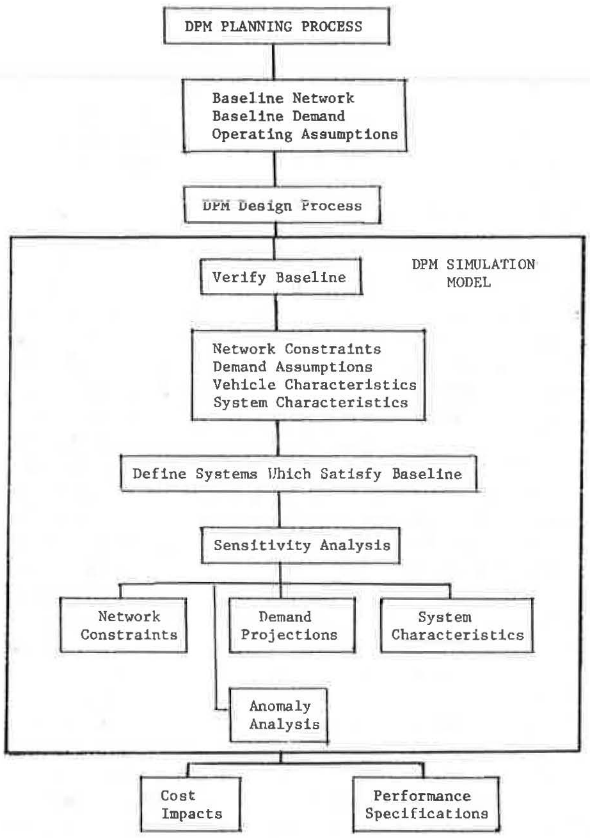
By use of the simulation input variables defined in Table 1, the system designer conducts a series of simulation experiments. The simulation model represents the movement of the vehicles on the guideway. Safe headway separation is maintained according to defined vehicle control strategies. Vehicles travel along predefined routes. The vehicles are dispatched from the stations according to the route headways and dispatch algorithms specified. Passengers arrive at their origin station at the time defined in the trip list and enter a boarding queue. When the simulated vehicle arrives, passengers deboard if the particular station is their destination and board if the vehicle has sufficient space and is headed toward their destination station. The model records the time at which individual passengers deboard and board all vehicles. The model keeps track of the current occupancy and time integral of occupancy for all passengers at each station and on each vehicle and for all vehicles at each station, on each link, and on each route of the network.

At user-specified intervals, samples are collected on passenger- and vehicle-related statistics. These sample statistics are reported at periodic intervals during the simulation process and stored for analysis during the output processing stage. The model also records statistics on each completed trip. During the output processing phase, the statistics of interest are displayed in a variety of ways, depending on user requirements. Three major types of reports can be generated: (a) a standard summary report; (b) time-series plots, histograms, summaries, or lists of sample values of key variables; or (c) station-to-station matrices for key travel time statistics.

A complete description of the model is provided in the DPMS program writeup (1). A computer-animated film has also been developed to illustrate the model's

key features. Both of these documentation sources are available from the Transportation Systems Center (TSC) of the U.S. Department of Transportation.

Figure 1. DPM planning and design process.



THE DPM MODELING PROCESS: THE LOS ANGELES DPM EXAMPLE

This section describes the modeling of the baseline Los Angeles DPM system. Variables that describe the Los Angeles DPM were supplied by the Los Angeles Community Redevelopment Agency. The variables were converted by TSC to DPMS inputs to model the Los Angeles baseline design. The Los Angeles DPM case studies are included in this paper to illustrate the use of the model.

Network Modeling

Network connectivity is modeled in the DPMS by a set of unidirectional links that are defined by in-nodes and out-nodes. The nodes are numbered, and contiguous links are defined by common node numbers. Since DPMS represents the actual movement of vehicles through defined physical areas, stations are also represented by unidirectional links. Figure 2 shows the DPMS representation of the Los Angeles DPM network.

In addition to the connectivity, the guideway links are defined by the following parameters: length (m), capacity (vehicles), average speed (m/s), nominal travel time (s), and minimum safe headway (s). The two most important variables for simulation are the nominal travel time and the minimum safe headway. These variables are used to schedule the completion of events in a vehicle's traversal of the guideway link. The minimum safe headway defines the time during which no other trains may enter a link. The nominal travel time defines the time required to complete the link travel event. If a train cannot enter a link due to capacity, headway, or failure conditions, the train is queued until the condition is cleared. The list below shows the Los Angeles network parameters.

Los Angeles DPM Baseline

1. Demand—evening peak period = 2 h, peak hour demand = 9200 passengers/h, and 4600 passengers are carried during the peak 20 min;

Table 1. Simulation variables related to ridership, cost, or performance specifications.

DPMS Variable	Ridership Variable	Cost Variable	Performance Specification
Input			
Station-to-station demand rate by time interval	In-vehicle and wait times		
Guideway speed limits	In-vehicle time		
Station configurations	In-vehicle time	Capital cost	
Station-to-station distance	In-vehicle time	Capital cost	
Vehicle speed	In-vehicle time	Vehicle cost	
Vehicle capacity	Wait time	Vehicle cost, guideway cost	Square meters per passenger, standee-seated ratio
Vehicle loading rates	Wait time		Deboard and board rate
Vehicles per train	Wait time	Station cost	
Minimum safe headway	Wait time	Vehicle cost	
Route headway	Wait time		
Transfer points	Transfer time		
Minimum and maximum door open times	Wait and in-vehicle times		Time to unload vehicle
Failure conditions	Wait and in-vehicle times	System cost	Maximum delay times
Output			
Wait-time distribution	Wait time reliability		Maximum wait
In-vehicle time distribution	In-vehicle time reliability		
Maximum station queue		Station cost	Platform size
Load factors	Fare		
Vehicle kilometers	Fare	Operating and maintenance cost	
Passengers served	Fare		
Peak vehicles	Fare	Capital cost, operating and maintenance cost	

Figure 2. Los Angeles DPM network.

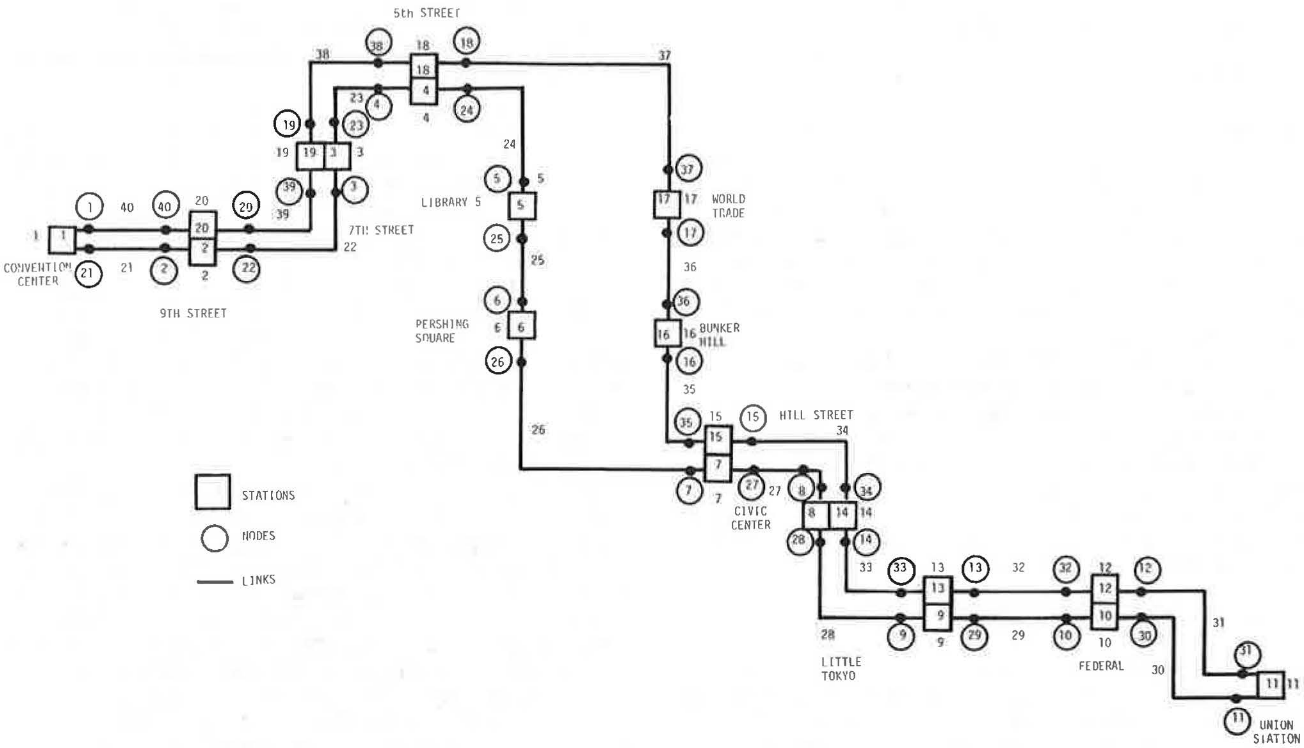
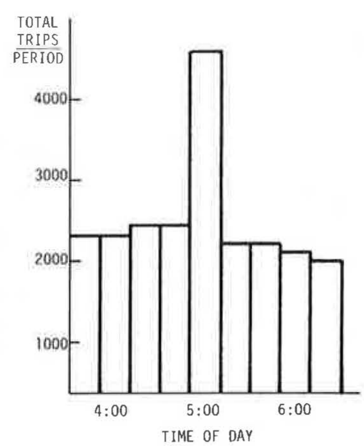


Figure 3. Los Angeles DPM 1990 evening peak profile.



2. Network—average speed on guideway including station input ramps and crossovers = 8.76 m/s, station configurations (all on line) are two end-of-line stations with inbound crossovers, four single-side one-way stations, and seven split-platform stations; average station-to-station distance = 436 m, network connectivity is closed loop, and door is open a minimum of 5 s and a maximum of 55 s;

3. Vehicle characteristics—vehicle speed = 8.76 m/s average, vehicle capacity = 41, number of seats per vehicle = 17, passenger loading rate per vehicle = 0.8 s/passenger, vehicle length = 8 m, and vehicle dispatch = midpoint; and

4. System characteristics—one route that stops at all stations, route headway = 106 s, vehicles per train = 4, transfer points = 2 (one at 10 s and one at 108 s), and average dwell time = 24.5 s (15 s with door closed + 5 s minimum + 4.5 s expected variable).

Station Modeling

In the DPMS, stations are modeled as a set of links and a set of events that occur on each link. Events include headway and travel times, which are similar to the guideway link events. In addition, deboard, board, and launch events can be specified. The deboard event removes passengers from vehicles and computes the time to deboard each vehicle in a train. The board event places trips from the station boarding queue in vehicles that have sufficient capacity and go to their destination and computes the time to board all passengers on each vehicle. Deboarding and boarding times may be constant or may vary between a minimum and maximum, depending on the number of passengers. The launch event is used to determine the time the vehicle is dispatched as a function of the current time and the dispatching algorithm in use. The modeling of the Los Angeles DPM stations was summarized in the above list.

Trip-List Generation

The DPMS generates a time-ordered sequence of passenger arrivals, called a trip list. The trip list contains the time of arrival and the origin and destination times of each trip. The trip list is generated from a station-to-station demand matrix and a set of scale and interval values. The model uses this information to determine an average arrival rate for each station pair. The trip list is generated as a series of Poisson arrivals based on these rates and a specified random number. Figure 3 shows the Los Angeles evening peak demand profile as a set of 20-min intervals with varying demand magnitudes.

Vehicle Characteristics

The vehicle characteristics that are specified include

capacity, length, speed, number of seats, and load and unload rates. The above list shows these variable values for the Los Angeles DPM baseline.

Network Operations

The final step in modeling the baseline Los Angeles DPM system is to specify the system operating conditions. These include a definition of the route (set of station stops), the operating headway on the route, the definition of which routes passengers board at each station for each destination, the points at which passengers transfer, the time it takes to transfer at each transfer point, the dispatching policy used, and the number of vehicles per train.

The model can calculate the desired route headway from the number of trains or vice versa based on the nominal round-trip time. Since the designer works within the constraints of the planned system, the route headway is usually input. If a given train capacity is known, the model can calculate the route headway and the fleet size based on a specified demand interval, such as the peak 20 min. The above list shows these variable values for the Los Angeles DPM baseline.

Los Angeles Baseline Results

Figure 4 shows one of the system summary printouts from the model, and the "base" column in Table 2 shows the significant statistics of this run. Passenger waiting times and trip times are key performance measures. The trip time is defined by the network speed constraints; the wait time is a measure of the service provided. The worst-case station (number 6, Pershing Square) had a maximum wait of 184 s. Figure 5 shows a time-series plot of maximum wait times at Pershing Square. This plot shows stable service with a slight degradation during the peak period. This plot also shows that vehicles are being dispatched from this station evenly, since each asterisk at a point greater than zero represents a dispatch in that 60-s sampling interval. We can also verify this stability from the minimum and maximum dispatch times shown in Table 2.

Other key statistics from Figure 4 include the system average load factor (0.315) and the maximum load factor on any link (1.00). The worst-case link is after the worst-case station. The queuing at station 6 is caused by a combination of high demand and trains that arrive at near capacity. The vehicle speed in the network and the ratio of planned to actual travel time vary due to the effects of variable dwell times.

Other key statistics shown in Table 2 include the 95 percent wait time, which is derived from a sort of the log of all completed trips; the standard deviation of trip time, which is derived from the trip log for each pair of stations and indicates predictability of service; the maximum door-open time, which indicates station dwells and affects minimum route headways; and the maximum number of waiting passengers at the worst-case station, which can be used to evaluate station platform capacity.

This simulation indicates that the system design described in the preceding list performs almost as ex-

Figure 4. System summary statistics.

DPMS STANDARD REPORT 2----SYSTEM SUMMARY STATISTICS

SYSTEM-WIDE MEASUREMENTS

	TOTAL	AVERAGE	MINIMUM	MAXIMUM
VEHICLE FLEET SIZE	-	56.000	56.000	56.000
SEAT CAPACITY	-	952.000	952.000	952.000
SEAT AVAILABILITY	-	335.753	101.832	450.340
VEHICLE METERS TRAVELLED	2365500.00	-	11116.000	26096.000
VEHICLE LOAD FACTOR	-	0.315	0.0	1.000
NUMBER OF PASSENGERS IN SYSTEM	-	879.233	606.000	1679.000
PASSENGER METERS TRAVELLED	29887632.0	-	108705.000	579067.000
PASSENGER WAIT TIME (SEC)	-	57.016	0.0	183.900
NUMBER OF PASSENGERS WAITING	-	112.165	37.000	324.000
PERCENT COMPLETED TRANSFERS	17.717	-	2.857	41.860
NOMINAL TRAVEL TIME / ACTUAL TRAVEL TIME	0.951	-	0.553	1.116
VEHICLE SPEED IN NETWORK--INCLUDING STATION TIME (M/SEC)	6.299	-	4.484	8.257
VEHICLE SPEED ON GUIDEWAY--EXCLUDING STATION TIME (M/SEC)	9.753	-	8.515	10.675

STATION MEASUREMENTS (BY STATION)

	TOTAL	AVERAGE	MINIMUM	MAXIMUM
--STATION 1--				
NUMBER OF VEHICLES	-	2.099	0.0	4.000
NUMBER OF VEHICLES QUEUED:				
INPUT RAMP	-	0.0	0.0	0.0
INPUT QUEUES	-	0.0	0.0	0.0
DOCKS	-	0.0	0.0	0.0
OUTPUT QUEUES	-	0.0	0.0	0.0
OUTPUT RAMP	-	0.0	0.0	0.0
STORAGE	-	0.0	0.0	0.0
VEHICLE TIME IN STATION (SEC)	-	55.669	51.500	69.400
NUMBER OF PASSENGERS:				
ENTERING	1424.000	-	4.000	30.000
EXITING	3515.000	-	0.0	100.000
TRANSFERRING	0.0	-	0.0	0.0
WAITING	-	10.223	0.0	47.000
PASSENGER WAIT TIME	-	53.960	0.0	112.900
VEHICLE LOAD FACTOR--IN	-	0.315	0.140	0.610
VEHICLE LOAD FACTOR--OUT	-	0.128	0.024	0.323
--STATION 2--				
NUMBER OF VEHICLES	-	0.829	0.0	4.000
NUMBER OF VEHICLES QUEUED:				
INPUT RAMP	-	0.0	0.0	0.0

Table 2. Los Angeles DPM year-2000 sensitivity analysis.

Characteristic	Base	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6
System							
Headway (s)	106	106	106	88	75	59	45
Train capacity	164	180	240	180	160	120	90
Number of trains	14	15	15	18	21	27	35
Evening peak demand	15 321	21 891	21 891	21 891	21 891	21 891	21 891
Evening peak time (h)	2	2	2	2	2	2	2
Throughput (places/h)	5570	6113	8151	7361	7680	7322	7200
Performance							
System maximum wait (s)	184	558	153	164	150	167	169
Station 6, 95 percent wait[a] (s)	129	404	122	114	95.2	99.7	103
System 95 percent wait[a] (s)	120	201	121	104	89.1	75.4	66
Trip time							
Mean (s)	400	433	423	413	409	399	394
Maximum SD[b]			38.7	57.8	25.8	29	37
Dispatch time							
Minimum	94	94	87	78	64	52	41
Maximum	119	119	125	102	87	67	51
Maximum door open, station 11	37	45	49.5	40.6	36.6	39	22
Maximum waiting passengers, station 6	77		102	103	95	100	116

[a]95 percent wait = 95 percent of trips wait less than this time.
[b]Maximum standard deviation of travel time between any pair of stations.

pected. However, it is significant to note that station 6 did experience some queuing. This potential problem would not have been predicted without the DPMS. The static peak link-load analysis indicated that a train capacity of 154 would be sufficient. Even though the simulation experiment was run by using a train capacity of 164 (a 7 percent increase), some passengers were forced to wait for another train.

One run of the simulation experiment, as discussed here, is not sufficient to base design conclusions. Since many random events interact with one another in the simulation, several experiments that use different random number seeds should be run to obtain the desired level of confidence. The experiments discussed in this paper are illustrative of the type of information available from the model.

LOS ANGELES BASELINE SENSITIVITY ANALYSIS

Once the baseline system has been modeled, the DPMS can be used to explore a range of demand and system characteristics and possible system anomalies. This section shows how the model was used to determine the combinations of route headway and vehicle capacity that would meet the baseline wait-time goals for an estimated demand of 100 000 persons/day in the year 2000 versus 72 000 persons/day in the baseline year.

In the absence of other information, we assumed that the spatial distribution in the evening peak would be the same in the year 2000 as in the baseline year, 1990. The increase in demand will be modeled by a change in the scale factor in the demand profile. The scale factor for each interval was multiplied by 1.38 to generate the trip list for the year-2000 case. This was derived from an increase from 72 000 to 100 000 persons/day.

The only other changes made to the baseline case were the removal of the switchback constraint at the end stations and an increase in the expected door-open time at the stations from an average of 10 s to an average of 15 s because of the higher demand. This change resulted in a total nominal travel time around the network of 1593 s, an increase of 109 s from the baseline case.

The simulation experiments examined system operation at the following nominal headways: 45, 59, 75, 88, and 106 s. Train capacities ranged from 80 to 240 passengers. To evaluate the performance of these various combinations of throughput, several measures of passenger wait time were calculated. These measures include the average, maximum, and 95 percent wait times. These measures were computed for the entire system and for each station from the station-to-station matrices, which were derived from the log of trips completed during the evening peak.

Results of Year-2000 Demand Study

Table 2 presents the results of a set of simulation runs for the demand in the year 2000. The left-hand column lists the key system and performance characteristics. The base run lists the statistics of the baseline (1990 demand) scenario discussed previously. Run 1 shows the results of using essentially the baseline system to try to serve the year-2000 demand. Even with an increase in train capacity from 164 to 180, the system is clearly saturated. Run 2 shows the results of a simulation experiment that uses a train capacity of 240 and a headway of 106 s. The statistics for this run indicate service comparable to the baseline case. The next four columns (runs 3, 4, 5, and 6) show the results of simulation experiments that use different headways and train capacities that offer comparable levels of service. In all cases the system average wait time was equal to one-half of the headway, and the average wait time at the worst-case station was slightly higher.

The performance characteristics for runs 2-6 are approximately equivalent in terms of the worst-case station (and system) maximum and 95 percent wait times. The system 95 percent wait time (as well as the system average wait time) decreases with decreasing headway. The next two lines show the average trip times for each run and the maximum of standard deviations for the station-to-station trip times. Trip time is measured from time of arrival at a station to time of completion. The average trip time is not affected significantly by the headway changes. In fact, the change shown is approximately one-half the headway change, as would be expected. The worst-case standard deviation of trip time is a measure of the predictability of service. In all cases shown here, this number is low. However, a decrease in the headway does not guarantee an improvement of this statistic. The average dispatch time between vehicles was equal to the planned headway in all cases. The range of dispatch times is shown here. These ranges indicate that the dispatching algorithm used by the model worked well. The maximum door-open time statistic at the end station can be used to evaluate the feasibility of using switchbacks at lower operating headways. The maximum number of waiting passengers can be used to evaluate the station design capacity.

Figure 6 shows the train capacities and headway com-

Figure 5. Time-series plot of maximum wait time at station 6.

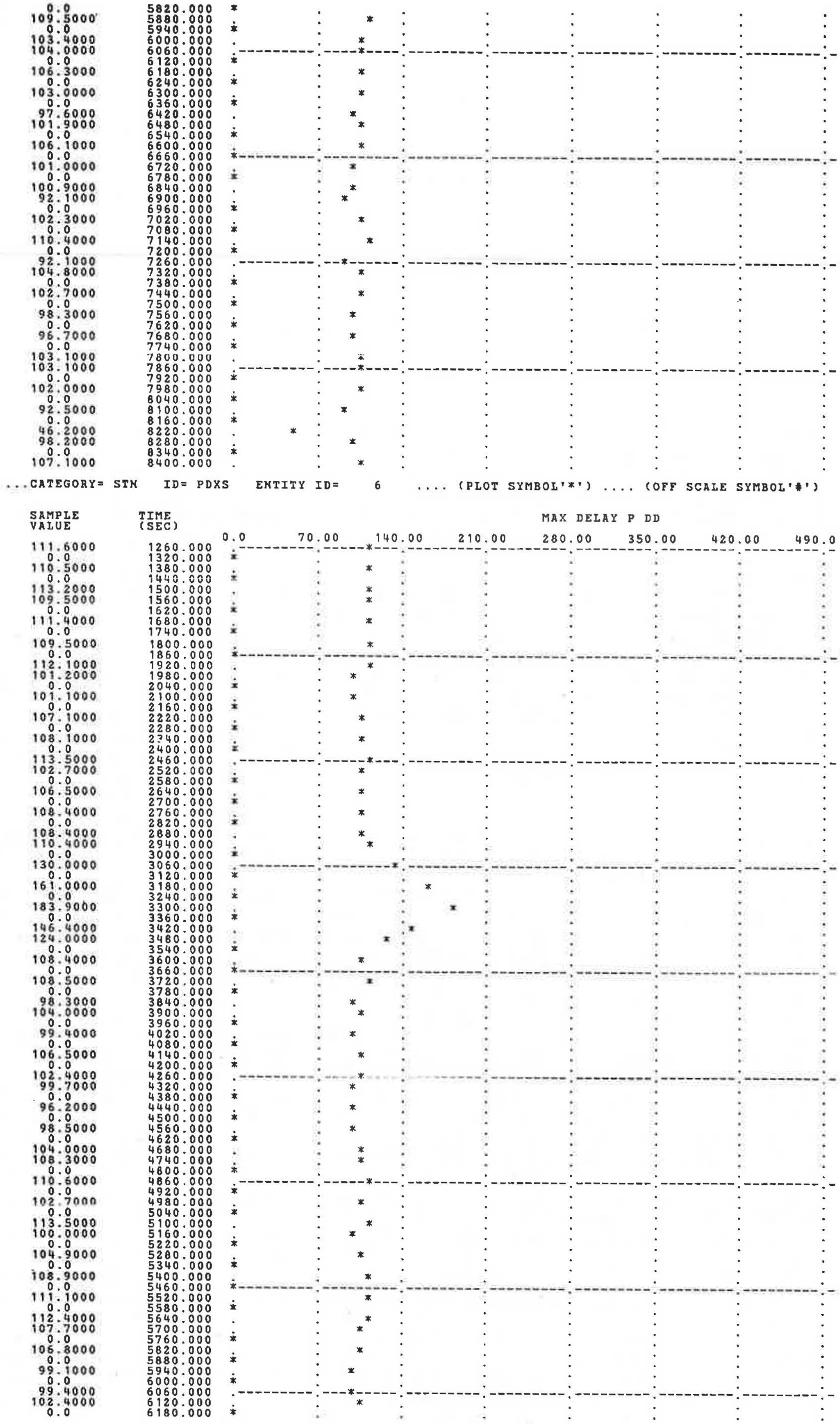


Figure 5. Continued.

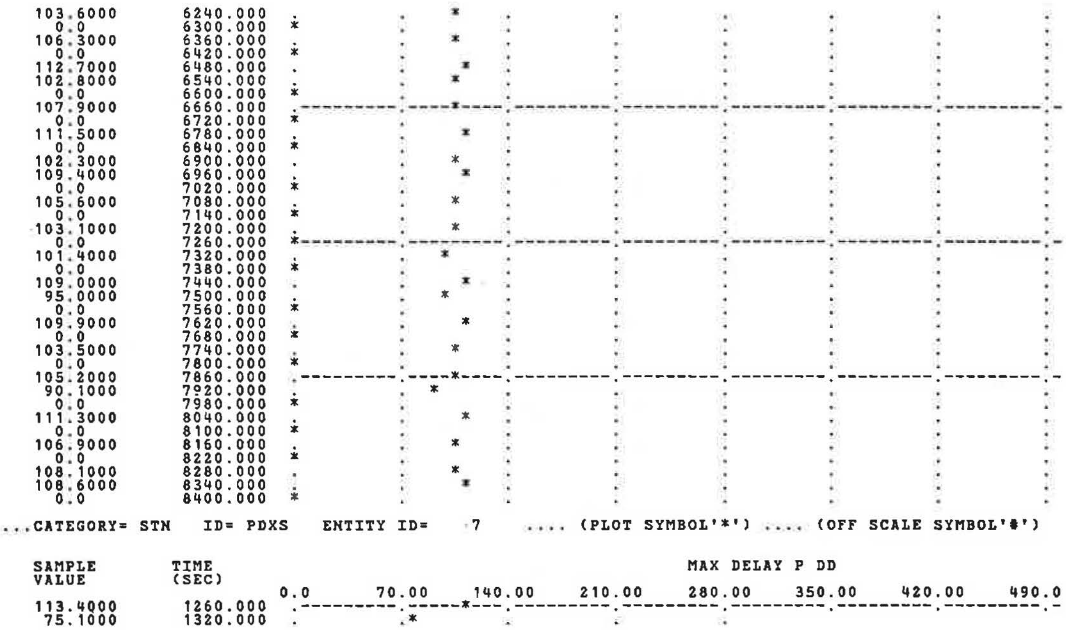
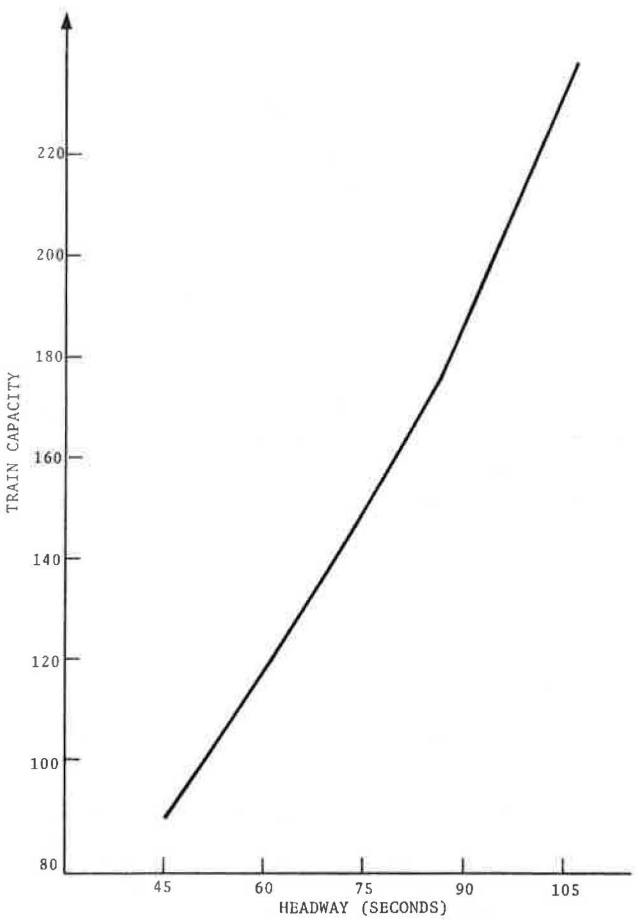


Figure 6. Headways and capacities needed to produce 95 percent wait time of 2 min at station 6 on the Los Angeles DPM in the year 2000.



binations needed to meet the 95 percent wait-time goal (2 min) at the worst-case station. These results were derived from a series of simulation runs.

Modeling System Anomalies

The DPMS was used to model system response to a vehicle stoppage on the link between stations 5 and 6 by inputting the location and time of the failure and recovery events. The failure is assumed to occur at the beginning of the peak 20 min and last for 5 min.

A vehicle stoppage is represented in the DPMS by a failure of a link exit, which then prohibits vehicles from leaving the specified link. The first vehicle that reaches the end of the defined link queues, as do all following vehicles. When the failure is removed, the guideway link model within the DPMS prompts the queued vehicles to start moving again. All passengers remain in the system and the failed vehicle is considered operational once the failure is removed.

Three failure experiments were conducted by using the same failure scenario and different dispatching algorithms. The dispatching algorithms affect the time the system takes to recover from the failure. The first algorithm dispatches vehicles as fast as they are loaded if they are far behind schedule. The second algorithm dispatches vehicles one predetermined headway behind the preceding vehicle. The third algorithm dispatches vehicles midway between the predetermined headway and the time vehicles are ready to be launched.

Analysis of the Failure Scenario

Table 3 presents the results of the simulation experiments. The system and performance characteristics shown are similar to those shown in Table 2. The first column shows the baseline results. The second column shows the results of the baseline system by using the as-soon-as-possible dispatching scheme. The increase in system maximum and system 95 percent wait times are dramatic, as expected. The average wait time over the 2-h peak period increases by only 15 s. The station that has the worst average and 95 percent wait times is now station 17. The range of dispatch times indicates the possibility of bunching.

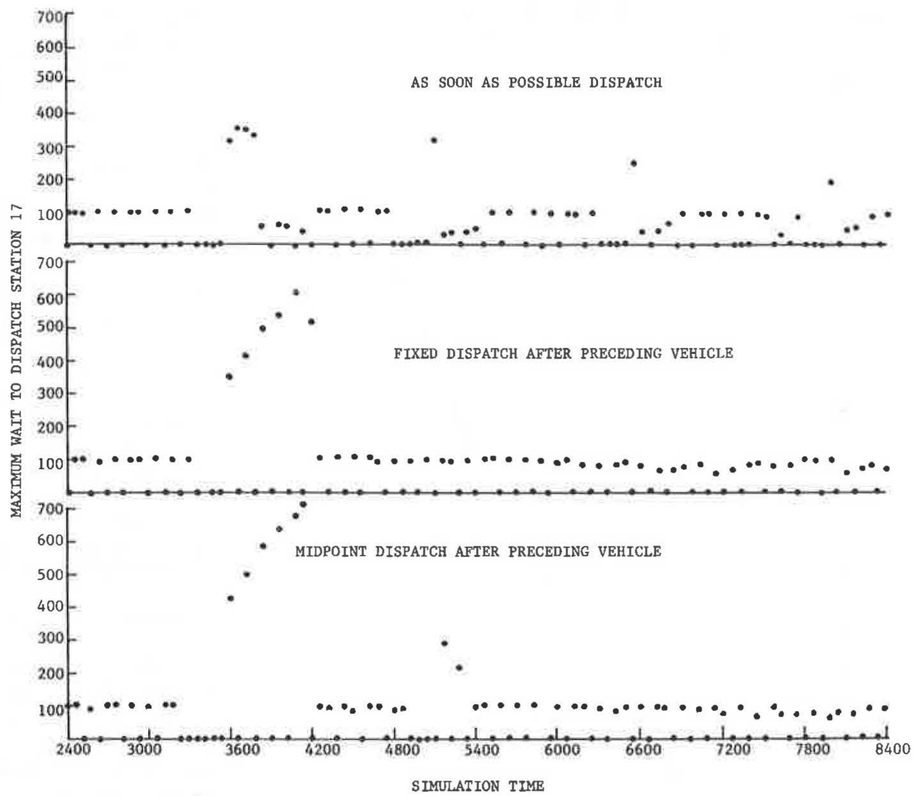
The results of the alternative dispatch algorithms can be explained by looking at Figure 7. This figure is a reproduction of the time-series plots of passenger maximum wait time at station 17. Samples have been taken every 60 s. A zero wait time means that no passengers were dispatched; a positive time indicates that at least one vehicle was dispatched in that interval.

