OT-RSPA-DPB-50-78-2 AUGUST 1, 1978

MASTER

CONF-761278--

EMERGING TRANSPORTATION PLANNING METHODS

DEPARTMENT OF TRANSPORTATION
Office of University Research
Research and Special Programs Administration
Washington, D.C. 20590

Technical Report Documentation Page

1. Report No. DOT-RSPD-DPB-50-78-2	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle
PROCEEDINGS OF THE SEMINAR ON EMERGING TRANSPORTATION PLANNING METHODS, HELD DECEMBER 1976 IN DAYTONA BEACH, FLORIDA

5. Report Date
March 15, 1978

6. Performing Organization Code

7. Author(s)
Editor, William F. Brown

8. Performing Organization Report No.

9. Performing Organization Name and Address
Office of University Research
Research and Special Programs Directorate
U. S. Department of Transportation
Washington, D. C. 20590

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

12. Sponsoring Agency Name and Address (Jointly)
Technology Sharing Program and
Office of University Research
U. S. Department of Transportation
Washington, D. C. 20590

13. Type of Report and Period Covered
Seminar -
Held December 5-9, 1976

14. Sponsoring Agency Code

15. Supplementary Notes

PLANNING STAFF:	William F. Brown	Alfonso B. Linhares	Edward Weiner
	Robert B. Dial	Michael J. Rabins	Wilbur Williams
	David S. Gendell	Donald E. Ward	

16. Abstract

17. Key Words

18. Distribution Statement
Document is available to the U. S. Public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report) UNCLASSIFIED	20. Security Classif. (of this page) UNCLASSIFIED	21. No. of Pages	22. Price

Form DOT F 1700.7 (8-72)   Reproduction of completed page authorized

DOT-RSPA-DPB-50-78-2 August 1, 1978

EMERGING TRANSPORTATION PLANNING METHODS

A Compendium of Papers on Transportation
Demand Forecasting Techniques,
Transportation Evaluation Methods and
Transportation/Land Use Interactions.
Based on a Seminar held in Daytona Beach, Florida
in December 1976.

Editors: William F. Brown, Principal
Robert B. Dial
David S. Gendell
Edward Weiner

OFFICE OF UNIVERSITY RESEARCH
RESEARCH AND SPECIAL PROGRAMS ADMINISTRATION
U.S. DEPARTMENT OF TRANSPORTATION
WASHINGTON, D.C. 20590

CONTENTS


Seminar on Emerging Transportation Planning Methods
Jointly Sponsored by the Technology Sharing Program and the University Research Program
Office of the Secretary
U.S. Department of Transportation
Washington, D.C. 20590

FOREWORD

The U.S. Department of Transportation has sponsored research on new transportation planning techniques for several years. A number of significant advances have been made both in manual and computer procedures. Many of the new techniques were developed at universities and have distinct advantages over traditional methods. In 1976 the Office of University Research organized and conducted a 4-day seminar to instruct transportation planners in the theory, utility, and application of emerging transportation planning techniques and to encourage their widespread use. The planning committee for the seminar consisted of representatives from the Office of the Secretary, the Federal Highway Administration, the Transportation Systems Center, and the Urban Mass Transportation Administration. Three basic subjects were chosen for indepth examination:

1. Transportation demand forecasting techniques;
2. Transportation evaluation methods; and
3. Transportation/land use interactions.

One full day was devoted to each of these topics. Experts were commissioned to prepare papers in their respective subject areas. Two speakers presented each subject in state-of-the-art lectures. In addition, each day the participants were divided into small workshops which were moderated by discussion leaders. In evening plenary sessions, the participants further explored any unresolved questions and issues. Because of the free-flowing nature of the discussions, no attempt was made to keep a minute by minute account of the proceedings. This book contains nine papers: six state-of-the-art presentations and three papers summarizing the workshop discussions and question-and-answer periods. The book has been prepared in a format similar to the seminar; however, the papers have been carefully edited. Some explanations have been expanded so that the text may be instructive as well as valuable as a reference. The remainder of the "Foreword" describes the three major sections of the book.

The scope of the first section is to explore the theory and practice of disaggregate demand models in forecasting usage of various modes of urban transportation.

Daniel McFadden, Professor of Economics, University of California at Berkeley, wrote "The Theory and Practice of Disaggregate Demand Forecasting for Various Modes of Urban Transportation," as an introduction to disaggregate behavioral forecasting. The paper outlines the concepts underlying this approach and contrasts behavioral with conventional forecasting methods. It also describes practical applications of disaggregate forecasting and illustrates some early findings.

The second paper by Paul O. Roberts, Professor of Transportation at the Massachusetts Institute of Technology, is entitled, "Disaggregate Demand Modelling: Theoretical Tantalizer or Practical Problemsolver?" The paper describes disaggregate modelling as a pioneering behavioral approach by delineating its points of departure from, and advantages over, the more traditional aggregate approach to transportation demand forecasting. Based on the consumer choices of a person or family, this new, but not untried, method is described as more policy responsive, more flexible, and suitable to more applications than previous models. The paper encompasses the disaggregate nature of travel, the philosophical underpinnings of the disaggregate model, and the prediction framework for using the model.

The scope of the second section is to explore methods for developing and presenting evaluative information to the transportation decisionmaker.

Joseph L. Schofer, Professor of Civil Engineering at Northwestern University, covers transportation evaluation methods in his paper, "Evaluating Transportation Alternatives." Noting a lack of standard approaches, Professor Schofer focuses on achievable improvements in evaluation methodology. He presents the strategic issues that must be considered in developing a strong foundation for specific, successful evaluation tasks. He also prescribes some key steps that should lead the planner to better evaluation. Evaluation is defined as the technical process that links decisionmaking with analysis, planning, and design. Professor Schofer poses several questions that can define and measure the success of this process.

In the second paper on transportation evaluation methods, Thomas B. Deen, Chairman of the Board of Alan M. Voorhees & Associates, discusses "Practical Considerations in Transportation Decisionmaking." He contends that transportation planners face "devilish" problems; the main characteristic is the lack of consensus existing on either the nature of the problem or how to determine whether the problem has been solved. Asserting that transportation planners probably cannot satisfy all the demands of their constituency, he offers several suggestions about how transportation planners can conduct themselves, including: shedding illusions of finding any "best" solution; continuing to improve methodologies while investing at least as heavily in improving communication with officials and citizens; and refraining from the assumption that planners are more "pure" than politicians or officials.

To complement the coverage of transportation evaluation methods, the summaries of the six workshops are presented. The topics are: treating distributional effects of uncertainty, the evaluation of project plans, evaluation of regional plans and programs, preparation and interpretation of evaluation results, strategic approaches to evaluation, and evaluation and decisionmaking.

The scope of the third section is to explore the integrated forecasting of transportation and land use.

Stephen H. Putman, Associate Professor of City and Regional Planning at the University of Pennsylvania, prepared the paper, "The Integrated Forecasting of Transportation and Land Use." He discusses two closely related advances in operational planning techniques that, when put together, make integrated transportation forecasting and policy analysis a reasonable operational analysis technique. The first of these advances was the demonstration of both the feasibility and superiority of an integrated transportation/land use model package; the second was the development of a more general form of urban land use model along with the procedures necessary for its calibration.

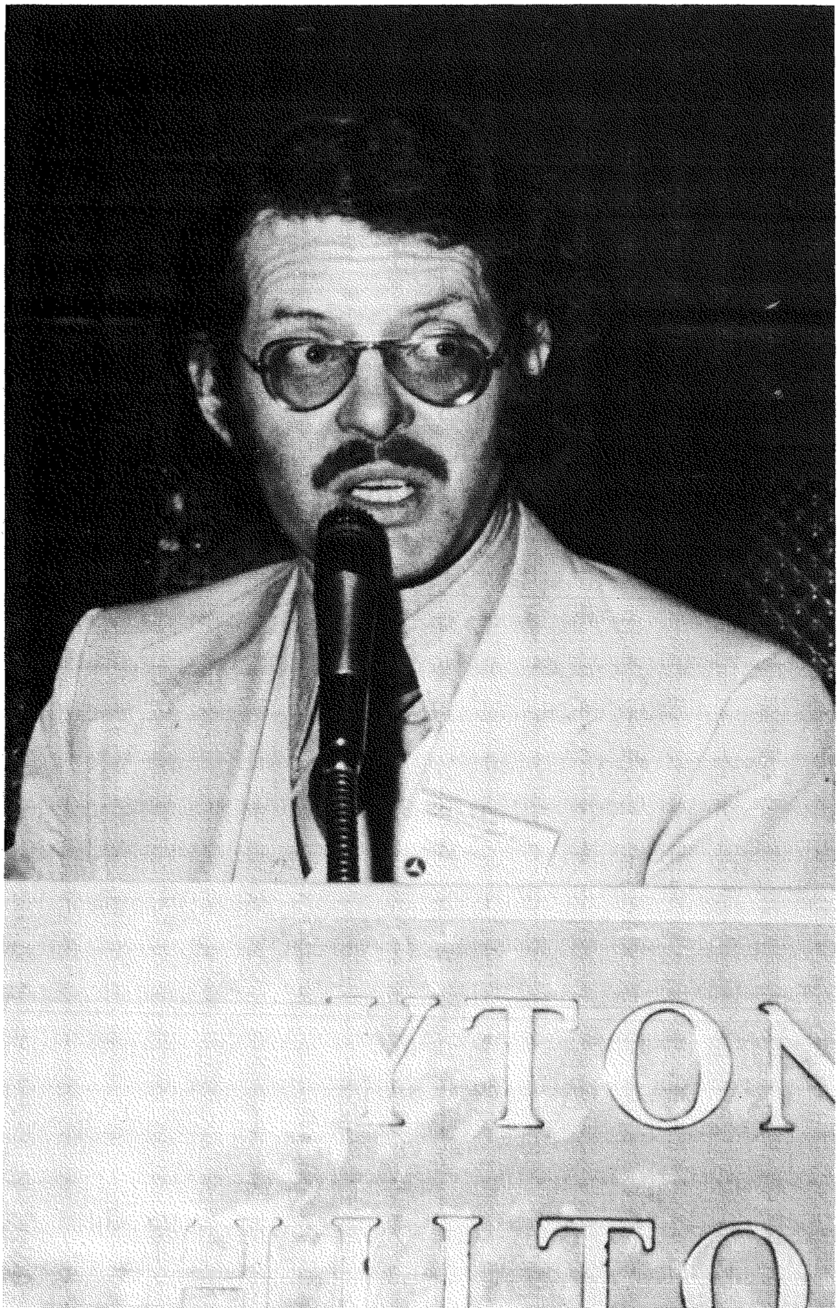
The second paper on transportation/land use interactions was written by Douglass Lee, Jr., Associate Professor at the University of Iowa, and is entitled, "Improving Communications Among Researchers and Planners in the Transportation and Land Use Field." This paper concerns primarily communication between and among professionals and researchers, as well as communication between the technical and political sides of transportation and land use. The report briefly documents discussions at two workshops that were based on case studies illustrative of how political decisionmaking takes place. The actual participants of the two case studies—including politicians—were involved in the workshop discussions. The case studies themselves, which are not presented here, concern the I–66/Metro Corridor and the Mt. Hood Freeway decision.

The third and final paper summarizes six workshop sessions: land use modelling, decisionmaking, and politics; land use/transportation modelling structures; application of land use models in comprehensive planning; short-range forecasting and land use impacts; land use/transportation forecasting at the community scale; and details of the integrated transportation/land use package.

The Seminar on Emerging Transportation Planning Methods, as judged by the participants, was very successful. This success, plus the long term value of the papers, led to the decision to publish this book. The Office of University Research, in order to effectively disseminate research results, plans to hold more seminars of this type in the future.

William F. Brown
Deputy Director
Office of University Research

I. TRANSPORTATION DEMAND FORECASTING TECHNIQUES



Daniel L. McFadden, Professor of Economics
University of California at Berkeley

The Theory and Practice of Disaggregate Demand Forecasting for Various Modes of Urban Transportation

Daniel L. McFadden

Professor of Economics
Massachusetts Institute of Technology

A major responsibility of transportation planners is to forecast those changes in travel demand induced by alternative transportation policies. In recent years, the range of analyzed policy alternatives and the range of considered policy questions have greatly expanded. Emphasis has shifted from long-run planning of highway networks to short-run planning and to management of integrated multimodal transportation systems. These shifts have placed considerable strain on conventional forecasting tools, which were originally designed to address problems of highway network design.

