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 alternative (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.