The first algorithm works best in this case, because it moves vehicles to the stations downstream of the fail-

Table 3. Los Angeles DPM failure scenario analysis.

		Dispatch		
Characteristic	Baseline	As Soon As Possible	Fixed After Preceding Vehicle	Midpoint After Preceding Vehicle
System				
Headway	106	106	106	106
Train capacity	164	164	164	164
Number of trains	14	14	14	14
Evening peak demand	15 321	15 321	15 321	15 321
Evening peak time (h)	2	2	2	2
Failure location		Link 25	Link 25	Link 25
Failure time (s)		300	300	300
Performance				
System average wait (s)	57	72	76	86
System maximum wait (s)	183	395	615	734
System 95 percent wait (s)	121	194	208	275
Worst station	6	17	17	17
Worst station average wait (s)	62	92	112	141
Worst station 95 percent wait (s)	128	256	403	509
Dispatch time				
Average	106	106	108	110
Minimum	94	20	106	63
Maximum	119	332	367	462
Passengers with excess travel time of				
300-600 s	0	37	288	351
More than 600 s	0	0	47	53

Figure 7. Los Angeles DPM baseline to 5-min failure on link 25.



ure quickly, which is important when the failure occurs in the peak hour. The effects of bunching are seen in the periodic peaking of the wait times later in the simulation. The second plot shows that the vehicles are spaced nicely but that they provide insufficient throughput to handle the large queues during the peak. This algorithm might be more effective for off-peak stoppages. The third algorithm moves the vehicles downstream only slightly faster (the sixth vehicle arrives sooner). Its poor performance results mainly from the fact that the initial gap between dispatches after the failure was 420 s, rather than 300 s for the first two alternatives. Another set of failure runs was made by using the same failure scenario (time and place) but a different random number. They showed that the midpoint dispatching algorithm was slightly more effective than the fixed-interval algorithm, but both were inferior to the first algorithm. From this result, we can conclude that the best algorithm for this failure scenario is the one that dispatches vehicles as fast as possible and that bunching is a secondary effect.

SUMMARY

The results presented illustrate the capability of the model to represent a variety of demand, network, system operation, and system event scenarios. The model-

ing of these changes from the baseline case is quite simple and the model provides the analyst with the information necessary to understand the results.

ACKNOWLEDGMENT

The development of the DPMS model was jointly sponsored by UMTA's Office of New Systems and Automation and Office of Planning Methods and Support. The specifications for the model were developed by me, Art Priver, Don Ward, and Sy Prensky of TSC and Gran Paules of UMTA. The software and documentation for the model were written by John Duke of General Motors Transportation Systems Division and Mark Handelman and Al Melgaard of IBM Federal Systems Division. The modeling of the Los Angeles DPM was done by me in conjunction with Gill Hicks and Foster Needles of the Los Angeles Community Redevelopment Agency. The year-2000 demand analysis was done in conjunction with George Scheck of Kentron Hawaii, Ltd., and Art Priver. The data and conclusions presented here are mine and do not represent the position of either UMTA or the Los Angeles Community Redevelopment Agency.

REFERENCE

1. Systems Operation Studies for Automated Guideway Transit Systems, Downtown People Mover Simulation (DPMS) Program Write-Up. General Motors Transportation Systems Division, Warren, MI, Rept. EP-79020, Feb. 1979.

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

Generating Alternatives for Alternatives Analysis

William S. Herald

Alternatives analysis is the planning process mandated by the Urban Mass Transportation Administration for the assessment of major transit investments. The alternatives analysis process is a means of ensuring comparability between rapid transit planning studies across the nation. Up to now, the focus of attention has been on the results or products of the process. Interest has centered on the selection of a recommended alternative and its costs and impacts. This paper examines an earlier stage in the planning process that has critical importance in the validity of alternatives analysis studies. The basic concern is with the ways that alternatives are derived and described. If alternatives are the central feature of the process, we should know more about what they are and where they come from. The investigation reviews a group of alternatives analysis reports to establish the state of the art in generating alternatives for major transit studies. Ten potential imputs to alternatives generation are identified. In addition, the paper assesses the use of specific techniques or methodologies for the generation of alternatives. Specifications for alternatives and the properties of alternative sets are reviewed. The paper includes an examination of the ways that transportation system management and baseline alternatives have been defined and used in past studies. Conclusions on the state of the art in alternatives generation and its expression in alternatives analysis studies are presented as the results of the investigation.

Alternatives analysis is the process mandated by the Urban Mass Transportation Administration (UMTA) for planning major rapid transit facilities. As the name implies, the central feature of this planning process is the comparative assessment of the costs and impacts of a set of alternative configurations of technologies and services. Paradoxically, although alternatives occupy a central position in this planning process, little attention has been focused on the alternatives themselves. The major focus of methodological interest has been on techniques to predict the impacts of alternatives on the urban transportation system and the environment.

This paper is concerned with such issues as where alternatives come from and what makes up an alternative. The review of the transportation literature for this investigation suggests that these basic issues are seldom articulated and that alternatives development may be an activity characterized by pervasive assumptions and a lack of structural approach. The purpose of this paper is to examine the state of the art in the identification of alternatives. We are interested in the answers to a number of questions:

1. When are alternatives generated in the planning process?
2. What are the inputs to alternatives generation?
3. What techniques are used to identify alternatives?
4. What characteristics have been used to define and describe alternatives? and
5. How have baseline and transportation system management (TSM) alternatives been considered?

To assess the actual experience of the alternatives analysis process with respect to the generation of alternatives, 15 sources were read and evaluated (1-15). These sources were a mixed collection of complete alternatives analysis final reports, single volumes from a series of rapid transit engineering studies, supplementary reports submitted in response to UMTA questions, and draft reports. Altogether the group of documents reviewed represents 11 alternatives analysis efforts. This represents roughly a 30 percent sample of the universe of 35-40 alternatives analysis studies identified in this investigation. Although this is not a random sample, it is assumed that the reviewed studies are reasonably typical of past and current alternatives analysis experience.

Some important limits on the analysis must be noted. The first of these is the evolutionary nature of the alternatives analysis process. Because the process has developed over time, each application has been treated somewhat differently. Thus, all the studies cannot be expected to be similar. Also important is the fact that the alternatives analysis process was imposed on several rapid transit planning efforts in midstream. It can be expected that these studies show significant differ-

ences from alternatives analysis studies that began at the beginning.

A final limit on the analysis is the problem of documentation. A frequent question raised was whether a particular issue was never addressed in the actual study or simply not documented in the report reviewed here. There are three components to this problem. The first is a potential authorship bias of current study participants against previous work and the contributions of others. This understandable bias may be reflected in the documentation of the study effort in the final report. The second component is that past alternatives analysis reports have been primarily oriented toward convincing the reader that the recommended alternative represents the best course of action. As such, the focus of the documents is outward, toward convincing the audience, not inward, toward the study process itself. As a result, the documents are not simple recordings of the alternatives analysis as it actually unfolded. The third component of the documentation problem is the editorial process. Alternatives analysis studies generate enormous volumes of data, information, and analysis. In preparing a final report, an author is obliged to be selective, to make editorial decisions about what is included or excluded or about what is emphasized and what is deemphasized. Although selectivity is necessary to the efficient exchange of information, the need to communicate effectively may obscure elements of interest to this investigation.

Finally, it should be pointed out that the intent of this paper is not to evaluate individual planning efforts or the alternatives analysis process as a whole. Differences in treatment are noted to shed light on the ways that alternatives are developed, not to label a particular study as good or bad. Indeed, the review of completed studies indicates that overall the studies were very similar and provided valuable information to decision makers in each metropolitan area.

METHODS FOR GENERATING ALTERNATIVES

Two key elements are of concern regarding the methods for generating alternatives:

1. Where in the planning process does the generation of alternatives occur, and what are the necessary inputs? and
2. Once the proper inputs are available, how are alternatives actually identified in a procedural or analytical sense?

Current practice must be examined with each of these elements in mind. The consideration of past performance of alternatives analyses with respect to the identification of alternatives must carefully note the caveats outlined earlier in the paper. The studies reviewed are believed to be representative of past practice but include only those final and draft submissions available for analysis. Moreover, assessment and description of a particular study's characteristics in these areas are based solely on the documentation provided in the report made available for review.

Timing and Location of Alternatives Identification

Review of the studies selected provides little information about when or where the identification of alternatives has taken place. This fact is significant in itself. Based on the reports reviewed, past alternatives analysis practice has had little awareness of the study as a process. None of the studies reviewed included a flow chart or diagram to indicate the process of alternatives analysis as it was understood and applied at that time and in that context. This lack of process or internal awareness is characteristic of many aspects of the studies with respect to the identification of alternatives. In general, the concern for methodology or a process that surfaces in these reports is directed toward the activities of impact prediction and evaluation. No doubt this lack of information on the timing and location of alternatives identification is attributable to both a lack of awareness and an editorial decision on material most appropriate to the final report.

The survey of alternatives analysis studies indicates that the timing of alternatives identification conformed to a roughly similar conception of national planning practice. This is to say that an after-the-fact diagram of the work flow in each study would bear some resemblance to a typical ideal planning process like those discussed in the transportation planning literature. As expected, the identification of alternatives has occurred at or near the commencement of the study. As such, it is generally among the first real future-oriented activities, often taking place concurrently with or following a documentation of existing conditions, definitions of goals, and projection of future conditions and needs. In most cases, however, the initial identification of alternative activities actually preceded the commencement of alternatives analysis. This was the case in 11 studies that used alternatives carried over from previous studies. The extreme example of this phenomenon was in Pittsburgh, where the identification of alternatives was fully accomplished before the consultant began work on the alternatives analysis study.

Inputs to Alternatives Identification

The first step in this investigation was to compile a comprehensive list of factors that can influence the identification of alternatives. Accordingly, a search of the relevant planning literature and alternatives analysis material produced the following list of potential inputs to the identification of transportation alternatives:

1. National goals and objectives,
2. Local and regional goals and objectives,
3. Assessment of needs,
4. Forecasts of future conditions,
5. Local opportunities and constraints,
6. External requirements and constraints,
7. Carry-over alternatives and prior analysis,
8. Catalogue of technologies,
9. Assessment of operational feasibility, and
10. Public involvement.

This list of potential inputs was used in an analysis of the selected alternatives analysis studies. The purpose of the analysis was to determine the incidence of use for each potential input. The matrix in Figure 1 displays the results for the sample of completed studies. Three different states are identified: documented presence of an input, implied use, and no apparent involvement.

National Goals and Objectives

This category of input was broadly defined for the purpose of this review to include not only consideration of national transportation and urban development goals but also UMTA-promulgated objectives and procedural guidelines for alternatives analysis. As a result, almost all the selected studies showed some evidence of national

Figure 1. Use of potential inputs to the identification of alternatives in past alternatives analysis studies.

Study	Potential Inputs to Alternatives Identification Process									
	National Goals and Objectives	Local Goals and Objectives	Assessment of Needs	Forecasts of Future Conditions	Local Opportunities Constraints	External Requirements Constraints	Carry-Over Alternatives	Catalogue of Technologies	Operational Feasibility	Public Involvement
Buffalo	●	●	—	—	○	○	●	○	—	●
Denver 1975	●	●	●	—	●	○	—	●	●	○
Denver 1976	●	○	—	—	●	○	●	○	—	●
Honolulu 1975	○	●	●	●	●	—	●	●	—	○
Honolulu 1976	●	●	●	●	●	○	●	●	●	○
Los Angeles	○	○	○	○	○	○	●	○	—	○
Miami	●	○	○	—	●	○	—	●	—	●
Pittsburgh	—	●	—	—	○	—	●	●	●	○
San Juan (Draft)	●	●	●	●	●	○	●	●	—	○
Washington F	●	○	—	—	○	○	●	—	—	●
Washington J-H	●	○	—	—	○	○	●	—	—	●
Washington K	●	○	—	—	○	○	●	—	—	●
West Shore (Draft)	○	—	—	—	●	○	●	●	●	●

● Documented Presence
○ Implied Presence
— No Apparent Involvement

goals as an input to the identification of alternatives. Nine reports provided documented presence of a consideration of national aims and three additional studies were judged to show at least an implied consideration of national objectives.

For the most part, the presence of national goals as an input to alternatives identification was in the form of UMTA guidelines and stipulations related to the alternatives analysis process. Very few of the completed studies included a discussion of national transportation goals and their significance to this effort. The 1976 version of the Honolulu study was noteworthy in its consideration of national aims and what meaning they had to the alternatives analysis study. Even this consideration, however, failed to include an actual linkage of national objectives to the process of identifying alternatives.

The Pittsburgh study is seen as reflecting no apparent involvement of national goals with the identification of alternatives. Ironically, however, the Pittsburgh South Hills report (6) includes the most complete discussion of UMTA specifications for a successful proposal. The distinction made here is that the consideration of national aims in this case clearly follows after the process of identifying alternatives and has no demonstrated relation to it. Instead, the view is taken in the study that the recommended alternative will be refined in the course of the planning process to reflect those national aims. The distinction is, therefore, a narrow one between the concepts of identification and refinement.

Local and Regional Goals and Objectives

This category of input was defined as any statement of local or regional aims or objectives concerning urban development and transportation services. Twelve of the reports reviewed incorporated either explicit or implied involvement of local goals in the identification of alternatives. The draft version of the West Shore alternatives analysis (13) showed no apparent involvement of local objectives in the generation of alternatives.

The treatment of local goals and objectives in most of the sample reports was much stronger than the consideration of national goals. Generally, detailed lists of transportation objectives and policy statements on desired attributes of future development were excerpted from the long-range plan or other goal-oriented planning efforts. Typically, however, these goal statements showed little direct relation to the identification of alternatives for analysis. The local objectives were often vague statements, designed to indicate a general direction of betterment. Despite the apparent gap between these ideals and the specific dimensions of alternatives, few of the studies provided any transition or linkage. The common practice was to include a section on local goals, followed by one on identification of alternatives. The reader is left to infer a relationship. Thus, for the most part, it is only possible to say that the alternatives identified responded to local goals insofar as the alternatives were perceived to be improvements or representations of a better transportation infrastructure.

Two studies, however, should be noted for different treatment of local aims. The discussion of local goals in Buffalo was limited to a typical presentation of goal statements from another source. The study was somewhat unique, however, in that it developed objectives for the alternatives analysis study and then related the identification of alternatives to those objectives. This reflected an awareness of the meaning and role of alternatives that was lacking in other reports. Honolulu's 1976 report (5) included an extensive discussion of local and regional goals. This discussion was then focused on the planning process and used to derive specific criteria categories useful for the identification and evaluation of alternatives in the alternatives analysis study.

Assessment of Needs

One element that frequently appears as an input to the identification of alternatives in theoretical presentations of the planning process is an assessment of needs. The

logic of needs as an input is clear: Future conditions will create certain demands that cannot be met by existing facilities and operations. Therefore, alternatives should be formulated to address those needs. The review of sample alternatives analysis reports indicated that needs assessment has played a minor role in the identification of alternatives. Only four studies showed a clear assessment of future needs as a direct input to the identification of alternatives. The 1976 Honolulu report (5) is noteworthy in this respect for it presents an extensive discussion of needs and anticipated deficiencies. This information is then used to develop specifications or criteria for the generation of alternatives. The draft San Juan report (3) also focuses on the needs of current and future travelers for improved transportation. Planning studies for Los Angeles and Miami both showed the influence of a needs assessment on the identification of alternatives, but the effect was implied rather than stated.

Forecasts of Future Conditions

This category of potential inputs was defined as alternative population, employment, income, urban development, and economic scenarios that influenced the identification of alternatives. Because of the future orientation of alternatives analysis, forecasts of future conditions are a fundamental part of every study. In most cases, however, those who perform alternatives analysis studies have consciously or unconsciously taken the point of view that future conditions affect the performance of alternatives rather than their development. Thus the forecasting of future conditions has been oriented toward producing estimates of travel demand to be accommodated by alternative technologies.

A few studies took a different perspective. In these instances, forecasts of future economic, social, and development conditions were considered as inputs to the identification of alternatives. As might be expected, the use of forecasts as an input was largely correlated with the use of needs assessment to develop alternatives.

The Honolulu study (6) was particularly noteworthy in that several different forecasts and growth assumptions were presented and discussed with respect to their implications on alternatives. The draft version of the West Shore report (13) also shows an awareness of the sensitivity of the alternatives analysis process to assumptions on future conditions. A number of evaluation measures are given for two different growth scenarios. There is, however, no apparent involvement of these scanarios in the identification of alternatives for analysis.

Local Opportunities and Constraints

This category is defined to include the activity of identifying local influences on alternatives to be studied. The focus is on both physical and institutional opportunities and constraints on the development of alternatives. As shown in the matrix, all of the sample studies incorporated this activity as an input to the identification of alternatives. Roughly one-half of the reports documented this search, but the activity is so fundamental that it was clearly present in all of the studies. Typically, local constraints and opportunities focused on physical options such as identification of potential rights-of-way along highway medians or abandoned railroad lines.

Although the consideration of local opportunities and constraints was universal, the level of documentation of this activity as an input to the identification of alternatives varied widely. In six of the reports, this activity is only implied as potential routes; alignments and station locations suddenly appear as part of an alternative without any indication of how they were developed. In several cases other study memoranda and interim reports are referenced for details on the search for alignment alternatives.

External Requirements and Constraints

This category is defined as legal and institutional influences on the generation of alternatives that originate outside the basic mandates for alternatives analysis. By far the most important of these external influences is the National Environmental Policy Act of 1969. This law and its accompanying guidelines for the preparation of environmental impact statements have a major effect on the identification of alternatives. The key influence is the requirement that a no-build or null alternative be considered in the planning process for major federal actions that have a significant effect on the environment. As a result, nearly all of the sample studies showed an external influence through the inclusion of a no-build alternative. Significantly, in all cases, this influence was not acknowledged in the reports. Typically, past alternatives analysis studies have included no-build alternatives, but the justification, if any, for their consideration has been vague notions of proper planning practice and the need for an evaluation benchmark. Two of the reports surveyed showed no apparent involvement of these external requirements. From the documentation provided, neither the first Honolulu effort nor the Pittsburgh South Hills study (4, 6) included a no-build alternative.

Several other requirements are also potentially significant to the identification of alternatives. One of these is Section 4(f) of the Department of Transportation Act of 1966, which constrains the use of parks and wildlife refuges for rights-of-way. Similar obstacles are imposed by the National Historic Preservation Act of 1966, concerning historic and cultural resources, and by a variety of other laws enacted to preserve natural resources. The implementation of the Clean Air Act of 1963 and its subsequent amendments has a potentially direct effect on alternatives identification. A number of laws also incorporate social considerations that could impact the generation of alternatives. These include the requirements for elderly and handicapped accessibility, requirements for minority group accessibility, and requirements related to provisions to be made for persons displaced by federal actions. The review of sample reports indicated, however, that none of these external requirements had any influence on the identification of alternatives in these studies.

Carry-Over Alternatives

This category of potential inputs was defined as including prior analysis efforts and alternatives recommended and carried over from earlier studies. The matrix indicates that this category was the most completely documented input to alternatives generation in the sample studies. Eleven of the reports indicated that prior rapid transit planning had influenced the development of one or more alternatives for the alternatives analysis study. Only the 1975 Denver study (15) and the Miami report (9) did not show the apparent use of an alternative carried over from earlier planning efforts. The near universality of the use of carry-over alternatives as an input and the clear documentation of their use indicates very strongly the evolutionary nature of past alternatives analysis practice.

Catalogue of Technologies and Actions

This category of inputs is defined as the potential significance of the existence of a variety of available technologies with certain operating characteristics. The availability of certain modes and technological hardware may stimulate a desire to include them in the analysis. This is especially true in alternatives analysis studies that consider alternatives to be only different technologies. Thus, each alternative needs a different technology and the list of available technologies becomes an input to the identification of alternatives.

Operational Feasibility

This category of input was defined as the idea that a preliminary assessment of whether a given element would actually work may enter the process of identifying alternatives. Thus, this input is somewhat related to alternative refinement rather than identification. The distinction is of little importance in this case because only four of the sample reports showed the consideration of operational feasibility in the development of alternatives. The 1975 Denver effort (15) included some preliminary network testing as an activity related to the formulation of alternatives. The 1976 Honolulu study (7) excluded a carried-over waterborne transit alternative from consideration in alternatives analysis because the concept lacked basic operational feasibility. In Pittsburgh and the West Shore studies (6, 13), adjustments were made to alignments to reflect operational characteristics of the heavy rail technology considered in those alternatives.

Public Involvement

The definition of this input was broad and included both participation and use of public hearings. For many years the legislation and policies that mandated the transportation planning process have called for effective citizen participation at a point early enough in the study to permit substantive involvement. The emphasis on early involvement suggests that citizen input to the identification of alternatives is appropriate to alternatives analysis as well as to planning efforts oriented toward implementation.

The review of sample reports indicates that all of the studies involved the public in the alternatives identification process. The level of public involvement varied from study to study. Similarly, the documentation of public involvement varied between studies. Roughly one-half of the sample reports documented the public's involvement in alternatives identification; the other half of the sample merely implied the participation of the public in the identification process.

Several reports documented noteworthy public involvement in the identification of alternatives. The Miami study (9) incorporated substantive citizen participation in the identification of alternatives through a series of community forums that resulted in the addition of 27 new alternatives to the study. The series of studies (10-12) done on subway routes in the Washington area make reference to what must have been a substantial public involvement effort that made extensive use of elected officials and appointed task forces to provide meaningful citizen participation to the alternatives identification process.

TECHNIQUES USED IN ALTERNATIVES GENERATION

Conceptual approaches are theoretical and give little emphasis to the practicalities of how the various planning activities are to be performed. Some of these approaches, particularly the most recent ones, explore in detail the methodologies available for impact prediction and evaluation. Very few of these theoretical models offer guidance on how alternatives are to be developed. In general, the conceptual approaches merely state that alternatives are devised by the planner from the prescribed set of inputs.

Perhaps this activity has been taken for granted because no specific definition has ever been advanced as to what constitutes an alternative. Indeed, the field is so broad that a meaningful specification could only be created within a limited area, such as alternatives analysis studies. Because no definite concept of the output exists, few techniques have been advanced to produce alternatives.

The literature does include some brief consideration of ways that alternatives may be generated. Most authors seem to agree with Dickey who states, "Few planners or designers can deny that most of the better solutions for problems (especially those in transportation) have come to the surface through a series of intangible and untractable steps that are the outgrowth of the creative force at work" (16, p. 33).

Some techniques, however, are identified as augmentation for the creative force. Several authors include some discussion of the application of various operations research techniques to the task of generating alternatives. These include mathematical techniques such as linear programming, nonlinear programming, and dynamic programming. Mathematical programming is useful because it entails the optimization of some objective function when certain constraints are evident. These techniques appear to have some utility to the generation of alternatives in limited applications where objectives and constraints are susceptible to reduction to mathematical analysis.

Other methodologies include search and experimental design. A search approach involves multiple iterations of variation of input variables on a model or set of models that replicate the system. According to the outputs of the preceding iteration, the variables are adjusted until the desired state is reached. Experimental design is a similar technique but involves experiments on the actual system, not on models.

The typical approach assumed by most of the theoretical models is one of creative trial and error. The process for generating alternatives generally assumes that, if all the information inputs are available, the planner can devise alternatives to meet a given set of criteria or specifications. If something is not right the first time, it is modified in subsequent refinement iterations.

The review of sample alternatives analysis reports indicates that the creative trial-and-error procedure is typical of past planning practice. This determination is a judgment based on procedures implied in the reports. Only one study provided documentation on the methods used to generate alternatives. In the draft version of the alternatives analysis for San Juan, the following steps are outlined:

1. Review regional goals and objectives;
2. Examine current and future characteristics of the study area such as location, size, growth, economy, and social structure;
3. Identify current transportation problems and opportunities;
4. Determine reasons for studying rapid transit;
5. Choose candidate corridors; and
6. Develop alternatives by using combinations of corridors and advantageous characteristics of each.

These steps are clearly related to the basic models of the planning process and the creative trial-and-error technique of alternatives generation discussed above. The steps identify a familiar list of inputs (such as goals, forecast, inventories, needs, and technologies) and bring them together in a single step of development that matches needs and analysis with likely potential solutions. Although other studies reviewed for this aspect did not outline such steps, the reports imply that similar procedures were followed in each of the sample studies. In each case the judgment of process participants was used to act on the informational inputs and produce alternatives from potentialities.

CHARACTERISTICS OF ALTERNATIVES

UMTA Specifications

Because UMTA is the motivating force behind alternatives analysis, UMTA must have some conception of what makes up an alternative. A review of published policy statements indicates that at least a loose idea of content or definition exists that is communicated in terms of characteristics that describe what an alternative is like. The table below presents a compilation of criteria or specifications for alternatives taken from two primary policy statements that set forth the basis for the alternatives analysis process. As shown in the table, all of these criteria are contained in the policy statement on major urban mass transportation investments (17). Three of the criteria are reiterated in the statement of policy toward rail transit (18).

Criterion	Major Urban Mass Transportation Investments	Policy Toward Rail Transit
An alternative should meet local transportation needs	X	
An alternative should promote local social, economic, environmental, and urban development goals	X	
An alternative should support national aims and objectives	X	
An alternative should be consistent with the long-range plan	X	
An alternative should incorporate public involvement in its definition and content	X	
An alternative should incorporate community feeder and collector and distributor services in addition to line-haul transit options	X	X
An alternative should include supportive TSM actions	X	X
An alternative should be defined in terms of corridor location, length of initial segment, technology, horizontal and vertical alignment, grade separation, station locations, and other relevant factors	X	
An alternative should reflect the need for incremental development	X	X

A set of nine criteria can be derived from the policy statements. Most of them are specifications of what an alternative must be and do. Some are concerned with how an alternative is developed. Overall, these criteria can be divided into three groups. The first group can be described as criteria oriented toward ensuring consistency with ongoing transportation planning efforts. These specifications are concerned with the relationship of an alternative to local needs, regional and national objectives, and the long-range transportation plan. The second group of criteria are concerned with the objectives and procedures appropriate to alternatives analysis. These two criteria direct alternatives to incorporate public involvement and the need for incremental development of fixed-guideway transit systems. The third group of criteria deal with the specific content or characteristics of an alternative. These three specify that an alternative (a) must include collector and distributor services in addition to line-haul transit, (b) must include supportive TSM actions, and (c) should be defined in terms of a list of relevant factors. These factors include

1. Corridor,
2. Length of segment,
3. Technology,
4. Horizontal alignment,
5. Vertical alignment,
6. Grade separation,
7. Station locations, and
8. Other factors.

The list of criteria discussed above are not actually presented as a set in the referenced policy statements. These criteria are stated in the text of the policy document as guidelines and indications of UMTA's intentions. As such, they are not organized and are presented in the policy statements as a set of criteria for alternatives, although a careful reading of the documents suggests that this is their intended function.

Alternatives Analysis Experience

The criteria or specifications discussed above constitute UMTA's public definition of what the content of an alternative should be. The central concern in this paper, however, is the past performance of alternatives analysis studies. This section examines the way alternatives have been defined and described in the sample reports.

In order to compare and summarize past experience, a long list of descriptive factors was generated from the UMTA criteria and other sources. These factors were then arrayed in a matrix format opposite the sample of alternatives analysis reports, and the incidence and type of usage for each factor was noted for each study. This matrix is presented in Table 1.

The review of the content and characteristics of alternatives in past studies shows that alternatives analysis studies have embodied different conceptions of what makes up an alternative. The matrix illustrates that the definition of alternatives in alternatives analysis studies has focused on facilities and service area factors. Description of alternatives in terms of technology, location, extent, alignment, and capacities has been nearly universal.

Level-of-service characteristics have received moderate use as definitional elements of alternatives. In general, levels of service are attached to technologies and rarely featured as the chief distinguishing element of a particular alternative. Forecasts of future conditions or scenario factors have received even less attention as components of alternatives. Very few studies formulated alternatives specifically designed to test the impacts of such variables as central business district (CBD) parking costs or complementary transportation or nontransportation actions. Such elements were frequently assumed to be background factors that were held constant for all alternatives.

Overall, the level of documentation with respect to the content or definition of alternatives in each study was low. Typically, alternatives have been described in terms of a few major factors early in the study, then fleshed out as the need arises during the stages of impact prediction and analysis.

Table 1. Descriptive factors of individualized transportation alternatives.

Factor	Buffalo	Denver 1975	Denver 1976	Honolulu 1975	Honolulu 1976	Los Angeles	Miami	Pitts-burgh	San Juan	Washington F	Washington J-H	Washington K	West Shore
Technology types	P	P	P	P	P	P	P	P	P	P	P	P	P
Designated corridor	P	P	P	F	F	P	P	F	P	F	F	F	F
Specified alignment	P		P	P	P	P	P	P	P	P	F	P	P
Vehicle capacity	P	P	P	P	P	P	P	P	P				P
Station locations	P		P	P	P	U		P	P	P	P	P	P
Yard and maintenance facility locations	P		U	U	U	U	U	U	P			P	U
Explicit cost targets			P										
Designated market segments													
Frequency													
Peak and base	P	P	P	U	U		P	P	P	U	U	U	P
Average daily	P	P	P	U	U	P	U	U	P	U	U	U	P
Fare level													
Average	P	U	U	U	U	U	U	U	U	U	U	U	P
Fully structured	P	U	U	U	U	U	U	U	U	U	U	U	P
Vehicle kilometers of service by type	P	P	U	U	U	U	U	U	P	U	U	U	U
Route kilometers of service by type	P	P	P	P	P	P	P	P	P	U	U	U	P
Travel times and speeds	P	P	P	U	U	P	P	P	P				P
Development controls													
Automobile use disincentives				U	U								
CBD parking costs	F	U	U	U	U	U	U	U	U	U	U	U	U
Population forecasts	F	U	U	U	U	U	F	U	F	U	U	U	P
Income forecasts	F	U	U	U	F	U	U	U	U	U	U	U	U
Employment forecasts	F	U	U	U	F	U	F	U	F	U	U	U	P
Automobile operating costs	F	U	U	U	U	U	U	U	U	U	U	U	U
Complementary or supplementary transportation actions				-[a]		-[b]	-[c]	-[d]		-[e]	-[f]	-[g]	
Complementary non-transportation actions													

Note: P = present; U = unstated; F = fixed for all alternatives.
[a] Express bus feeder.
[b] Fixed TSM bus improvements.
[c] Arterial express bus.
[d] South PATWAY.
[e] Parking facilities, bus priority, and conventional bus.
[f] I-95 express bus.
[g] Express bus feeders.

UMTA Criteria for Sets of Alternatives

The structure of the alternatives analysis process indicates that alternatives have a dual nature. Each alternative has significance both by itself and as a member of a set of alternatives selected for analysis. Thus, just as individual alternatives are defined in terms of certain descriptive factors, the set of alternatives is defined in terms of certain collective properties.

UMTA has promulgated some criteria or standards for alternatives through the two major policy statements that outline the alternatives analysis process. In addition to the specifications provided for the content of individual alternatives, the policy statements have also included a loose set of standards for the collective set of alternatives. These criteria for the set of alternatives are shown in the table below. A series of six criteria statements have been extracted from Major Urban Mass Transportation Investments Policy (17) and Policy Toward Rail Transit (18).

Criterion for Set of Alternatives	Major Urban Mass Transportation Investments	Policy Toward Rail Transit
Reflect a range of options	X	
Cover a range of technologies including rapid transit, light rail, busway, people-mover, and TSM actions		X
Be defined and presented at the same level of detail	X	
Include a TSM alternative	X	
Include a light rail alternative		X
Reflect the need for incremental development	X	X

Basically, the criteria establish necessary or desirable elements of the set and indicate a concern for comparability and the realities of implementation. Most of the standards for the set of alternatives deal with the idea of range. The set of alternatives must reflect a range of unspecified options. Some of these options, however, are established to be technological. In fact, the emphasis apparent in the interpretation of range is on technology, with particular stress on use of TSM technologies and light rail transit. Other criteria indicate that all alternatives should be developed at similar levels of detail for the integrity of the comparative analysis and that the set of alternatives should reflect a range of fixed guideway lengths that are suitable to identify a potential initial increment.

Because the above table deals only with specifications set forth by UMTA, an important standard for the set of alternatives is not shown. This is the external requirement imposed on the planning process by the National Environmental Policy Act of 1969, that the consideration of environmental impacts for major federal actions such as those contemplated under alternatives analysis include consideration of a no-build or baseline alternative.

Baseline and TSM Alternatives

The baseline or no-build and TSM alternatives constitute a special type of alternative. As seen above, both types of alternatives are required components of the alternatives set as a result of federal law and policy. The significance of these alternatives to the validity and utility of the alternatives analysis process is clear. The purpose of the baseline alternative is to provide a benchmark for comparison of the various rapid transit alternatives to assist in the identification of a best course of action. The purpose of the TSM alternative is to investigate the level of benefit that can be captured by using

low-capital techniques to improve transportation services through better management of existing resources and facilities.

Table 2 presents information on the consideration of baseline and TSM alternatives in the sample alternatives analysis reports. Three different types of baseline alternatives can be distinguished. The first type is termed benchmark-maintenance. This is defined as the baseline alternative recommended in a National Cooperative Highway Research Program (NCHRP) report that calls for only the continued maintenance of existing facilities and services (19, p. 13).