Flexible demand forecasting methods have consequently been sought, particularly those capable of incorporating the behavioral forces linking individual transportation decisions. The resulting behavioral disaggregate methods expand the policy sensitivity of forecasts. Tests and practical experience with these methods indicate that they are comparable or superior to conventional forecasting techniques in terms of data gathering and computational requirements and forecast accuracy. They provide, in short, a useful way of tackling the expanded list of contemporary planning questions.

Most conventional forecasting models were originally developed to address problems of highway design, and were conceived using analogies with physical systems—with traffic flows described in terms of hydraulic or gravity flow models. Different model components in conventional modeling are not developed from a unified framework. For example, a trip generation model may be developed quite independently of a model of modal split. Another deficiency of conventional models is that they often involve costly and time consuming data gathering and computational requirements, and they are not easily adapted to short-run planning and transportation system management. In particular, they are poorly adapted to pencil-and-paper or quick-response policy analyses planners need.

In contrast to conventional methods, disaggregate behavioral forecasting methods are based on a unified conceptual framework. They start from the idea that all travel demand is generated by individual choice behavior, and more specifically (in the current generation of disaggregate behavior models) generated by maximization of preferences or utility. An advantage of disaggregate forecasting is that it is not based on one model; it is an approach or system for building models, and as such it can provide the planner with a method of dealing with a variety of problems as they occur. It is possible to build complex disaggregate behavioral model systems on a scale approaching or exceeding that of conventional models. On the other hand, it is possible to use these techniques to do "quick-and dirty" planning, using "back-of-the-envelope" calculations, without extensive data collection requirements. In general, the use of behavioral models greatly conserves data collection costs relative to conventional models in both the calibration phase and the forecasting phase. A major advantage of disaggregate models is that they allow the planner to address questions, such as the demand for a new mode, which are difficult to answer in a conventional framework. The current generation of disaggregate models have accuracies comparable to or better than that of conventional models. Disaggregate models have proved practical and successful in a number of applications. I emphasize that the state-of-the-art of disaggregate modelling is evolving rapidly; current models are not the final answer, and have some undesirable features. There are many unexpected characteristics of disaggregate models, and many uncharted pitfalls for the user. Disaggregate models are valuable now for solving some planning problems. In the future, as better disaggregate models evolve, the list of effective applications will grow.

The rich, poor, healthy, and handicapped are rarely homogenous and the aggregate forecasting methods which treat them as such make a specification error. More importantly, these methods preclude the possibility of answering questions such as who benefits and who pays for policy changes. These shortcomings are frequently corrected in part by segmenting the zone population by income class in conventional models. Further segmentation by those socioeconomic characteristics other than income that influence travel patterns would be useful. Conventional calibration of an aggregate model for numerous market segments requires an often unobtainable quantity of data. Pursued to a logical conclusion, each segmented market in an aggregate model should contain a sub-population with identical socioeconomic characteristics and identical transportation environments. This segmentation would amount in practice to distinguishing *each individual* as a "market segment." Aggregate forecasts would then be regarded as the sum of the travel demand of *individuals*, which is a disaggregate forecasting procedure. Disaggregate demand modelling is, then, essentially market segmentation carried to an extreme and is one end of a continuum, with aggregate demand forecasting at the opposite extreme. Consider for example, the mode split for work trips from an origin zone to a destination zone. The aggregate share of a mode is by definition the sum over market segments of the share of the mode in each market segment, weighted by the proportion of the total origin-zone population contained in this market segment. If the segmentation is complete, then one has the formula shown below, with each homogenous market segment having a share for the particular mode. The aggregate share is the weighted average of the shares in the homogenous market segments.

$$
\begin{aligned}
\begin{bmatrix}\text{Aggregate}\\ \text{Share of}\\ \text{A Mode}\end{bmatrix} &= \begin{bmatrix}\text{Share of Mode}\\ \text{in First Market}\\ \text{Segment}\end{bmatrix} \times \begin{bmatrix}\text{Proportion of First}\\ \text{Market Segment in}\\ \text{Population}\end{bmatrix}\\
&+ \begin{bmatrix}\text{Share of Mode}\\ \text{in Second}\\ \text{Market Segment}\end{bmatrix} \times \begin{bmatrix}\text{Proportion of Second}\\ \text{Market Segment in}\\ \text{Population}\end{bmatrix} \qquad (1)\\
&+ \ldots + \begin{bmatrix}\text{Share of Mode}\\ \text{in Last Market}\\ \text{Segment}\end{bmatrix} \times \begin{bmatrix}\text{Proportion of Last}\\ \text{Market Segment in}\\ \text{Population}\end{bmatrix}
\end{aligned}
$$

This formula is one that will recur several times.

An axiom of behavioral disaggregate, choice theory is that the individual is the basic decisionmaking unit, choosing from available alternatives the most desirable. The desirability—or *utility*—of a choice depends upon its attributes and upon the characteristics of the individual. Suitably modified to take account of the psychological phenomena of learning and perception errors, this theory has been used widely and successfully in analyzing and forecasting economic consumer behavior, of which transportation behavior can be viewed a part.

Let us first clarify what transportation behavior is. A complete definition of a transportation alternative for an individual includes the total pattern of travel: location of residence and job; purchases of vehicles; frequency of work, shopping, personal business, recreation and other trips; destination of trips; scheduling of trips; mode choice; and route choice. In practice, travel demand models concentrate on certain dimensions of travel behavior such as mode choice, taking as given other aspects such as scheduling of trips or location of residence. (A great deal of the *behavioral* theory of disaggregate modelling which will not be presented explicitly here deals with how these decisions can be broken apart.)

An alternative's attributes include the transportation level of service variables associated with its pattern of travel. The individual's utility of an alternative is a function of level-of-service variables for the alternative. Utility also depends on the individual's tastes and background—or socioeconomic characteristics. Examples of level-of-service variables are travel time and travel cost. Examples of socioeconomic characteristics are income and family size. An individual chooses among the available alternatives the one which maximizes utility.

Some socioeconomic characteristics and level-of-service variables are observed by the transportation planner. Others are unobserved. For example, income and in-vehicle travel time are usually observed or calculated, while attitudes towards privacy or vehicle noise level are usually not observed.

Consider a group of individuals with similar observed backgrounds and decision environments, characteristics, and observed level-of-service variables for the alternatives. This could be called a *homogeneous market segment*. The frequency of choice for an alternative within a homogeneous market segment is determined by the number of members of this group whose unobserved level-of-service and socioeconomic variables, operating in tandem with the observed variables, give this alternative the highest utility. For example, if an individual's observed travel times on alternative modes, in combination with unobserved attitudes towards privacy, lead him to a higher utility for bus than for auto, then he will choose the bus. Other people with the same observed travel time, and therefore in the same homogeneous market segment, may have different attitudes towards privacy, and as a result may take the auto.

A disaggregate choice model is defined by specifying a probability distribution of the unobserved variables affecting utility, given the values of observed variables in a homogeneous market segment. This probability distribution then determines the *choice probabilities*—the proportions of the group with maximum utility for each alternative.

In summary, a disaggregate behavioral model is specified by forming a concrete individual utility function, a probability distribution of the unobserved variables, and a share of each market segment in the population. Examples of specific utility functions and probability distributions are given below. Using the formula in equation (1), once a concrete utility function is formed and the distribution of unobserved variables specified, each of the shares in a homogeneous market segment is specified. Knowing the proportions of the population in the various market segments, one can then compute the average share.

I will define the *mean utility* of a homogeneous market segment to be the average of the utilities of all the individuals in this segment. Mean utility depends on the observed level-of-service and socioeconomic variables, and on other determinants of the *distribution* of unobserved variables.

Assuming a concrete probability distribution for the unobserved components of utility leads to a concrete formula for the choice probability. Unfortunately, most distributions of unobserved components yield computationally forbidding choice probability *formulae,* making them difficult to use in practical calibrations and forecasting. One exception is the multinomial logit model, which has choice probabilities of the form shown below. ("Exp" denotes exponentiation.)

$$\begin{bmatrix}\text{Share of}\\ \text{the i-th}\\ \text{Alternative}\end{bmatrix} = \frac{\exp\begin{bmatrix}\text{mean utility of}\\ \text{i-th } \textit{alternative}\end{bmatrix}}{\exp\begin{bmatrix}\text{mean utility of}\\ \text{the}\\ \textit{first alternative}\end{bmatrix} + \ldots + \exp\begin{bmatrix}\text{mean utility of}\\ \text{the}\\ \text{last alternative}\end{bmatrix}} \tag{2}$$

The multinomial logit model has the following characteristics: first, it can be interpreted as a disaggregate behavioral model with special assumptions on the probability distribution of the unobserved variables which will not be detailed here. Second, a multitude of possible disaggregate travel demand models can be formulated in the multinomial logit framework, with the form of the mean utility function depending on the application. Third, the multinomial logit model has the mathematical form of share models used in conventional travel demand forecasting systems, such as the gravity or intervening opportunity models. For example, consider a singly constrained aggregate gravity model for distribution,

$$N_{kj} = O_k A_j / T_{kj}^{\,h} \,, \tag{3}$$

where N_{kj} = number of trips from zone k to zone j;
A_j = attraction of zone j;
T_{kj} = impedance between k and j;
O_k = scale factor to equate trips distributed from zone k to trips originating in zone k.

Then, the *share* of trips from zone k to zone i satisfies

$$P(i|T_{k1}, \ldots, T_{kj}, A_j) = \frac{A_i/T_{ki}^{h}}{\sum_{j=1}^{J} A_j/T_{kj}^{h}}. \tag{4}$$

This is a multinomial logit functional form in equation (2) with mean utility = log A_i − h log T_{ki}. Hence, the multinomial logit form is not new to planners, but has been widely used in one form or the other, although perhaps not widely recognized. As the example makes clear, the multinomial logit form can be used in ways which are quite different in motivation than the principles of disaggregate behavioral theory. In the special case of two alternative modes, the multinomial logit model is termed the (binomial) logit mode split model. This case gives a response curve to a type familiar to every planner in which the share of a particular mode is plotted against the relative desirability of the modes, as in Figure 1. If desirability is measured in terms of relative impedance or more generally relative disutility, standard mode split models can be interpreted as behavioral models.

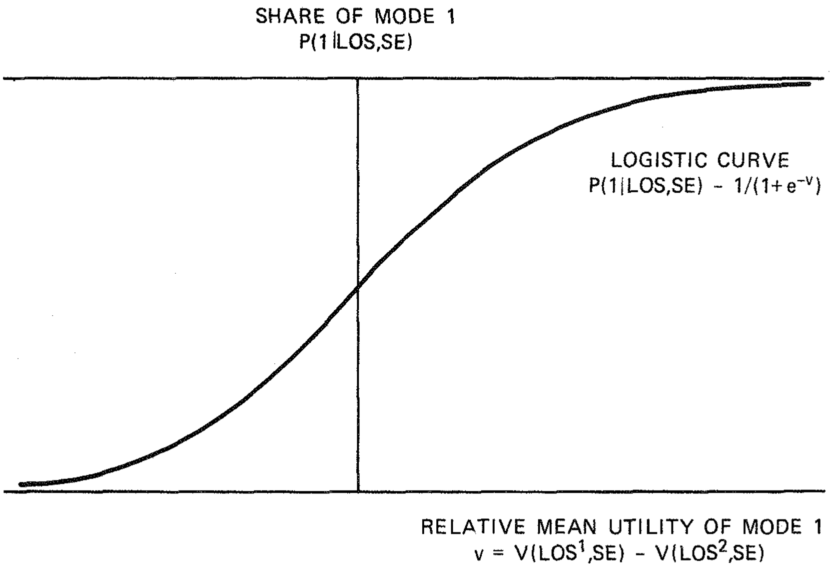


Figure 1.—Binary logit response curve.

What then are the primary differences between traditional aggregate share models and the multinomial logit disaggregate models? First, the structure of the mean utility function in the multinomial logit model is based on economic and psychological regularities in individual behavior. As a consequence it will have a similar form in models of different aspects of transportation choice such as generation, scheduling, distribution, and mode split. For example, if one can determine the variables that matter in mode split, it should be the case that similar variables matter in trip distribution. Second, the calibration and utilization of the model are carried out at the disaggregate level for homogeneous market segments rather than applied to aggregate data.

A successful forecasting model, behavior disaggregate or otherwise, must assess correctly the impact of level-of-service changes on demand. This requires in calibration that the effects on demand of variations in level-of-service be sorted out from the effects of non-transportation variables. For example, suppose large families with small children locate disproportionately in the suburbs where walk time to transit is high, and workers in large families are disposed to transit because of competing needs of household automobiles. Then a mode split model which fails to control family size and attributes the pattern of transit usage to variations in walk time will understate the onerousness of walk time and yield faulty predictions of the impact on transit patronage of policies influencing transit walk time. The problem is corrected by including family size as an explanatory variable in the model.

Disaggregate calibration methods allow inclusion of a more extensive list of level-of-service and socioeconomic variables than do most aggregate methods, improving the possibility of untangling the effects of level-of-service and other variables. It should be noted, however, that it is possible to develop simple disaggregate models using only conventional variables familiar to planners, such as travel time and travel cost. Empirical tests suggest that the introduction of variables other than conventional components of impedance in a disaggregate mode choice model may improve only marginally the ability of the model to "explain" observed choices in a calibration data base, but may significantly improve forecasting accuracy.

It should be emphasized that the disaggregate behavioral approach is a systems approach to modelling, not a specific model. Disaggregate models can be developed to meet the specific needs of the individual planner. In particular, one can build disaggregate models which are completely analogous to conventional models in terms of data used and types of variables employed, such as travel time and cost. Alternately, one can expand on these models by expanding the description of level-of-service attributes, thereby increasing the ability of the models to be responsive to expanded policy questions. Or, one can expand the socioeconomic description of these models to take into account correlations between level-of-service variables and socioeconomic variables which may have been leading planners to spuriously impute impacts to level-of-service variables. Finally, even though when one thinks of forecasting models, one usually thinks in terms of concrete and well-defined variables such as travel time and travel cost which can themselves be forecast from networks under alternative policy scenarios, it is also possible to develop models which depend on survey data on perceptions or attitudes. Although a distinction is sometimes made between attitudinal models and behavioral models, the disaggregate systems approach to building models incorporates both.

Disaggregate models are relatively parsimonious in terms of data requirements. A typical mode choice model, for example, can be calibrated on a sample size of 300 to 3,000 individuals with quite tolerable levels of accuracy. Socioeconomic variables are normally available at an individual level from household surveys. Transportation level-of-service variables are much harder to provide at the level of the individual traveler. Typically these data are obtained from transportation networks, which can provide data only at a traffic analysis zone level. Studies have shown that it is usually reasonable to approximate level-of-service variables for the individual by zonal averages. One exception is walk time to transit, where significant improvements in forecast accuracy can be obtained by segmenting zones geographically and recording individual walking distances. A final point is that the accuracy of any model, conventional or disaggregate, which uses network data is limited in forecasting accuracy by the accuracy of the network. There are many subjective elements and assumptions that go into coding of networks, and one has to be careful in applying forecasting models to understand how these assumptions interact with model calibration.

In principle, the mean utility function in a disaggregate behavioral multinomial logit model can be a very complex function of personal characteristics and level-of-service variables. In existing practical models, however, these variables have appeared in a simple form, usually linearly, with an "importance weight" attached to each variable. For example, mean utility might equal the negative of the sum of travel time and travel cost deflated by wage, each weighted by an importance weight. It may be useful to describe how such a specification can be related to an underlying theory of individual behavior. I will outline a very simple model which provides this special structure. Let us assume that the utility or desirability of an alternative depends on the amount of goods an individual consumes; the amount of leisure he has available; the hours spent traveling (a "bad" rather than a "good" for most people); amenities of various travel destinations; and unobserved factors. The alternatives available to the individual in this model are destination and mode choice, so this is a joint destination and mode choice model. The option of no trip is included as an alternative, so that the model includes trip generation as well.

Each individual is assumed to be constrained by a budget which requires his total expenditures, equal to expenditure on goods plus expenditure on travel, be equal to his total income, which in turn equals wage income plus other income. Time is allocated between leisure, labor, and travel in a way to maximize utility. First, for any travel alternative a mix of labor and leisure is chosen to maximize utility. If the individual considers the alternative of taking the bus to a particular shopping destination, then the optimal amount of time worked, adjusted to take into account the choice of this alternative, will be determined. At this optimal mix, the marginal utility of goods (defined to be the amount of additional utility obtained from one additional unit of goods) multiplied by the wage rate equals the margin utility of leisure. Second, the travel alternative actually chosen is the one maximizing utility, taking into account the labor-leisure adjustment above for each alternative. The features of this model are summarized in Table 1.

The preceding argument provides a justification—from the economic theory of utility maximizing behavior—for the entry of travel time and travel cost divided by wage as linear variables in the mean utility function. Generalization of this model is possible in several directions.

Time, cost, and other attributes of alternatives may have subcomponents. Time, for example, can be partitioned into on-vehicle time under congested or non-congested conditions, walk time, and wait time. Costs can be divided into overhead, indirectly charged per-trip costs such as fuel and maintenance, and daily out-of-pocket costs such as tolls. These components can be given separate coefficients in equation (*) of the preceding table; the relative weights of components can then be determined as part of the calibration of the model, which is preferable to assigning traditional weights.

The coefficients b_T, b_C, and b_A may depend on observed socioeconomic variables. For example, the weight b_T associated with the walk time component of travel time may be a function of an individual's age and health status, or of those neighborhood characteristics correlated with safety. If this association is expressed in a linear-in-parameters form, then the mean utility function (*) is linear

TABLE 1. A Simple Behavioral Utility Model

- Utility depends on goods, leisure, hours spent traveling, amenities at various travel destinations, unobserved factors
- Alternatives describe destination and mode choice, including the no trip option
- Each individual is constrained by a budget:

$$\begin{matrix}\text{Expenditure}\\ \text{On Goods}\end{matrix} + \begin{matrix}\text{Travel}\\ \text{Cost}\end{matrix} = \begin{matrix}\text{Wage}\\ \text{Income}\end{matrix} + \begin{matrix}\text{Other}\\ \text{Income}\end{matrix} \qquad (5)$$

- Time is allocated between leisure, labor, and travel
- For any travel alternative, the mix of labor and leisure is chosen to maximize utility. At this mix, the marginal utility of goods, multiplied by the wage rate, equals the marginal utility of leisure
- The chosen alternative maximizes utility

$$\begin{bmatrix}\text{UTILITY}\\ \text{OF } i^{th}\\ \text{ALTERNATIVE}\end{bmatrix} = -\left[\begin{bmatrix}\text{TRAVEL}\\ \text{TIME}\end{bmatrix} + \frac{[\text{TRAVEL COST}]}{[\text{WAGE}]}\right] \times \begin{bmatrix}\text{MARGINAL}\\ \text{UTILITY}\\ \text{OF LEISURE}\end{bmatrix}$$

$$+ \begin{bmatrix}\text{TRAVEL}\\ \text{TIME}\end{bmatrix} \times \begin{bmatrix}\text{MARGINAL UTILITY}\\ \text{OF TRAVEL TIME}\end{bmatrix}$$

$$+ [\text{AMENITIES}] \times \begin{bmatrix}\text{MARGINAL UTILITY}\\ \text{OF AMENITIES}\end{bmatrix}$$

$$+ \begin{bmatrix}\text{UNOBSERVED}\\ \text{ATTRIBUTES}\end{bmatrix} \times \begin{bmatrix}\text{MARGINAL UTILITY OF}\\ \text{UNOBSERVED ATTRIBUTES}\end{bmatrix}$$

$$\begin{bmatrix}\text{MEAN UTILITY}\\ \text{OF THE } i^{th}\\ \text{ALTERNATIVE}\end{bmatrix} = -b_T \times \begin{bmatrix}\text{TRAVEL}\\ \text{TIME}\end{bmatrix} - b_C \times \frac{[\text{TRAVEL COST}]}{[\text{WAGE}]} \qquad (*)$$

$$+ b_A \times [\text{AMENITIES}]$$

b_T, b_C, AND b_A ARE PARAMETERS

in these parameters, and the calibrated model will describe both the importance weight attached to walk time and the variation of this weight with socioeconomic factors.

There are a number of methods available to calibrate disaggregated behavioral multinomial logit models. The technique which is most commonly used is maximum likelihood estimation. From the user's point of view, this method is comparable to regression analysis—the inputs and outputs of computer programs which carry out this calibration resemble closely the inputs and outputs of regression programs, and require the same skills from the user as do regression analyses. Therefore, any planning organization which currently has the capacity to do regression analyses also has the potential ability to calibrate multinomial logit models.

There are good statistical computer programs available for multinomial logit analyses using the maximum likelihood method. One available to many planners is the ULOGIT programs in the UTPS package. There are several other stand alone logit programs available with options not included in ULOGIT. QUAIL, a flexible data management and multinomial program developed by McFadden and his colleagues, is available in versions suitable for use on CDC or IBM machines. Multinomial logit programs for IBM machines are also available from Cambridge Systematics, Inc. and from Charles Manski at Hebrew University. All these programs are available at the cost of reproducing tapes and manuals.

In addition to maximum likelihood estimation, there are several other techniques for fitting multinomial logit models. One technique, currently available only on QUAIL, is non-linear least squares. This method has an advantage relative to maximum likelihood estimation in that it is less sensitive to data measurement errors, an important consideration given the nature of transportation data. Finally, there is an estimation technique called the Berkson-Theil method which requires grouped data rather than individual observations. If data is collected by individual, it must be grouped to use this method. On the other hand, the method requires only a standard regression program, and hence is readily available to most planners. When data can be grouped easily, the Berkson-Theil procedure is recommended. It has good statistical properties, and is considerably less expensive than maximum likelihood estimation.

Let us next consider a simple calibrated disaggregate multinomial logit model with work trip mode choice. The model in Table 2 was calibrated by the maximum likelihood technique on a sample of 771 commuters in the San Francisco Bay Area in 1973, before the inauguration of BART Trans-Bay service. The explanatory variables in this model are the level-of-service attributes commonly used to define impedance in conventional models, on-vehicle travel time, excess or out-of-vehicle time, and cost divided by wage. The model contains four alternatives: auto drive alone, auto shared with someone else (either family or non-family carpool), and bus, subdivided by access mode. Auto access to bus includes "kiss-ride" and "park-ride." Alternative-specific dummy variables are introduced to capture the average influence of unobserved attributes of each mode. The number of dummy variables is one less than the number of alternatives, as the coefficient of the bus-with-walk-access dummy is normalized to zero. One such arbitrary normalization is necessary.

A negative coefficient for a variable indicates that an increase in this variable for a mode will lower the mode's choice probability. For example, the coefficient of excess time is negative. If excess time rises for a particular mode—say, bus-with-walk-access—then the mean utility of this mode will fall, and as a consequence the choice probability for this mode in a homogeneous market segment will fall. The T-statistics on the right-hand-side are indicators of the precision of the parameters. Values less than two indicate that the parameters cannot be reliably distinguished between zero. This particular model indicates that individuals react strongly to transportation level-of-service variables. The average effects of unobserved variables, reflected in the coefficients of the alternative-specific dummy variables, are important.

TABLE 2. A Simple Work Trip Mode Choice Model, Estimated Pre-BART

(Mode 1—Auto Alone; Mode 2—Bus, Walk Access; Mode 3—Bus, Auto Access; Mode 4—Carpool)

Model: Multinomial Logit, Fitted by the Maximum Likelihood Method

(The Variable takes the described value in the alternatives listed in parentheses and zero in non-listed alternatives)

Independent Variable	*Estimated Coefficient*	*T-Statistic*
Generic		
Cost divided by post-tax wage, in cents divided by cents per minute (1–4)	– .0412	7.63
On-vehicle time, in minutes (1–4)	– .0201	2.78
Excess time, in minutes (1–4)	– .0531	7.54

TABLE 2. A Simple Work Trip Mode Choice Model, Estimated Pre-BART—*(Continued)*

Specific		
Auto alone dummy (1)	− .892	3.38
Bus with auto access dummy (3)	−1.78	7.52
Carpool alternative dummy (4)	−2.15	8.56

Likelihood ratio index	.1499
Log likelihood at zero	−1069.0
Log likelihood at convergence	− 717.7
Percent correctly predicted (by maximum probability)	58.50 (compared with 39.42 by chance)

Value of time saved as a percent of wage (t-statistics in parentheses):

On vehicle time	49 (2.68)
Excess time	129 (5.16)

All cost and time variables are calculated round-trip. Excess time is defined as the sum of walk time, transfer wait time, and half of initial headways. Dependent variable is alternative choice (one for chosen alternative, zero otherwise).