The second type of baseline alternative found in past alternatives analysis practices is termed extended service. This type of baseline is commonly defined as the extension of existing transit services to cover areas of new population growth anticipated between the base and analysis years. Viewed in another way, this type of baseline posits the same per capita level of transit service in the analysis year as exists today.

The third type of baseline alternative is termed improved service. This definition hypothesizes that the benchmark alternative should represent what the transit system would be like in the analysis year if it developed according to current trends, excluding major capital improvements.

Table 2 shows that only the Los Angeles study (1) incorporated a strictly defined benchmark alternative that corresponds to the NCHRP guidelines. Three studies—Buffalo, Denver, and Honolulu (2, 4, and 15)—used an extended service concept as the baseline for planning analysis. The draft San Juan report and the series of Washington studies (3, 10-12) made use of an improved service concept as the baseline. In San Juan the existing routes and schedules were perceived as antiquated and revisions were developed as the basis for the horizon-year service. The Washington studies examined no-build alternatives that offered improved express bus services as a replacement.

In the Honolulu studies (4, 5, 7), the baseline alternative received a different treatment than in other sample reports. In this case, the baseline alternative was evaluated in a preliminary exercise to determine whether the no-build option would meet the projected needs. After the no-build alternative was rejected as insufficient, the study proceeded to the generation and evaluation of other high-capacity alternatives. Thus the conclusion was reached to do something early in the study, and this directed the subsequent analysis and presentation of the results.

It should be noted that three studies did not include an apparent baseline alternative. These are Miami, Pittsburgh, and the West Shore alternatives analyses (6, 9, 13). The Miami study (9) was already in the preliminary engineering phase when the analysis of alternatives was required, and it is possible that the analysis of a baseline alternative was viewed as an entirely separate effort related only to preparation of the draft environmental impact statement (EIS). The reasons for the lack of a baseline alternative in the Pittsburgh South Hills study (6) are also unclear and it may be possible that the analysis of a baseline alternative was perceived to be only appropriate to the separate activity of preparation of the draft EIS.

The treatment of a baseline alternative in the draft West Shore report (13) may signal an emerging trend. In this study a TSM alternative was developed and used as the benchmark alternative for the analysis year. This represents a borderline type of approach to fulfilling both the baseline and TSM requirements. The point of view, however, has considerable validity, as reflected in the NCHRP guidelines that call for the assumed completion of committed projects for short-term analysis. In the West Shore case, the planned and committed projects were TSM improvements. Thus the baseline and TSM alternatives were seen to virtually coincide and were subsequently consolidated into a single alternative.

Table 2 also shows that three types of TSM alternatives can be identified. The first of these is termed a distinct TSM alternative. This is a single alternative that represents the application of TSM technologies or actions to the forecasts of travel demand. The second type of TSM treatment is the integration of a fixed set of TSM actions into all or several of the alternatives. The third type of TSM consideration is the integration of different TSM actions in different alternatives designed to complement and support the major features of the alternative.

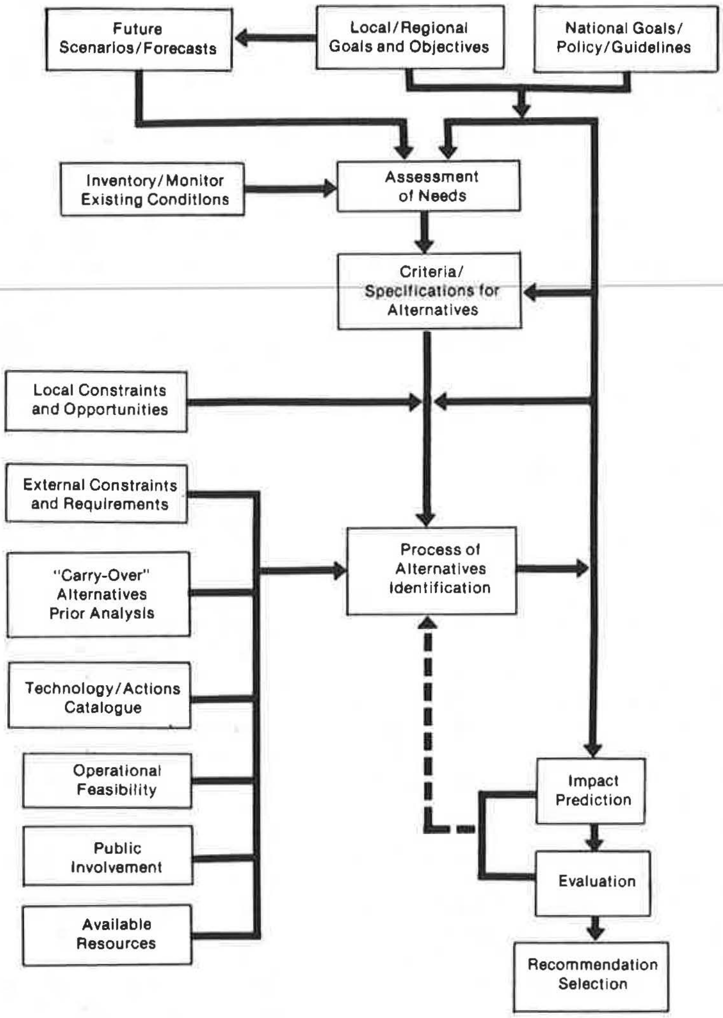
It should be noted that several of these studies were completed or in midstream when consideration of TSM improvements were mandated in September 1975. The Buffalo and 1975 Denver reports (2, 15) are examples of this timing situation and consequently neither shows any consideration of specific TSM improvements. The Pittsburgh and the draft San Juan reports (3, 6), however, were published well after September 1975 and show no consideration of TSM. This can probably be attributed to delays in incorporating TSM into the alternatives analysis process.

The most common treatment of TSM in the other studies was as a distinct alternative. One study, the 1975 Honolulu effort (4), appears to have developed a fixed set of TSM actions that was incorporated into several alternatives. In the studies in Los Angeles, Miami, and the F-corridor in Washington (1, 9, 10), a number

Table 2. Characteristics of baseline and TSM considerations in past alternatives analysis studies.

	Baseline Alternative			TSM		
					Integrated	
Study	Benchmark-Maintenance	Extended Service	Improved Service	Distinct Alternative	Fixed	Variable
Buffalo		X				
Denver						
1975		X				
1976				X		
Honolulu						
1975					X	
1976		X		X		
Los Angeles	X			X		X
Miami				X		X
Pittsburgh						
San Juan			X			
Washington						
F			X			X
J-H			X	X		
K			X	X		
West Shore				X		

Figure 2. Factors that influence the identification of alternatives.



of alternatives incorporated different TSM actions as complements to other features of the alternative. The use of the TSM alternative as a baseline in the draft West Shore report (13) was noted above.

CONCLUSIONS

This paper has presented the results of an investigation into the process of generating alternatives for analytical studies of major urban mass transit investments. Among the questions considered were the following:

1. What are the inputs to alternatives?
2. What techniques have been used to generate alternatives?
3. How have alternatives been defined?

The initial portion of the paper examined the actual use of inputs to the alternatives generation activity within the planning process. The analysis also focused on the timing or location of the alternatives generation activity. Figure 2 brings the results of these two investigations together. This diagram shows the factors that influence the identification of alternatives. Because the flow chart is a composite of a variety of conceptual and actual approaches, it represents a description of the universe of possible inputs and their interrelationships.

Although the intent of this diagram is not to prescribe how planning and alternatives generation should be accomplished, the flow chart constitutes, in a sense, a simple generalized model of the planning process. Activity begins with the formulation of local and national goals and objectives. On the basis of local objectives and growth trends, a series of alternate forecasts of future socioeconomic, transportation, and land use conditions are generated and input to an assessment of needs. Needs are determined on the basis of these future forecasts and the inventory of current conditions and facilities. From these stated needs, a set of criteria may be developed that specify what characteristics and properties an individual alternative and the collective set of alternatives must possess. These criteria for output are then fed into the alternatives identification process. A set of seven new potential inputs may also be tapped. These include local constraints and opportunities, external constraints, carry-over alternatives, the spectrum of available technologies, an assessment of operational feasibility, public involvement, and available resources. The final potential input to alternatives generation is the feedback or refinement loop that originates in the activities of impact prediction and evaluation of prior iterations of the planning cycle.

The flow chart in Figure 2 is a depiction of the state of the art in the area of the organization of and inputs to the task of generating alternatives. A number of other specific conclusions can be drawn from the investigation.

Overall, the activity of generating alternatives has not received the attention given to other elements of the planning process such as impact prediction and evalua-

tion. As a result, both the conceptual approaches and actual alternatives analysis experience show a marked lack of documentation in this area. The generation of alternatives has been largely glossed over or taken for granted. The sample of alternatives analysis reports shows little internal awareness of the planning process and almost no attention to the ways by which alternatives are developed.

The relationship between goals or objectives and the development of alternatives was relatively weak. Although many of the sample reports listed local and national goals and objectives, little use was made of them in the task of developing alternatives. In only one or two studies were actual linkages present that linked a specific alternative to a specific objective.

The conceptual approaches make an assumption of a clean slate that is seldom found in actual practice of alternatives analysis. The review of sample reports indicated that alternatives analysis studies were often significantly influenced by prior analysis efforts. The prevalence of carry-over alternatives indicated that many external forces shape the technical work of the planning and analysis cycle.

The search for local opportunities and constraints was among the most prevalent inputs to alternatives generation in past alternatives analysis studies. Other major inputs were alternatives carried over from previous studies. Goals, objectives, and public involvement were widely documented but failed to exert much substantive influence on alternatives generation. The weakest or least-cited inputs were assessments of operational feasibility needs and forecasts of future socioeconomic, population, and land use configurations.

Significantly, the review of sample studies showed that almost none of the completed efforts developed criteria or specifications for individual alternatives and the set of alternatives. As a result, few studies exhibited a clear idea of what an alternative actually was. The product of the alternatives generation activity was never specified in advance.

Overall, public participation in the identification of alternatives was weak in completed alternatives analysis studies. Although public involvement was widespread, it did not penetrate to the actual formation of alternatives for analysis. Public involvement was largely viewed as a review and comment function after a basic set of alternatives had already been devised.

In general, the sample reports indicate a lack of alternatives screening and refinement. Miami and the draft West Shore reports document a process of developing a small set of detailed alternatives from a large number of preliminary alternatives. Other studies appeared to only generate a small number of alternatives and variations on them for input to the analysis. Some alternatives analysis studies may have included multiple iterations and refinements but failed to document the process in the final report. The Pittsburgh South Hills study documents a set of criteria for such a refinement process.

Although some have suggested the application of operations research techniques (such as mathematical programming) to the task of alternatives generation, there is no widespread recognition that these techniques are particularly effective or cost saving. Rather, the general conception appears to be that these techniques are simply too limited and too complex for general use in generating alternatives.

The most common technique for generating alternatives in both the conceptual approaches and the alternatives analysis experience is a loosely structured creative trial-and-error method. This technique constitutes a default option and is seldom articulated or documented as the procedure to follow in a given instance.

The concept of an alternative has not been fully defined. UMTA policy statements have provided some criteria for individual alternatives but, on the whole, UMTA has refrained from issuing a rigid specification of what an alternative must incorporate. As might be expected, alternatives analysis studies have shown a variable conception of the content or characteristics of alternatives. Despite the existence of guidelines for the implementation of the National Environmental Policy Act of 1969, the sample alternatives analysis reports show no common perception of what a baseline alternative is. Indeed, not even all the studies included baseline alternatives although it is a clear requirement of the regulations. Three different interpretations of the baseline alternative were identified.

The sample studies display similar varying treatments of TSM concepts. Some studies predated TSM requirements, others interpreted the need to consider the TSM approach in three different ways. The Los Angeles study provided the most thorough discussions of TSM potential.

Future problems may arise in distinguishing between baseline and TSM alternatives. If the strictly defined benchmark alternative should include planned and committed projects, then areas that have active TSM improvement programs will find little to distinguish a baseline from a TSM alternative.

Overall, alternatives analysis experience has corresponded closely to the general state of the art in the identification of alternatives. In large part, this may be due to the fact that the state of the art has not developed much beyond an intuitive judgmental status. It is apparent that past alternatives analyses have ignored several potential inputs to the generation of alternatives. More important has been the lack of internal or process awareness exhibited in the draft and completed alternatives analysis reports. Little consideration has been given to the activity of identifying alternatives and, therefore, its significance and the importance of how the process is carried out have gone unrecognized.

ACKNOWLEDGMENT

This paper reports the results of research sponsored by the Transportation Systems Center and the Urban Mass Transportation Administration of the U.S. Department of Transportation.

REFERENCES

1. Alan M. Voorhees and Associates, Inc. A Summary of the Technical Report of the Study of Alternative Transit Corridors and Systems. Southern California Rapid Transit District, Los Angeles, Oct. 1974.
2. Alan M. Voorhees and Associates, Inc. Metro for Buffalo: Transit Alternatives for the Buffalo-Amherst Corridor. Niagara Frontier Transportation Authority, Buffalo, NY, June 1976.
3. Consultores Tecnicos Asociados; Alan M. Voorhees and Associates, Inc. Metro for San Juan: A Study of Transit Alternatives for the Metropolitan Area of San Juan. Puerto Rico Department of Transportation and Public Works, San Juan, June 1979.
4. Daniel, Mann, Johnson, and Mendenhall. Alternative Transit Concepts Analysis. Department of Transportation Service, City and County of Honolulu, July 1975.
5. Daniel, Mann, Johnson, and Mendenhall. Analysis of Transit Alternatives. Department of Transportation Services, City and County of Honolulu, April 1976.
6. DeLeuw, Cather and Company. Comparative Analy-

sis Study of Alternative Transit Systems—South Hills Corridor. Port Authority of Allegheny County, Pittsburgh, Nov. 1976.
7. Additional Technical Information for Alternatives Analysis. Department of Transportation Services, City and County of Honolulu, April 1976.
8. Letter to Honorable Robert E. Patricelli, Administrator, Urban Mass Transportation Administration. Joint Regional Planning Program, Regional Transportation District, Denver, CO, April 1976.
9. Kaiser Engineers and others. Draft Milestone 1 Report: General System Concept and Criteria. Dade County Transportation Improvement Program, Miami, FL, Aug. 1974.
10. Peat, Marwick, Mitchell, and Company. Metrorail Alternatives Analysis: F Route. Metropolitan Washington Council of Governments, Washington, DC, Oct. 1977.
11. Peat, Marwick, Mitchell, and Company. Metrorail Alternatives Analysis: J-H Route. Metropolitan Washington Council of Governments, Washington, DC, Oct. 1977.
12. Peat, Marwick, Mitchell, and Company. Metrorail Alternatives Analysis: K Route. Metropolitan Washington Council of Governments, Washington, DC, Oct. 1977.
13. Seelye, Stevenson, Value, and Knecht; Alan M. Voorhees and Associates, Inc. West Shore Corridor Transit Alternatives Study. Tri-State Regional Planning Commission, New York, Draft Rept., Oct. 1978.
14. Summary Report: Evaluation of Transit Alternatives for the Los Angeles Starter Line Corridor. Southern California Rapid Transit District, Los Angeles, July 1976.
15. System Management Contractor. Long-Range Transit Development Analysis: Transit Concept Comparison. Regional Transportation District, Denver, CO, Draft Rept., April 1975.
16. J. W. Dickey. Metropolitan Transportation Planning. Scripta Book Co., Washington, DC, 1975.
17. Major Urban Mass Transportation Investments Policy. Federal Register 41, 185, Sept. 22, 1976, pp. 41512-41514.
18. Policy Toward Rail Transit. Federal Register 43, 45, March 7, 1978, pp. 9428-9430.
19. J. S. Lane, L. Grenzeback, T. Martin, and S. Lockwood; DACP, Inc. Impact Assessment Guidelines: The Role of the No-Build Alternative in the Evaluation of Transportation Projects. NCHRP, Proj. 8-11, 1977.

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

Abridgment

Factors Affecting Willingness to Conserve Gasoline

Catherine E. Meyers

This paper explores the role of travel behavior, demographics, and attitudes toward energy conservation to explain consumer willingness to conserve gasoline. Data from a telephone survey of 500 New York households were analyzed on several measures of willingness to conserve by using a statistical analysis procedure called automatic interaction detection. It was found that willingness to conserve gasoline is generally independent of demographics, travel behavior, and other attitudes toward energy. The factors of mild importance included residence location (New York City residents were more willing to conserve) and attitudes toward use of energy in the United States. Generally, those New Yorkers most willing to conserve were those who had (a) the least to lose if gasoline were curtailed, (b) the most flexibility in current travel behavior, and (c) the most additional service options available. A brief review of other recent surveys shows that the public is consistently receptive to policy that encourages gasoline conservation by increasing travel options and offering incentives for their use. Punitive or restrictive measures are met with strong disfavor. Based on the results of this and other studies, policy suggestions include increased perception of travel options and their true costs, increased transit services, and stabilized fares.

In the spring and summer of 1979, reduced availability and rapid price increases of gasoline disrupted the mobility-oriented American life-style. More than 50 percent of the petroleum consumed in the United States is used by the transportation sector; half of that is used by cars. Obviously, conservation in this sector would contribute significantly to improving the future outlook of our energy situation. If we are to make headway, we must understand how to encourage consumers to conserve gasoline. The purpose of this study is to determine which factors (including travel behavior, demographics, and attitudes toward energy) can be used to predict willingness of travelers to conserve gasoline. A complete description of the analysis and conclusions can be found elsewhere (1).

Numerous surveys and polls have been conducted in recent years concerning Americans' views of energy use and conservation. The National Assessment of Education Progress conducted a nationwide survey of 26- to 34-year-olds in the summer of 1977 (2). More than 90 percent of those surveyed believed that energy and gasoline shortages would be serious problems at some time in the future. The large majority stated that they take energy into consideration when they purchase a car or home, travel to work, and vote (level of education was the most significant variable highlighting differences between groups). Similarly, a statewide poll in Wisconsin conducted in June 1977 (3) showed education to be the only significant differentiating factor: Level of education and awareness of energy issues were directly correlated. Although energy was the most frequently mentioned problem that confronts the state, less than one person in five believed energy shortages would become a permanent feature of American life.

The New York State Department of Transportation conducted a public opinion survey on energy and transportation in the fall of 1977 (4). Almost three-quarters of New York's residents thought that the energy problem was at least fairly serious. The most favored proposal

for cutting gasoline use was encouragement of transit use within cities and between cities. The approaches that increase travel options and offer incentives for saving energy were greatly preferred over forced or travel-restrictive programs.

A nationwide poll conducted for the U.S. Department of Transportation (DOT) in 1978 (5, 6) showed the most favored proposals for reducing gasoline consumption to be as follows:

1. Enforcement of 88-km/h (55-mph) speed limit,
2. Allocation of funds to improve transit, and
3. Encouragement of carpooling.

The least acceptable proposals were gasoline rationing, increased gasoline tax, and restriction of downtown parking. The most frequently cited advantages of commuting by automobile were speed, convenience, and flexibility. Rapid, steep price increases (like those that occurred in the winter of 1973-1974 and spring and summer of 1979) tend to slow down gasoline consumption, whereas gradual, incremental increases do not. Almost all of the changes occurred in discretionary travel and very little in commuting habits. A February 1979 Gallup poll (7) showed that Americans believed energy to be the nation's third-most-important problem, after inflation and foreign policy. (In February 1973, energy did not rank on the list of top 15 problems.) Gasoline rationing that would require a one-quarter reduction in driving was opposed 52 percent to 40 percent; 8 percent had no opinion. Only one-third expected a repetition of the long gasoline lines of 1973-1974. Not quite one-half of the respondents indicated that they were making an effort to save energy in the transportation sector.

In March 1979, 70 percent of those polled by Associated Press-NBC News (8) believed the imminent gasoline shortages were a hoax designed by the major oil companies to force higher prices. This figure dropped by only 5 percent in June (9), even after severe shortages were evident. A survey conducted by the New York Times (10) in May 1979 showed that changes occurred in summer vacation plans in the face of reduced gasoline availability and increased prices. Such changes included shorter automobile trips (particularly those that exceeded a tankful of gasoline), increased use of public transportation, and an increase in travel and lodging reservations.

STUDY DESIGN

The data for this study were collected by the 1977 public opinion survey (4). The data consist of a random telephone survey of 500 households in New York selected in multistage fashion from four geographic strata: New York City, large metropolitan areas, small metropolitan areas, and rural areas. Within strata, sampling is proportional to population.

The question used to derive attitudes toward gasoline conservation consists of 13 proposals to conserve gasoline and was as follows:

Do you agree or disagree with each of the following ways to encourage New Yorkers to cut gasoline use?

1. Law limiting new car size.
2. Tax on new big cars.
3. Enforce 88-km/h (55-mph) speed limit.
4. Increase gasoline prices through taxes.
5. Ration gasoline to families, with certain exemptions.
6. Commuter and parking taxes.
7. Incentives for carpooling.
8. Rebate for buyers of little cars.
9. Encourage people to plan trips better.
10. Encourage people to shop closer to home.
11. Encourage people to shop less frequently.
12. Improve traffic flow on city streets.
13. Build more superhighways.

An answer of agree = 1, neutral = 2, and disagree = 3.

Index = Σ responses/13   range = 1.00-3.00

Consolidation of the 13 responses into one average score yields an attitude index toward gasoline conservation that is used as the dependent variable to be explained by predictor variables.

A statistical grouping procedure known as automatic interaction detector (AID) was used for the analysis. The technique has recently seen much application in transportation planning (1, 11-13). AID produces a series of mutually exclusive binary splits that maximize the between sum of squares (BSS) and minimize the total sum of squares (TSS) between each successive pair. The process continues until no more splits can be made in accord with the given criteria for significance. The BSS for each split divided by the TSS of the sample equals the percentage of variation explained by a split. The BSS and TSS for each split are summed to yield the total amount of variation explained. This statistic, analogous to R^2, defines the power of AID to predict the dependent variable with the given independent variables.

Three series of tests were made, each using the same dependent variable (mean index). The first uses demographics and behavioral data as the predictors (note that the first five listed are the best predictors):

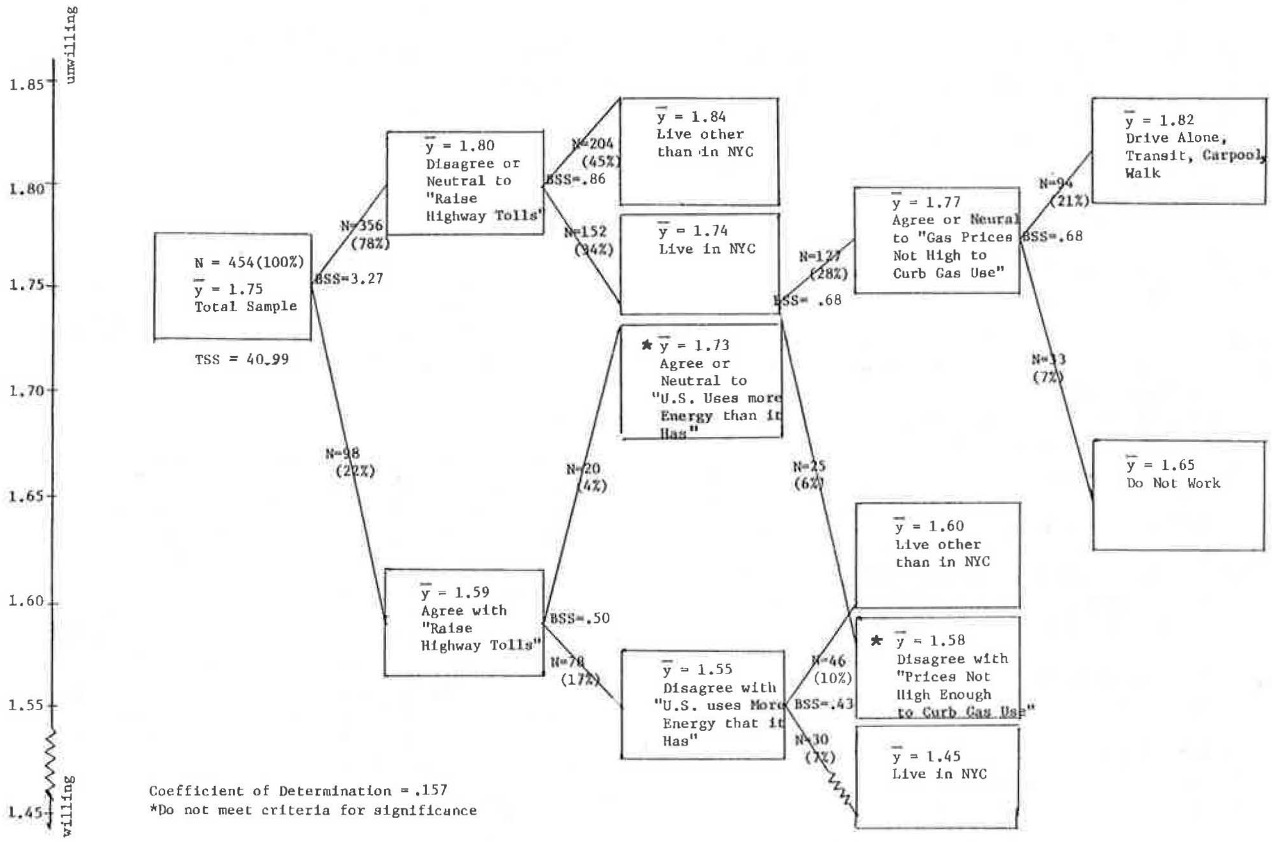
1. Stratum (New York City, large urban, small urban, or rural),
2. Mode to work of principal wage earner,
3. Family size,
4. Number of work or school trips made by members of household yesterday,
5. Number of cars per household,
6. Distance from home to work,
7. Number of shopping or recreation trips made by members of household yesterday,
8. Number of personal business trips made by members of household yesterday,
9. Age of respondent, and
10. Sex of respondent.

A second test uses attitude information toward other aspects of energy as the predictors (note that all but 6c and 6d are the best predictors):

1. Belief in energy crisis,
2. Price of gasoline,
3. "United States uses more energy than we have,"
4. Where should New York conserve energy,
5. How should we encourage transit use within cities, and
6. How should we encourage transit use between cities: (a) higher automobile tolls, (b) lower train and bus fares, (c) higher air passenger tax, or (d) tax break to companies that urge public transportation for business travel.

A third AID analysis was also performed by using the best of both sets of predictors (the first five in the first list and all but 6c and 6d in the second list), meaning those predictors in the first two tests that explained the most variation of the sample data.

Figure 1. Analysis of best behavioral and attitude factors influencing willingness to conserve gasoline.



RESULTS

Overall AID results are shown in Figure 1. In none of these tests were the predictor variables under consideration able to explain more than 15 percent of the variation in attitude toward willingness to conserve gasoline. The results show clearly that willingness to conserve gasoline is apparently independent of many primary demographic, behavioral, and attitudinal descriptors. This result is disconcerting since it suggests that clearly defined consumer segments are not identifiable; therefore, the task of encouraging conservation will not be easy.

The best of behavioral and attitude variables (Figure 1) show that those most agreeable to gasoline conservation were those who feel that raising automobile tolls will encourage gasoline conservation and disagreed that excess U.S. consumption has caused the energy crisis. Those who live outside of New York City and do not agree with raising highway tolls are least willing to conserve.

Location

All our tests showed New York City residents to be more agreeable to gasoline conservation than residents of the rest of the state. This is reasonable since New York City has greater availability of existing transit options and averages fewer cars per household than do other areas of the state. Less dependence on private automobile means New York City residents would incur fewer disruptions of current life-style if gasoline conservation policies are enacted.

Work Travel

The general tendency separates those households whose members work (regardless of mode of travel) from those households that make no work trips; the latter are more agreeable to gasoline conservation. Absence of the daily work trip means that the proportion of discretionary travel is much higher. Discretionary travel lends itself to greater flexibility in timing and ordering of activities that require travel, which increases the probability of finding suitable travel options.

Raise Highway Automobile Tolls

The most significant attitude factor was response to a proposal to raise highway automobile tolls to encourage energy savings between cities. Those who agree with the statement are more agreeable to gasoline conservation than those who are neutral or disagree. Apparently, those who agree view increased travel cost as a tradeoff for retaining travel independence. Furthermore, they may not encounter tolls regularly and would then be insensitive to policy actions.

Awareness of Energy Problem

Not surprisingly, those who did not believe an energy problem exists are least agreeable to conserving gasoline. Although the number of disbelievers is small (N = 24), this observation was congruent with the literature and supports the view that perception of the energy situation as a problem is one underlying factor that contributes to an attitude agreeable to energy conservation.

Lowering Train and Bus Fares

Those who are most agreeable to gasoline conservation find reduced fares to be an incentive for transit use between cities but exclude them as an incentive for transit use within cities. Obviously, situational factors require consideration here. The impact of lowering fares between cities would be primarily on discretionary travel. Lowering fares within cities is intended to attract the daily work trip.

SUMMARY AND POLICY IMPLICATIONS

Willingness to conserve gasoline is generally independent of demographics, patterns of travel behavior, and other attitudes toward energy. Only 15 percent of the variation in willingness to conserve gasoline could be explained in any of the tests made. Those New Yorkers who have options available to them, whose travel behavior patterns are flexible, and who currently consume relatively low amounts of gasoline are most agreeable to gasoline conservation. The public favors policies that increase travel-restrictive and punitive measures. No matter what their potential, policies cannot be effective unless they have public support.

Three component parts interact to effectively increase travel options:

1. Increased service availability,
2. Increased cost availability, and
3. Increased perception of travel options.

The daily commute to work and holiday travel are two separate issues. The daily work trip is nondiscretionary; any changes in services or costs have a high potential impact on daily activity patterns. Vacation travel can be adjusted to accommodate any service or monetary limitations.

With respect to increased services, it is obvious that, unless services are available, they will not be used. Of particular importance is transit service. Planners need to give careful consideration to travel patterns, possible additional or revised transit routes, and increased frequency. Since speed and convenience are often cited as advantages for driving alone, express transit services from areas of high population density, such as apartment complexes, to central business districts must also be considered. Park-and-ride lots located at large shopping malls would minimize required capital investment and provide needed services in some communities. Better transit services between cities, connected to convenient local transit services from depots, would make intercity transit a more attractive alternative for long-distance travel.

Concerning awareness of energy matters, much more can be done to provide practical knowledge about gasoline conservation. Emphasis on conservation in school curriculum is a long-term approach. Fair treatment of energy issues by news media, publicity of available travel options by the advertising and travel industry, and dissemination of information on available options should all be used to provide the public with a better-founded perspective on energy matters for basing travel decisions.

Our studies show that the policies most favored by the public to promote savings in gasoline consumption are those that increase the alternatives for gasoline conservation. Options must be available before they can be chosen. But it is evident that gasoline conservation will not be effectively promoted by any single approach; therefore, an array of travel options is needed. Presentation should be made in such a way that several advantages are perceived (i.e., monetary savings, speed, convenience, and safety). By such means, willingness to conserve gasoline can be increased and gasoline use reduced.

ACKNOWLEDGMENT

I wish to extend sincerest thanks to David T. Hartgen, Fred Neveu, and Wayne Ugolik for their academic guidance, enduring patience, and moral support and to Diane E. David and Wilma C. Marhafer for the preparation of the manuscript. I, of course, assume responsibility for any errors of fact and omission.

REFERENCES

1. C. E. Meyers. Factors Affecting Willingness to Conserve Gasoline. Planning Research Unit, New York State Department of Transportation, Albany, Prelim. Res. Rept. 167, July 1979.
2. Energy Knowledge and Attitudes: A National Assessment of Energy Awareness Among Young Adults. National Assessment of Educational Progress, Denver, Rept. 08-E-01, Dec. 1978.
3. R. D. Hedlund, K. E. Hamm, and R. M. Stein. Public Attitudes Toward Energy and Its Conservation: A Statewide Survey of Public Opinion. Urban Research Center, Univ. of Wisconsin, Milwaukee, June 1977.
4. A. J. Neveu. Public Opinion Survey on Energy and Transportation. Planning Research Unit, New York State Department of Transportation, Albany, Prelim. Res. Rept. 135, Dec. 1977.
5. Through Their Eyes, Part 2: The People Speak. Office of the Assistant Secretary for Governmental Affairs, U.S. Department of Transportation, March 1978.
6. Peter D. Hart Research Associates, Inc. A Survey of American Attitudes Toward Transportation. Office of the Secretary, U.S. Department of Transportation, Jan. 1978. NTIS: PB 280-164/5G1.
7. Gallup Opinion Index. Mood of U.S. Public: Their Views on Inflation, Jobs, Energy. American Institute of Public Opinion, Princeton, NJ, Rept. 164, March 1979.
8. M. Muskal. Poll: Citizens Think Oil Lack Merely a Hoax. Albany Times Union, March 23, 1979, p. 1.
9. Poll: Americans Doubt Gas Shortage Is for Real. Albany Times Union, June 4, 1979, p. 1.
10. R. Blumenthal. Year of Gasoline Shortage Causing Changes in Vacation Plans. New York Times, May 13, 1979, p. 1.
11. P. B. Schary, D. J. Brown, B. W. Becker. Consumers as Participants in Transportation Planning. Transportation, Vol. 6, 1977, pp. 135-148.
12. C. B. Leutze and W. R. Ugolik. Who Pays the Highest and Lowest Per-Mile Transit Fares? Planning Research Unit, New York State Department of Transportation, Albany, Prelim. Res. Rept. 136, Feb. 1978.
13. 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, Prelim. Res. Rept. 129, Aug. 1977.