Number of people in sample who chose

Auto alone	429
Bus with walk access	134
Bus with auto access	30
Carpool	178
Total sample size	771

I will expand further on the nature of the variables entering this model, and specifically on the alternative-specific dummy variables. Socioeconomic variables which influence the mean utility of every alternative in exactly the same way have no influence on choice probabilities. They change both the numerator and the denominator of the multinomial logit formulae by a factor which cancels out. Hence, there is interest only in those socioeconomic variables which interact with level-of-service variables to affect the mean utility of different alternatives differently. For example, income can matter only if, when income changes, it increases the attractiveness of one of the alternatives relative to a second. Travel cost divided by wage is one example of interaction. A second example is a variable which takes the value of one for an alternative which requires driving a vehicle when the individual has a driver's license, and is zero otherwise. The variable in this example is the product of a socioeconomic variable which is one if the individual can drive and zero otherwise, and a level-of-service variable which is one if the alternative requires driving and zero otherwise. In the model in Table 2, an alternative-specific dummy variable for an alternative is one for this alternative and zero for all other alternatives. Mean utility may be included in alternative-specific dummy variables appearing alone, or in interaction with other variables. The coefficient of an alternative-specific dummy variable can be interpreted as reflecting the impacts of an alternative's unmeasured level-of-service attributes that are not captured in the remaining variables. For example, the auto-alone dummy variable is one for the auto-alone mode and zero otherwise. (The number following the name of the variable indicates for which alternatives it is non-zero.) The coefficient −.892 can be interpreted as representing the average impact of unmeasured characteristics of the auto alternative relative to the bus-with-walk-access alternative.

A variable which is the result of interaction between an alternative-specific dummy variable and another variable is termed an alternative-specific variable. An example of an alternative-specific variable would be one which gives the value of out-of-vehicle travel time for the bus with auto access alternative and zero for all other alternatives. The coefficient of this variable compared with the coefficients of other alternative-specific travel times would reflect the impact of specific attributes of auto-accessed transit on the onerousness of transit travel time. A generic, or homogeneous-effect, variable is one which does not incorporate interaction with alternative-specific dummy variables. An example is a variable which gives out-of-vehicle travel time for each alternative, uninfluenced by the name of the alternative; i.e., an out-of-vehicle time of fifteen minutes is treated the same whether it is auto access time or transit wait time. In this model the level-of-service variable—cost, on-vehicle travel time, and access time—were all generic or homogeneous-effect. Each of these variables has values for each of the four alternatives. For example, travel time in auto has the same importance weight as travel time in transit.

Individual utility, expressed as a function of observed and unobserved variables, should depend on only generic variables. The reason for this is behavioral—individual utility depends on the constellation of physical experience associated with an alternative, and cannot depend on labels such as "auto," "transit," or "CBD"—attached to alternatives by the planner. Mean utility on the other hand may depend on alternative-specific variables which mimic or act as proxies for the influence of unobserved generic variables. For example, suppose individual utility depends on generic on-vehicle travel time weighted by a generic index of comfort. Suppose the comfort index is unobserved, but varies between alternatives. Then the mean utility for an alternative will have a coefficient of on-vehicle time which reflects the average comfort index on this alternative. It will then appear to the planner that mean utility depends on alternative-specific travel times. Alternative-specific variables in a multinomial logit model are evidence of failure to observe generic variables which are influencing behavior. A long-run objective of behavioral demand analyses is to improve model specification and data collection to the point where alternative-specific variables are not needed. Models based solely on generic variables are also desirable from the point of view of forecasting. Coefficients of alternative-specific variables do not isolate behavioral sources of variation across alternatives, or establish that alternative-specific effects will be stable or extendable to new situations when forecasting. In the current state-of-the-art of disaggregate demand analyses, alternative-specific effects do capture the impacts of variables not observed in standard transportation data sets; their omission would bias the importance weights associated with other variables.

In the lower half of Table 2 are several summary statistics which give some notion of the goodness-of-fit of this model to the calibration data base. The likelihood ratio index is an analog of the multiple correlation coefficient in regression analysis. Empirically, its values run lower than typical values for a multiple correlation coefficient. A value of .2 to .3 indicates a good fit. A second measure of goodness-of-fit is the ability of the model to forecast accurately. In this particular sample, 39% of the choices of individuals would be predicted correctly by change, whereas the model predicts 58% correctly. A third method commonly used to assess the merit of models is to compute the implicit values of time implied by the model. This is a potentially misleading measure of goodness-of-fit, both because these statistics tend to be very unreliable and because there is some tendency to accept or reject models on the basis of consistency with earlier result in the literature, which could perpetuate errors in the assessment of time evaluations. On the other hand, the critical role of value of time tradeoffs in policy applications makes it necessary to compute these values. Value of time calculations in the multinomial logit model are determined from the ratio of time and cost coefficients. These calculations assume that within a homogeneous market segment, the value of time is uniform. Note that this is not necessarily a good assumption. For the model in Table 2, on-vehicle time is valued at half the wage rate and access time at 130% of the wage rate. Table 3 describes a more complex multinomial logit modal split model.

TABLE 3. Work Trip Mode Choice Model, Estimated Pre-BART

(Mode 1—Auto Alone; Mode 2—Bus, Walk Access; Mode 3—Bus, Auto Access; Mode 4—Carpool)

Model: Multinomial Logit, Fitted by the Maximum Likelihood Method

(The Variable takes the described value in the alternatives listed in parentheses and zero in non-listed alternatives)

Independent Variable	*Estimated Coefficient*	*T-Statistic*
Cost divided by post-tax wage, in cents divided by cents per minute (1–4)	— .0284	4.31
Auto-on-vehicle time, in minutes (1,3,4)	— .0644	5.65
Transit on-vehicle time, in minutes (2,3)	— .0259	2.94
Walk time, in minutes (2,3)	— .0689	5.28
Transfer wait time, in minutes (2,3)	— .0538	2.30
Number of transfers (2,3)	— .105	0.776
Headway of first bus, in minutes (2,3)	— .0318	3.18
Family income with ceiling of $7,500, in $ per year (1)	— .00000454	0.0511
Family income minus $7,500 with floor of $0 and ceiling of $3,000, in $ per year (1)	— .0000572	0.430
Family income minus $10,500 with floor of $0 and ceiling of $5,000, in $ per year (1)	— .0000543	0.907
Number of persons in household who can drive (1)	—1.02	4.81
Number of persons in household who can drive (3)	— .990	3.29
Number of persons in household who can drive (4)	— .872	4.25
Dummy if person is head of household (1)	— .627	3.37
Employment density at work location (1)	— .00160	2.27
Home location in or near CBD (2 = in CBD, 1 = near CBD, 0 otherwise) (1)	— .502	4.18
Autos per driver with a ceiling of one (1)	5.00	9.65
Autos per driver with a ceiling of one (3)	2.33	2.74
Autos per driver with a ceiling of one (4)	2.38	5.28
Auto alone alternative dummy (1)	—5.26	5.93
Bus with auto access dummy (3)	—5.49	5.33
Carpool alternative dummy (4)	—3.84	6.36

TABLE 3. Work Trip Mode Choice Model, Estimated Pre-BART—*(Continued)*

Likelihood ratio index	.294
Log likelihood at zero	−1069.0
Log likelihood at convergence	− 595.8
Percent correctly predicted (by maximum probability)	67.83 (compared with 39.42 by chance)

Values of time saved as a percent of wage (t-statistics in parentheses):

Auto on-vehicle time	227 (3.20)
Transit on-vehicle time	91 (2.43)
Walk time	243 (3.10)
Transfer wait time	190 (2.01)

Value of initial headways as a percent of wage: 112 (2.49)

All cost and time variables are calculated round-trip. Dependent variable is alternative choice (one for chosen alternative, zero otherwise).

Number of people in sample who chose

Auto alone	429
Bus with walk access	134
Bus with auto access	30
Carpool	178
Total sample size	771

One way of judging the effectiveness or the accuracy of a disaggregate demand model is to compute what is called a prediction success table. Table 4 is a prediction success table for the model in Table 3. Each column corresponds to a predicted alternative and each row corresponds to an actual choice. The number 296.6, for example, is the number of persons who were predicted to take auto alone who did in fact choose this alternative, and 29.0—the next number below it—is the number predicted to take auto alone who in fact took bus with walk access. Predictions in this table are based on the choice probabilities of individuals. For example, the entry 29.0 is the sum of the predicted choice probabilities of auto alone, taken over the set of all individuals who actually chose bus-with-walk-access. This prediction success table summarizes goodness-of-fit of the model to its calibration data base. This table has the property that the average observed shares (56% auto, 17% bus/walk, 4% bus/auto, and 23% carpool in this sample) coincide with the predicted values. This is a consequence of calibration, and says nothing about the accuracy of the model. A notion of how well the model fits is obtained by looking at the percent correctly predicted

TABLE 4. Prediction Success Table for Pre-BART Model and Calibration Data Base

Actual Alternatives	*(1) Auto Alone*	*(2) Bus/ Walk*	*(3) Bus/ Auto*	*(4) Carpool*	*Row Total*	*Observed Share (%)*
	Predicted Alternatives					
(1) Auto alone	296.6	29.4	10.0	93.1	429.0	56
(2) Bus/Walk	29.0	75.1	6.6	23.3	134.0	17
(3) Bus/Auto	9.8	5.9	6.7	7.6	30.0	4
(4) Carpool	93.6	23.7	6.7	54.0	178.0	23
Column Total	429.0	134.0	30.0	178.0	771	100
Predicted Share (%)	56	17	4	23	100	
Percent Correct	69.1	56.1	22.3	30.4	56.0	
Success Index	1.23	3.30	5.58	1.32		

The equality of predicted and observed shares is a conseqence of the calibration process.

in aggregate for each alternative. For auto alone, 69% of our predictions are correct, while for the bus with auto alternative, only 22% are predicted correctly. These figures illustrate that it is much easier to be successful when you are predicting demand for a highly used mode than when you are predicting demand for a little used mode. This observation applies throughout travel demand modeling, including conventional models. An index of prediction accuracy for an alternative can be obtained by dividing the percent correctly predicted by the percent correct you could achieve by chance. The higher this prediction success index, the better the model. In terms of the prediction success index, the model in Table 3 has the most difficulty distinguishing between auto alone and carpool, and does reasonably well in predicting transit usage.