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

Park-and-Ride Planning for Energy Conservation: An Optimization Methodology

John G. Schoon

A linear programming approach is used to allocate the location and size of park-and-ride facilities to minimize energy use. The basic objective is to minimize the use of energy by estimating the minimum vehicle kilometers traveled for specific corridor or areawide park-and-ride programs. Parameters such as vehicle occupancy, travel costs, constraints that include transference of core-area parking to fringe areas, and community-imposed limits on parking in suburban areas are considered. Thus, the model is made responsive to areawide energy conservation, program costs, transportation system management and transportation control planning actions, and community policy concerns. This normative approach to park-and-ride planning is performance oriented; the results of other demand models are used as inputs to the linear programming process. It can assist also in contingency planning for energy savings by defining the allocation of park-and-ride facilities to meet specific levels of energy use subject to related system capabilities and constraints. The use of standardized linear programming routines permits rapid and relatively inexpensive evaluation of alternative scenarios by planners and decision makers. The paper describes the modeling techniques used and areas of approximation in the techniques and provides a hypothetical example to illustrate the type of results available. It concludes with a brief discussion of areas that warrant further investigations to assist in developing the techniques discussed.

Extensive interest is currently focused on park-and-ride facilities because of their beneficial impacts such as fuel savings, reduced air pollution and downtown traffic congestion, and more limited use of the private automobile in general.

Also, specific policy actions such as transportation system management (TSM) and transportation control planning (TCP) measures have emphasized public transportation and, therefore, the need for park-and-ride facilities. These factors, in addition to federal, state, and local involvement in the planning process, require a rational, coordinated planning effort.

Accordingly, this paper describes investigations and initial results of a transportation analysis approach, by using mathematical programming techniques to minimize the fuel needs of park-and-ride users consistent with TSM and other appropriate public policy concerns.

ENERGY SAVINGS AND PARK-AND-RIDE LOCATION

If the locations and number of spaces at specific park-and-ride facilities can be arranged to minimize the vehicle kilometers traveled by commuters who drive to main-line transit facilities, the maximum amount of fuel savings will result from this category of user. At present, however, many commuting motorists do not drive to the nearest (or any) park-and-ride facility. Their choice is affected by mode selection factors, which may include the following park-and-ride-related concerns:

1. Access problems from adjacent arterials and freeways;
2. Certain lots being full before others;
3. Cost of using specific lots;
4. Level of attractiveness (fare, travel time, or convenience) of associated main-line transit;
5. Restrictions by specific municipalities on the use of lots within their jurisdiction (often due to environmental considerations) and community concern about traffic impacts;
6. Comparative advantage of using automobile versus main-line transit for specific route segments; and
7. Security conditions at specific parking lots.

An areawide or corridor plan that details the locations and sizes of park-and-ride facilities to minimize vehicle kilometers of travel consistent with the above policy and operational concerns can provide a guide for planners faced with coordinating transit (rail and bus) and related programs. The plan becomes all the more useful if changes in inputs (such as costs, vehicle occupancies, and development constraints) can be readily accommodated in the analysis process.

PLANNING METHODOLOGIES

Current Approaches

Various methods of mode-of-access analysis have been proposed to rationally estimate the demand for park-and-ride spaces for motorists who use public transportation for the major portion of their commuting trip. Mode of access means the mode of transportation used between home and the change-of-mode (in this case park-and-ride) facility. These methods may be categorized as descriptive (deterministic or probabilistic) demand models or normative (performance-oriented) models.

Examples of recent deterministic models are those of Abdus-Samad and Grecco (1, 2), which use linear regression analysis based on experience at existing facilities. A similar mathematical basis is provided by Keck and Liou (3), primarily based on motorists' travel times and costs. Probabilistic methods of estimating demand include those used for Altrincham, England (4), and a probit analysis used for facilities in Washington, D.C. (5).

Regarding normative approaches to park-and-ride planning that specify or define the performance of a given plan for the allocation of parking spaces in terms of criteria such as energy saving and increased transit ridership, relatively few methodologies have evolved. One example, proposed by Schneider and others (6) for use in Seattle, uses interactive computer graphics and a worth score of travel characteristics to estimate, iteratively, a preferred plan. A linear programming approach to the allocation of parking spaces for minimum vehicle kilometers of travel and minimum cost objectives subject to community and other constraints on the number of spaces and associated TSM and TCP measures has been proposed by Schoon and others (7, 8).

Mathematical Programming Approach for Park-and-Ride Planning

Development of normative park-and-ride planning models can assist in transportation planning by providing es-

timates of systems performance in terms of vehicle kilometers of travel and cost, subject to transportation operations and management programs and to public policy constraints.

The analysis outlined in this paper, therefore, is the result of ongoing efforts at Northeastern University to formulate a methodology for developing park-and-ride plans, specifically in terms of energy savings and monetary costs, that are also responsive to community concerns and transportation operations efforts. The methodology is being designed to be as direct as possible in its inputs, analysis processes, and outputs to enable planners and decision makers to explore a full range of scenarios and policies with maximum flexibility.

Some of the more significant capabilities of the mathematical programming approaches to park-and-ride planning are as follows:

1. Determination of the absolute minimum total vehicle kilometers of travel and the allocation of lots and spaces can be made, consistent with any given set of parameters and constraints. This may also be done manually for a limited number of park-and-ride facilities; however, it becomes tedious or impractical when a large number of facilities are planned.
2. Determination of the allocation of lots and the number of parking spaces within each lot to provide the absolute minimum cost for any given set of parameters and constraints can be determined.
3. TSM actions can be tested to see how effective they may be in terms of energy use on an areawide or corridor basis. This can be done by varying the input parameters such as facility locations, vehicle occupancy, and facility costs.
4. Effects of community-related requirements can be explored by varying input constraints such as the maximum number of available spaces at a given location. For instance, if the likelihood of obtaining parking spaces at one lot is very low, the effects on areawide vehicle kilometers of travel due to a potential redistribution of park-and-ride facilities can be determined.
5. Under contingency circumstances, it may be necessary to induce motorists to park at specific park-and-ride facilities in order to limit fuel consumption to predetermined levels. A mathematical programming approach can determine which park-and-ride location is preferable for each motorist in order to attain this objective.

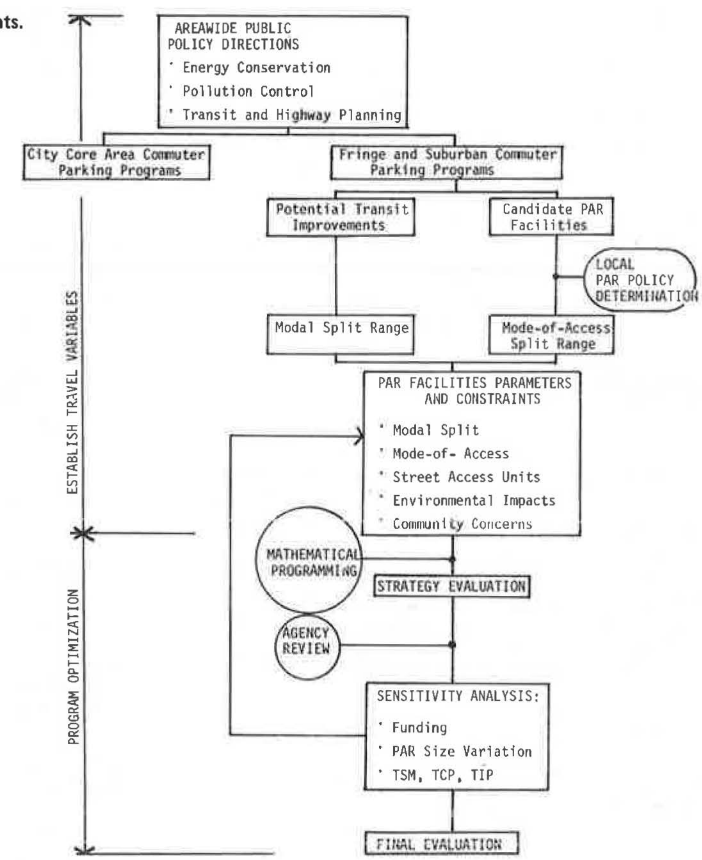
Thus, the mathematical programming approach can be considered as an extension of demand analysis. By using this concept, demands that result from uncontrolled or controlled variables can function as input parameters and constraints that affect the attainment of objectives by means of a specific park-and-ride plan.

PRINCIPAL FEATURES OF THE METHODOLOGY

Overall Process

Overall concepts associated with formulation of an integrated, areawide or corridor plan for park-and-ride facilities by use of mathematical programming techniques can be illustrated as shown in Figure 1. This diagram summarizes the technical analysis and public policy factors that lead to a final presentation of alternative strategies or plans for the location and size of park-and-ride facilities in fringe or suburban areas. The process is divided into two principal stages: (a) the initial establishment of travel variables such as modal split and vehicle occupancy and (b) the analysis process that leads

Figure 1. Park-and-ride planning process—principal elements.



to initial results, which can be used as guidelines for the park-and-ride plan and as inputs to a number of iterations for progressively refining the plan and incorporating potential operational and policy options. The major features of this two-stage process are described below.

Establish Travel Variables

The essential steps in the identification of zone-specific variables are as follows. First, identify line-haul public transportation facilities within the area, including service frequency, capacity, level of service, usage determinants, and a detailed investigation of park-and-ride and other facilities associated with public transportation. The investigation includes an inventory of facilities, user origin patterns, fares, and usage levels. Next, identify existing and potential park-and-ride locations and capacities. Those at shopping centers, highway interchanges, and other locations can be identified at this time, as well as the more usual rail-transit-related park-and-ride lots.

Based on analysis of the extent of detail required, determine the extent and boundaries of line-haul transit station influence areas or analysis zones. Then, formulate station influence area or zone-specific constraints, including those for community concerns, modal split, mode of access, and parking-lot capacities. Formulation of the appropriate limits for use in the constraints will result from factors such as carpooling, extent of likely feeder bus, kiss-and-ride, dial-a-ride service, and other mode-of-access determinants.

Progressive Optimization and Evaluation

This stage of the process is concerned with the optimal capacity of the park-and-ride lot and the evaluation, iteration, possible modification, and assessment of the implications of each scheme. The optimization process for each of the strategies mentioned previously is conducted, with the necessary iterations, as follows:

1. Conduct initial optimization for each alternative by using the measure of effectiveness defined from an initial assessment of vehicle kilometers of travel and costs by using the constraints defined earlier;
2. From the initial allocation of park-and-ride facilities, reassess vehicle kilometers of travel for each zone and conduct the first iteration optimization of each strategy;
3. Conduct further iterations by modifying vehicle kilometers of travel inputs until the final optimization is achieved;
4. Conduct sensitivity analyses to establish implications of varying levels of investment and other determinants;
5. Present findings and implications of each of the strategies under the defined constraints; and
6. Modify public policy emphasis, funding levels, or other constraints if implications of the initial master plan are unacceptable or if policy options require modification.

Although linear programming is the specific optimization method described here, potential may also exist for other forms of analysis such as goal programming or dynamic programming.

Linear Programming Applications

Linear programming is often used as a tool for selecting a course of action given a quantitatively defined objective and associated constraints. Land use and transportation planning applications of linear programming to determine the optimum location of land uses related to transportation facilities approaches have been described by Herbert and Stevens (9), Harris (10), and Blunden (11). Recent investigations in network planning, which involved linear programming techniques (12-14), have indicated a potential for its use, although effective applications have often been hampered by a lack of truly quantitative data and difficulties in controlling levels of the factors involved.

A formal mathematical statement of the general linear programming problem may be stated as follows: Find $x_1, x_2, \ldots, x_n$, which maximizes (or minimizes) the linear function

$$z = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n \tag{1}$$

subject to the restrictions

$$a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n \leq b_1 \tag{2}$$

$$a_{21} x_1 + a_{22} x_2 + \ldots + a_{2n} x_n \leq b_2 \tag{3}$$

$$a_{m1} x_1 + a_{m2} x_2 + \ldots + a_{mn} x_n \leq b_m \tag{4}$$

where $x_1 \geq 0$, $x_2 \geq 0$, $\ldots$, $x_n \geq 0$ and a_{ij}, b_i, and c_j are given constants.

The above model, interpreted in terms of park-and-ride planning, states that, given n competing activities, the decision variables $x_1, x_2, \ldots, x_n$ represent the levels of these activities (the number of parking spaces in each of n park-and-ride lots). If each activity is the formation of units of the jth product, c_j is the increase in the overall measure of effectiveness (vehicle kilometers of travel or cost) that results from production of each unit of a corresponding product. The number of relevant scarce resources is m, and each of the m linear inequalities expresses a restriction (constraint such as available land for park-and-ride spaces) on one of these resources. Each b_i is the amount of resource i (such as total program vehicle kilometers of travel or cost) available to the n activities, and a_{ij} is the amount of resource i consumed by each unit of activity j. The total usage of the respective resources is given by the left side of these inequalities. The nonnegativity restrictions (x_j, 0) express the fact that a negative quantity of an activity cannot exist.

EXAMPLE: PARK-AND-RIDE FACILITIES IN A RADIAL CORRIDOR

Problem

A frequent situation is the problem of allocating park-and-ride spaces in lots throughout a specific corridor served by various forms of main-line transit for the major portion of the commuting trip between home and the city core area. The hypothetical, simplified example presented here illustrates the main features of a linear programming approach.

The following features are assumed in the corridor analysis:

1. Five stations (referred to as Q, R, S, T, and U) constitute the possible park-and-ride stations in a corridor (see Figure 2);
2. The average vehicle kilometers of travel per vehicle associated with each station is such that the distances between users' homes and park-and-ride facilities (mode-of-access distance) are greater the farther the station is from the core area;
3. Average cost per park-and-ride space associated with each station tends to decrease the farther the station is from the core area;

4. The constraints on the number of parking spaces (modal split, mode of access, street access, and community concerns) have been consolidated to provide one upper and one lower level of parking space constraints for each station;

5. If the number of available parking spaces at a given park-and-ride facility is reduced, it is assumed that park-and-ride users will divert to the next park-and-ride location nearest the core area; and

6. For purposes of simplifying the example and for clarifying the essential relationships, the effects of kiss-and-ride and feeder bus users have not been included.

The major strategies to be examined will be called strategy 1, minimization of vehicle kilometers of travel, and strategy 2, minimization of system costs. These two strategies represent the extreme points of the relationship of cost and vehicle kilometers of travel for the specific constraints and, as such, assist in defining the cost and vehicle kilometers of travel domain within which possible variations in the master plan can be formulated.

Formulation of the model, indicating the strategy and alternative breakdowns and the associated constraint levels, is shown in Figure 3. This summarizes the essential components described above and provides the inputs for the linear programming analysis. Key parameters, constraints, and other features of the problem formulation are also shown in this figure.

The major outputs of the analysis for strategies 1 and 2, by using a standard computer linear programming package, are summarized in Table 1. The principal points of note in this table concerning vehicle kilometers of travel and total costs are that the minimum attainable vehicle kilometers of travel for the commuters in the example corridor for the specified parameters and constraints is 32 536 km. The corresponding cost for implementation of this plan is \$52 621. When the minimum cost of implementing a plan consistent with the specified parameters and constraints is the objective, the total cost can be reduced to \$45 208. However, under this plan the corresponding vehicle kilometers of travel increases to 54 628 km.

Differences that correspond to the above can be seen in terms of average vehicle kilometers of travel and cost per park-and-ride space. A check on the allocated park-and-ride spaces at each of the stations indicates that all of the specified constraints are met.

The basic output data shown in Table 1 also provide overall assessment of each strategy and establish relationships between key variables that can be adjusted to investigate sensitivity and general relationships.

Sensitivity Analysis

The sensitivity of the cost versus vehicle kilometers of travel relationship to changes in vehicle occupancy, unit costs, and reductions in core parking is shown in Figure 4. This illustrates, for specified changes in these parameters, how the cost and vehicle kilometers of travel

Figure 2. Example of layout of park-and-ride facilities.

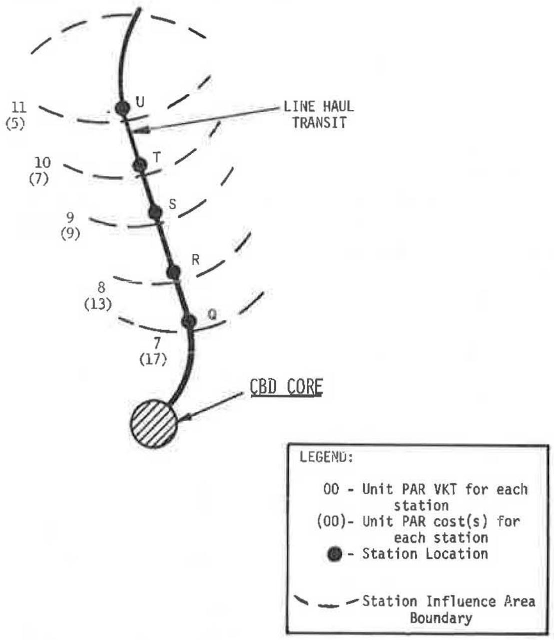
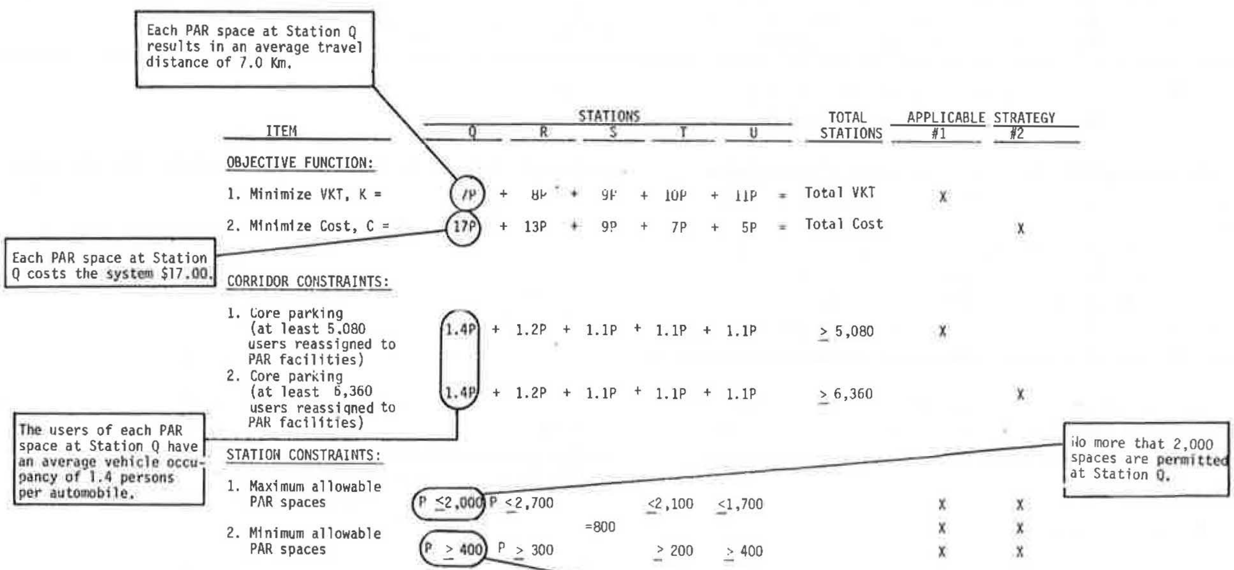


Figure 3. Example of problem formulation.



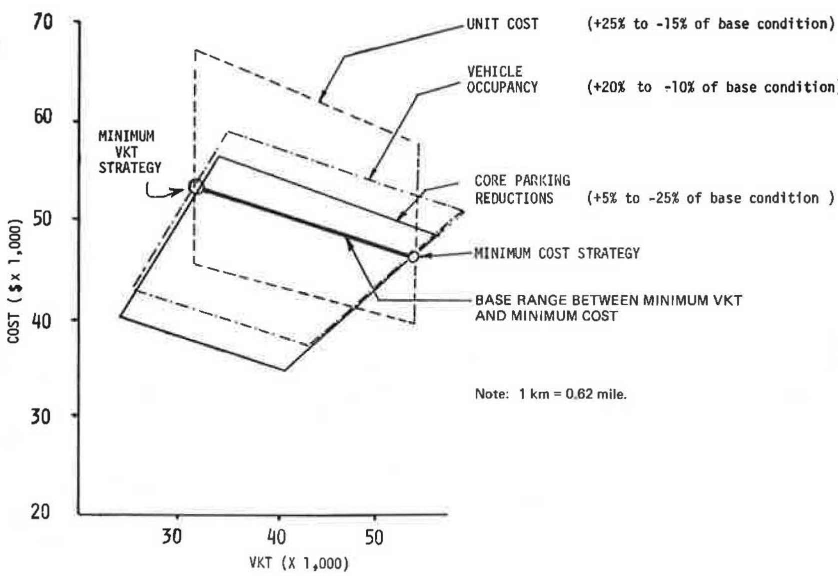
Notes: 1 km = 0.62 mile.
See Figure 2 for VKT and cost parameters.
P = number of PAR spaces.
For corridor constraints, total reassigned spaces may differ due to areawide PAR program requirements.

Table 1. Park-and-ride optimization plan—summary of results.

Station	Park-and-Ride Spaces	Vehicle Occupancy	Persons	Distance (km)	Vehicle Kilometers Traveled (km)	Unit Cost ($)	Cost ($)
Strategy 1							
Q	2000	1.4	2800	11	14 000	17	34 000
R	617	1.2	740	13	4 936	13	8 021
S	800	1.1	880	15	7 200	9	7 200
T	200	1.1	220	16	2 000	7	1 400
U	400	1.1	440	18	4 400	5	2 000
Total	4017		5080		32 536		52 621
Strategy 2							
Q	400	1.4	560	11	2 800	17	6 800
R	616	1.2	740	13	4 928	13	8 008
S	800	1.1	880	15	7 200	9	7 200
T	2100	1.1	2310	16	21 000	7	14 700
U	1700	1.1	1870	18	18 700	5	8 500
Total	5616		6360		54 628		45 208

Notes: 1 km = 0.62 mile.
For strategy 1, the average vehicle kilometers traveled to a park-and-ride space is 8.1 and the average cost is $13.1.
For strategy 2, the average vehicle kilometers traveled to a park-and-ride space is 9.7 and the average cost is $8.1.

Figure 4. Relationship between cost and vehicle kilometers of travel and example of selected changes in input parameters and constraints.



will change. It substantiates and quantifies the intuitive analysis that

1. Reductions in vehicle occupancy will simultaneously reduce costs and vehicle kilometers of travel and vice versa.
2. Increases in unit costs will increase total costs but will not increase vehicle kilometers of travel provided that no upper total cost is imposed on the program.
3. An increase in parking restrictions in the city core will require a greater provision of park-and-ride facilities. Hence, a greater increase in total cost and vehicle kilometers of travel associated with park-and-ride facilities will result. A decrease in the core parking requirements will have the opposite effect.

The sensitivity relationships illustrated in this example show changes in the basic condition when changes are made in only one parameter at a time. However, different combinations of parameters and constraints can be changed simultaneously. Also, the linear programming method would result in step functions rather than the generalized straight-line relationships shown in Figure 2.

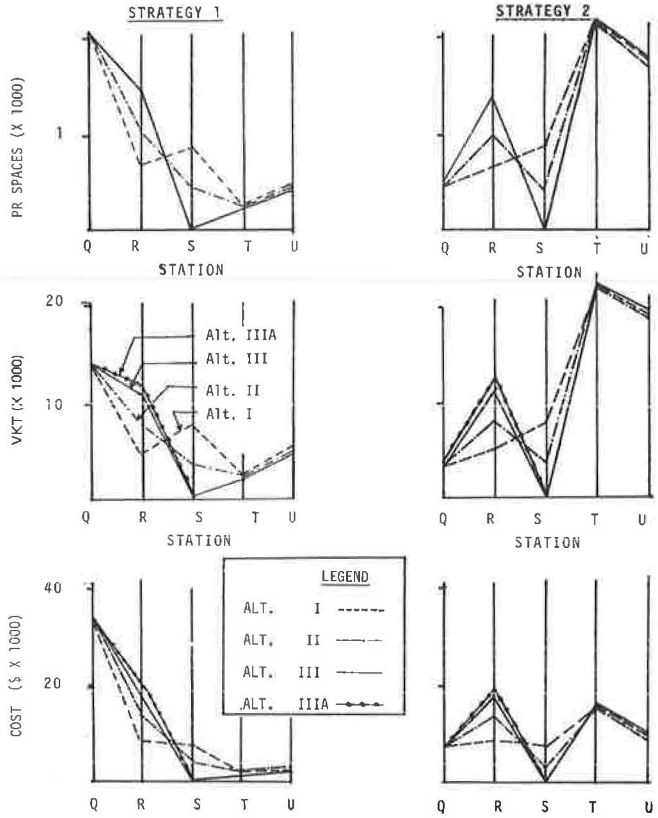
Another important area of interest in sensitivity analysis is in exploring the effects of varying the acceptable number of parking spaces at a specific station. Consider, for example, the following four alternatives associated with station S:

Alternative 1—800 park-and-ride spaces must be provided at station S;
Alternative 2—upper limit of 400 park-and-ride spaces must be provided at station S;
Alternative 3—0 spaces allowed at station S; and
Alternative 3A—same as alternative 3, but with increased vehicle kilometers of travel for users from influence area of station S due to their diversion to station R.

Based on substitution of each revised constraint in the problem format and rerunning the program, Figure 5 indicates that for strategy 1 (minimization of vehicle kilometers of travel)

1. As the number of available park-and-ride spaces at station S is reduced (alternatives 2 and 3), the excess spaces are allocated to station R because allocation to

Figure 5. Sensitivity to changes in park-and-ride spaces at stations.



station R is compatible with the vehicle kilometers of travel minimization objective, and station R has a sufficiently high potential capacity to accomodate the excess park-and-ride users diverted from station S. Had there been insufficient capacity at station R to accommodate the diverted users, they would either continue to station Q or would be lost to the park-and-ride system (i.e., would select a different destination or would have traveled the entire distance between home and the downtown by automobile).

2. Vehicle kilometers of travel and costs reflect the allocation of spaces to each station.

3. The adjustment in vehicle kilometers of travel due to the larger influence zone of stations R plus S (alternative 3A) indicates a relatively small difference in the park-and-ride space, vehicle kilometers of travel, and cost apportionment.

For strategy 2 (minimization of costs), the apportionment of park-and-ride spaces for alternatives 2 and 3 emphasizes a reassignment primarily to stations T and R consistent with the cost-minimization strategy. Also, the park-and-ride spaces are assigned to station R when the upper constraints on station T are reached. As with strategy 1, additional average vehicle kilometers of travel due to combining influence areas for stations R and S is relatively minor.

The linear programming approach offers considerable opportunity to conduct further sensitivity analyses. A summary list of how these analyses can be approached in response to TSM concerns is as follows:

TSM Element	Approach to Inclusion of TSM Element in Park-and-Ride Planning Methodology
Improved vehicular flow	Affects access time between home and park-and-ride facilities; travel time and cost for line-haul transit vehicles include mode-of-access, modal-split, and cost parameters
Preferential treatment for high-occupancy vehicles	Improved travel times and routes of line-haul vehicles included in modal split, time, and cost estimates
Reduced peak-period travel	Included in estimates of access time and demand levels at park-and-ride facilities for specific time periods; also will affect transit costs
Parking management	Center-city parking restrictions will result in a transference of spaces to park-and-ride facilities; increased number of park-and-ride parking spaces must be allocated to provide minimum vehicle kilometers of travel
Promotion of high-occupancy and nonvehicular travel modes	Increased vehicular occupancy is included in basic linear programming format (see also example)
Transit and paratransit service improvements	Increased transit use will affect modal split and total passenger volumes on specific routes; paratransit between home and park-and-ride locations will lower the need for park-and-ride parking spaces
Transit management efficiency measurements	Reductions in transit operating costs and improvements in operating efficiency will affect modal split and transportation system costs

CONCLUSIONS

Current provision of park-and-ride facilities attempts to achieve maximum use of line-haul public transportation from city centers. Park-and-ride planning described here can generate information on a range of options to form the basis for decision making.

Areas in which further data and research are desirable in this context include consideration of trip-making characteristics of park-and-ride users, estimation of park-and-ride-related travel costs, mode-of-access analysis, and variations in the mathematical programming approach.

Linear Approximation of Variables

The use of linear programming techniques implies that relationships used in the model, such as travel cost versus vehicle kilometers of travel, behave in a linear fashion; however, such relationships are rarely linear. Investigations should be made to determine the effects of such approximations on the accuracy of the results, within the typical limits likely to be obtained in practical situations.

Trip-Making Characteristics of Park-and-Ride Users

Although we can assume that most park-and-ride users will travel to the park-and-ride facilities nearest to their trip origin before they begin the line-haul portion of the trip (thus simplifying the estimation of average mode-of-access trip distances), investigations are warranted into the effects of imposing limits on certain park-and-ride facilities on the average trip length and the choice of park-and-ride facility.

Note that users from the same zone often have a different choice of park-and-ride facilities, depending on the time at which the journey is started, due to changes

in occupancy of facilities, temporary access deficiencies, and congestion points that fluctuate in extent and severity.

Urban Travel Patterns and Modal-Split Analysis

Probably the major effects of restrictions on private automobile use to and from city core areas and the associated park-and-ride facilities program will be the result of an imposition of a ceiling level on the use of a particular mode, which will thus distort the free demand level. This will also occur at certain of the park-and-ride facilities. In particular, two factors arise from the policy of imposing limits on parking space:

1. Economic discrimination between user categories if pricing policies are used as the mechanism for reducing parking needs (those users more able to meet increased costs will have a corresponding mobility advantage) and
2. Decreased advantage to some captive automobile users whose schedules may conflict with restricted availability of parking spaces.

Also, as mentioned earlier, the determiners of modal split frequently are not linear functions and each specific case would have to be assessed to determine the actual effects on modal split.

Mathematical Analysis Variations

The analysis approach described earlier to provide optimal master plans for park-and-ride facilities that have alternative objectives and constraints exhibits two distinct mathematical stages. First is the estimation stage in which the parameters such as average vehicle kilometers of travel per zone per vehicle and vehicle occupancy were estimated and also in which predictions about the range of modal split, mode of access, and other constraints were made. Second is the prescriptive stage in which the estimates of the first stage were assembled to provide boundary conditions within which levels of the variables could be determined in order to attain defined (or prescribed) objectives. In this second stage, a linear programming methodology was used and, as shown in the example, the linear approximations to the nonlinear functions were investigated.

In addition to the basic linear programming format, a number of refinements and variations could prove advantageous in more detailed studies. The most likely applications in this regard are the use of dynamic programming or of linear programming under uncertainty—involving either stochastic programming or chance-constrained programming. Also, goal programming, where each potential objective is ranked in terms of its priority, offers the potential for future applications.

ACKNOWLEDGMENT

I would like to acknowledge the comments and assistance given by J. C. Falcocchio, L. J. Pignataro, and W. McShane of the Polytechnic Institute of New York during earlier work on the subject research and the assistance of the Department of Civil Engineering, Northeastern University, during the current investigations. However, I am solely responsible for the content of the paper.

REFERENCES

1. U. R. Abdus-Samad. Change of Mode Parking Facilities. Purdue Univ., Lafayette, IN, Joint Highway Research Project, No. 1, Feb. 1972.
2. U. R. Abdus-Samad and W. L. Grecco. Sensitivity Analysis of Community Savings Due to Change-of-Mode Operations. TRB, Transportation Research Record 557, 1975, pp. 1-11.
3. C. A. Keck and P. S. Liou. Forecasting Demand for Peripheral Park-and-Ride Service. TRB, Transportation Research Record 563, 1976, pp. 63-74.
4. G. R. Niblett. The Altrincham Line Demonstration Project. Traffic Engineering and Control, Vol. 15, No. 4/5, Aug.-Sept. 1973, pp. 196-199.
5. Wilbur Smith and Associates; London Transport Executive; Main Lafrentz and Company; Allen T. Eaton. Mode of Access Model. Washington Metropolitan Area Transit Authority, Washington, DC, Memorandum Rept. No. 15, 1974.
6. J. B. Schneider, D. G. Miller, and T. W. Friedman. Locating and Sizing Park-and-Ride Lots with Interactive Computer Graphs. Transportation, Vol. 5, No. 4, 1976, pp. 389-406.
7. J. G. Schoon, J. C. Falcocchio, L. J. Pignataro, and W. R. McShane. Public Policy and Optimal Public Transportation Planning Strategies. TRB, Transportation Research Record 614, 1976, pp. 14-20.
8. J. G. Schoon. Change-of-Mode and Optimal Planning Strategies. Polytechnic Institute of New York, Brooklyn, Ph.D. thesis, 1977.
9. J. D. Herbert and B. H. Stevens. A Model for the Distribution of Residential Activity in Urban Areas. Journal of Regional Science, Vol. 2, Fall 1960.
10. B. Harris. Linear Programming and the Projection of Land Uses. Penn-Jersey Transportation Study, Philadelphia, Penn-Jersey Paper 20, 1962.
11. W. R. Blunden. The Land-Use/Transportation System. Pergamon Press, Oxford, England, 1971.
12. M. L. Manheim. Search and Choice in Transport System Analysis. HRB, Highway Research Record 293, 1969, pp. 54-82.
13. F. Ochoa-Rosso and A. Silva. Optimum Project Addition in Urban Transportation Networks via Descriptive Traffic Assignment Models. In Search and Choice in Transport System Planning, MIT Press, Cambridge, MA, Vol. 5, 1968.
14. E. K. Morlock, N. L. Nihan, and R. F. Sullivan. A Multi-Mode Transportation Network Design Model. Transportation Center, Northwestern Univ., Evanston, IL, Research Rept., 1970.

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

Abridgment

Comparison of Multicriteria Optimization Methods in Transport Project Evaluation

Terry L. Friesz, Francisco A. Tourreilles, Anthony Fu-Wha Han, and J. Enrique Fernandez

The evaluation of transport projects frequently requires consideration of multiple criteria other than or in addition to economic efficiency. Nonetheless, few of the important methodological advances of recent years in the areas of multicriteria decision making and multicriteria optimization have been discussed in the literature about transport project evaluation. This paper compares certain key multicriteria optimization methods and illustrates their relative strengths and weaknesses through application to a hypothetical transport project in a developing country. In particular, the results of an earlier paper that compares the weighting method suggested by Zadeh and Marglin to the iterative preference-incorporation method of Geoffrion, Dyer, and Feinberg are combined with new analyses based on the constraint method suggested by Marglin and the yes-no iterative algorithm of Zionts and Wallenius. Conclusions are drawn concerning the relative attractiveness of these solution methods and the characteristics of an ideal multicriteria optimization algorithm for the evaluation of transport projects.

That the evaluation of transport projects involves the consideration of multiple criteria or objectives that are noncommensurable is increasingly recognized by planners and engineers. Although the importance of multiple criteria has been recognized, agreement as to an appropriate methodological approach for handling noncommensurable criteria in transport project evaluation has not been reached.

A useful vehicle for examining the role of multiple criteria in transport project evaluation is provided by rural roads in developing countries. The problem of investment in rural roads may be used to illustrate the application of standard multicriteria evaluation tools and to compare the attractiveness of solution methods.

Multicriteria evaluation problems are frequently most naturally articulated as vector mathematical programming problems, as the rural roads example of this paper will illustrate. Methods for the solution of vector mathematical programming problems can be divided into two categories [Cohon (1)]:

1. Generating methods that identify all efficient solutions and
2. Preference-incorporation methods that use decision-maker preferences to examine only a subset of all efficient solutions.

We will give a detailed illustration in subsequent sections of the characteristics of the so-called constraint method, one of the most widely known generating techniques, which is generally attributed to Marglin (2), and the iterative preference-incorporation technique of Zionts and Wallenius (3). The constraint method offers certain computational advantages over another standardly employed generating technique, the so-called weighting method, which was also suggested by Marglin (2). Similarly, the Zionts and Wallenius technique provides, at least theoretically, certain advantages over the preference incorporation of Geoffrion and others (4), which is sometimes considered the prototypical preference-incorporation method. For comparison we will also review the results of applying the weighting method and the method of Geoffrion and others.

IDENTIFICATION OF THE RURAL ROAD PROBLEM

The lack of adequate transportation facilities has been a major determining factor of rural underdevelopment in developing countries, particularly through its constraining effect on the agricultural sector. In order to analyze alternative rural transport project evaluation techniques, we will work with a simplified model based on the following assumptions:

1. Transport cost savings caused by a road investment project are fully transferred to agricultural producers in the form of correspondingly higher prices,
2. Total cultivable land area is fixed,
3. Agricultural producers are price takers,
4. Agricultural producers are profit maximizers, and
5. The total amount of the homogeneous agricultural product of concern is marketed only after transport over the road system.

Let us further suppose that we have two agricultural regions, 1 and 2, that follow the assumptions above and that are connected through roads R_1 and R_2, respectively, to market A where their output is sold. Both regions draw on common fixed resources. Improvement to road R_i will decrease the market transport cost and will thus increase the product price perceived by the producer. As a result, agricultural production in region i will expand. On the other hand, a production increase in one region will lead to a production decrease of smaller magnitude in the other region since the common resource's share to the latter will diminish as a consequence of the production expansion in the former. A transport investment program is being considered that will improve both roads R_1 and R_2.

Before the implementation of the R_1 improvement the perceived price is p_0, and total output is determined by $p_0 = MC_0$ (price equals marginal cost of production) at q_0, which is the profit-maximizing output. After the implementation of the R_1 improvement, the product demand curve (as perceived by the farmers) shifts upward to $p = p_1$ due to transporation cost reduction. If the improvement of R_2 is also implemented, production in region 2 will increase and, as a result, the marginal cost curve of region 1 will shift upward from MC_0 to MC_1. Final equilibrium output in region 1 will be q_2, which is larger than q_0 but less than q_1, the level that could have been achieved had no R_2 investment been made.

We assume that the objective of the investment program is to maximize the vector of regional agricultural

productions and we make the following definitions:

$Z = [Z_1, Z_2]$,
Z_i = agricultural production level of region i,
X_i = amount of investment in road R_i ($000 000s),
$\alpha, \beta, \gamma, \delta > 0$,
$\beta, \gamma < 1$,
b_1 = total national (both regions) budget devoted to rural road improvements ($000 000s),
b_2 = a maximal allowable excess of region 2's transport investment share over region 1's (a nonnegative number) ($000 000s),
b_3 = rural road improvement budget of region 1 ($000 000s), and
b_4 = rural road improvement budget of region 2 ($000 000s).

Then our problem becomes

$$\text{MAX } Z = [Z_1, Z_2] \quad Z_1 = \alpha X_1 - \beta Z_2 \quad Z_2 = -\gamma Z_1 + \delta X_2 \tag{1}$$

subject to:

$$X_1 + X_2 \leq b_1 \quad -X_1 + X_2 \leq b_2$$
$$X_1 \leq b_3 \quad X_2 \leq b_4 \quad X_1, X_2 \geq 0.$$

The constraints are purposely written to allow the implicit bounding $X_1 - X_2 \leq b_3$; this simplifies the exposition. The more general constraint $|X_2 - X_1| \leq b_2$ may be introduced if desired with relatively simple modifications of the numerical examples discussed subsequently. In addition, we will assume that a bicriterion welfare function is defined and given by

$$U = U[Z_1(X), Z_2(X)] \tag{2}$$

where X denotes the vector of decision variables (X_1, X_2). The function $U(\cdot)$ is a monotonically increasing function of the objectives with convex isoquants in objective space.

Solution of an Example Problem

In this section, we show that the rural road problem can be reformulated as a (linear) multiobjective programming problem. We then assign some numerical values to the coefficients and solve the example problem by a standard generating method, the constraint method [Marglin (2)], and by the iterative preference-incorporation technique of Zionts and Wallenius (3). These results are then compared to those obtained from the weighting method [Marglin (2)] and the iterative procedure of Geoffrion and others (4). Friesz and others (5) show that

$$Z_1 = aX_1 - bX_2 \tag{3a}$$

and

$$Z_2 = cX_1 + dX_2 \tag{3b}$$

where the coefficients are all positive. To illustrate the solution process, we assign some arbitrary positive integers as the coefficients of expressions 3a and 3b and of the constraint set stated in problem 1. The example problem with two objectives and two decision variables is thus as follows:

$$\text{Maximize } Z(X_1, X_2) = [Z_1(X_1, X_2), Z_2(X_1, X_2)] \tag{4}$$

where

$$Z_1(X_1, X_2) = 6X_1 - 1X_2 \quad Z_2(X_1, X_2) = -1X_1 + 4X_2$$

subject to:

$$-X_1 + X_2 \leq 3 \quad X_1 + X_2 \leq 8 \quad X_1 \leq 6 \quad X_2 \leq 4 \quad X_1, X_2 \geq 0.$$

We now discuss techniques for solving the example problem 4. Throughout the discussion that follows, the general solution of a multicriteria or vector optimization problem is taken to be the set of all efficient or noninferior alternatives. A noninferior alternative is a feasible alternative such that it is impossible to improve one objective (criterion) without causing a degradation in at least one other objective. The noninferior alternative that maximizes some aggregate social welfare measure (in our two-objective case, the welfare function, Equation 2) is termed the best compromise solution.

Application of the Constraint Method

We can apply the constraint method [see Cohon (1)] to generate the noninferior set for Equation 4 in the following steps.

1. Step 1. Construction of a payoff table; solve individual maximization problems to find the optimal solution for each of the objectives. If there is more than one optimum for any of these problems, then choose the noninferior solution from among the alternative optima.
2. Step 2. Construction of a single objective problem; transform the multiobjective optimization problem into a single objective-constrained problem.
3. Step 3. Define lower bounds for the objectives placed in the constraint set during step 2.
4. Step 4. Solve the constrained problem of step 2 for the lower bounds determined in step 3.

The results of the application of the constraint method to the example problem with n = 4 are summarized below.

Bound (L_2)	Optimal Value of Objective Function (Z_1^*)	Optimal Solution: In Decision Space (X_1^*, X_2^*)	Optimal Solution: In Objective Space (Z_1^*, Z_2^*)
-6	36	(6, 0)	(36, -6)
1	34.25	(6, 1.75)	(34.25, 1)
8	25.6	(4.8, 3.2)	(25.6, 8)
15	2	(1, 4)	(2, 15)

Application of the Zionts and Wallenius Method

Zionts and Wallenius (3) developed an iterative preference-incorporation procedure by which the decision maker is requested to provide yes-no answers to questions regarding certain trade-offs. The procedure leads to an approximation of the best compromise solution. It is assumed in this method that all relevant criteria or objective functions are concave functions to be maximized and that the constraint set is convex. The overall utility function is assumed to be unknown to the decision maker, but it is implicitly assumed to be a linear function or, more generally, a concave function of the individual scalar objective functions.

We are interested in the underlying problem of maximizing a multiattribute utility or welfare function defined over K objectives, which can be denoted symbolically as

$$\text{Maximize } U[Z_1(X), Z_2(X), \ldots, Z_K(X)] \tag{5}$$

subject to

$$X \epsilon F_d \quad (6)$$

where X is a vector of decision variables and F_d is the feasible region of the problem. The basic structure of the Zionts and Wallenius algorithm that addresses this problem is as follows:

1. Step 1. Initialization. Determine a set of initial weights λ_i^0 for each objective i = 1,...,K. Let $\lambda_i = \lambda_i^0$ for all i.

2. Step 2. Consider the linear program

$$\text{Maximize } U = \sum_{i=1}^{K} \lambda_i Z_i(X) \quad (7)$$

subject to

$$X \epsilon F_d \quad (8)$$

where F_d is the feasible region and X is the vector of decision variables. Solve this linear program to obtain solution X^0 and the reduced cost w_{ij} of each nonbasic variable X_j with respect to objective i.

3. Step 3. Partition the set of nonbasic variables into two subsets: those that when introduced into the basis lead to efficient adjacent extreme points in objective space and those that do not. Call these, respectively, efficient and inefficient variables.

To check if a nonbasic variable X_ℓ is efficient, consider the following linear program:

$$\text{Minimize } Z = \sum_{i=1}^{K} w_{i\ell} \lambda_i \quad (9)$$

subject to

$$\sum_{i=1}^{K} w_{ij}\lambda_i \geq 0 \quad j \neq \ell \text{ and } X_j \text{ nonbasic} \quad (10)$$

$$\sum_{i=1}^{K} \lambda_i = 1 \quad (11)$$

$$\lambda_i \geq 0 \quad i = 1, \ldots, K \quad (12)$$

Solve this linear program where the w_{ij}'s are reduced costs obtained in step 2. If Z is negative, X_ℓ is efficient; otherwise X_ℓ is inefficient. If no efficient variables are found, stop. The best compromise solution has been reached. Otherwise go to the next step.

4. Step 4. For each efficient variable X_j, the decision maker is asked to respond yes or no to a potential trade-off among objectives. These trade-offs are described by the reduced cost w_{ij}. (There will be at least one negative and at least one positive w_{ij} for each efficient variable X_j.) For each yes response construct an inequality of the form:

$$\sum_{i=1}^{K} w_{ij}\lambda_i \leq -\epsilon \quad (13)$$

where ϵ is a sufficiently small positive number. For each no response, construct an inequality of the form:

$$\sum_{i=1}^{K} w_{ij}\lambda_i \geq \epsilon \quad (14)$$

5. Step 5. If this is the first iteration, go to step 6a. For the second and successive iterations, if the overall utility function is assumed to be a linear additive function, go to step 6a; if the overall utility function is assumed to be a general concave function of objectives that are, in turn, linear functions of the decision variables, go to step 6b.

6. Step 6. (a) Put the constraints that are generated in step 4 together with all previously constructed constraints, including Equations 11 and 12. Find a feasible solution λ_i^*, i = 1,...,K that satisfies all constraints of the form of Equations 11-14. Let $\lambda_i = \lambda_i^*$ for all i and go to step 2.

Next, (b) query the decision maker to determine whether the new solution is preferred to the old. If so, all old responses are purged. This implies that only the constraints generated in step 4 of the current iteration together with constraints 11 and 12 will be used to generate new feasible weights λ_i. Find a feasible solution λ_i^*, i = 1,...,K that satisfies all these constraints. Let $\lambda_i = \lambda_i^*$.

If the old solution is preferred to the new solution or if the decision maker is indifferent, the analyst presents the decision maker with all the efficient solutions adjacent to the old solution in objective space. If any of the adjacent efficient solutions are preferred to the old solution, the procedure is continued from one of these preferred solutions. That is, find the efficient (nonbasic) variables associated with the preferred adjacent solution and then go to step 4. If none of these adjacent solutions is preferred to the old solution, stop. The optimal solution lies in a neighborhood of the old solution comprised of that portion of the noninferior set defined by the old solution and its adjacent efficient solutions in objective space.

To apply the Zionts and Wallenius algorithm to the example problem we first assume that the multivariate utility function can be expressed in product form $U[Z_1(X), Z_2(X)] = Z_1Z_2$ where, of course, $X = (X_1, X_2)$. The algorithm terminates, having determined that the best compromise solution lies in a neighborhood $X^2 = (4,4)$. This neighborhood is that portion of the noninferior set defined by line segments that join $X^0 = (6,2)$, $X^1 = (4,4)$ and (1,4) in decision space. This corresponds to that portion of the noninferior set defined by line segments that join $Z^0 = (20,12)$, $Z^1 = (34,2)$ and (2, 15) in objective space.

Overview of Results Obtained

Friesz and others (5) applied both the weighting method and the method of Geoffrion and others to the sample problem. The table below summarizes the results for the former method.

Weight		Optimal Value of Objective Function	Optimal Solution	
			In Decision Space	In Objective Space
w_1	w_2	(Z^*)	(X_1^*, X_2^*)	(Z_1^*, Z_2^*)
1	0	36	(6, 0)	(36, -6)
1	1	36	(6, 2)	(34, 2)
1	2	44	(4, 4)	(20, 12)
1	3	56	(4, 4)	(20, 12)
0	1	15	(1, 4)	(-2, 15)

They also showed that, under a hypothetical decision environment, the method of Geoffrion and others correctly converges to the unambiguous best compromise solution $X^* = (4,4)$ with $U^* = Z_1Z_2 = (20)(12) = 240$ when applied to the example problem.

CONCLUSION

We have shown through a hypothetical problem of investment in rural roads that multicriteria optimization methods may be used to assist the decision maker in evaluating transport projects. The methods considered—the generating technique known as the constraint method and the yes-no, iterative preference-incorporation method of Zionts and Wallenius—differ substantially in the types of information required by the decision maker and the degree of interaction between the decision maker and the analyst. The constraint and weighting methods, prototypical of the methods usually classified as generating methods, strive to approximate the noninferior (or efficient) set, under the implicit assumption that knowledge of this set will allow the decision maker to select a best compromise solution. The iterative techniques of Zionts and Wallenius and Geoffrion and others are preference-incorporation techniques and they seek to identify the best compromise solution without generating the entire noninferior set by soliciting preference information from the decision maker. The Zionts and Wallenius method does not lead to a definitive statement of the best compromise solution unless the overall utility measure is assumed to be a linear additive function of the multiple objectives. The method of Geoffrion and others by contrast always leads to a definitive best compromise solution. This precision is at the expense of requiring more sophisticated preference information from the decision maker.

An ideal multicriteria optimization algorithm would accommodate yes-no preference information and be able to treat discrete alternatives (i.e., integer variables). This last capability can only be introduced in the methods discussed here with a severe loss of computational efficiency.

REFERENCES

1. J. L. Cohon. Multiobjective Programming and Planning. Academic Press, New York, 1978.
2. S. Marglin. Public Investment Criteria. MIT Press, Cambridge, MA, 1967, 103 pp.
3. S. Zionts and J. Wallenius. An Interactive Programming Method for Solving the Multiple Criteria Problem. Management Science, Vol. 22, 1976, 652 pp.
4. A. Geoffrion, J. Dyer, and A. Feinberg. An Interactive Approach for Multi-Criterion Optimization with an Application to the Operation of an Academic Department. Management Science, Vol. 19, 1972, 357 pp.
5. T. Friesz, F. Tourreilles, and A. Han. Multicriteria Optimization Methods in Transport Project Evaluation: The Case of Rural Roads in Developing Countries. Proc., Transportation Research Forum, Vol. 20, Oct. 1979.

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

Abridgment

Goal-Programming Approach to Multiobjective Highway Network Design Model

Jossef Perl

A new approach to the highway network design model is presented that allows comparisons of networks on the basis of multiple incommensurable objectives with different degrees of importance. The goal-programming approach is not only capable of solving the multiobjective network design problem in a relatively efficient manner, but it can also be used to generate the multidimensional trade-off curve that provides additional important information to that provided in the two-dimensional curve derived by using the linear programming model. An example illustrates the application of the linear goal-programming model with four objectives.

Decision problems in general, and particularly those associated with transportation systems, are made in the context of multiple conflicting objectives. Decisions about transportation systems should be weighted against the social, economic, environmental, and aesthetic needs of the community. Among the most important decisions in the transportation planning process are those regarding the structure of transportation networks and the level of service to be offered on them.

In recent years there has been interest in the application of advanced analytic techniques to the search for good alternative transportation networks. A class of models known as network design models has been developed to solve the following problem: Given an existing network, a list of improvement options for various links, and projected increase in demand between various origin and destination pairs, select the optimal set of links to be improved or added to the existing network. The models perform two tasks simultaneously: (a) they choose the optimal set of links to be improved or added and (b) they assign the projected traffic to the new network.

This paper deals with the extension of a continuous Highway Network Design Model (HNDM) developed by Agarwal (1-3). The approach adopted by Agarwal follows that used by Hay and others (4) in an urban context and Morlok and others (5) for the Northeast Corridor Transportation Project. The HNDM is a linear programming model developed as a sketch planning tool.

There are two ways to incorporate multiple objectives in linear programming—as elements of the objective function or as constraints. In the first method, various objectives are collapsed into a single objective by using

Table 1. Link characteristics for example network.

Link Segment ij	m	Average Daily Travel Time, C^{m}_{ij}(h)	Average Daily Capacities, k^{m}_{ij}(vehicles/day)	Total Cost, B_{ij}[($1000/vehicle)/day]	Link Length, V_{ij}(miles)	Dwelling Units to Be Relocated per Capacity Unit (000s)
12	1	0.121 43	15 000	600 400	2.0	2
	2	0.805 80	5 000			
14	1	0.040 0	40 500	252 800	1.0	3
	2	0.246 73	12 500			
23	1	0.064 29	15 000	376 200	1.4	4
	2	0.396 27	5 000			
34	1	0.057 14	15 000	347 600	1.4	4
	2	0.376 67	5 000			
45	1	0.100 00	15 000	126 400	1.7	10
	2	0.077 87	5 000			
35	1	0.021 43	15 000	505 600	0.6	3
	2	0.139 27	5 000			

Table 2. Demand matrix for example network.

From Node	To Node 1	2	3	4	5
1		15 000	4516	5 222	7000
2	7685		6218	10 414	1122
3	5517	3 229		711	2217
4	6224	3 429	1704		858
5	1333	2 129	3209	5 817	

some kind of weighting scheme (6). The difficulty with this approach is that it is not possible to collapse objectives that are significantly different to a common unit without a great deal of arbitrary judgment. The problems with the second approach are both conceptual and practical. In linear programming terminology, an objective represents a target and a constraint represents a restriction on feasibility. In practice, if all the constraints cannot be satisfied, the problem is termed infeasible. Such a conclusion is obviously misleading when objectives are formulated as constraints. The inflexibility of linear programming does not allow one to consider some, if not all, of the constraints as not absolutely binding.

The purpose of this paper is to demonstrate an approach to the multiobjective HNDM. The proposed goal-programming approach has significant advantages over the interactive programming techniques for solving multiobjective problems previously applied by Agarwal (5). The goal-programming approach can generate trade-off curves in their full dimensionality. The interactive approaches are very time consuming because they require a continuous interaction with the decision maker. Goal programming can solve multiobjective problems while still employing the simplex algorithm (on a modified basis). This allows exploration of a much larger number of alternatives in a given time period than does an interactive programming technique.

In the process of using the HNDM, information that describes the trade-offs between the achievement of various objectives is perhaps more valuable than the actual point solution. This information is presented by a trade-off curve. A serious issue in the multiobjective trade-off analysis is the display of the multidimensional trade-off curve in two-dimensional space. This paper will demonstrate an approach to the presentation of multidimensional trade-off curves.

GOAL-PROGRAMMING MODEL

A discussion of linear goal programming is beyond the scope of this paper and can be found in Ignizio (7). Linear goal programming can be defined as (7) "a systematic methodology for solving linear, multiple objective problems wherein preemptive priorities and weights are associated with the objectives."

The linear goal-programming model presented here includes the following objectives:

G_1 = flow conservation objective,
G_2 = flow definition objective,
G_3 = link capacity objective,
G_4 = budget objective,
G_5 = level-of-service objective,
G_6 = household relocation objective, and
G_7 = vehicle-miles-of-travel objective.

Since G_1, G_2, and G_3 can be legitimately viewed as absolute objectives (objectives that must be satisfied), they are assigned a top priority. In the following formulation of the linear goal-programming model, the household relocation objective is assigned to priority level two, the level of service objective to priority three, the budget objective to priority four, and the vehicle-miles-of-travel objective to priority five. The linear goal program for a network with N nodes, a set of L undirected links, and a set of S origin and destination nodes can be written as follows:

Find: ${}^{s}x_{ij}$, x^{m}_{ij}, k_{ij} so as to minimize

$$\bar{a} = [(n_1 + p_1 + n_2 + p_2 + \ldots n_H + p_H + \ldots n_q + p_q + \ldots N_Q + P_Q + P_{Q+1} + \ldots P_r + \ldots P_R), (P_d), (P_t), (P_b), (P_v)] \quad (1)$$

such that for G_1,

$$\sum_{k \epsilon A_j} {}^{s}x_{jk} - \sum_{k \epsilon B_j} {}^{s}x_{kj} + n_h - p_h = {}^{s}D_j \; \forall i, j \epsilon N \quad \forall s \epsilon S, h = 1, \ldots, H \quad (2)$$

For G_2,

$$\sum_{m=1}^{M_{ij}} x^{m}_{ij} - \sum_{s=1}^{S} ({}^{s}x_{ij} + x^{s}_{ji}) + N_q - P_q = O \forall ij \epsilon L \quad s \epsilon S, q = H + 1, \ldots, Q \quad (3)$$

For G_3,

$$x^{m}_{ij} - F^{m}_{ij} k_{ij} + n_r - p_r = K^{m}_{ij} \; \forall ij \epsilon L_p \quad \text{or} \quad x^{m}_{ij} + n_r - p_r = K^{m}_{ij} \; \forall ij \epsilon L_E \quad (4)$$

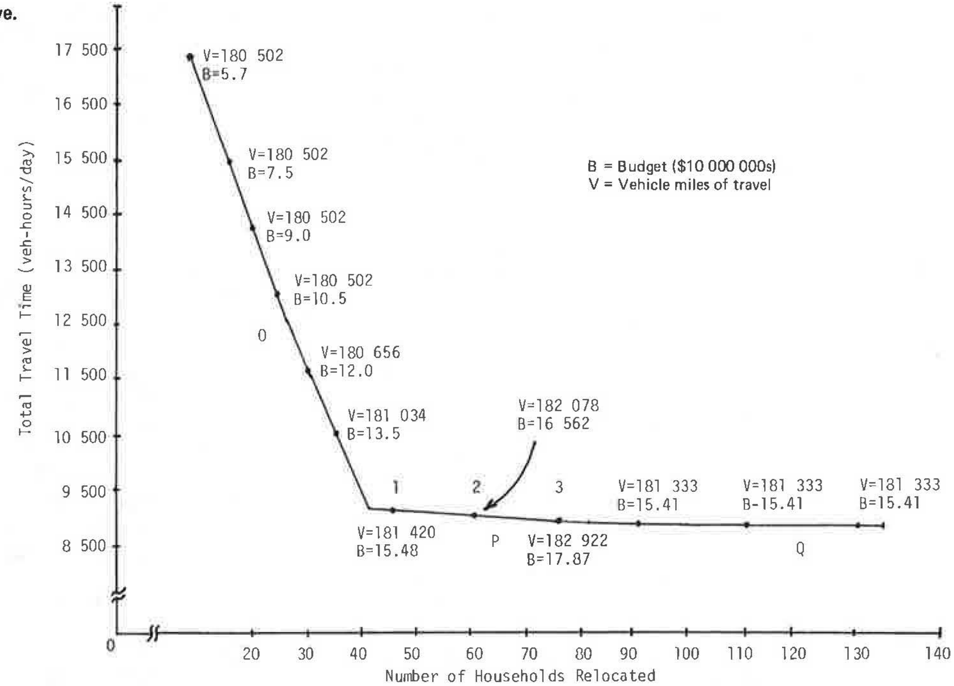
For G_4,

$$\sum_{ij \epsilon L_p} N_{ij} k_{ij} + n_d - p_d = W \quad \forall ij \epsilon L_p \quad (5)$$

For G_5,

$$\sum_{ij \epsilon L} \sum_{m=1}^{M_{ij}} c^{m}_{ij} x^{m}_{ij} + n_t - p_t = T \quad (6)$$

Figure 1. Multiobjective trade-off curve.



For G_6,

$$\sum_{ij \epsilon L_p} B_{ij}k_{ij} + n_i - p_i = B \quad (7)$$

For G_7,

$$\sum_{ij \epsilon L} V_{ij} \sum_{n=1}^{M_{ij}} x_{ij}^m + n_v - p_v = V, \quad {}^s x_{ij}, x_{ij}^m, k_{ij} = 0 \quad (8)$$

where

n_i = a negative deviation from the aspiration level in objective i,
p_i = a positive deviation from the aspiration level in objective i,
${}^s X_{jk}$ = flow on arc jk going from mode j to destination s,
${}^s D_j$ = demand from origin j to destination s,
A_j = the set of modes after j,
B_j = the set of modes before j,
H = S(N-1),
x_{ij}^m = flow on the m th segment of the total travel-time curve of link ij,
M_{ij} = number of segments on the total travel-time curve of link ij,
Q = S(N-1) + L,
k_{ij} = capacity added to link ij,
K_{ij}^m = existing capacity on segment m of the total travel-time curve of link ij,
L_p = the set of links that can be improved,
L_E = the set of links that cannot be improved,
F_{ij}^m = the portion of capacity added to link ij assigned to the m th segment of its total travel-time curve,
N_{ij} = number of households to be relocated per capacity unit added to link ij,
W = a desirable limit on the total number of households relocated as a result of network improvements,
c_{ij}^m = average travel time on the m th segment of the total travel-time curve of link ij,
T = a target value on the total daily travel time in the network,
B_{ij} = cost per additional unit of capacity on link ij,
B = total budget available,
V_{ij} = length of link ij, and
V = a desirable limit on the total daily vehicle miles of travel.

The data used for demonstrating the application of the model are presented in Tables 1 and 2 (2). The multidimensional trade-off curve derived by using the model is shown in Figure 1. This curve shows the sensitivity of total travel time, investment, and total vehicle miles of travel to changes in the desirable limit on the number of households that can be relocated.

If we ignore for a moment the objectives in priority levels four and five, we obtain in Figure 1 the two-objective trade-off curve normally generated by a linear programming model. The two-objective trade-off curve shows the change in level of service as a result of a change in the limit on the number of relocated households. It is unlikely that the alternatives in region 0 are desirable, for a small increase in the number of relocated households will bring about a large reduction in travel time if the decision maker is willing to move to region P. Similarly, the alternatives in region Q are probably not attractive because a large increase in the number of relocated households to move from P to Q results in a relatively small reduction in total travel time. A good solution probably lies in region P. However, since within region P the reduction in travel time that results from allowing relocation of additional households is constant, the decision maker may have difficulties in selecting the most desirable network. The information on the level of achievement of other objectives provided by the multiobjective trade-off curve can help in selecting the preferred network.

In the multiobjective curve, the level of achievement of the objectives at lower priority levels are displayed next to each efficient point. The achievement level of a more-important objective is closer to the curve. Looking at Figure 1, suppose the achievement of total travel time of 9912 vehicle-h with the displacement of 55 households, total travel time of 9104 vehicle-h with the displacement of 70 households, and total travel of 9024

vehicle-h with the displacement of 85 households are considered by the decision maker as equally desirable (all in region P). However, by using the information provided by the multiobjective trade-off curve, he or she can realize that the first of these three alternatives (point 1) is dominant with respect to the other two objectives and is therefore superior to the other two.

CONCLUSIONS

A new approach to the HNDM that allows comparison of network alternatives on the basis of multiple incommensurable objectives with different degrees of importance was presented. The goal-programming approach is capable of solving the multiobjective highway network design problem in a speedy and efficient manner. Furthermore, it could be used to generate multidimensional trade-off curves that provide additional important information to that provided by two-dimensional trade-off curves derived from the linear programming model.

The goal-programming approach was shown to overcome some serious limitations of linear programming. In linear programming, a solution that violates one or more of the constraints is termed infeasible. It is easy to realize that this type of conclusion provides no useful information and can often be considered misleading. For example, a basic assumption of the HNDM is that interzonal demands are given with certainty. In reality, predicted demands are subject to great uncertainty. Consequently, certain combinations of prediction errors can result in an infeasible solution and no further information is provided to the decision maker.

The budget objective is formulated in the linear programming model as a constraint. There are two serious problems with such a formulation. First, the decision maker does not have an a priori knowledge of the investment required to satisfy the predicted demands. In fact, he or she would probably expect to obtain this information from the model. If the budget is set too low, it may result in infeasibility. Setting the budget too high to avoid infeasibility would lead to overconstruction and an unrealistic flow pattern. Second, the budget is not independent of the level of service in the network. In fact, the budget is determined to achieve a desired level of service or certain levels of other impacts. The goal-programming approach avoids these problems because aspirations about the level of service and other impacts can be specified and we are allowed to consider the budget as a nonabsolute objective.

REFERENCES

1. S. K. Agarwal. Optimizing Techniques for the Interactive Design of Transportation Networks Under Multiple Objectives. Department of Civil Engineering, Northwestern Univ., Evanston, IL, Ph.D. dissertation, 1973.
2. E. K. Morlok and others. Development and Application of a Highway Network Design Model. Transportation Center, Northwestern Univ., Evanston, IL, Final Rept., Vol. 1, 1973.
3. S. K. Agarwal and J. L. Schofer. A Transportation Network Design Model for Regional Development Planning. International Technical Cooperation Center Review, Vol. 4, No. 3, July 1975, pp. 43-61.
4. G. Hay, E. K. Morlok, and A. Charnes. Toward Optimal Planning of a Two-Mode Urban Transportation System: A Linear Programming Formulation. HRB, Highway Research Record 148, 1966, pp. 20-38.
5. E. K. Morlok, E. N. Thomas, and others. The Development of a Geographic Transportation Network Generation and Evaluation Model. Transportation Center, Northwestern Univ., Evanston, IL, Final Rept., 1969.
6. M. Bruynooghe. An Optimal Method of Choice of Investment in a Transportation Network. Proc., Planning Transport Research and Computations, 1972.
7. J. P. Ignizio. Goal Programming and Extensions. Lexington Books, D.C. Heath and Company, Lexington, MA, 1976.

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

Abridgment

Development of Year-2000 Alternative Transportation Plans for the Delaware Valley Region

Thabet Zakaria

This paper discusses the concept and methodology used to develop long-range alternative transportation plans for the Delaware Valley Region. Four year-2000 alternative plans, including the no-build alternative, were formulated for simulation and evaluation. After a comprehensive evaluation of these alternatives, one of them will be selected and modified to be the year-2000 transportation plan. The alternatives were developed to achieve a set of regional goals prepared to deal with transportation issues and problems. The regional development pattern, travel demand and system deficiencies, short-range plans and programs, financial resources, administrative and legal requirements, and governmental and citizen recommendations were the major criteria considered in the formulation of the alternative plans. The alternatives were developed through an open two-way communication process between the staff of the Delaware Valley Regional Planning Commission and the various governmental agencies and private citizens involved in transportation planning. This process, which resulted in economical, feasible, practical, and implementable alternatives, could be applied successfully to any urban region in the country.