A more interesting test of the accuracy or validity of a disaggregate model is to examine its ability to predict on a data set different than the calibration data set. Recall that the model in Table 3 was fitted to 1973 data, prior to the inauguration of Trans-Bay BART service. To test the validity of the model, we used it to forecast mode split in 1975, including full BART service. This was done by comparing the actual mode choices of a 1975 sample with the choices predicted by the model in Table 3 when the 1975 set of alternatives and level of explanatory variables were substituted for each individual. The prediction success table for these forecasts is given in Table 5. The columns correspond to predictions using the 1973 calibrated model. Recall that the 1973 model has no BART alternatives, only auto-alone or shared-or bus-with-walk or auto access. From these alternatives we wish to predict the patronage on two new models, BART with auto access and BART with walk access. The model in Table 3 contains some alternative-specific variables, and it was necessary to make judgments about what form those alternative-specific variables would have in the post-BART situation. We assumed that BART with auto access has the same unobserved characteristics as bus with auto access, so that their alternative-specific variables would enter with the same coefficients. Analogously, we assumed that BART with bus access has the same characteristics as

TABLE 5. Prediction Success Table for Pre-BART Model and Post-BART Data

	Predicted Alternatives					
Actual Alternatives	*(1) Auto Alone*	*(2) Bus/ Walk*	*(3) Bus/ Auto*	*(4) BART/ Bus*	*(5) BART/ Auto*	*(6) Carpool*
(1) Auto alone	255.1	22.21	6.362	1.513	13.72	79.07
(2) Bus/Walk	11.56	36.43	2.988	1.679	1.421	13.92
(3) Bus/Auto	1.249	2.811	.687	.0066	1.625	2.622
(4) BART/Bus	.858	1.934	.120	1.391	.258	1.440
(5) BART/Auto	8.898	3.149	1.756	.695	8.828	9.674
(6) Carpool	74.68	12.43	3.305	1.357	7.497	37.73
Column Total	352.4	78.97	15.22	6.642	35.35	144.4
Predicted Share (%) (standard error)	55.8 (11.4)	12.5 (3.4)	2.4 (1.4)	1.0 (.5)	5.3 (2.4)	22.9 (10.7)
Percent Correct	72.4	46.1	4.5	21.0	26.5	26.1
Success Index	1.30	3.69	1.88	21.0	5.0	1.14
Predicted Share less observed share	–4.1	1.7	1.0	0.05	0.1	1.2
Actual Share (%)	59.9	10.8	1.4	.95	5.2	21.7

Totals		
Sample Size	631	
Percent Correct	53.9	(42.0 by chance)
Success Index	1.28	

bus-with-walk-access, with alternative-specific variables entering with the same coefficient. An overall judgment from Table 5 is the disaggregate model in Table 3 is relatively successful in predicting demand for a major new transportation mode. The model forecast a BART mode share of 6.3 percent, compared with an observed share of 6.2 percent. A caveat is necessary, however. The statistical imprecision of the calibrated coefficient of the pre-BART model would lead one to expect forecasts for modes with low aggregate shares, such as BART, to have relatively large percentage errors. The actual prediction accuracy here is better than one could expect by change, given the size of these standard errors of the forecasts. Further, disaggregate models in the form in Table 3 tend to be quite sensitive to the selection of variables entering the mean utility function, and to the definition and measurement of explanatory variables. For example, one of the problems which appears in this table is an overforecast of bus usage. An explanation can be found in the network calculation of bus access time. To construct these times, we used a 1980 Bay Area network which was constructed assuming 1980 bus service levels. The network was scaled back to 1975 by dropping bus links which did not exist in 1975, but the 1980 walk times which were shortened because the assumed 1980 transit service remained at the 1980 levels. As a result, walk time from our network calculations under-estimate true bus access time. This data measurement problem seems to be the major source of prediction error in Table 5. However, disaggregate models such as the one in Table 3 exhibit some anomalies when calibration samples are partitioned by location, family composition, or choice-alternative definition, suggesting that there are factors influencing travel demand which the current models do not capture adequately.

A statistical test of whether the post-BART data was in fact explained by the pre-BART model failed. That is to say, from a statistical point of view there are post-BART factors which are not explained adequately by the 1973 model despite the fact that it does a reasonably good job of forecasting aggregate BART patronage. In short, disaggregate demand forecasting has the flexibility and the potential accuracy to meet current planning needs, but the field of disaggregate demand forecasting is relatively uncharted, offering many potential pitfalls to the planner.

One property of the multinomial logit model which has gained some notoriety is called the *independence from irrelevant alternatives* (IIA) condition. This is a feature of the model which occurs when the mean utility of an alternative depends only on the attributes of that alternative and on the characteristics of the decisionmaker, and not on the attributes of other alternatives. In this case, the IIA property requires that the relative share of any two alternatives is independent of the attributes of the remaining alternatives. The terminology is due to the psychologist Duncan Luce, who first proposed the IIA property as an axiom for behavior in psychological choice.

The IIA property is a blessing and a curse for the multinomial logit model. It has some significant advantages. First, it allows calibration without having to consider all possible alternatives. For example, if one wants to carry out a study of destination choice, it is possible to calibrate the model with data on a selected number of destinations rather than having to consider the full set of destinations. This can substantially reduce data collection requirements. Second, IIA permits quick determination of the effects of introducing a new alternative, because the forecast of mode share for a new alternative mode can be obtained by including one additional term in the denominator of the multinomial logit formula.

The IIA property also has some major disadvantages. It fails to allow for different degrees of competition or similarity between alternatives. Consider the following example. Suppose that individuals initially have a shopping choice between the central business district (CBD) and a shopping mall—call it East Mall; and suppose that they initially split 50–50 between these two destinations. For simplicity, assume all individuals have exactly the same observed explanatory variables; i.e., they represent a homogeneous market segment. Suppose now that a new situation is introduced in which a North Mall is constructed. Suppose the North Mall and East Mall are equally far away for these individuals, with equal amenities. Then one would expect individuals who previously chose to shop in the CBD to continue to do so, and individuals who previously went to the East Mall to now split evenly between the East and North Malls. Hence, one would expect in this situation to observe a split of 50% CBD, and 25% for each of the two Malls. On the other hand, a multinomial logit model will predict a one-third split for each of the alternatives. The reason it does so is that it assumes that the relative odds of choosing between CBD and East Mall will be unchanged when an additional alternative is introduced—the North Mall. In other words, the multinomial logit model is unable to take account of the fact that the new North Mall will be more competitive with the East Mall than it will be with CBD shopping.

Let us pursue this example one step further. Suppose that we could break down the "homogeneous" market segment further, into, say, males and females, and that there were very strong differentials in shopping characteristics for these two socioeconomic groups. Suppose before the construction of North Mall the female segment divides 95–5 in favor of shopping at East Mall, while the male segment divides 95–5 in favor of CBD destinations. The aggregate share for the two segments is 50–50. Suppose now one applies the multinomial logit model to forecast destinations after North Mall is built, with separate forecasts for males and females. Then, the predicted splits for the female segment will be 48.7% for each Mall and 2.6% for CBD destinations; for the male segment, 4.8% for each Mall and 90.5% for CBD destinations; and finally an aggregate mode split of 46.5% for CBD destinations and 27% for each of the two Malls. Compare this to the observed split which is 50% for CBD destinations and 25% for each Mall. Then, the error introduced by the failure of IIA is small when market segmentation is effective in dividing the market.

In summary, the IIA property is extremely useful for practical planning. Its limitations are a more serious problem in aggregate modelling than in disaggregate modelling, where refining market segments can minimize errors. Although much of the discussion of the IIA property in the literature is concentrated on its logical possibility, a much more important consideration for the practicing planner is its empirical validity. If the disaggregate multinomial logit model having the IIA property can be shown to fit calibration data sets well and to forecast accurately in a particular application, then it is a useful tool for the planner.

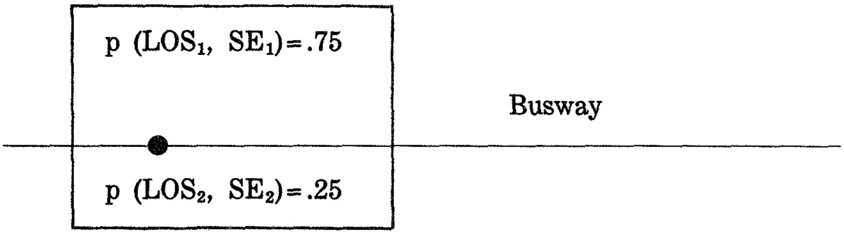
Specific statistical tests for the IIA property applicable to transportation data sets have been developed by McFadden, Tye, and Train. These tests can be used to investigate various specific sources of failure of IIA. Tests of IIA have been applied to a seven-alternative work trip data set for the San Francisco Bay Area. Because of the multiple transit alternatives (we have three BART, two bus, auto alone, and carpool alternatives) with common main-mode characteristics for alternative access modes, one would expect this data set to provide a rather stringent test of the IIA property. The multinomial logit model tested was of the same general form as the model in Table 3. The hypothesis that the model satisfied the IIA property was accepted for all the tests performed, with two exceptions which tended to point to data specification problems rather than IIA problems. Hence, this empirical study suggests that although IIA is an unpalatable logical restriction from the standpoint of the general theorist, it may be inconsequential from the standpoint of practical planning. At the very least, satisfaction of IIA is an empirical question, not a question of doctrine.

What should a planner do about the IIA property, given that its validity is a matter of concern in the profession? First, carry out diagnostic tests of the validity of the property for the specific data set you are using. If you reject the IIA property, try to refine the specification of your model by a more detailed market segmentation, improving data definition, or by adding variables to the models. If necessary, replace the multinomial logit model with one allowing patterns of substitution between alternatives.

The multinomial logit model is a special case of a disaggregate model, and not in any sense the end of the line in terms of realism and accuracy. However, it is the only disaggregate model which I believe is of current widespread practical useability.

I have described the process of defining and calibrating disaggregate behavioral models. Now I will discuss how these models are applied in forecasting. First, one must translate policy questions into specific technological features of the proposed transportation service. For example, suppose the policy question posed is "How much more transit service can we provide with a $1,000,000 block grant?" The question must be first translated into specific operating proposals for headways, route density, and so forth. Then, network or manual calculations, or an idealized supply model, must be used to provide the level-of-service variables resulting from a proposal. These variables must be provided for each homogeneous market segment for the level of segmentation at which the analysis is being carried out. Next, the size of each homogeneous market segment must be determined. In the short-run, one can normally assume population demographics continue to hold. For long-run forecasting, one must make projections of land use and demographic trends, and factor these forecasts into the segmentation. Finally, one must use the basic aggregation formula in equation (1) to predict changes in aggregate shares. Information on homogeneous market segments can be used to calculate the distributional consequence of proposals if this information is needed. Patronage and revenue calculations for the homogeneous market segments can be carried out, and aggregated to give totals. These figures, along with the capital and operating costs of alternative proposals, determine their feasibility. Among those proposals forecast to be feasible, a selection can be made using the evaluation criterion employed by the planning agency.

Consider the following example of the use of this procedure. Assume in Figure 2 that the square box at the top represents a traffic zone. Assume that the traffic zone is bisected by an express busway, and that one busway station denoted by the black dot serves the zone. The population densities within the zone are such that 75% of the people live north of the busway, and the remaining 25% live south of it. Suppose there is no parking provided at the busway station; hence, the people either walk, take feeder bus, or are driven to the station. Suppose current feeder-bus headways are twenty minutes on both the north and the south side and that the modal shares to the busway station are as follows: on the north side 52% walked; 18% take the bus; 30% are driven. On the south side 10% walk; 10% take the bus; 80% are driven and in total in this zone 41% walk; 16% take the feeder bus; 43% are driven.



Current Modal Shares

	Walk	*Bus*	*Driven*	*Proportion in Population*
North	.52	.18	.30	.75
South	.10	.10	.80	.25
Total	.41	.16	.43	1.0

Figure 2.—An example: The impact of improved feeder bus service.

The Planning Commission is contemplating improving the feeder bus service on the north side by reducing the headway from twenty minutes to five minutes, but leaving it unchanged in the low density area south of the busway. The consequences of this policy are calculated using a multinomial logit model with mean utility function at the bottom of Table 6. The mean utility is 3.11 times a variable which is 1 if the person walks and 0 otherwise, plus .495 times a variable which is 1 if the individual takes the bus, zero otherwise, minus .11 times travel time, minus .08 times headway, minus .2 time cost, plus .672 times the number of drivers (if the person is driven) and 0 otherwise. Here are the changes in mode shares calculated from the multinomial logit model when north-side feeder bus headways are reduced: North of the segment, the walk share goes down by .15, the bus share rises by .24, the number driven goes down by .09. Summed over the zone, the impact then is a .18 increase in the feeder bus share, a .11 decrease in the walk share, and a .07 decrease in the share of persons driven.