This paper discusses the concept and methodology used

to develop year-2000 alternative transportation plans for the Delaware Valley Region, which includes nine counties and three large cities in Pennsylvania and New Jersey and has a population of more than 5 million. The year-2000 planning effort of the Delaware Valley Regional Planning Commission (DVRPC) will replace the 1985 land use, water, sewer, open space, and transportation plans that were adopted in 1969 (1). Only a few of the facilities proposed in the 1985 transportation plan have been constructed and opened to traffic. This plan is currently under major review due to social, economic, and environmental changes that occurred in the last decade.

METHODOLOGY FOR THE DEVELOPMENT OF ALTERNATIVE PLANS

Rational planning requires that a wide range of alternative plans be developed, tested, and evaluated to determine their impacts on land use, traffic patterns, and the natural and social environments. After a comprehensive evaluation of these alternatives, one is usually selected for programming and implementation (2-4).

Many alternatives can be formulated on the basis of a given set of alternative land use plans and transportation projects and policies. When the number of transportation projects and policies becomes large, a great number of possible combinations can be obtained. Because of the limited budget and time available to simulate travel demand and to evaluate alternative plans, DVRPC developed only four alternatives. The no-build alternative is considered one of these alternatives.

A discussion follows of the activities performed by DVRPC to develop the alternatives, which were endorsed unanimously by the DVRPC board in December 1978 (5). The following activities were accomplished.

Discussion and Definition of Transportation Problems and Needs

To obtain a general consensus on regional needs and preferences, DVRPC conducted a conference on regional planning issues in 1975. The participants were citizens, political leaders, technicians, university professors, and public and private officials of diverse backgrounds.

The conference centered on planning issues pertaining to society, land use, environment, and transportation. Examples of such issues and problems are safety and security, traffic congestion, provision of public transportation service, air and noise pollution, parking, energy and future technology, capital cost and financial resources, and system operating and maintenance cost.

Consideration of Governmental Recommendations

As in other metropolitan areas, the state, county, and city governments are involved in transportation planning and operation. The governments, which are represented by the 18 members of the DVRPC board, make the final decisions on DVRPC plans and programs.

To obtain government inputs on the alternative plans, DVRPC staff met individually and collectively with city and county planning commissions, transit operating agencies, state departments of transportation, and turnpike and bridge authorities. These meetings and additional correspondence between DVRPC staff and these agencies resulted in a set of specific transportation policies and facilities for each of the alternative plans.

Figure 1 shows the decision-making process followed in the development of year-2000 alternative plans. This process is open and has two-way communication between the concerned parties or agencies.

Consideration of Citizens' Recommendations

Citizen participation is a prerequisite for successful transportation planning and implementation. The DVRPC citizen participation program has been expanded to accommodate citizen inputs in the various steps of planning. As shown in Figure 1, citizen inputs into the alternative plans were obtained from the county citizen forums and the year-2000 regional transportation advisory committee, which has more than 100 members and meets monthly at the DVRPC offices.

County citizen forums are held periodically in each of the region's nine counties. The forums involve representatives of county and local citizen groups and interested individuals. At the regional and county citizen forums, citizens were asked to review and comment on DVRPC work and to provide their recommendations concerning any specific transportation policy and facility that should be included in the alternative plans.

Preparation of a Set of Transportation Planning Goals

After the definition of transportation issues and problems, a set of goals and objectives for the development and evaluation of alternative plans was prepared. It was decided to use the goals for developing the alternatives for testing and evaluation. These goals reflect not only transportation concerns but also socioeconomic and environmental considerations expressed in the statement of transportation problems and issues. The goals and objectives were discussed with technicians and citizens at the regional and county levels.

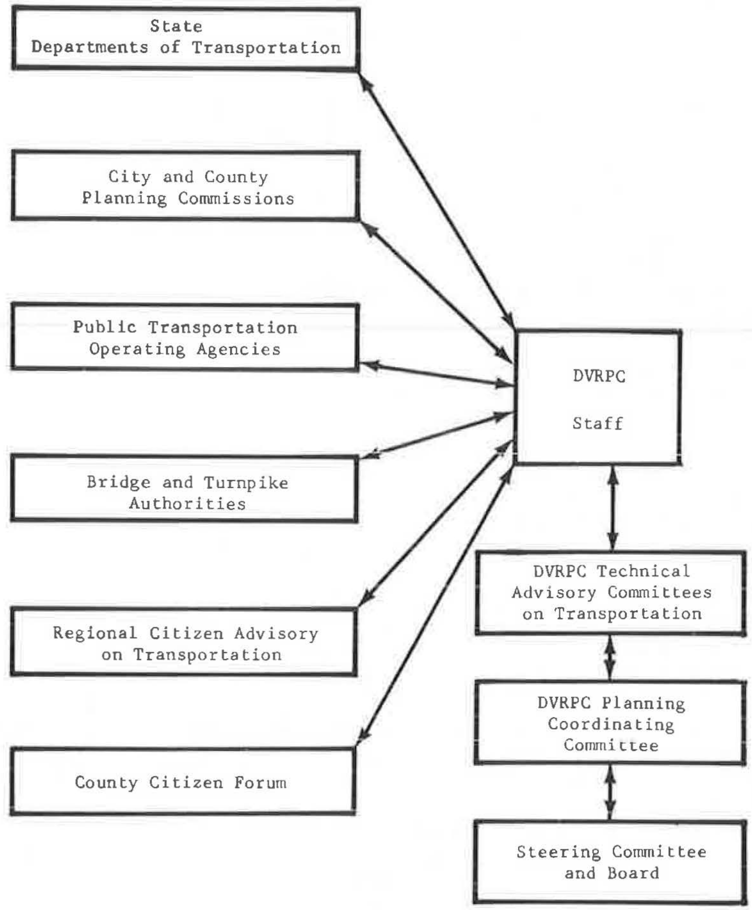
Consideration of the Regional Development Guide

The regional development guide (RDG) includes the policies adopted by the DVRPC board concerning the magnitude and location of future regional development (6). The RDG was adopted after 10 scenarios of regional growth had been screened by the DVRPC board and citizens. Each scenario framed a set of policies that responded to the regional issues. After considerable discussion, the DVRPC board selected four scenarios (futures) for further study.

These scenarios were studied for their implications and consequences for land use, open space, housing, transportation, water quality, air quality, energy, and environmentally sensitive areas. In the transportation study, sketch-planning models were used to study their impacts. Each scenario was tested for its effect on the existing transportation system. Areas in which congestion could be expected were identified, and approximate measures of the cost required to eliminate these traffic problems were prepared.

After considerable discussion of the implications of the four scenarios, the DVRPC board adopted the RDG in 1977. Moderate rates of growth in the regional population and employment are projected in the RDG. Future growth is encouraged to locate in and around the existing urban and suburban centers, which have been declining. Some recentralization of land uses within the urban area is also encouraged. The expansion and improvement of the existing transportation system is encouraged to serve the type of land use development described in the RDG.

Figure 1. Decision-making process for the development of year-2000 alternative plans.



Analysis of Current and Future Travel Demand on the Existing Transportation System

The analysis of current and future traffic on the existing system (no-build alternative) indicates the location and magnitude of highway and transit deficiencies. The information obtained from this traffic simulation is used to develop alternative highway and transit solutions to such traffic problems.

The analysis of highway volumes and capacities resulted in the identification of highway links that are, or will be, congested by the year 2000. The analysis of future transit demand indicates whether there is a need for additional public transportation facilities and service. Such a service will be required if transit trips increase significantly due to regional growth or if there is no service at the present time (7).

Coordination of Year-2000 Alternatives with Short-Range Plans and Programs

Although the target year for the alternative plans is the year 2000, DVRPC staff gave increased emphasis to short-range projects that are included in the transportation system management element (TSME) and the transportation improvement program (TIP). Such short-range projects have been under consideration for a number of years and have been studied many times. Some of them have reached the implementation stage and are under construction or have been committed for construction in the near future. Most of these projects, however, are small, low-capital-intensive projects intended to improve the existing transportation service. These projects and other missing highway or transit links that complete the regional transportation system are generally committed for construction and are included in the year-2000 alternative plans. In addition, DVRPC considered the recommendations of past transportation studies in the formulation of year-2000 alternative plans so that no major conflict between the regional plan and local plans will arise in the implementation stage.

Analysis of Cost and Financial Resources

The analysis of financial resources is perhaps the most important factor that should be considered in the development of alternative plans. This analysis at DVRPC has indicated that capital for constructing and operating transportation facilities will continue to be increasingly scarce in the future (8). Discussion with federal, state, and local officials responsible for transportation programming and budgeting pointed out that the funds anticipated to finance long-range transportation projects will fall short of the region's expected needs. They based their recommendations on the fact that the escalation of construction cost of the programmed facilities and the cost committed to operate and maintain the existing system will consume all of the anticipated funds. Further, other urban services will be competing highly with transportation for scarce financial resources.

Consideration of Legal and Administrative Requirements

Legal and administrative requirements are important in

the development of alternative plans since they indicate the types of transportation facilities and policies that are feasible. Federal, state, and local requirements are especially critical to the implementability of the planning program. Many federal requirements, mandated by the U. S. Department of Transportation (DOT) and U. S. Environmental Protection Agency (EPA), limit the choices of transportation alternative plans and projects. DVRPC analyzed such regulations and guidelines so that the alternatives recommended will be feasible and implementable.

CHARACTERISTICS OF THE ALTERNATIVE PLANS

In addition to the no-build alternative, DVRPC staff developed three alternative plans based on the factors outlined previously in this paper. Before they were endorsed by the DVRPC board, the alternatives were defined, mapped, and presented to the various technical and policy committees and citizens shown in Figure 1. Some modifications were made to the initial alternatives after the first round of discussion and coordination. The alternatives endorsed are briefly described below (2).

No-Build Alternative

This alternative assumes no major new projects and consists of the existing facilities that were open to traffic in 1978. To operate the existing transportation system at a minimal level of service, some essential improvements are provided, such as the following:

1. Highway and bridge replacement and rehabilitation,
2. Restoration and resurfacing,
3. Transit vehicle replacement, and
4. Minor station improvements.

These improvements are also included in the other alternatives. This alternative does not provide an acceptable level of service. It is analyzed mainly for the purpose of formulating and evaluating the other alternatives.

Alternative 1

This alternative includes the transportation facilities included in the recent editions of the TIP. This alternative assumes that the TIP improvements and recommendations may take up to 22 years to complete rather than the 6 or 12 years scheduled in the TIP. This conservative view is generated from the current analysis of anticipated funding to build and operate new facilities. TIP projects are included in this alternative because they complete the missing links in the regional transportation network and improve the operation of the existing system.

Alternative 2

This alternative includes small projects to manage and improve the existing transportation system for servicing the year-2000 travel demand. It also includes the committed major facilities to complete the missing highway and transit segments that are essential to traffic flow and passenger service. Most of the facilities in this alternative are proposed to improve the operation of the existing highway and transit systems. Alternative 2 includes fewer new highway and transit facilities than alternative 1 because low-capital-intensive projects are used to serve future travel demand. Therefore, the level of transportation service provided by this alternative will be lower than that provided by alternative 1.

Alternative 3

This alternative is designed to provide alternative transportation service to that recommended in the other alternatives at some areas and corridors. Alternative 3 replaces some major facilities recommended in alternative 1 with less-extensive projects. It also includes other needed highway or transit facilities not recommended in the other two alternatives. This alternative will result in testing and evaluation of the impacts of these new facilities on the transportation system. Like alternatives 1 and 2, this alternative includes the committed facilities and other small projects to improve the efficiency of the existing transportation system.

Summary

Each of the alternative plans consists of five elements:

1. Existing transportation system in 1978, except for those facilities proposed for abandonment;
2. New freeway and transit facilities that are under construction or recommended for construction;
3. Recommended highway and transit improvements, such as arterial widening, construction of new arterials, electrification of rail lines, improvements to transit stations, and purchase of new transit cars;
4. Other improvements, such as rehabilitation and restoration of highways, repairs of critical bridges, fringe parking facilities, Traffic Operations Program for Increasing Capacity and Safety (TOPICS) projects, highway safety projects, replacement of old buses and automobiles, improvement to rapid-transit-line power equipment, and replacement of old railroad cars (most of these high-priority projects are not listed in year-2000 plans, but a total sum of money is allocated for such small improvements on the basis of the TIP cost estimates; these improvements are considered an essential element of any plan); and
5. Additional regional highway and transit improvements that are limited to turnpikes, toll bridges, bicycle projects, and National Railroad Passenger Corporation (Amtrak) lines and stations.

Table 1 (2, tables 3 and 4) provides a summary of the total capital cost for highways, transit facilities, and various transportation improvements for each alternative plan. It also shows the annual operation and maintenance cost required to maintain the highway and public transportation system and to provide adequate transportation service for each alternative plan. The capital cost of all alternatives was found to be within the range of anticipated financial resources.

The development of DVRPC alternatives was a lengthy process. It took nine months and involved many meetings and discussions with the states, counties, funding agencies, transit operating agencies, and citizens. Each group provided specific suggestions and recommendations for inclusion in the alternative plans after discussion with DVRPC about transportation problems, goals, future trends, and other important factors. Staff reviewed all inputs and incorporated the recommendations that fit the criteria for plan formulation. All suggestions and recommendations not included in the alternatives were listed with explanations for their rejection and were discussed with the concerned organization before arriving at a final decision.

The alternatives were developed through an open and

Table 1. Capital and operating-and-maintenance costs of year-2000 alternative plans.

Item	Alternative[a] ($000 000)			
	No-Build	1	2	3
Capital cost of highway improvements				
New freeways		1431	674	1417
Arterials		370	251	644
Other	618	838	838	838
Total	618	2639	1763	2899
Capital cost of public transportation improvements				
New facilities		1647	181	1093
Facility renovation		489	464	492
Other	672	973	973	973
Total	672	3109	1618	2558
Total capital cost	1290	5748	3381	5457
Annual operating and maintenance cost	345	393	389	397
Capital cost of additional region-wide improvements				
Bicycles	5.7	5.7	5.7	5.7
Turnpikes		70.7	137.6	70.7
Calhoun Bridge				16.6
Amtrak	210.8	210.8	210.8	210.8

[a]Costs are in 1977 dollars.

participatory two-way communication process between DVRPC staff and the various groups that are affected by transportation decisions. DVRPC prepared the documents and guidelines that are essential for conducting meaningful, rational, and organized discussions among the various groups involved in the planning process.

The reader may have observed that all DVRPC alternative plans include the following common elements.

Intensive Improvements of the Existing Transportation System

All alternatives (as shown in Table 1) include many small projects to improve the existing highway and transit systems. Such projects are given high priority and will be recommended in the final plan. Without such improvements the existing system will not function in the Delaware Valley because it is old and deteriorating. Also, it is incomplete and uncoordinated.

All Modes of Surface Transportation

None of the alternatives is highway- or transit-oriented, as was usually the case in past transportation planning studies. The alternatives developed are balanced and provide all modes of transportation service to all population groups in the region. All alternatives place some emphasis on the efficiency of the transportation system and energy consumption.

Small Transportation Projects as Well as Large Ones

As shown in Table 1, a sum of money is allocated for small projects although they are not specifically listed in the alternative plans. This is a departure from past regional transportation plans, including the 1985 DVRPC plan, which dealt mainly with major highway and transit facilities. As indicated in Table 1, the total cost of small (or other) improvements is considerable and should be accounted for in any long-range plan.

Short- and Long-Range Projects

Short-range projects should be included in the long-range plan because they are often as important as the long-range projects. Further, short-range plans and programs must be coordinated with long-range planning to obtain consistency in the transportation planning process and to recommend a reasonable implementation program.

Subregional and Regional Projects

Although they ensure system continuity and compatibility, the alternative plans are actually aggregations of subregional and regional improvements intended to solve local and county transportation problems. When accumulated, local problems become regional in scope. Generally, the negative impacts of transportation projects and improvements affect people at the local level. This does not mean that DVRPC alternatives are big municipal plans. Rather, local projects are considered in the development of regional plans that function systemwide. The impacts of any regional project will be considered in the evaluation of the alternative plans. This evaluation will result in recommending the most feasible and effective projects for inclusion in the final year-2000 plan.

Specific Projects and Flexible Solutions

Although the alternatives include many specific and well-defined projects, they are flexible and could be adapted to future changes in social and economic conditions and transport technology. The long-range plan that will be recommended on the basis of the evaluation of these alternatives will include specific projects and flexible strategies. The fixed or specific facilities are only those currently under construction or committed for construction in the near future to complete the missing links of the regional transportation network.

CONCLUSIONS

As part of the planning process to develop a year-2000 plan, DVRPC formulated a limited number of alternatives for testing and evaluation since it is impossible to predetermine the optimum plan. The regional transportation alternative plans were developed in cooperation with citizens, public transportation authorities, and planning departments of the various county and state governments.

The DVRPC alternatives were developed through an open and participatory two-way communication process between DVRPC and the various groups that are involved in the planning process. The role of DVRPC staff was to assist citizens and private groups and governments at all levels in defining transportation issues, problems, and goals and in developing alternative courses of action that are practical, economical, feasible, and implementable.

The alternative transportation plans were coordinated with land use and environmental plans to provide adequate service for all groups of the population. The alternatives consider short- and long-range solutions to the various transportation problems, within the constraints of financial resources and legal and administrative requirements. Further, the alternatives were defined to be specific with respect to projects and policies and yet adaptable to future changes in social and economic conditions that impact the transportation system.

The DVRPC approach to long-range planning can be applied successfully to any urban region throughout the country.

ACKNOWLEDGMENT

This paper was financed in part by the Federal Highway Administration and the Urban Mass Transportation Administration of the U.S. Department of Transportation and by the Pennsylvania and New Jersey Departments of Transportation. The contents reflect my views, which do not necessarily reflect the official views or policy of the funding agencies.

REFERENCES

1. 1985 Regional Transportation Plan: Technical Supplement. Delaware Valley Regional Planning Commission, Philadelphia, Plan Rept. 5, 1969.
2. Alternative Transportation Plans for the Delaware Valley. Delaware Valley Regional Planning Commission, Philadelphia, Year 2000 Rept. 7, Feb. 1979.
3. W.G. Hansen and S.C. Lockwood. Metropolitan Transportation Planning: Process Reform. TRB, Transportation Research Record 582, 1976, pp. 1-13.
4. M.L. Manheim and others; Massachusetts Institute of Technology. Transportation Decision Making: A Guide to Social and Environmental Considerations. NCHRP, Rept. 156, 1975, 135 pp.
5. H.S. Cohen, J.R. Stowers, and M.P. Petusilia. Evaluating Urban Transportation Alternatives. U.S. Department of Transportation, Rept. DOT-P-30-78-44, Nov. 1978.
6. The Year 2000: A Policy Statement on Regional Development. Delaware Valley Regional Planning Commission, Philadelphia, Year 2000 Rept. 5, 1977.
7. Preliminary Year 2000 Travel Forecasts and Assignments on the Current (1977) Transportation Systems. Delaware Valley Regional Planning Commission, Philadelphia, Summary Rept., June 1978.
8. Capital Funding of Transportation in the Delaware Valley Region. Delaware Valley Regional Planning Commission, Philadelphia, Draft Rept., Proj. 31.17, June 1978.

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

Transportation Planning for Small Communities: Western Canadian Experience

S. Teply

The paper discusses the principles, constraints, and objectives of transportation planning in small communities. It compares some of the basic relationships derived in the United States with those found in several western Canadian communities. A synthetic planning process called the four-purpose trip generation and distribution model is described in detail. It uses an analogy approach by starting with estimated data found applicable in similar communities. In this way it avoids the costly and time-consuming data collection stage. The model is verified and calibrated after data processing. Computer traffic volumes are compared with traffic counts and, if necessary, the input values are adjusted. Sensitivity of the model to errors in the initial estimated data is analyzed in relation to the basic zonal land use characteristics (i.e., population and employment). A set of graphs is presented to expedite the calibration process. They relate the size of the unit outcome error (i.e., the difference between the computer and the surveyed traffic volume) to the required adjustment of initial estimates of trip purpose distribution.

Small communities in western Canada must determine the directions of their future development. Oil, gas, lumber, agriculture, and initial industrialization form the basis for a dynamic economy, especially in Alberta. The towns and cities have experienced a period of steady growth and strive to maintain a balanced development in all aspects of urban life in the future. This goal creates a need to plan ahead in order (a) to influence the demand, (b) to provide and control the supply of facilities, or (c) to do both. Planning in small expanding communities is rather difficult because even small unforeseen facilities, activities, or policies may have dramatic effects. The decision of a single industrial company to move into the area and locate at an opportune (yet at the planning stage unconsidered) location may make previous transportation plans invalid. The range of effects of such uncertainties is much more pronounced than in large, established cities.

THE PROBLEM

In the past 10 years, the trend in urban transportation planning has been to recognize the specifics of small communities and to adjust procedures and models accordingly. Identifying features of small communities may be listed as follows:

1. Population size of up to about 100 000 inhabitants [several research studies dealt with smaller (up to 50 000) or larger (up to 250 000) communities];
2. Economy usually pivots around several key activities;
3. Life-style in smaller communities is simpler;
4. Scenarios that are easily identifiable can cause significant migration into or out of the area;
5. Civic governments, both in elected and administrative portions, lack the expertise or resources for solving unusual problems (i.e., those that exceed day-to-day operations); and
6. Strain on financial resources, especially when considered on a per capita basis, is usually much larger because small communities cannot use the luxury of economy of scale.

Size of a community seems to be only the most visible indication of the other community characteristics that should be considered in their full context. The economic balance of the community may be influenced drastically by outside forces. Social activities in small communities are less diversified and identifiable than are those in larger urban centers. In some instances, the governing bodies of smaller communities lack the necessary political maturity and stability for efficient decision making.

The purpose of transportation plans for small communities may be one or more of the following:

1. Identification of the consequences of major development alternatives;
2. Assistance in selection of the best land use master plan;
3. Definition of the transportation requirements of a specific land use plan; and
4. Identification and testing of specific solutions, such as new bridge locations.

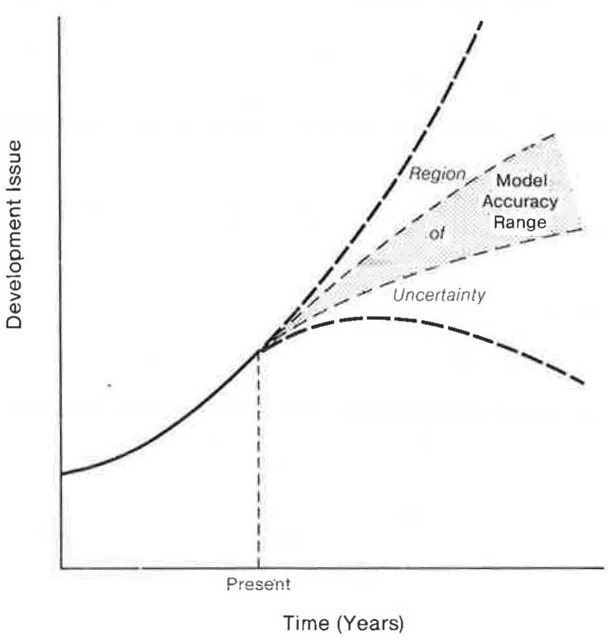
Constraints of comprehensive transportation plans for small communities usually are the following:

1. Minimum cost,
2. Minimum time,
3. Reasonable accuracy,
4. Large number of time frames,
5. Studies are more goal oriented, and
6. Population interest.

Traditional transportation planning methodology, common in large communities, is prohibitively expensive for small communities, both for their data collection and for modeling. The marginal per capita cost of such studies may be several times higher in small communities than in large cities.

Although small communities have recognized the benefits of a continuous planning process and plan evolution, the nature of community life and population expectations demands expedient answers. In addition, time spent on a study, even during its inactive stage, increases cost.

Figure 1. Symbolic representation of model predictive ability.



The concept of model accuracy in transportation planning is based on the predictability of independent variables and on model cost. In small communities, the economic future and consequent transportation development is less predictable than in larger, more diversified communities (Figure 1). Highly sophisticated, accurate models, even when coupled with a scenario or alternative future approach, may not perform better than their simplified versions or other crude models. Also, in a small community, achievement of the same accuracy goal as that required for a large city would be proportionately more expensive. In simple terms, a high degree of accuracy is not required.

Models used should be easily applicable to a variety of problems that may range in time from immediate needs to those that are 10 or more years in the future. In past studies, much time, money, and attention were spent on data collection, analysis, and model development. On the other hand, in some instances, generation of solution alternatives and their evaluation left much to be desired. Small communities, however, are usually more goal oriented than large communities.

In a small community, a major transportation study may quickly become a matter of everyday community talk. Lack of competing issues as well as community expectations generate public interest and, for that reason, require a reasonable openness of the planning process.

MODEL CHARACTERISTICS

The traditional structure of the four-step transportation model (i.e., trip generation, modal split, trip distribution, and trip assignment) has a special appeal for small communities in that it is straightforward and easily understandable. In addition, individual steps are relatively independent, can be treated at different levels of sophistication, and can be easily verified or calibrated individually.

Bates (1) postulated that the primary assumption in transportation planning (i.e., consistency in time of relationships between travel demand and certain social, economic, and physical parameters) can be logically extended to consistency in space. This seems to be a reasonable assumption, especially when cultural and behavioral differences between communities are also considered. Based on this premise, Bates reasoned that it may be feasible to formulate models that can synthesize results of field data collection and thus eliminate the necessity for this expensive and time-consuming phase of transportation planning.

Recognition of the analogy approach to transportation planning in small communities is widespread today. It is a simplified but valid approach, whose capabilities are in line with the constraints of the planning process.

The following is a brief review of planning characteristics described in the literature and applied in several western Canadian studies for small communities.

Trip Generation

Jefferies and Carter (2) list specific characteristics for trip generation models for small communities as follows. Data used to develop independent variables should be (a) easily obtainable from existing records or simple and inexpensive surveys, (b) capable of explaining future as well as existing trip-generating characteristics, and (c) able to reflect the influence of the change on trip generation rates. Reliable and inexpensive methods should be available for forecasting data used as independent variables. The final trip generation equation should be easy to use and should contain the fewest variables

consistent with the amount of accuracy required.

A condition of an easy validation and calibration of the relationships between the independent and dependent variables can be added. The types of independent variables used in various studies for both trip production and attraction include population, dwelling units, automobiles, employment, retail employment, industrial employment, other than industrial employment, government and finance employment, and school enrollment. Dependent variables used include total trips, home-based work trips, home-based nonwork trips, home-based shopping trips, and non-home-based trips.

Figure 2 shows the relation between population size and total automobile trips based on an analysis of 14 small U.S. communities (3-12). Other sources also indicate that the variability of the total number of trips in small communities is larger than that for larger cities. In general, the number of trips per capita per working day ranges from 1.6 to 3.0; a major cluster of values falls between 1.8 and 2.2. The list of values found applicable to seven western Canadian small communities is shown in Table 1 and interpreted in Figure 2. The U.S. studies were taken from Hajj (13). The higher value for Fort McMurray may be attributed to the dynamic nature of the town; many young people were attracted by the opportunities of the Alberta oil sands regions (14).

Figure 2. Relationship between population size and total vehicle trips.

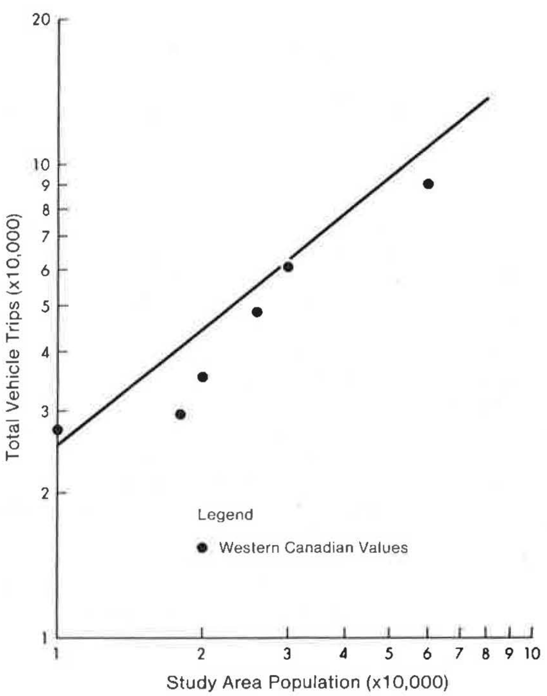


Table 1. Trip generation rates determined in the studies.

Community	Year Studied	Population	Trips per Capita per Day
Fort McMurray, Alberta[a]	1974	10 000	4.20[b]
Medicine Hat, Alberta	1977	30 000	3.25[b]
Medicine Hat, Alberta[a]	1971	26 000	2.60[b]
Kelowna, British Columbia[a]	1973	20 000	2.60[b]
Grande Prairie, Alberta	1977	18 000	2.50[b]
Prince George, British Columbia[a]	1974	60 000	2.35[b]
Murray, Kentucky	1967	10 000	2.80[c]
Joplin, Missouri	1960	16 000	2.93[c]
Clarksville, Tennessee	1965	11 000	2.29[c]
Gainesville, Florida	1960	21 000	1.94[c]
Pittsburg, Kansas	1961	22 000	2.38[c]
Staunton, Virginia	1965	32 000	1.69[c]

[a] Studies used in sensitivity analysis.
[b] Person trips per capita per day.
[c] Vehicle trips per capita per day.

Car ownership may also be a partial explanation for higher than usual values of trip production in some small communities. For example, the average provincial ratio of population to vehicles in private use in Alberta in 1978 was about 1.5 persons/vehicle. For rural areas and small communities the ratio may be considerably higher than the average.

Although a number of regression equations for trip production that use many of the previously mentioned independent variables have been developed elsewhere (1, 2, 13, 15-17), a formula that is based solely on population size and trips per capita per day has been found sufficiently accurate for western Canadian communities. The problem of predicting future population still remains. Refined methods, such as the cohort survival method, do not work well in western Canadian small communities because of the migration phenomenon. For that reason, economic models based on industrial or agricultural potential have been applied in some instances.

A more complex situation exists in trip attraction models, although regression equations for trip attraction quoted in the literature employ independent variables that are similar or identical to those for trip production. Western Canadian models generally use only population, employment, and retail employment.

The problem at hand is data collection. Large-scale interviews are costly and small samples in small communities do not always guarantee the required statistical significance. This is one of the reasons why the analogy approach has become increasingly popular.

Modal Split

With few exceptions, public transit does not play a major role in small western Canadian communities. If transit is provided at all, its share usually does not exceed 5 percent or, in a few cases, 10 percent. In small communities, roadway capacity is plentiful at almost any time of day, and high automobilization and a neighborhood spirit exist. A private automobile, therefore, can provide the least expensive transportation service from a community point of view. Nevertheless, as a community grows, the need for transit or for more transit emerges.

In those studies in which trip production and trip attraction have been determined in person trips (as opposed to a direct determination of vehicular trips), a conversion must be made. Usually, when trips per day are used, an overall daily average vehicle occupancy factor is applied (around 1.5 in western Canada). For those cases where peak-hour trips are determined, differentiation of occupancy among various trip purposes is necessary.

Trip Distribution

The shopping list of models available for small communities includes growth factor and gravity approaches. Both have been applied successfully.

An interesting variation of a Fratar method was described by Rhodes and Hillegass (16). It combines both the trip distribution and trip assignment stage in that pseudoattraction factors are used to determine the distribution table and, after trip assignment, growth factors for individual network links are determined.

Although full versions of gravity models have been applied elsewhere, it is interesting to observe the effect

Figure 3. Relationship between population size and the maximum trip length.

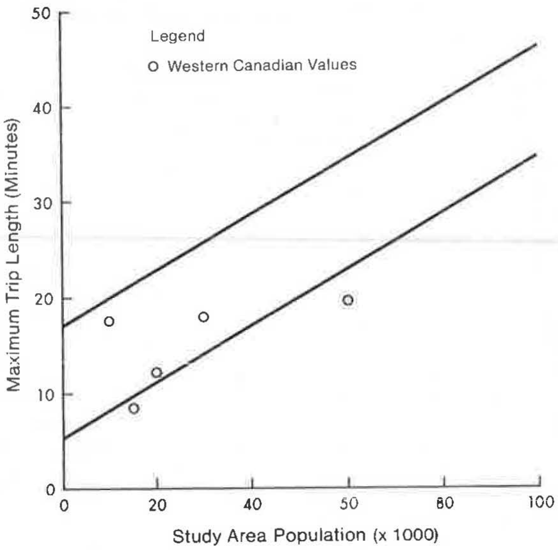


Table 2. Calibrated trip purpose distributions.