So far I have discussed the calculation of the effects of policy change on a homogeneous market segment. It is necessary to in general combine results for homogeneous market segments into an aggregate prediction for the population as a whole. If the segmentation is extremely detailed, then it may not be practical to carry through the aggregation by summing over all homogeneous market segments. There are a number of short-cuts or approximations to the aggregation process which can be used. I will mention four. First, one can approximate the empirical distribution of homogeneous market segments in

TABLE 6. Change in Mode Shares when North Side Feeder Headway is Cut from 20 Minutes to 5 Minutes

Segment	*Walk*	*Bus*	*Driven*	*Proportion in Population*
North	−.15	+.24	−.09	.75
South	0	0	0	.25
Total	−.11	+.18	−.07	

Mean Utility = 3.110 (if walk) + 4.950 (if bus) − .110 (traveling time) − .080 (headway) − .200 (cost) + .672 (no. of drivers, if driven)

the population with a mathematical distribution for which the expectation, or average, can be calculated analytically, possibly after a transformation of variables. Second, one can approximate the empirical distribution of socioeconomic variables and level-of-service variables in the population by a histogram, with each cell in the histogram corresponding to a fairly homogeneous market segment. Then the aggregate forecast is approximately equal to a sum over these market segments. This segmentation can be as coarse or as fine as desired; the finer the structure, the more accurate the segmentation. If a very coarse segmentation is used, then the method is close to an aggregate procedure. Third, one can approximate the empirical distribution of attributes of homogeneous market segments by using series expansions in terms of statistical moments, so that aggregate shares are written as functions of choice probabilities at average arguments and moments of the distribution of explanatory variables. Fourth, one can sample randomly from the empirical distribution of characteristics of homogeneous market segments, and form the sample expectation as an approximation to the population expectation. The first and third methods require information on moments of the distribution of explanatory variables. The second requires data on the size of market segments, and the fourth requires a representative sample from the population. The first method is not feasible except in special cases. Segmentation method two is feasible, and simple to apply for quick, rough answers when the number of explanatory variables is not too large. The third method does not converge rapidly, or perhaps not even at all, unless the distribution of explanatory variables is relatively concentrated. The fourth method is the most flexible. The required data for this method can be supplied from a calibration data base provided that the base is representative of the population, or from other data sources such as U.S. Census data, provided these sources contain the variables used in the forecasting model. In contrast to calibration, forecasting requires no data on actual transportation choices. Those are predicted by the model. Hence one can utilize socioeconomic data sets which are not specifically transportation-oriented to provide explanatory variables. A method of synthesizing socioeconomic data from Census data has been developed by Cosslett, Duguay, Jung, and McFadden (1977).

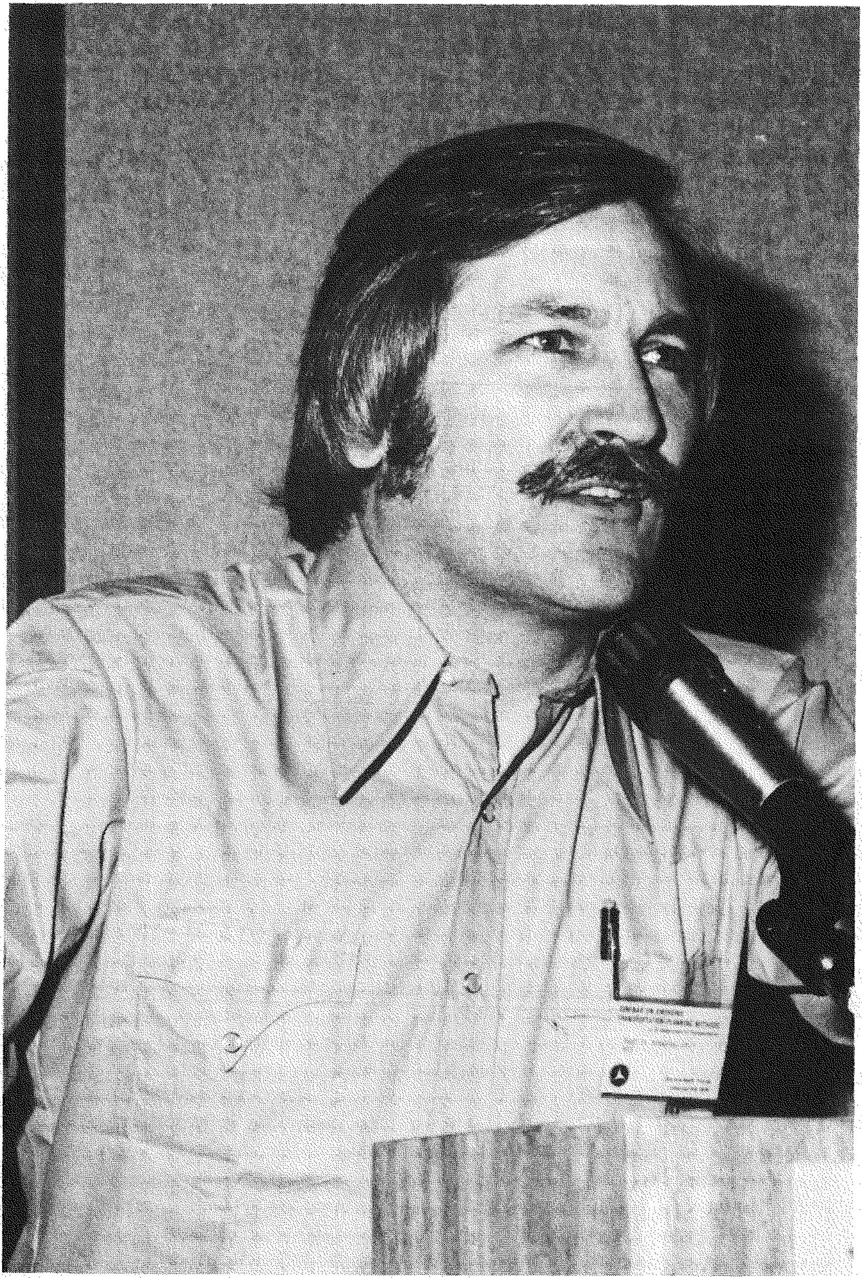
In summary, the sampling method of approximating statistical expectations is the most flexible tool for aggregate forecasting from disaggregate models. The method can be combined with survey or synthesized data to provide aggregate forecasts at reasonable cost.

The basic principles of behavioral disaggregate modelling, in summary, are that aggregate travel demand can be expressed as the sum of the demands of homogeneous market segments, and that the demand within a homogeneous market segment has a structure determined by behavioral regularities that are stable over time and space. How different are disaggregate and aggregate models in concept? They differ primarily in degree. Disaggregation carries market segmentation to the extreme. It emphasizes the regularity of individual choice behavior, in contrast to conventional modelling which emphasizes the physical regularity of aggregate flows. Aggregate and disaggregate models differ significantly in the number and form of explanatory variables, consistency across different aspects of travel behavior, calibration methods, and forecasting techniques. These differences are, however, primarily technical; the result of historical development and the practical limitations of data compilation and computation. Behind every good aggregate model stands a disaggregate model, and vice versa. The discovery of empirically valid regularities which simplify and extend forecasting methodology, and the relaxation of empirically invalid restrictions, should be a goal of every transportation analyst. From this point of view, disaggregate behavioral forecasting is a natural evolution of traditional aggregate demand analysis.

Calibration of behavioral disaggregate models requires less data than aggregate model calibrations. In forecasting, disaggregate models need to consider both the explanatory variables for each homogeneous market segment, and the computation of each segment's mode split. Fortunately, a variety of analytic or statistical methods, or a coarse market segmentation, can provide forecasts of aggregate mode shares. The range of answerable policy questions is limited by the extent of level-of-service variables affecting the choice probability. The planner's ability to translate policy changes into level-of-service changes is another potential limitation.

Aggregation predictions in disaggregate models can be adapted to comprehensive analysis of large-scale transportation system changes, or to "quick and dirty" analysis of limited aspects of travel behavior and incremental policy changes. In short, the behavioral disaggregate forecasting methodology can provide a multi-channel forecasting system. The theory of individual behavior provides a blueprint for the construction of disaggregate models. The methodology has the flexibility to meet the varied policy analysis needs of the planner.

It must be stressed that disaggregate behavioral analysis is neither a model nor model system; it is an approach to the development of model systems. There will never be "best" or "final" disaggregate models. Model systems will continue to evolve as experience accumulates. Not all model systems developed from behavioral principles will be "good." The method is open to abuse and misuse, as are aggregate model systems. Given that the analytic and statistical methods employed in disaggregate behavioral modelling will be new to many planners, and given that many planners are not well-grounded in the "folk theory" of behavioral modeling from economics and psychology, one can predict the unsuccessful disaggregate models will outnumber the successful ones. On the other hand, there is now a track record of success with these models. They have proved that they can provide accurate and flexible forecasts, and that used with judgment, they can provide a useful tool for organizing and systematizing policy analysis.



Paul O. Roberts, Jr., Professor of Civil Engineering
Massachusetts Institute of Technology

Disaggregate Demand Forecasting: Theoretical Tantalizer or Practical Problemsolver?

Paul O. Roberts, Jr.

Professor of Transportation
Massachusetts Institute of Technology

PIONEERING A NEW APPROACH

Before coming to this conference, you probably had already heard of disaggregate demand modelling, though you may have puzzled a bit over exactly what it was. You might have also inferred the existence of aggregate demand modelling since there can hardly be disaggregate modelling if there is not an aggregate form. You may have learned of the existence of both aggregate and disaggregate models by finding yourself in the middle of a heated argument between two of your more esoteric colleagues discussing it at the top of their lungs. You may have wondered why the subject seems to engender such passion. The reason is that disaggregate modelling is relatively new and aggregate modelling is, by comparison, old. And, since advocates of the tried and true can be expected to argue against being abruptly displaced by some untried new approach, it is not surprising that disaggregate modelling would raise some controversy.

The question for you is, does the new approach deserve your attention? Can it be used cost-effectively to answer the questions you are called upon to answer? Like any new frontier, disaggregate modelling has its pioneers. You had the opportunity to hear this morning Dan McFadden, the 'Daniel Boone' of disaggregate modelling, presenting the fundamentals. Before the conference is over, you will also hear from the Lewis and Clark, the Davy Crockett, the Zebulon Pike, the Kit Carson, and the Jesse James of disaggregate modelling. These are real pioneers of the disaggregate approach—men with a clear vision of the future and the courage and strength to make that future come true.

The trouble is, you must be thinking, 'I don't want to have to fight off all those Indians, clear all those trees, and plough all that ground to get a few, meager results. After all, this is supposed to be the space age. I thought these fellows you mentioned were part of the new frontier, not the old frontier. Isn't there going to be any practical fall-out of the space-age research we heard described this morning? Can it actually be applied? Does it do anything for you? or, is it just a case of the mathematical overkill so prevalent in Academia?'

I am here as a speaker at this dinner today as your modern-day Neil Armstrong to answer your questions about disaggregate models in the affirmative. There is more to this disaggregate thing than initially meets the eye. We are not just talking about a new way to do trip generation or distribution. What the disaggregate approach really implies is a whole new way to look at the demand side of transportation forecasting. I would like to begin my presentation by pointing out some of the subtle and perhaps even some of the not so subtle distinctions that come out of a closer look at the disaggregate approach. I would like to start out by discussing three major aspects of this subject. These are:

1. The disaggregate nature of travel itself;
2. The philosophical underpinnings of the disaggregate model; and
3. The prediction framework within which the model is used.

I will treat each in turn.

DISAGGREGATE NATURE OF TRAVEL

The basic concern of passenger travel demand modelling is "the trip." Each trip has an origin and a destination. It also involves the number and the characteristics of each of the travelers, the mode, route, time of travel, purpose, etc. There is an equivalent set of information for freight shipments. We will, however, concern ourselves with passengers and leave implied a basic similitude for most of what is said for freight travel.

There are literally thousands of trips taken every day over the United States as a whole—urban trips, inter-city trips, shopping trips, long trips, short trips (too many to look at individually and make anything of them). In order to understand what is going on, one must focus on an area or a system of interest and aggregate the individual trips to obtain something that can be perceived, manipulated, and associated with relevant policy questions. For example, we typically aggregate travel by mode and trip purpose between zones and feel that this furnishes us with the basic movements to be used in our studies. The result of such an aggregation of the more basic data, such as home interviews, is typically a table of origin to destination traffic flows. Such O/D traffic flows have, in fact, become the starting point for most urban design studies.

We should note, however, that the basic process of aggregation inherently loses information. We lose information about spatial detail when we aggregate into zones. We lose time detail when we aggregate trips over the day, and we lose modal detail when we fail to distinguish between auto passenger, auto driver or taxi, in the case of automobile; or local bus, express bus, subway, or dial-a-ride, in the case of public transportation. Once the data has been aggregated, it is impossible to disaggregate it to obtain true origin, destination, mode, or route information which has been thrown away. Only by returning to the original source documents (if they have not been destroyed) can the basic data be reaggregated and factored up to give an aggregate picture of the flows with the particular detail that is desired.

To summarize, travel is a very disaggregate process and therefore the raw data concerning travel is typically disaggregate. Ordinarily, the data must be aggregated in some fashion to answer policy questions. Different policy questions may require different aggregations of the raw data.

Our question is, then, should we model the process at the disaggregate level using a so-called disaggregate model? Or, alternatively, should we wait until after aggregation and build our model on the aggregated data (obviously an aggregate model)? There is, of course, a trade-off that must be considered. To get at this trade-off, we will have to look at the other two aspects—which I said I would discuss.

THE PHILOSOPHICAL UNDERPINNINGS

There are several major ways in which the disaggregate model itself is different from the aggregate models that transportation analysts are already familiar with. First, the disaggregate model is a model of an individual decisionmaking unit, a person, or a family, not a zone. I am sure you have heard the cliche: zones don't make trips—people do. People, as families, decide where they're going to live, how many autos to own, what kind of house they will live in, how they will get to work. Disaggregate models attempt to replicate this set of joint decisions that are evidenced by what people do.

A second point of departure is that there is an underlying theory—consumer theory—which postulates how people facing alternatives will decide between them. The theory postulates that they will tend to maximize their overall utility in their choices. This is unlike the gravity model, that most-widely used aggregate model, which argues, unconvincingly, through similitude that "bodies will be attracted in proportion to their mass and, inversely, as the square of the distance between them." Consumer theory argues the existence of a utility function for a consumer that is behaviorally-based. The utility function is nothing more than an expression incorporating all those variables which are important, along with coefficients to be estimated. The utility function will consist of both transport level of service variables describing the tranport system and socioeconomic variables describing the traveller. For example, under *transport level of service* attributes such things as:

1. Schedule delay,
2. Time and reliability,
3. Waiting time,
4. Walking time,
5. Privacy,
6. Personal safety, and
7. Out-of-pocket charges,

are typically cited. Under *socioeconomic* variables describing the traveller and his family we may include:

1. Income,
2. Family size,
3. Age,
4. Sex,
5. Education,
6. Number of workers, and
7. Number of drivers, etc.

All relevant variables can be included.

Disaggregate models are typically probabilistic over the choices addressed by the model. That is, each choice is predicted with some probability instead of with certainty. For example, the probability predicted by the model increases that I will choose the bus as the fare goes down. I still may not ride the bus on any given trip. The original observations used to estimate the coefficients of the model are, however, discrete choices over the set of alternatives considered. Thus, the basic data is typically zero or one in nature. This suggests a mathematical form such as the multinomial logit model which will handle this kind of raw data. By aggregating the sample observations over the area as a whole, the probability of making a particular choice tends to approach the figure predicted by the model. The probability times the number of trips represented by the sample, factored up, produces an estimate of aggregate flows useful for planning.

The model can be used to predict the probability of choosing between a wide variety of alternatives available to the decisionmaker—different modes, different travel times, different levels of auto ownership, different places to shop, and/or combinations of these probabilities—just like in the real world. A disaggregate model can incorporate choices beyond or instead of transportation since it is, in fact, a general choice model which is theoretically sound across the entire range of human endeavors. It can therefore be used to address issues involving trade-offs between transportation and other activities (i.e., housing, recreation, or other uses of family funds). Since the model has the values associated with various socioeconomic classes built into the parameters, the models are theoretically transferable from one place to another. This feature has been tested in practice by calibrating models for two separate areas and, comparing the coefficients, it has also been shown that a model calibrated for one area can predict flows in another area at a reasonable level of accuracy. See the reference by Atherton and Ben-Akiva [5].

This transferability feature carries with it data collection economies. Unlike the conventional aggregate models (e.g., the gravity model), which are fundamentally tied to the aggregation structure by which they were created, disaggregate models are associated with the individual decisionmakers and, once the coefficients are developed for a truly-representative sample of decisionmakers, they can be used elsewhere as long as the basic nature of the decisionmakers remains unchanged. They are, therefore, not tied to a specific area like aggregate models. In view of the lack of money for data collection and modelling at this time generally, this turns out to be an important argument in favor of disaggregate models.

The basic behavioral nature of the model and its relationship to the individual decisionmaking unit means that it is much more policy-responsive than aggregate models, i.e., a wide range of policy alternatives can be posed through their representation in the transportation level of service attributes input to the model.

Finally, since the model considers individual decisionmakers, the distribution of the impact of a given policy can be determined. It is possible, for example, to pick out of the sample a particular subgroup for observation. We may want to pay particular attention to the impact of a transit fare increase on elderly people living in various parts of the city. Or, we might be interested in its impact on former taxi users or even perhaps on the shopping habits of travelers from a particular affluent section of the city. This is easily done with disaggregate models, since it is easy to apply the model to the individuals in any group that we would like to address. In fact, to do this we merely identify the observations in the sample that have the desired characteristics and not the difference between the base case choices made by these individuals and those of the same group made in the policy case. A similar kind of thing cannot be done with aggregate models, since individuals are not identified at all. The only policies which can be addressed are those policies for which the model was constructed in the first place.

THE PREDICTION FRAMEWORK

We noted earlier that a disaggregate model works at the level of the basic decisionmaking unit. The individual choices of decisionmakers are replicated. You may be wondering just what all this detail does 'to us' as opposed to 'for us'. You may note that the output of the disaggregate model is as hard to use as the original observations. That is precisely the case—the results of the model have to be aggregated just like observations in the original data. On the other hand, this has tremendous advantages because they can be aggregated in any number of ways just like the original raw data observations. This does require that we have an aggregation scheme which can be used on the predicted choices made for each observation by the model.

There are several ways aggregation can be done. First, if the population is small, then the model can be run for each observation in total enumeration. This becomes impractical when the population becomes large, so a second approach is to draw a disaggregate random sample or a stratified random sample from the population. If this sample is truly representative of the population of interest, the model can be run on each observation of the sample and the results aggregated and factored up to produce the total flows. Note that the representative disaggregate sample used for policy analysis could, in some cases, be the original data used to estimate the model, if it is the same as the population of interest.

A third approach, employed where a real population is difficult or expensive to sample, uses a synthesized random sample. Such a sample can be drawn from available public data such as the census files, the county tax roles, etc. The actual behavior of the sample may not be known. Thus, the base case may be simulated in the same way that the policy cases are simulated.

Another generalized approach is to classify the population into groups with different basic socioeconomic characteristics. The model is then run for each group at the group mean. The results are then aggregated and factored up by multiplying by the number in the full population that falls within each group.

Finally, pivot point methods provide a practical and easy means to do policy analysis for many situations where the flows are already known for the base case. The pivot point method uses the elasticities within the model to estimate changes in the flows that will occur as the result of changes in the level of service variables. Thus, there are a variety of ways in which the models can be used, of which the approaches mentioned here are only illustrative.

The basic steps underlying the use of the models within the prediction framework are easy to understand. These steps are:

1. Select transport policies of interest;
2. Identify the population impacted by the policies;
3. Develop a sample which is representative of the population of interest;
4. Apply the previously estimated disaggregate model to each observation in the sample under the conditions of the base case;
5. Factor up the results to give a base case answer;
6. Change the transport level of service offerings to reflect the impact of the transport policy changes;
7. Rerun the estimated model on each of the revised observations in the sample;
8. Factor up the results to give the policy case result;
9. Compare the base case to the policy case to determine the impacts; and
10. Trace out the full distributional implications of each of the alternative policy cases.

As you can see, the approach described here is extremely flexible and can be applied to a wide range of policy situations. Since the models are designed to replicate behavior at the individual level, they can be applied to any number of different policies. However, the aggregation scheme itself and the sample used for any particular set of policies may have to change as the policy changes. The fundamental difference between disaggregate models and aggregate models is that the disaggregate model requires that a new aggregation scheme be developed for *each* use of the model. By contrast, the aggregate model does not require a new aggregation scheme to be developed but it is inherently less adaptable to the same wide range of policies. The aggregate models are also limited in their range of location since a new aggregate model must be developed every time the location is moved.

You have heard the good news; now for the bad news (or so the aggregate modelers would have you believe). ''Agreed!'' they would say. "The model is better specified, more readily estimated, on a smaller data base, more easily transferred, etc. But you have got to forecast all those myriads of variables in the model and you must also aggregate the results. For these reasons, it is better to specify an aggregate model with fewer variables and address the policy questions at the time the model is being developed. Then, a simplified model, which predicts aggregate results, can be used to do policy analysis quickly and easily!" Or so his argument would go.

It certainly is true that aggregate models, once developed, can be used more directly for policy analysis. However, the world has had a disconcerting habit of changing policies of interest over the past few years. Our emphasis has changed from long-range planning of infrastructure to short-range, low-capital projects. The concerns are now for pollution and energy conservation. The same aggregate models that used to be applied to infrastructure planning are totally inappropriate when applied to these new concerns. I would prefer to have a model which would replicate the results of an individual's response and have to cope with the problem of developing a new aggregation scheme for use with the model so that I could properly address the particular policy problems at hand rather than attempt to apply my aggregate models, developed for a former era and a former set of policies, to the policy question at hand and find that they no longer applied.

With respect to the comments about forecasting all those myriads of variables, if you have the disaggregate data for one point in time, one only needs to forecast the important changes. If the data doesn't change, the original input will suffice. Conversely, if the variable was important and is changing in the future, there is no way, even using an aggregate model, that you can ignore the change and leave it out of the model without grave concern for the predicted results.

My final comment would be to ask a rhetorical question. Have you ever calibrated and used the conventional model systems for urban transportation planning? Complete with trip generation, network routing, trip distribution, modal split and assignment, and maybe even involving a land-use forecasting submodel as well? Can you honestly say that the use of such a model is simple and straightforward? I challenge any approach to be any more involved or extensive.

I think I can read your minds. If you face a problem with the use of disaggregate modelling it is over having to stop what you are presently doing and learn a new set of techniques. Your attitude probably is, "If the present set of techniques works, why not continue to use it? Why take time off from a busy schedule to learn a complex, theoretical, and esoteric new approach?" Certainly, to do it, the results must be worth the effort.

THE MODELS AVAILABLE

I would like, if I can, to show you that the use of disaggregate models is actually more flexible and adaptable than conventional aggregate models, easier to use, not harder, and they can be used in more applications and in ways that are considerably more policy-responsive than previous models. Two things are required to demonstrate this:

1. A repertoire of estimated models; and
2. A set of schemes for aggregating the results.

We have already discussed several possible schemes for aggregation. Let me, therefore, turn directly to the question of a repertoire of estimated models. For urban passenger planning, the following models have been developed and used in various policy studies:

1. Work place choice,
2. Mode to work and auto occupancy,
3. Mode for non-work trips,
4. Frequency for non-work trips,
5. Destination choice for non-work trips, and
6. Time of day for non-work trips.

Many of the applications require models which can be used to make joint or simultaneous choices. For example:

1. Joint auto ownership/mode choice to work,
2. Joint frequency/destination/mode choice for shopping and sociorecreational travel,
3. Joint frequency/destination/mode choice for weekend travel, and
4. Joint residential location/housing type/auto ownership/choice of mode to work.

These joint choice models can be thought of as conditional, based on previous decisions made by the decisionmaking unit. The last example is virtually a form of land-use model. Work on incorporating time of day of travel, length of trips, and carpooling into these models have all been completed and are either available or now being written up. The use of pivot-point analysis has also been completed and an example of its uses will be presented at this conference.

The models I have described are typically choice specific, but are usually abstract with respect to level of transport service attributes. To apply the available models, one must be able to represent the changes in the transport level of service variables that will occur as the result of the policy change being contemplated, (i.e., parking, surcharge, carpooling policy, transit fare change, construction of a new facility, or a modal service offering, etc.). Thus, the full set of level of service variables in the models is not just desirable, but necessary.

Finally, as we mentioned previously, a scheme for applying the models to some representative portion of the appropriate population and aggregating the results is necessary. This is part of the process which requires the most departure from previous practice. But, it is also the point at which there is the most room for experimentation and creativity. A variety of short-cut methods for accomplishing this have been developed and more are imminent. Obviously, a considerable amount of judgment has to be used in deciding how to represent the population of interest. The selection of a sampling method and a sampling frame are important. Methods for synthesizing samples will eventually become commonplace. There is, typically, a considerable amount of data available concerning the distribution of population and industry so that it ought not be difficult to find a common data base for practically every area and set of concerns. Available real-world data concerning flows, service levels, etc., can all be used in the calibration and testing process.