Community	Vehicle Trips by Purpose (%) Home-Based Work	Home-Based Shopping	Home-Based Other	Non-Home-Based
Fort McMurray, Alberta[a]	22	24	27	27
Medicine Hat, Alberta, 1977	40	15	25	20
Medicine Hat, Alberta, 1971[a]	37	13	28	22
Kelowna, British Columbia[a]	30	15	31	24
Grande Prairie, Alberta	40	15	25	20
Prince George, British Columbia[a]	35	15	26	24
Leduc, Alberta	39.7	24.2	17.8	18.3
Murray, Kentucky[a]	14.9	48.1		37.0[b]
Pine Bluff, Arkansas	13.5	48.6		37.9[b]
Fort Smith, Arkansas	22.5	38.7		38.8[b]
Kingsport, Tennessee	20.4	42.9		36.7[b]
Greenville, South Carolina	16.9	39.6		43.5[b]
Memphis, Tennessee	21.2	41.5		37.3[b]
Pulaski, Arkansas	21.8	47.3		30.9[b]

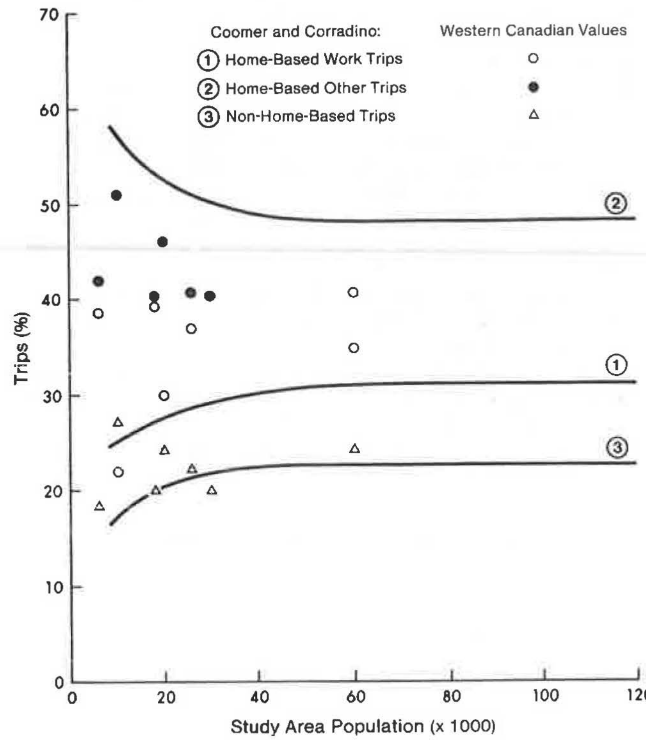
[a]Studies used in sensitivity analysis. [b]Includes truck and taxi trips.

of trip impedance on trip distribution. Small communities are not only limited in area; they usually have fewer problems with traffic congestion. Trip lengths are therefore rather short in terms of time.

Hajj (13) documented a sharp distribution for Madisonville. An average trip, regardless of its purpose, was between 6 and 7 min long. The relationship between the longest trip and community size were determined by Bates (1), as shown in Figure 3, and compared with approximate values for several western Canadian communities (4-12). Bates also observed that the ratio of average and longest trips is relatively stable for all small communities and depends, to a certain degree, on trip purpose. For home-based work trips the ratio is 0.362; for home-based other trips, 0.317; and for non-home-based trips, 0.287. Graphs by Hajj (13) indicate ratios of 0.380, 0.397, and 0.402, respectively, where the non-home-based trips include trucks and taxis.

In view of the physical size of small communities and the fact that average trip duration is only about one-third of that of the longest trip, trip impedance does not seem to play a significant role in trip distribution. An indirect confirmation of this phenomenon can be found in the fact that several small western Canadian communities (such as Prince George, British Columbia—population 60 000) experience the highest and directionally most diversified traffic peak of the day (up to 13 percent) around noon, when employees go home for lunch. Jefferies and Carter (2) also observed that accurate trip productions and attractions, not the distance, appeared to be the key to gravity model calibration.

Figure 4. Relationship between population size and trip purpose distribution.



Because of the insignificance of distance, the trip distribution model can be simplified and trips from one zone distributed only in proportion to the relative attraction of the other zones. This is the basis of the proportionate trip distribution model popular in the Canadian west. Four internal trip purposes (home-based work trips, home-based shopping trips, home-based other trips, and non-home-based trips) are usually considered from the attraction point of view. The values used in seven western Canadian studies are listed in Table 2. The U.S. data were taken from Hajj (13). Figure 4 indicates their relation to the values determined by Coomer and Corradino (3-12).

All studies and literature indicate a separate treatment of external-internal relationships, because they may be subject to different laws.

Validation and calibration of trip distribution models are carried out in the usual way (i.e., by aggregating zones and comparing resulting volumes across selected screen or cordon lines). In strictly synthetic models, which do not employ any interviews and use analogy data from similar communities, this is also the opportunity to adjust trip generation or originally estimated occupancy values. A discussion of some of the aspects of calibration follows.

Trip Assignments

The number of available route choices in small communities is small and capacity constraint does not usually enter the problem in a significant way. Individual routes

are easy to recognize, especially if the planner has some insight into population habits and network operation. For these reasons, either a standard all-or-nothing approach or assignment of trips between each pair of zones to a number of empirically determined available routes is favored.

In both cases, computer programs can be used. Nevertheless, resulting volumes for certain links may be unrealistic. Assignments must therefore be reviewed and corrected. For this reason, some planners prefer to perform the assignment stage manually to provide themselves with better understanding of network load performance. Some adjustments may then be incorporated as the analysis proceeds.

FOUR-PURPOSE PROPORTIONATE TRIP GENERATION AND DISTRIBUTION MODEL

This model evolved in the late 1960s and, for reasons mentioned before, has become very popular in western Canada. It was described in a 1972 internal manual of the Alberta Department of Highways (18). A case study of the town of Fort McMurray, Alberta, was presented by Sargious and Morrall in 1975 (14). The following description is based on a refined version of the model as used in a computer program developed at the University of Alberta by Kondo in 1975 (19).

The process involves the following steps:

1. Divide the study area into zones.
2. For each zone, identify the population, total employment, and retail employment for the base year and for each of the future scenarios.
3. Conduct screen and selected link classification counts, as well as an external cordon count or a roadside interview survey. Include occupancy counts for selected locations.
4. By using an analogy with similar communities, estimate the initial present trip generation rates per capita per working day. Although rates per evening peak hour have been employed in several instances, they are more difficult to use. For communities of larger size, a morning peak-hour base seems to be a more promising approach.
5. Estimate the current modal split.
6. Determine the current average automobile occupancy; if counts are not available, estimate them.
7. To determine vehicle trip generation for each zone, use

$$G_i = [P_i t(1 - r)/v] + (P_i trB/b) \quad (1)$$

where

G_i = trip generation of zone i (two-way, i.e., single vehicular trips),
P_i = population of zone i (people),
t = trip generation rate (trips/day),
r = transit ridership,
b = average bus occupancy (passengers),
v = average automobile occupancy (passengers), and
B = passenger-car equivalent for buses.

8. By using an analogy with similar communities and estimated adjustments, determine the current trip purpose percentages, i.e., internal home-based work (HBW), internal home-based shopping (HBS), internal home-based other (HBO), and internal non-home-based trips (NHB).
9. Distribute the trips generated in each zone in proportion to the relative attractiveness of other individual zones, expressed in their shares of population, total employment, and retail employment. In the model used by the University of Alberta, the central business district (CBD) is subject to a slightly different treatment because it was found that a strictly proportionate distribution underestimated its attractiveness. In mathematical terms, the initial distribution of internal trips to the CBD (zone 1) can be expressed as follows:

$${}_1D_{i1} = G_i(HBW)(E_1/\Sigma E_j) + G_i(HBS)(R_1/\Sigma R_j) + G_i(HBO)k_1 \quad (2)$$

where

${}_1D_{i1}$ = trips to CBD from zone i,
HBW = ratio of home-based work trips,
HBS = ratio of home-based shopping trips,
HBO = ratio of home-based other trips,
E_j = zonal employment,
R_j = zonal retail employment, and
k_1 = CBD coefficient for attraction of home-based other trips (0.1 was found to be a reasonably accurate value for four of the studies tested).

The initial distribution to the other zones can be written as follows:

$${}_1D_{ij} = G_i(HBW)(E_j/\Sigma E_j) + G_i(HBS)(R_j/\Sigma R_j) + G_i(HBO)(1 - k_1)(P_j/\sum_{j\neq 1} P_j) \quad (3)$$

The final distribution of internal trips to all zones is computed by using the proportion of trips already attracted as a distribution rule for non-home-based trips:

$${}_2D_{ij} = {}_1D_{ij} + G_i(NHB)(\sum_i {}_1D_{ij}/\sum_i \sum_j {}_1D_{ij}) \quad (4)$$

where ${}_2D_{ij}$ = exchange of trips internal to the study area between zone i and j and NHB = ratio of non-home-based trips.

10. Distribute the external-internal trips from points on the external cordon according to the external origin-destination survey, if available. If such a survey was not conducted, distribute the number of external-internal trips determined on cordon counts in proportion to the attraction of the zones for internal trips; that is, the final distribution table involves values as follows:

$${}_3D_{ij} = {}_2D_{ij} + T_k(\sum_i {}_2D_{ij}/\sum_i \sum_i {}_2D_{ij}) \quad (5)$$

where

${}_3D_{ij}$ = exchange of all trips between the zones,
i = a zone that includes external cordon count point k, and
T_k = number of vehicles that enter the study area at point k.

If external trips include buses, such as those used to transport workers to an external industrial plant, passenger trips and occupancies must also be considered.

11. Aggregate zones on each side of screen lines and compare their volumes with traffic counts. If all computed volumes are higher or lower than the counts, adjust the estimated input values of trip generation and vehicle occupancy. If the computed volumes fluctuate on both sides of the counted volumes, adjust trip purpose distributions.
12. Assign the distribution table to the present network by using a full all-or-nothing model or split the values into the number of available routes.
13. Compare the computed link volumes with traffic counts where available and with perceived volume ranges

for less important links that were not counted. Adjust the assignments in order to obtain the best fit.

14. Find those links for which a good volume fit can be achieved only by illogical assignments and identify zones that contribute most of the volumes on such links. By using the zonal values and graphs described in the next section, adjust the estimated present trip purpose percentages. Repeat the process until the inconsistencies are rectified.

15. By using future population, total employment, retail employment, the values of trip generation, and vehicle occupancy adjusted to express future trends, apply the developed model to future situations. Adjustment of trip purpose distribution is also possible but not advisable in more standard cases, since the model would lose much of its transparency.

MODEL SENSITIVITY TO LAND USE CHARACTERISTICS

The sensitivity of the input-output values of the four-purpose trip generation and distribution model was analyzed at the University of Alberta during 1975-1978. The objective was to determine how initial input estimates influence the output trip distribution table with respect to different zonal land uses.

Four western Canadian communities that had recent transportation studies available were investigated: Medicine Hat, Alberta (1971); Kelowna, British Columbia (1973); Prince George, British Columbia (1974); and Fort McMurray, Alberta (1974) (4-7). Independent variables were population, employment, and retail employment. The rate of change of the zonal attractiveness was used as the dependent variable. For each zone, the following relation was examined:

$$y = cx + b \qquad (6)$$

where

- y = change of trips attracted to the zone (i.e., number of trips calculated from the given trip purpose distribution divided by the number of trips calculated from the calibrated trip purpose distribution minus 1.0) expressed as a percentage;
- x = error in home-based work, home-based shopping, or home-based other trip percentages;
- c = coefficient that expresses the rate of change of trips attracted to the zone for a unit error in trip purpose distribution (this coefficient was used as the dependent variable in the final sensitivity analysis); and
- b = constant (found to be negligible).

Since the objective was to find the relation between c, which may be called the sensitivity coefficient, and some zonal land use characteristics, the study concentrated on finding suitable independent variables based on population, employment, and retail employment. In order to normalize the zonal characteristics, ratios of zonal population to total population, zonal employment to total employment, and zonal retail employment to total retail employment were first used. They did not, however, produce very consistent results. Several other measures, such as the ratios of zonal population to zonal employment or to zonal retail employment, were also tried. They, too, did not bring about any reasonable consistency. The best results were obtained by using the ratio of the ratio between zonal population and total population to the ratio of zonal employment to total employment, that is,

$$k = (P_i/\Sigma P_j)/(E_i/\Sigma E_j) \qquad (7)$$

where

- k = land use measure,
- P_i = zonal population,
- E_i = zonal employment,
- ΣP_j = total population, and
- ΣE_j = total employment.

All the values obtained were examined by regression analysis and, with one exception, the results were found significant at the 95 percent confidence level.

The distribution of trip purposes, as calibrated and used in the studies, was assumed to be correct. In order to simplify the analysis, four combinations of modified trip distribution were analyzed. They are shown in Table 3.

The combined results of the analysis for all four communities are graphically shown in Figures 5-8. They may be interpreted as follows.

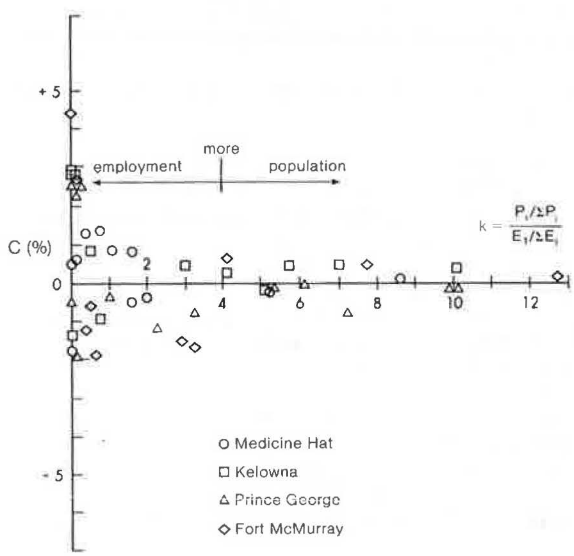
In Figure 5, overestimation of the share of home-based work trips and underestimation of the share of home-based shopping trips (or vice versa) by 1 percent produces about ±1 percent error in the attractiveness of residential zones and between +4 and -2 percent error in the attractiveness of zones that have pronounced employment characteristics. This means that, for an initial error of 5 percent of overestimation of the percentage of home-based work trips and 5 percent of underestima-

Table 3. Combinations of trip purpose distributions used in sensitivity analysis.

	Modification of Trip Purpose Distribution[a] (%)			
Data	Home-Based Work	Home-Based Shopping	Home-Based Other	Non-Home-Based
1	+5 to -5	-5 to +5	0	0
2	+5 to -5	0	-5 to +5	0
3	+5 to -5	0	0	-5 to +5
4	+5 to -5	-2.5 to +2.5	-2.5 to +2.5	0

[a]Deviation from the calibrated value (see Table 2).

Figure 5. Sensitivity of trip attraction to zonal population and employment and to errors in the share of home-based work and home-based shopping trips (data 1).



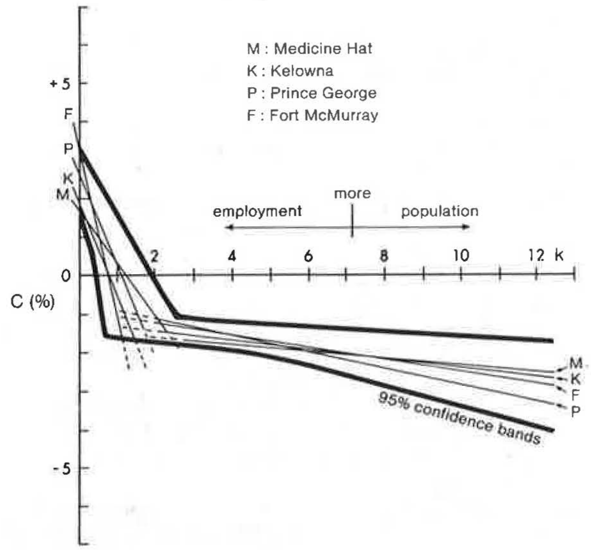


Figure 6. Sensitivity of trip attraction to zonal population and employment and to errors in the share of home-based work and home-based other trips (data 2).

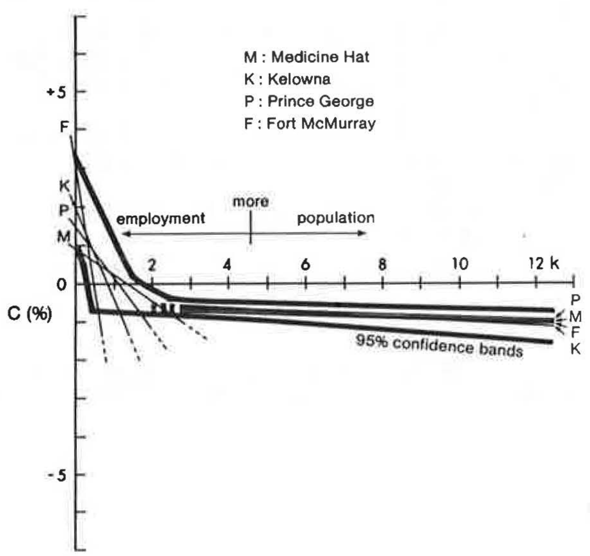


Figure 7. Sensitivity of trip attraction to zonal population and employment and to errors in the share of home-based work and non-home-based trips (data 3).

tion of the percentage of home-based shopping trips, the magnitude of the outcome error in the attractiveness of the individual zones will be in the range of ±5 percent for residential zones and between + 20 percent and -10 percent for employment zones. The sign of the outcome error in relation to the sign of the initial error in home-based work trips cannot be predicted.

In Figure 6, overestimation of the share of home-based work trips and underestimation of the share of home-based other trips by 1 percent produces about 2-3 percent of the underestimation of the attractiveness of residential zones and up to about 3 percent of overestimation of the attractiveness of employment zones. Conversely, underestimation of home-based other trips produces similar magnitudes of overestimation of residential attractiveness and underestimation of employment attractiveness.

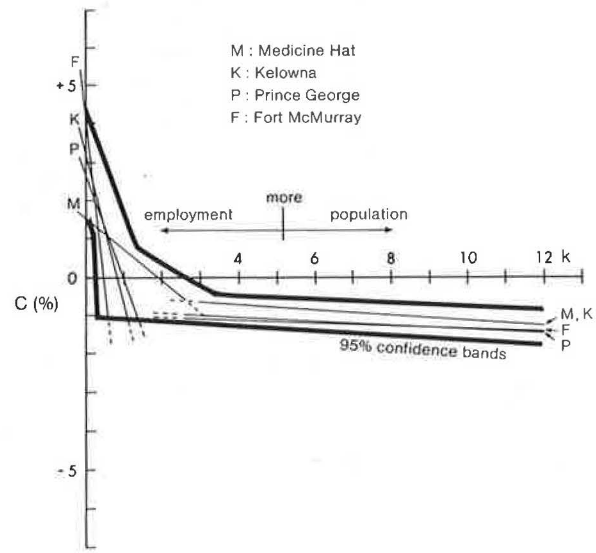


Figure 8. Sensitivity of trip attraction to zonal population and employment and to errors in the share of home-based work, home-based shopping, and home-based other trips (data 4).

In Figure 7, overestimation of the share of home-based work trips and underestimation of the share of non-home-based trips by 1 percent produces about 1 percent of the underestimation of the attractiveness of residential zones and up to about 3 percent of overestimation of the attractiveness of employment zones and vice versa.

In Figure 8, overestimation of the share of home-based work trips by 1 percent and simultaneous underestimation of home-based shopping and home-based other trips by 0.5 percent each produces about 1 percent of the underestimation of the attractiveness of residential zones and up to about 4 percent of the overestimation of the attractiveness of employment zones and vice versa.

In Figures 6-8, zones that are combined residential and employment in character, for which the independent variable k is about 1.0 (i.e., their percentage of the area population is about the same as their percentage of the area employment), are almost indifferent to errors in trip purpose distribution. Shopping trips (Figure 6) are an exception. An interesting observation in Figures 5-8 is that a dynamic community (Fort McMurray) generates a higher degree of error instability (especially for employment zones).

Practical implications of this sensitivity analysis can be inserted into the following steps of the four-purpose trip generation and distribution model as follows:

Step 3, addition—The screen lines should preferably be defined in such a way that the residential and employment zones are separated, or that the number of trips to and from selected residential or employment zones or zone clusters can be identified.

Step 11, addition—The counted number of trips to and from the selected zones or zone clusters should be compared with those computed. Determine the percentage error. For the zones involved, determine their k value. By using Figures 5-8, estimate realistic values of the sensitivity coefficients c. Divide the percentage error by the sensitivity coefficient, thus obtaining an approximate value of adjustment of the share of home-based work trips. By using the character of the analyzed zones, decide which other trip purposes should compensate for the adjustment of the home-based work trips. Repeat the process until a satisfactory fit is found.

CONCLUSIONS

The experience with the synthetic four-purpose trip generation and distribution model in small communities in western Canada has been good. It suits the needs and constraints of the communities and provides an appropriate level of accuracy at a reasonable cost. The model's sensitivity analysis helps to speed up the calibration and may increase the planner's confidence in the validity of the process and its results. There is sometimes a need, however, to carry out a full transportation analysis of a selected community in order to verify the basic data that can be used in other communities.

ACKNOWLEDGMENT

I wish to express gratitude to the National Research Council of Canada and the Alberta Research Council for their financial support and to many practicing colleagues for their support and comments. Special thanks belong to T. Kondo and B. Stephenson for the analysis and comparison of the Canadian studies.

REFERENCES

1. J. W. Bates. Synthetic Derivation of Internal Trips for Small Cities. TRB, Transportation Research Record 526, 1974, pp. 93-103.
2. W. R. Jefferies and E. C. Carter. Simplified Techniques for Developing Transportation Plans: Trip Generation in Small Urban Areas. HRB, Highway Research Record 240, 1968, pp. 66-87.
3. D. B. Coomer and J. C. Corradino. Trip Generation-Distribution in Small Urban Area: An Efficiency Analysis. Traffic Engineering, June 1973, pp. 61-67.
4. Traffic Planning and Parking Study for the City of Medicine Hat. Stanley Associates Engineering, Ltd., Edmonton, Alberta, 1971.
5. City of Kelowna: Traffic Planning Study. Stanley Associates Engineering, Ltd., Edmonton, Alberta, 1973.
6. Traffic Planning Study: City of Prince George. Stanley Associates Engineering, Ltd., Edmonton, Alberta, 1974.
7. General Plan for the Town of Fort McMurray: Transportation Planning. TES Research and Consulting, Ltd., Calgary, Alberta, 1974.
8. A Public Transportation Study for the City of Grande Prairie. N. D. Lea and Associates, Ltd.; L. G. Grimble Associates, Ltd., Edmonton, Alberta, 1975.
9. City of Medicine Hat Traffic Planning Study. Schaeffer Consultants, Ltd., Calgary, Alberta, 1977.
10. Grande Prairie Transportation Study. Stanley Associates Engineering, Ltd., Edmonton, Alberta, 1977.
11. City of Drumheller Transportation System Study. Underwood, McLellan and Associates, Ltd., Edmonton, Alberta, 1972.
12. Leduc Transportation Study. Alberta Transportation, Edmonton, Alberta, 1975.
13. H. M. Hajj. Synthesis of Vehicle Trip Patterns in Small Urban Areas. Transportation Research Record 369, 1971, pp. 181-198.
14. M. A. Sargious and J. F. Morrall. Simplified Traffic Forecasting Process for Small Cities. Canadian Journal of Civil Engineering, Sept. 1975, pp. 314-320.
15. W. L. Greco, F. J. Wegman, J. A. Spencer, and A. Chatterjee. Transportation Planning for Small Urban Areas. NCHRP, Rept. 167, June 1976, 71 pp.
16. D. Rhodes and T. Hillegass. Demonstration of a Simplified Traffic Model for Small Urban Areas. TRB, Transportation Research Record 677, 1978, pp. 76-81.
17. L. R. Goode. Evaluation of Estimated Vehicle Trip Productions and Attractions in a Small Urban Area. TRB, Transportation Research Record 638, 1977, pp. 21-25.
18. Travel Forecasting Techniques Used in the Province of Alberta. Alberta Transportation, Edmonton, Alberta, internal paper, 1972.
19. T. Kondo. Sensitivity of the Proportionate Synthetic Model for Trip Generation and Distribution in Small Cities. Univ. of Alberta, Edmonton, Alberta, M.Sc. thesis, 1976.

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

Transferability of Trip Generation Models

Lawrence C. Caldwell III and Michael J. Demetsky

Cross-classification, disaggregate regression, and aggregate regression trip generation models and trip rates were developed for three cities. The model for each of the cities was transferred to the two other cities and comparisons were made. The comparisons revealed that models calibrated on aggregate zonal data perform better than models calibrated with disaggregate household data when aggregate data are used for forecasting. However, if cross-classification models are acceptable, they can be transferred between cities if good judgment is used to select cities that are similar enough for this purpose. Recommended is the establishment of a standard procedure for data collection and trip generation analysis in selected studies of the near future so that the transferability question can be properly addressed. The emphasis should be on the development of new prototype models for application in groups of cities.

Trip generation is that phase of the urban transportation planning process that establishes relations between urban activity and travel. In the past, each transportation study has calibrated its own set of trip generation procedures based on origin-destination (O-D) data from home interview surveys. Data collection through O-D surveys is costly, especially in small cities where a high sample rate is required. Accordingly, the Federal Highway Administration (FHWA) has advocated planning methods that reduce data collection requirements (1). In this regard, the goal of FHWA is to develop a travel simulation procedure that is based on using information

and experience from one locality to develop trip generation and trip distribution models that can be applied in other areas.

There is little documented experience concerning the application of the synthetic trip generation analysis procedures advanced by FHWA. Therefore, there was a need to test the transferability of these trip generation models and the adequacy of the prescribed method. Also, there was a need to determine the suitability of the use of the FHWA method for transportation planning in Virginia. This determination included calibrating and transferring models and the availability of forecasting data.

TECHNICAL ISSUES

Before the trip generation methodology is explicitly considered, certain technical issues that influence the performance of the modeling procedures being examined, but currently unresolved, are identified. The specific considerations addressed are area classification strategies, local versus synthetic models, and aggregate versus disaggregate data.

Area Classification

Applied methods of classifying cities for the purpose of aiding transportation planners in transferring trip generation models between cities were examined. One approach classifies cities by population and automobile availability (2). Automobile availability was found to be highly correlated with trips per person; however, areas that have high automobile ownership rates generated fewer trips per person than areas that have low automobile ownership rates (3). Consequently, the validity of this classification scheme is questionable because one would expect greater automobile availability to lead to a greater number of trips.

Another method of classifying cities is based on their structure (4). This technique measures city structure by the time distribution of job opportunities within the metropolitan area. This classification method appears to be more applicable to trip distribution than to trip generation because it is potentially useful for classifying cities in order to transfer gravity-model friction factors.

A third method classifies cities according to their dominant economic activity (5-9). This classification scheme is based on the percentage of the labor force employed in various industries and is a good measure of the distribution of total land use in the city. Since trip attraction rates generally depend on land use type, this classification method is sensitive to the trip attraction intensity and distribution in the area.

A fourth method of classifying cities (10, 11) uses factor and cluster analyses to group cities according to selected characteristics (variables) input to the factor analysis. The factor analysis groups similar single measures (individual variables) into factors and rates each city according to the set of generated factors. The cluster analysis then involves forming groups such that the within-group variances are minimized while the between-group variances are maximized. Factor analysis appears to be the most comprehensive method; however, it is complicated and the inclusion of extraneous variables may confound the results of the classification scheme.

None of the methods described for classifying cities has been applied for the specific purpose of identifying urban characteristics that associate directly with differences in trip generation activity. Consequently, no method is available for making a strong case for transferring models between selected pairs of cities.

In-depth testing of the city classification-model transferability issue was limited in the study because of (a) the incompatibility between the models developed here with Virginia data and those given in the FHWA report and (b) the limited number of cities for which models were collected. The problem of model transferability can be properly addressed only when a large number of models that are based on the same parameters are available from a wide range of cities.

Local Versus Synthetic Models

A number of problems are encountered in the transfer of transportation forecasting models. For instance, in order to test the validity of a transferred trip generation model, the productions and attractions must be processed through the trip distribution, mode choice, and traffic assignment phases to show link volumes that are comparable with traffic count data. This procedure must be employed with caution because of the multiple sources of potential error that affect the projected volumes. In addition to the possibility of erroneous estimates of trip ends, the flows that result from borrowed parameters for the trip distribution model or the route assignment role that is employed may be wrong. Thus, if the simulated flows do not agree with the observed values, the exact source of error is nearly impossible to specify.

Another possible problem in borrowing models is that the variables used in the borrowed model must be available locally and must be easily forecast. For example, a large number of cross-classification models use automobile ownership and income as the independent variables, but these models were difficult to use in Virginia because income information was not directly available.

Aggregate Versus Disaggregate

Another technical consideration in this study concerns the application of a disaggregate model to aggregated data. Cross-classification curves are calibrated on disaggregate household data, but generally aggregated zonal averages are used for forecasting. Much of the variance in data that can be accounted for in the disaggregate model is lost when data are aggregated to the zonal level. When zonal averages are used with the cross-classification model, one must assume that the number of trips produced by the average household in a zone is equal to the average number of trips produced by the households in the zone. This assumption was found to be false. For example, in Roanoke the average household size was 3.23 persons and the trip rate that corresponds to this household size was 8.21 trips/household; however, the average number of trips per household was only 7.60, which resulted in 8 percent overestimates of the total trip productions. When a disaggregate model is used with aggregated data, some measure of the distribution (e.g., standard error) should be given so that the magnitude of this estimation error can be determined.

TRIP PRODUCTION MODELS

In order to test the transferability of cross-classification procedures, models were developed for selected cities in Virginia. These cities were chosen on the basis of certain similarities and on the availability of data. Two pairs of cities were selected for study. These cities along with the selected characteristics are listed in Table 1. Originally Lynchburg and Roanoke were

selected as pair 1, but the necessary data were not available for Lynchburg. Therefore, the study concentrated on Roanoke, Harrisonburg, and Winchester.

The two explanatory variables were selected to be automobile ownership and household size. Although income is highly recommended by FHWA as one of the independent variables, it was not used in this study because information on incomes is not available in Virginia. These cross-classification trip generation models are exemplified by the curves developed for Roanoke as shown by Figures 1-3.

Regression equations were also calibrated with the household data from the O-D survey. These equations used household size and automobile ownership as the independent variables so that they could be compared with the cross-classification trip rates. The household regression equations obtained are as follows:

In Harrisonburg,

$$\text{Trips/household} = -1.48 + 1.85 \times \text{household size} + 3.35 \times \text{automobiles/household} \quad (1)$$

Table 1. Characteristics of selected cities.

City	Population in 1970	Persons per Household	Automobiles per Household	Per Capita Income ($)
Lynchburg	70 842[a]	3.02	1.140	2906
Roanoke	156 621[a]	2.97	1.224	3085
Harrisonburg	14 605	2.79	1.120	2742
Winchester	14 643	2.80	1.090	2954

[a] Urbanized area population.

Figure 1. Roanoke percent households by automobile ownership and household size.

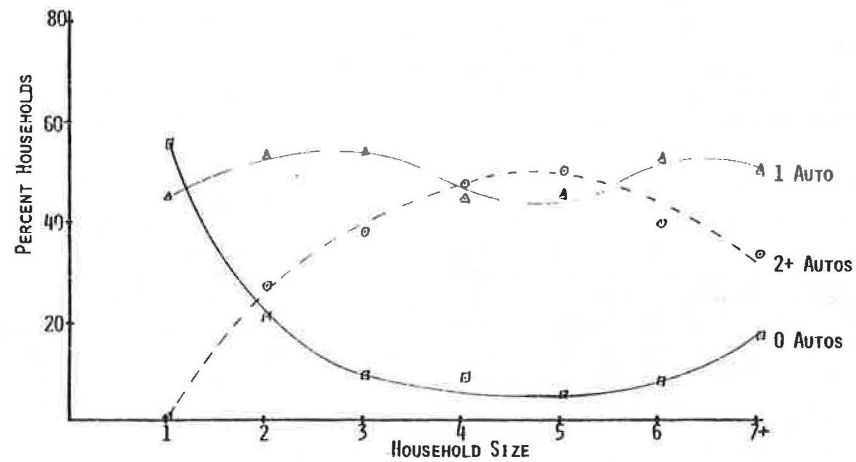


Figure 2. Roanoke trip rates by automobile ownership and household size.

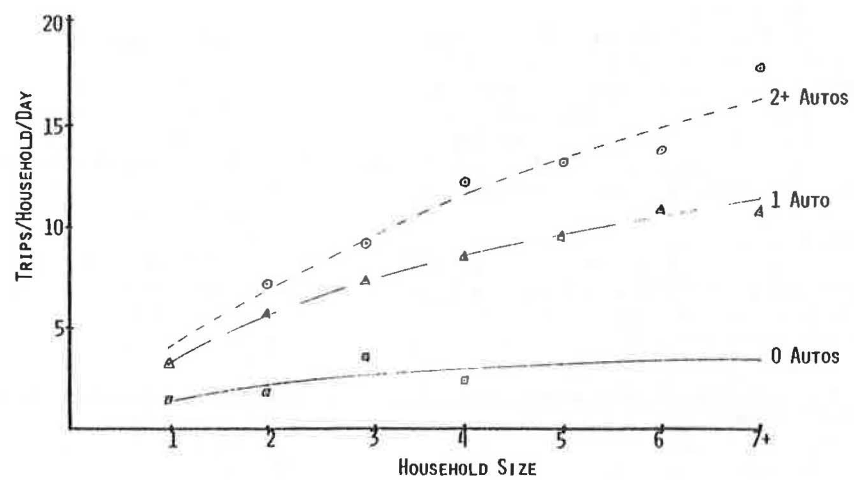


Figure 3. Roanoke percent of trips by trip purpose and household size.

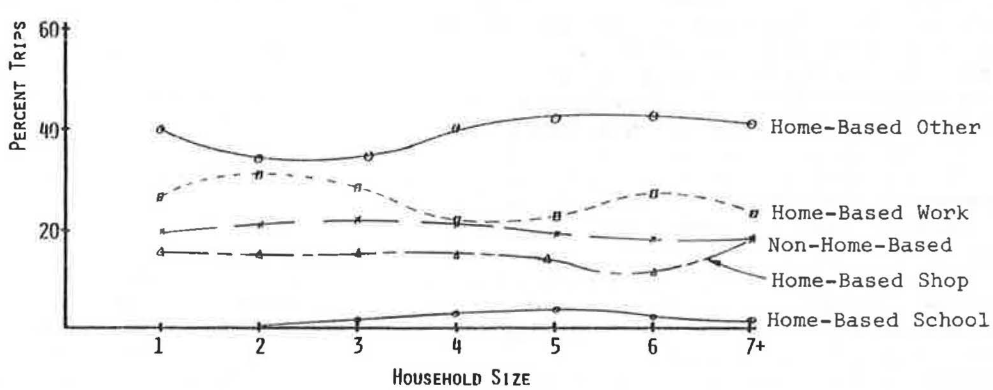


Table 2. Calibrated model predictions.

City	Trips Reported	Trips Predicted: Cross-Classification Total	Cross-Classification χ^2	Disaggregate Regression Total	Disaggregate Regression χ^2	df	χ^2 Required for Significance at $\alpha = 0.05$
Disaggregate data							
Roanoke	17 340	17 332	1171	17 333	1359	160	190
Winchester	11 298	11 295	405	11 298	460	50	67
Harrisonburg	9 174	9 117	599	9 346	819	96	119
Aggregate data							
Roanoke	17 340	18 658	3157	17 250	1333	160	190
Winchester	11 298	11 801	1004	11 274	453	50	67
Harrisonburg	9 174	9 091	1767	9 299	798	96	119

Table 3. Transferred model predictions by using disaggregate data.

City	Cross-Classification: Roanoke n	Roanoke χ^2	Winchester n	Winchester χ^2	Harrisonburg n	Harrisonburg χ^2	Disaggregate Regression: Roanoke n	Roanoke χ^2	Winchester n	Winchester χ^2	Harrisonburg n	Harrisonburg χ^2	df	χ^2 Required for Significance at $\alpha = 0.05$
Roanoke	17 332	-	18 260	91	19 554	368	17 333	-	18 399	70	19 706	336	160	190
Winchester	10 707	46	11 295	-	11 977	80	10 674	40	11 298	-	12 051	65	50	67
Harrisonburg	8 127	172	8 646	49	9 117	-	8 355	112	8 892	30	9 346	-	96	119

$R^2 = 0.6903$; standard error of estimate of the mean (SE) = 0.7709.

In Roanoke,

$$\text{Trips/household} = -0.28 + 1.27 \times \text{household size} + 3.06 \times \text{automobiles/household} \quad (2)$$

$R^2 = 0.5076$; SE = 0.7947.

In Winchester,

$$\text{Trips/household} = -0.66 + 1.35 \times \text{household size} + 3.54 \times \text{automobiles/household} \quad (3)$$

$R^2 = 0.5909$; SE = 0.7562.

In addition to the cross-classification matrices and disaggregate regression equations, regression equations were calibrated for Roanoke on aggregated data used in the Roanoke area transportation study of 1965. The available data for each traffic zone included number of total persons, number of occupied dwelling units, number of passenger cars, school attendance by zone of school (excluding college), and total blue-collar and white-collar employment by zone of work (TOTEMP).

The home-based work, home-based other, and non-home-based productions and attractions by traffic zone for the base year (1965) were available as vehicle trips. They were then converted to person trips by using automobile occupancy rates by trip purpose developed for Charlottesville. The total person trip productions computed by this expansion method compared very well with the total trips reported in the Roanoke area transportation study (353 493 as compared with 353 385).

The following regression equations were then calibrated for each of the three trip purposes:

Home-based work productions,

$$\text{HBW P} = 14.77 + 1.25 \times \text{automobiles} \quad (4)$$

$R^2 = 0.88$; SE = 0.353.

Home-based other productions,

$$\text{HBO P} = -84.59 + 3.07 \times \text{automobiles} \quad (5)$$

$R^2 = 0.89$; SE = 0.377.

Non-home-based productions,

$$\text{NHB P} = -5.75 + 1.25 \times \text{TOTEMP} + 1.09 \times \text{dwelling units} \quad (6)$$

$R^2 = 0.90$; SE = 0.728.

Total production,

$$\text{Tot P} = -47.63 + 5.12 \times \text{TOTEMP} \quad (7)$$

$R^2 = 0.93$; SE = 0.628.

MODEL EVALUATIONS

Trip productions were predicted by the cross-classification and regression models by using both household data (from the O-D survey) averaged for the traffic zones (aggregated data) and summing predictions for the individual household observations (disaggregate data) for each zone. In both cases, in all three study areas the distribution of predicted trips was statistically found to be significantly different from the reported trips. The total predictions and chi-square values are shown in Table 2. These results show that, even while using disaggregate data with the disaggregate models, the models do not perform accurately.

In addition to comparing actual trip productions reported in the O-D survey with trip productions predicted by using the calibrated models, the calibrated models were compared with models transferred from the other two study areas. The results of transferring the models are shown in Table 3. The Winchester cross-classification model transferred acceptably to both Roanoke and Harrisonburg and the Roanoke cross-classification model transferred adequately to Winchester. These results show that cross-classification models can be transferred between cities; however, care should be taken in selecting similar cities.

The cross-classification models were also evaluated with expanded base-year planning data aggregated to the zonal level. These results are shown in Table 4. The expectation that the disaggregate cross-classification models would not perform as well when aggregated data were used was found to be true, as can be seen in the

higher chi-square values (compare Tables 3 and 4).

A comparison of the predictive ability of the cross-classification model versus the aggregate and disaggregate regression models for Roanoke is found in the table below. As can be seen from this table, the aggregate regression model predicts trip productions better at the traffic zone level (lowest chi-square value) and on the citywide level (best total productions).

Roanoke Productions	Productions	Chi-Square (compared with actual)
Actual	353 493	0
Cross classification	383 318	93 003
Aggregate regression	353 494	30 505
Disaggregate regression model	353 957	57 294

It is more difficult to forecast data at the household level than at the zonal level and the accuracy of forecast houshold data has to be determined before a direct comparison can be made. Overall, models calibrated on aggregated zonal data appear to perform better than models calibrated on disaggregate household data when used with forecast aggregated zonal data.

Table 4. Transferred cross-classification models by using aggregate data.

City	Model	Total Productions	χ^2
Roanoke[a]	Calibrated	383 318	
	Winchester	389 857	578
	Harrisonburg	436 390	9776
Winchester[b]	Calibrated	62 822	
	Roanoke	63 139	60
	Harrisonburg	68 619	672
Harrisonburg[c]	Calibrated	46 306	
	Roanoke	43 521	320
	Winchester	43 553	362

[a]df = 160; χ^2 required for significance at $\alpha = 0.05 = 190$.
[b]df = 50; χ^2 required for significance at $\alpha = 0.05 = 67$.
[c]df = 119; χ^2 required for significance at $\alpha = 0.05 = 119$.

Table 5. Roanoke trip attraction rates.

Trip	Trips per Household	Trips per Employee	Trips per $1000 Retail Sales	Trips per Student
Home-based work		1.471		
Home-based shopping			0.195	
Home-based school				[a]
Home-based other	1.046	1.321		
Nonhome based	0.331	0.412	0.126	

[a]No school enrollment data available.

Table 6. Land use trip rate calculations for Roanoke study zones.

Item	Trip Rate	Trip Ends
Traffic zone 1		
286 dwelling units	10.0 trips/dwelling unit	2 860
425 hotel rooms	11.3 trips/room	4 802
14 461 m^2 office space	125.78 trips/1000 m^2	1 819
7637 m^2 manufacturing space	44.12 trips/1000 m^2	337
3 gasoline stations	748 trips/station	2 244
Total		12 062
Traffic zone 67		
1083 dwelling units	10.0 trips/dwelling unit	10 830
1465 m^2 restaurant space	1768.94 trips/1000 m^2	2 591
2048 m^2 warehouse space	53.91 trips/1000 m^2	110
5401 m^2 office space	125.78 trips/1000 m^2	679
710 elementary students	0.51 trips/student	362
Total		14 572
Traffic zone 151		
19 080 m^2 shopping center	536.92 trips/1000 m^2	10 244

Note: 1 m^2 = 10.76 ft^2.

TRIP ATTRACTION ANALYSIS

For the analysis of trip attractions, three techniques were considered. The first was the standard method recommended by FHWA (1), which uses general trip rates. A second method used the linear regression method, and the third employed specific land use trip rates to predict trip ends.

Trip Rate Procedures

The method for predicting trip attractions suggested by FHWA (1) is described as "a simplified approach... based upon the development of trip rates with a matrix."

Trip attraction rates for the Virginia test cities were calculated from available planning data. In computing the rates, the numbers of trips produced for the types of trip purposes were summed for all traffic zones to give a citywide total of trip attractions by each purpose. Total trip attractions for each purpose were set equal to total trip productions because trip production prediction methods are generally considered to be more accurate than trip attraction prediction techniques. The values of the socioeconomic variables (total employment, retail sales, number of households, and number of students) were also summed for all traffic zones to get citywide totals. The trip attraction rates were then computed to provide the matrix of trip rates for Roanoke in Table 5.

Regression Procedures

Attraction regression equations developed for Roanoke are shown below with the index of determination (R^2). Calibrated regression equations for trip attractions are as follows:

Home-based work attractions,

$$\text{HBW A} = -9.72 + 1.19 \times \text{TOTEMP} \tag{8}$$

$R^2 = 0.97$

Home-based other attractions,

$$\text{HBO A} = 123.82 + 1.39 \times \text{TOTEMP} + 1.18 \times \text{automobiles} \tag{9}$$

$R^2 = 0.68$

Non-home-based attractions,

$$\text{NHB A} = -5.75 + 1.26 \times \text{TOTEMP} + 1.09 \times \text{dwelling units} \tag{10}$$

$R^2 = 0.90$

Total attractions,

$$\text{TOT A} = 196.22 + 3.81 \times \text{TOTEMP} + 2.47 \times \text{dwelling units} \tag{11}$$

$R^2 = 0.88$

Land Use Trip Rates

Another method of predicting trip attractions involves the use of rates based on specific land uses. Some of these rates have been published by the Institute of Transportation Engineers (ITE) and by the Arizona Department of Transportation (12-14). In this method the number of units in each land classification is multiplied by the trip rate for that particular land use and summed for the analysis area to predict trip ends (both productions and attractions).

Table 7. Comparison of actual productions and attractions with predictions by using cross-classification, regression, and land use trip rates.

Study Zone	Actual Productions and Attractions[a]	Land Use Trip Ends[b]	Error (%)	Cross-Classification Productions and Attractions[a]	Error (%)	Regression Productions and Attractions[a]	Error (%)
Roanoke							
TZ 1	10 414	12 062	15.8	17 587	68.9	15 918	52.8
TZ 67	14 397	14 572	1.2	15 405	7.0	15 003	4.2
TZ 151	8 519	10 244	20.2	1 718	-79.8	1 646	-80.7
Lynchburg							
TZ 11	4 165	4 489	7.8				
TZ 20	3 478	1 780	-48.8				
TZ 34	25 862	25 960	0.4				

[a] Data are from 1965. [b] Roanoke data are from 1966; Lynchburg data are from 1968.

Available rates for Virginia were compared with the rates published by Arizona and ITE to determine which rates to use in this study. The rates from all three studies appeared to be very similar, with the exception of the rates for small shopping centers.

The trip rates published by ITE were used in this study because the land use classifications were slightly easier to use than those from the Arizona study. This method of estimating trip ends was performed on three traffic zones in Roanoke and on three zones in Lynchburg. We assumed that Lynchburg would be included in the study at the time this analysis was performed. A large number of units could not be classified from the aerial photograph alone, but when a city directory was used in conjunction with the aerial photograph, all units could be classified. Thus, the need for an on-site study was eliminated. An example of the land use characteristics and trip end calculations for traffic zones is shown in Table 6 (12). The floor areas were measured on the area photographs by using the scale of the photograph, the dimensions of the building, and the number of floors. This procedure is rather tedious and approximately 32 person-h were required to classify the six traffic zones.

Evaluation of Procedures

It is not necessary to transfer standard trip rates from one city to another because the trip rates can be computed from the planning data available for each city. The trip rates are based on total predicted trip productions so that total productions and attractions will balance. If trip rates were transferred from another city, the production totals and attraction totals would disagree. By using rates calculated from total productions and the city's own planning data, this problem is avoided. Because the rates are based on total trip productions, separate rates have to be computed for each trip production model transferred to the city. For example, if the Roanoke and Harrisonburg cross-classification models were transferred to Winchester, separate rates would have to be computed for each of the two transferred models because the total trip production varies with the transferred trip production model.

A comparison of the Roanoke actual attractions and estimated attractions was made by using the standard trip rate method for each traffic zone. Although the total attractions predicted by using the standard trip rates were 8.43 percent greater than the actual attractions, the zonal predictions were in error an average of 74.4 percent. The total attractions for Roanoke predicted by using the regression equations were 0.95 percent greater than the actual attractions; however, the zonal predictions were in error an average of 92.6 percent.

These results imply that the procedures for predicting trip attractions can be very inaccurate. Either revision in these methods or a new method of predicting trip attractions is needed. The use of specific land use trip rates is one such new method. The trip ends predicted by using the land use trip rates were compared with the actual productions and attractions for both Lynchburg and Roanoke. The Roanoke predictions were also compared with the regression and cross-classification predictions. These comparisons are shown in Table 7. The high percentage of error for traffic zone 20 of Lynchburg would indicate the possibility that the land use changed drastically between 1965 and 1968. Except for this zone, the land use trip rate method predicted trip ends reasonably well. It predicted much better than either the cross-classification or the regression method did in Roanoke. One disadvantage of the land use trip rate method is that it requires a directional factor to split trip ends to productions and attractions. A second disadvantage is that the method requires very specific land use forecasts.

CONCLUSIONS

An investigation of the application in Virginia of the trip generation procedures described by FHWA (1) revealed certain findings that can be interpreted to show how trip generation analysis procedures can be improved and standardized. Models calibrated on aggregate zonal data performed better than models calibrated on disaggregate household data when they were used with forecasts of aggregated data. The average rates given by a cross-classification table that are applied at a disaggregate level were not sensitive to locational (zonal) variations. It was also shown that cross-classification models can be transferred between cities; however, care should be taken in selecting similar cities between which the models are to be transferred.

The trip attraction forecasting methods based on areawide trip rates are not sensitive to specific site characteristics. Procedures for forecasting trip attractions that are based on land use trip rates appear to warrant consideration because of their sensitivity to specific land uses.

The reported findings led to certain recommendations for improving the synthetic trip generation analysis procedures. Initially, all states should be encouraged to use standard procedures for trip generation analysis and data collection for a period in the future so that information and results can be properly compared. The emphasis should be on the development of prototype models for application in groups of cities. A determination should be made of how accurately disaggregate household data can ultimately be forecast. If disaggregate data cannot be forecast accurately, then the study of disaggregate models is futile because they do not per-

form as well as aggregate models that use aggregate data.

ACKNOWLEDGMENT

This research was conducted at the Virginia Highway and Transportation Research Council in Charlottesville and was sponsored by the Virginia Department of Highways and Transportation and the Federal Highway Administration. We acknowledge that the research could not have been accomplished without the assistance of the personnel of these agencies, particularly Jack Page, Alan Lassiter, Grady Ketron, and Jerry Korf. The opinions, findings, and conclusions expressed in this paper are ours and not necessarily those of the sponsoring agencies.

REFERENCES

1. Trip Generation Analysis. Urban Planning Division, Federal Highway Administration, Aug. 1975. GPO: 050-001-00101-2.
2. W. H. Bottiny and B. T. Goley. A Classification of Urbanized Areas for Transportation Analysis. HRB, Highway Research Record 194, 1967, pp. 32-61.
3. Projection of Urban Personnel Transportation Demand. Peat, Marwick, Livingston, and Company, New York, March 1968.
4. A. M. Voorhees and S. J. Bellomo. Urban Travel and City Structure. HRB, Highway Research Record 322, 1970, pp. 121-135.
5. H. J. Nelson. A Service Classification of American Cities. Economic Geography, Vol. 31, 1955, pp. 189-210.
6. V. Jones and A. Collver. Economic Classification of Cities and Metropolitan Areas. In The Municipal Year Book, International City Management Assn., Washington, DC, 1960, pp. 67-79, 89-90.
7. V. Jones and R. L. Forstall. Economic and Social Classification of Metropolitan Areas. In The Municipal Year Book, International City Management Assn., Washington, DC, 1963, pp. 31-44.
8. V. Jones, R. L. Forstall, and A. Collver. Economic and Social Characteristics of Urban Places. In The Municipal Year Book, International City Management Assn., Washington, DC, 1963, pp. 85-157.
9. R. L. Forstall. Economic Classifications of Places over 10 000, 1960-1963. In The Municipal Year Book, International City Management Assn., Washington, DC, 1967, pp. 30-65.
10. B. J. L. Berry. Latent Structure of the American Urban System with International Comparisons. In City Classification Handbook: Methods and Applications, Wiley, New York, 1972, pp. 11-60.
11. T. F. Golob, E. T. Canty, and R. L. Gusta. Classification of Metropolitan Areas for the Study of New Systems of Arterial Transportation. Transportation Research Department, General Motors Corporation, Warren, MI, Aug. 1, 1972.
12. Trip Generation. Institute of Traffic Engineers, Arlington, VA, 1976.
13. Trip Generation Intensity Factors. Transportation Planning Division, Arizona Department of Transportation, Phoenix, July 1, 1976.
14. Trip Generation by Land Use. Maricopa Association of Governments, Maricopa County, AZ, April 1, 1974.

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

Abridgment

Analysis of Intercity Travel Markets in New York State

Robert J. Zerrillo and Alfred J. Neveu

This paper presents the results of an analysis of intercity travel market segments in New York State's Empire Corridor (New York City-Albany-Buffalo). The intercity travel data were obtained from a stratified random sample of Empire Corridor residents that was taken in the spring of 1979. The survey collected information on respondents' intercity travel habits and modal awareness, familiarity, and accessibility. Detailed analyses were performed on the Empire Corridor nonbusiness travel market. Tables were developed to show the demographic distributions and mean trip rates of nontravelers, light travelers (one to five trips per year), and heavy travelers (more than five trips per year). A multivariate statistical procedure, automatic interaction detector, was used to attempt to uncover the variables that best explain the variation in trip making. The results indicate that geographic stratum is the best travel segmentation variable. Other variables that have an important influence on intercity travel include age, household size, and knowledge of National Railroad Passenger Corporation (Amtrak) service (a measure of the information level of corridor residents). The variables collected in the survey had different effects on each geographic stratum, which supports the assumption that the intercity travel market is heterogeneous.

In response to the growing use of intercity public transportation in New York State, the New York State Department of Transportation performed a study to aid in the planning and marketing of intercity rail passenger service. A telephone survey of rail corridor (New York City, Albany, and Buffalo) residents was conducted in June 1979 to determine the characteristics of the intercity travel market in New York State.

Because different travel market segments are likely to vary in demand potential, intercity travelers were divided into appropriate markets for more detailed study. A literature search of market segmentation in intercity travel revealed few studies in the analysis of intercity travel market segments. Most of the research has been oriented toward intercity (or corridor) travel demand models, with travel segmentation by purpose and mode. Detailed market segmentation has been limited, for the most part, to intraurban travel [see, for example, Dobson and Tischer (1)]. This paper presents the results

of the analysis of the telephone survey of residents in New York State's Empire Corridor (New York City, Albany, and Buffalo) and describes the characteristics of intercity travelers in that corridor.

SURVEY DESIGN

The study focused on residents in New York State's Empire Corridor and Adirondack Corridor (Albany and Montreal). These heavily traveled routes connect the upstate New York major urban areas with each other and with New York City. Both corridors are served by highway (automobile and bus), rail, and air passenger facilities.

A sample of 600 respondents was selected and stratified into four geographic groups as follows: New York City (100 samples), metropolitan New York City counties (100), upstate urban areas that have rail stations (200), and all remaining cities and towns in the two corridors (200).

The stratum sample sizes shown were chosen in order to ensure statistical reliability for stratum 4 results (the stratum with the smallest population) while keeping the total sample at an acceptable level. Within strata 1, 2, and 3, all places were selected. In stratum 4, 40 cities and towns were selected randomly from a list of all the places in the stratum. Within each stratum, the number of samples allocated to each place is proportionate to the population. A more detailed description of the survey design can be found elsewhere (2).

The main objective of the survey was to collect information for the analysis of intercity travel market segments. The questionnaire covered the following topics:

1. Trip-making characteristics of intercity travelers;
2. Residents' perceptions of modal accessibility, availability, and familiarity with intercity travel;
3. Attitudes and opinions of residents concerning travel characteristics, travel problems, and reasons for mode choice; and
4. Background demographic information.

The demographic variables collected in the survey were factored to 1970 stratum populations and compared with the 1970 state census proportions. Survey values for age, family size, and automobiles per household compare favorably with the 1970 census distribution and even more closely with a 1977 New York State Department of Transportation telephone poll (3). However, the sample contained a greater proportion of women and more multicar households than did the census, even though much of the interviewing was done at night to increase the likelihood of reaching working men. The difference in automobile ownership over the 1970 census was expected, because the proportion of multicar households in the state has been increasing since 1970. Overall, the survey is considered representative of Empire Corridor residents.

For the remaining analysis, the sample was factored by stratum to represent 1980 population projections in the corridor.

SURVEY RESULTS

Intercity Travel

The analysis of intercity travel was performed on the Empire Corridor sample of residents (583 of the 600 interviews). The survey contained a number of questions on intercity trip making. For this study, an intercity trip was defined as any noncommuter trip of at least 80 km (50 miles) in length made within the past year. The 80-km minimum trip length was chosen based on the distance between major corridor cities.

Table 1 reveals that stratum 1 (New York City) contains almost half of all intercity travelers and 41 percent of the Empire Corridor nonbusiness trip makers. Stratum 1's size makes it the largest single intercity travel market. New York City's relative importance in percentage of trips decreases when we look at corridor trips only (i.e., stratum 1 residents do most of their intercity travel outside the Empire Corridor). Strata 2 (New York City metropolitan counties) and 3 (upstate cities with rail stations) have shares about equal to or greater than stratum 1's when corridor nonbusiness trips are viewed. Another interesting occurrence is the rate at which stratum 4 (remaining cities and towns) residents make corridor nonbusiness trips. Stratum 4 contains 7 percent of the trip makers, but they make 14 percent of the corridor's trips. Of the 79 percent of corridor residents who make intercity trips, 40 percent make trips within the Empire Corridor and 37 percent make nonbusiness trips.

The survey showed that automobile was the most widely used mode for intercity nonbusiness travel in the Empire Corridor. About 75 percent of the travelers used automobile for 85 percent of all corridor trips. Bus, airline, and rail travel shared the remaining travel market about equally.

Modal Accessibility, Availability, and Familiarity

Several questions were asked to determine corridor residents' knowledge of the various intercity travel modes in the Empire Corridor. The results showed that familiarity with corridor travel modes is generally very high in all areas of the state. However, the National Railroad Passenger Corporation (Amtrak) was consistently rated lower than the other modes.

In a recent nationwide intercity travel survey (4), 73 percent of the sampled residents of standard metropolitan statistical areas (SMSAs) knew the location of the intercity train station; however, 91 percent knew the location of the intercity bus station. In this study, 79 percent of the survey respondents know the location of the nearest Amtrak station, but 87 percent know where the intercity bus station is located. The presence of all intercity modes of transportation in the Empire Corridor accounts for the high percentage of knowledgeable residents. Intercity terminals were perceived as at least

Table 1. Distribution of trips and trip makers.

	Total Intercity (n = 480)		Total Empire Corridor (n = 286)		Empire Corridor Nonbusiness (n = 262)	
Stratum	Trip Makers (%)	Trips (%)	Trip Makers (%)	Trips (%)	Trip Makers (%)	Trips (%)
1	49	42	40	27	41	29
2	20	25	21	25	21	23
3	25	26	31	31	31	34
4	6	7	8	17	7	14

adequately accessible by most corridor residents surveyed.

For the remainder of this research, Empire Corridor nonbusiness trips are used because nonbusiness trips are discretionary and such travel may be more readily analyzed by using the variables collected in the survey. Corridor business trips were not used due to the low number of business travelers interviewed. Moreover, intercity business travel would be influenced by variables other than those collected in this survey.

ANALYSIS OF INTERCITY TRAVEL MARKETS

The total intercity travel market was divided into definable groups in order to be able to estimate the trip-making propensity of these more homogeneous groups with greater accuracy and to develop and implement more effective policies to deal with these travelers. For the purposes of this study, it was hypothesized that intercity travelers can be divided into three distinct groupings based on past trip-making behavior. These groups are nontravelers (no intercity trips in the past year), infrequent travelers (one to five intercity trips in the past year), and frequent travelers (more than five intercity trips in the past year).

The primary purpose of this analysis was to characterize the different hypothesized groups of intercity travelers by their demographic background and their perceptions of intercity service in the Empire Corridor. To accomplish this task, two procedures were employed. The first is a cross-tabulation of several demographic characteristics by the three travel groups. The second step was to determine the factors, both demographic and attitudinal, that may best divide the entire sample into low and high travel categories. To perform this step, a multivariate statistical procedure, automatic interaction detector (AID), was used. AID is similar to analysis of variance in that it attempts to explain the variation in a dependent variable through the use of independent variables (factors). The AID algorithm divides the total sample into a mutually exclusive series of subgroups through a series of binary splits. Each split attempts to maximize the between-group sum of squares. The result is a tree whose branches constitute a particular market group. No assumptions are made concerning linearity, correlation, or mathematical form of the relationship between variables(5).

Table 2 indicates that men and women have a nearly equal propensity for intercity travel. This is not surprising because many intercity nonbusiness trips are made by families. The distribution of trip makers by age reveals that younger and middle-aged respondents have a greater propensity to make intercity trips than do the older respondents. Almost three-quarters of the older respondents made no nonbusiness intercity trips in the past year. Young travelers (18-34) traveled at the greatest rate.

As expected, the amount of intercity travel rises as automobile ownership increases; households that own two or more automobiles traveled much more often than did other families. However, although households that owned two or more vehicles contain more travelers than single-car households, the latter group makes slightly more trips per traveler, as evidenced by the higher trip rate.

The distribution of trip makers by household size shows little variation, except in the frequent-travel group. In that instance, larger households are more inclined to make a larger number of nonbusiness intercity trips than the smaller households.

The geographic distributions show that stratum 1 contains the lowest average trip rate per traveler for Empire Corridor trips. This result is not surprising because most intercity travelers from New York City make their trips outside the Empire Corridor (2). Stratum 4 residents travel at a rate of 2 trips/year more per person than do residents of either stratum 2 or stratum 3.

Table 2 also shows that infrequent and frequent travelers constitute 31 and 6 percent of the corridor residents, respectively. The importance of this small number of frequent travelers (more than 5 trips/year) in the corridor is revealed by the fact that, of all corridor travelers surveyed, those that travel frequently (14 percent of all trip makers) account for 53 percent of all corridor nonbusiness trips. Diversion of a small portion of automobile users in this group could significantly increase intercity public transportation use in the corridor.

The second phase of the analysis used the multivariate statistical tool known as AID. The independent variables (factors) used in the AID analysis included the demographics previously employed as well as the respondent's awareness and perceptions of the modes available in the corridor and their attributes.

Each stratum was analyzed individually for several

Table 2. Demographic distribution by traveler group.

Item	Nontravelers (%)	Infrequent Travelers (%)	Frequent Travelers (%)	Mean Trip Rate per Traveler
Sex				
Male (n = 215)	64	31	5	4.23
Female (n = 368)	62	32	6	3.83
Age				
18-34 (n = 226)	62	32	6	4.54
35-54 (n = 184)	53	42	5	3.66
55+ (n = 167)	75	21	4	3.53
Automobiles per household				
0 (n = 60)	79	21	0	1.96
1 (n = 207)	71	24	5	4.90
2+ (n = 316)	47	45	8	4.05
Household size				
1 (n = 68)	69	29	2	4.61
2 (n = 193)	68	28	4	4.03
3-4 (n = 216)	52	41	7	3.97
5+ (n = 99)	66	27	7	3.41
Stratum				
1 (n = 100)	71	28	1	2.93
2 (n = 100)	59	32	9	4.49
3 (n = 185)	52	39	9	4.41
4 (n = 198)	48	34	18	6.58
Total (n = 583)	63	31	6	3.98

Table 3. AID results.

Stratum	Variation Explained (%)	Variables That Explain Intercity Travel	
		Low Rate	High Rate
1	24.6	Age (18-34, 55+), do not know how to use Amtrak	Age (35-54), easy access to airport
2	34.4	Married	Not married, age (25-44)
3	7.3	Low automobile ownership, large household size	High automobile ownership
4	18.4	Do not know how to use Amtrak	Know how to use Amtrak, cost of travel unimportant

reasons. Evidence from the previous analysis indicates that there are very real differences in intercity trip making among the geographic areas of New York State. In addition, it proved to be infeasible to use the factored sample in the AID analysis due to computational difficulties. AID is not now programmed to handle these factored data, although future improvements will remedy this shortcoming.

Table 3 shows the AID results for each stratum. The variables that affect travel are listed in order of importance. The difference in the effect of each variable on each stratum and the difference in the total travel variation explained by the AID analysis supports the conclusion that real travel differences exist between the geographic groups.

The AID analysis reveals several points:

1. Several of the variables are important in explaining the travel variation in two or more strata (e.g., age or knowledge of Amtrak),
2. The differences in percentages of variation explained by the AID analysis for each stratum reveal that the variables used are significantly better segmentors of travel in some strata than others, and
3. The relatively small amount of travel variation explained in each stratum indicates that intercity travel variation is not easily explained by the variables collected in the survey.

CONCLUSIONS AND IMPLICATIONS

From this analysis, several conclusions can be drawn with regard to nonbusiness intercity travel in the Empire Corridor. First, the results indicate that travel differences exist between each stratum. The analysis provides evidence to show that geographic stratum may be the best individual travel segmentor of corridor residents. Both the cross-tabulations and the AID analysis support this conclusion.

The AID results also corroborate much of the information given in Table 2. Residents of smaller cities (stratum 4) travel at a higher rate within this corridor than do residents of larger urban areas. Younger respondents travel more often than do older ones. Smaller households (one and two persons) travel at a higher rate across each stratum where household size appears (strata 1, 3, and 4). Knowledge of upstate Amtrak service (a measure of the information level of intercity travelers) appears in two strata (1 and 4) where the awareness of Amtrak service was lowest. The difference in AID trees between strata and the low amount of trip-making variation explained by the trees supports the assumption that intercity travel is a heterogeneous phenomenon. Overall, the results show little demographic difference among Empire Corridor travelers.

This analysis points to several interesting policy implications. Marketing campaigns to increase the public's awareness of the availability and location of the various intercity services and terminals, especially the energy-efficient modes, could lead to increased patronage on these travel modes. A system of group or family fare packages, now used by the airline industry, could have a strong effect on increasing the number of travelers who use the bus and rail modes. Also, improved access to intercity public transportation terminals from outlying areas (by charter or local bus systems, for example) might induce future travelers from these areas to alter their mode choices.

Because reported rail travel was a small share of total intercity travel, few policy implications concerning actual rail passenger travel in New York State can be drawn from the telephone survey. However, once the characteristics of corridor rail travelers are determined (from a train on-board survey), the total intercity travel market can be further analyzed to determine rail's penetration and market potential.

The analysis conducted points out the need for further investigation into issues of intercity travel market segmentation. A more detailed and in-depth survey of respondent's behavior and attitudes in this area is particularly warranted. Also, information on travelers' destinations, trip length, and repeat use would be extremely useful. The results obtained from this study should provide an adequate basis for future research efforts.

REFERENCES

1. R. Dobson and M. L. Tischer. Perceptual Market Segmentation Technique for Transportation Analysis. TRB, Transportation Research Record 673, 1978, pp. 145-152.
2. R. J. Zerrillo. An Analysis of Intercity Travel: Results of the New York State Rail Corridor Telephone Survey. Aviation and Rail Planning Unit, New York State Department of Transportation, Prelim. Statewide Planning Rept. 20, 1979.
3. A. J. Neveu. Public Opinion Survey on Energy and Transportation. Planning Research Unit, New York State Department of Transportation, Prelim. Res. Rept. 135, Dec. 1977.
4. M. J. Puma and R. B. Walters. Survey of the Attitudes of Intercity Automobile Travelers Toward Intercity Public Transportation. Applied Management Science, Inc., Silver Spring, MD, Dec. 1977.
5. J. A. Sonquist and J. N. Morgan. The Detection of Interaction Effects. Survey Research Center, Institute for Social Research, Univ. of Michigan, Ann Arbor, Monograph 35, 1970.

Publication of this paper sponsored by Committee on Statewide Multimodal Transportation Planning.