SOME EXAMPLE APPLICATIONS

The disaggregate approach is new but not untried. A variety of applications have been already accomplished successfully in the real world. In order to understand the state of development of the new techniques and their application, let me describe some example uses of the models which I introduced in the previous section. (These applications are taken from a recent paper describing the status of disaggregate modelling by Moshe Ben-Akiva, Steven R. Lerman, and Marvin L. Manheim entitled, "Disaggregate Models: An Overview of Some Recent Research and Practical Applications.")

A Traditional Traffic Study in the Netherlands—A traditional transportation planning process was used and the disaggregate models were incorporated for choice of mode to work and shopping. The modes involved included walk to work, bicycle, and moped, in addition to auto and transit [37] [38].

Energy Conservation Measures for the Federal Energy Administration—This study examined a series of carpooling incentive policies. A disaggregate random sample was used as representative of all households. The results of each policy were represented in the level of service offerings to represent each of the plans. The results were then aggregated over all the observations and expanded to the entire population. The results of many of these developments are being synthesized into a variety of simple procedures to be used by local officials in planning energy conservation measures. This approach was particularly interesting in that it involved the development of worksheets for hand calculator application of the models. The results included a workbook and example applications of the models for a variety of situations [3].

A Study of Auto-Restricted Zones for the Urban Mass Transit Administration—Sketch planning versions of the models were applied using pivot-point analysis to show the changes that would result from various restrictions in the Central Business District of selected cities as part of the process of selecting sites, and implementation strategies for a federally-sponsored demonstration program.

Parking Restrictions for the Federal Energy Administration Office of Contingency Planning—The disaggregate models were applied to an analysis of parking restrictions within the downtown areas of large metropolitan areas [19]. The purpose of the analysis was to determine the impact on energy conservation of a variety of parking and transit options.

A Study of Demand-Responsive Transportation for the Urban Mass Transit Administration—Disaggregate models of work and non-work trips were used to estimate the potential for demand-responsive transit [17]. This study involved developing hypothetical level of service vectors for a potential new mode and their use in the demand models. Data from Haddenfield, New Jersey, and Rochester, New York, were used.

An Integrated Transportation Planning Model System for the San Francisco Metropolitan Transportation Commission (MTC)—A series of integrated urban transportation planning models was developed for use in the computer planning packages being employed by the MTC. The MTC model system represents an extreme in the application of disaggregate techniques (i.e., the development and implementation of an entirely new framework of analysis). This model system has become a prototype for the application of disaggregate models within a total planning framework.

A Policy Study for the U.S. Federal Energy Administration on Carpooling—The effects of alternative programs of incentives to carpool (shared use of autos for work trips) were examined. Washington, D.C., and Birmingham, Alabama, were used as prototype cities. For the Birmingham analysis, the Washington model specifications were reestimated on Birmingham data. The Random Sample aggregation method was used [15].

Research and Development Priorities for Urban Travel Modelling, for the U.S. Department of Transportation—A system of disaggregate models has been implemented in UTPS, together with versions of the conventional 4-step models. Using Washington as a case study, forecasts were made of the effects on travel of a variety of auto and transit policies, to determine in what ways the alternative approaches produce similar or different forecasts [43].

A Mode-Split Model for the Southern California Association of

Governments—As a part of the Los Angeles Transportation Study, a mode split model was estimated using the Los Angeles data [19].

Choice of Mode and Parking Location—For the agency planning a "people-mover" system for internal circulation within the Los Angeles central business district, models have been developed for predicting, for peak-period trips, choice of parking lot and egress mode (travel from parking to destination) if arrival by auto, and egress mode if arrival by transit; and for noon-hour trips, frequency destination and mode of within-CBD trips. Modes include walk, minibus, and the use of the people-mover system [7].

Thus, disaggregate models have been and are being applied to a broad range of problems in a large number of places.

In addition to the projects described above, which are all completed, a number of research initiatives are currently underway. These include:

The development of models which explicitly represent trip-chaining decisions for non-work trips—All the models currently developed focus on simple home-destination-home trip chains. In contrast, actual travel behavior is often much more complex. Work by Adler [1] is considering the entire set of daily trip tours as potential travel alternatives, thereby allowing for possible consolidation of trips.

The design of more efficient sampling procedures for disaggregate model estimation—Lerman and Manski [30] have explored a potentially less costly sampling technique which takes data from decisionmakers based on the decision they actually made. This procedure, termed choice-based sampling, requires some extensions of the econometric theory underlying the estimation of choice models.

The inclusion of time allocation in the traveller's decision process—The allocation of an individual's time to various activities during the day acts as both an impetus to travel and a potential constraint on total travel time. The models developed by Bain [6] focus on the individual's decision of whether or not to participate in an activity on a given day and the expected level of participation.

The development of Policy Sensitive Models of Freight Demand—Terziev, Ben-Akiva, and Roberts [39] are currently exploring the potential for applying logical analogues to recent passenger travel demand modelling advances to problems in the freight demand area.

The development of methods for estimating multiple alternative choice models other than the logit—Cambridge Systematics, Inc., is currently developing efficient computational procedures for estimating the multinomial probit model. This model, discussed by Hausman and Wise [24], permits a relaxation of the independence of irrelevant alternatives properly inherent in the logit model, as well as allowing for explicit representation of random variation in tastes in the modelled population.

In the policy analysis area, work is currently underway to put many of the recent model improvements into actual practice. There are two major examples:

The development of a model for central government planning of urban transportation resource allocation—MIT is currently developing a procedure for utilizing disaggregate choice models to predict the aggregate travel demand for an entire city. This procedure is to be part of a larger model system which, when applied to all urban areas in the country, will help guide national-level transportation investment strategy [11].

The development of a model system for predicting the patronage of demand-responsive transportation systems—Cambridge Systematics [17] is currently combining a set of disaggregate travel demand models which include so-called demand-responsive transportation (i.e., dial-a-ride service) as a mode option, with performance prediction models as part of a larger forecasting system.

The development of land use forecasting methods—Worms [42] is exploring the potential of using joint, disaggregate models of residential location, housing, auto ownership, and mode to work models as the demand component of a comprehensive urban land use model. This work has used a modified version of the supply sector of the NBER Urban Simulation (Ingram [25]) as a basis for a new experimental land use model.

SUMMARY AND CONCLUSIONS

I would like to summarize my comments by saying that to me there is no question about the usefulness or the applicability of disaggregate models. The new approach is:

1. Issue oriented, rather than technique oriented;
2. Easily-tailored to the problem at hand;
3. Flexible in its use of both short-range as well as long-range planning problems, and applicable to small areas as well as large areas and to partial as well as comprehensive problems;
4. The methods are policy-sensitive and, perhaps most important of all; and
5. They are impact specific.

Thus, the problem is not whether to apply them, but how can we get them applied more quickly and easily.

REFERENCES

1. Adler, Thomas J., *Modelling Non-work Travel Patterns,* Ph. D. Dissertation, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1975.
2. Adler, Thomas J. and Moshe Ben-Akiva, "A Joint Frequency Destination and Mode Model for Shopping Trips", *Transportation Research Record 569,* Transportation Research Board, Washington, D.C., 1975.
3. Alan M. Voorhees and Associates, Inc., *Auto Restricted Zones: Background and Visibility,* Phase I Report, prepared for Urban Mass Transportation Administration, 1976.
4. Atherton, T. J., *Approaches for Transferring Disaggregate Demand Models,* S.M. Thesis, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, February, 1975.
5. Atherton, Terry J. and M. Ben-Akiva, "Transferability and Updating of Disaggregate Travel Demand Models", *Transportation Research Record 610,* Transportation Research Board, Washington, D.C., 1976.
6. Bain, Joseph H., *Activity Choice Analysis Time Allocation and Disaggregate Travel Demand Modelling,* M.S. Thesis, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, June, 1976.
7. Barton-Aschman Associates, and Cambridge Systematics, Incorporated, *Los Angeles Central Business District: Internal Travel Demand Modelling,* prepared for Community Redevelopment Agency, Los Angeles, California, 1976.
8. Ben-Akiva, Moshe, *Structure of Passenger Travel Demand Models,* Ph.D. Dissertation, Massachusetts Institute of Technology, Department of Civil Engineering, Cambridge, Massachusetts, 1973.
9. Ben-Akiva, Moshe, "Structure of Passenger Travel Demand Models", *Transportation Research Record 526,* Transportation Research Board, Washington, D.C., 1974.
10. Ben-Akiva, Moshe, and T. Adler, *Social Recreation Joint Choice Model,* Working Paper #8, Center for Transportation Studies, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1975.
11. Ben-Akiva, Moshe, Koppelman, Frank, and T. Watanatada, *Development of an Aggregate Mode of Urbanized Area Travel Behavior,* Phase I Report, prepared for the Office of the Secretary, U.S. Department of Transportation, 1976.
12. Binder, Robert H., "Major Issues in Travel Demand Forecasting", Brand, Daniel L. and Marvin L. Manheim (eds.), *Urban Travel Demand Forecasting,* Special Report 143, Washington, D.C., Highway Research Board, 1973.
13. Bouchard, Richard J., "Relevance of Planning Techniques to Decisionmaking", Brand, Daniel L. and Marvin L. Manheim (eds.), *Urban Travel Demand Forecasting,* Special Report 143, Washington, D.C., Highway Research Board, 1973.

14. Cambridge Systematics, Inc., *A Study Design for Dual Mode Transit Planning Case Studies,* prepared for Urban Mass Transport Administration, Department of Transportation, 1975.

15. Cambridge Systematics, Inc., *Carpooling Incentives: Analysis of Transportation and Energy Impacts,* prepared for Federal Energy Administration, 1976.

16. Cambridge Systematics, Inc., *Introduction to Urban Travel Demand Forecasting,* prepared for Software System Development Program, Urban Mass Transport Administration, Department of Transportation, 1974.

17. Cambridge Systematics, Inc., *Methods for Estimating Patronage for Demand Responsive Transit,* Draft Final Report, prepared for the Transportation Systems Center, U.S. Department of Transportation, 1976.

18. Cambridge Systematics, Inc., *The Development of a Disaggregate Behavioral Work Mode Choice Model for the Los Angeles Metropolitan Area,* prepared for the Los Angeles Regional Study and the Southern California Association of Governors, 1975.

19. Cambridge Systematics, Inc., *The Use of Disaggregate Travel Demand Models to Analyze the Energy Conservation Potential of Parking Restrictions within the Metropolitan Area,* prepared for the Federal Energy Administration, Office of Contingency Planning, May 1976.

20. Cambridge Systematics, Inc., *Travel Model Development Project,* Phase I Final Report, prepared for the Metropolitan Transportation Commission, Berkeley, California, 1975.

21. Charles River Associates, Inc., *A Disaggregated Behavioral Model of Urban Travel Demand,* Federal Highway Administration, U.S. Department of Transportation, 1972.

22. Cosslett, Steven, "Trip Timing Analysis: the Effect of Congestion Time on Scheduling Trips to Work by Auto", draft paper, Travel Demand Forecasting Project, University of California, Berkeley, California, 1976.

23. Domencich, T., and D. McFadden, *Urban Travel Demand: A Behavioral Analysis,* North-Holland, Amsterdam, 1975.

24. Hausman, Jerry A., and David Wise, *A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Inter-dependence and Heterogenous Preference,* Working Paper #73, Department of Economics, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1976.

25. Ingram, Gregory K., John F. Kaine, and J. Royce Ginn, *The Detroit Prototype of the NBER Urban Simulation Model,* National Bureau of Economic Research.

26. Koppelman, Frank S., *Travel Prediction with Disaggregate Choice Models,* Ph.D. Dissertation, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, June, 1975.

27. Lave, C. H., "A Behavioral Approach to Modal Split Forecasting", *Transportation Research, 3,* No. 4, December, 1969.

28. Lerman, Steven R., *A Disaggregate Behavioral Model of Urban Mobility Decisions,* Ph.D. Dissertation, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1975.

29. Lerman, Steven R., and M. Ben-Akiva, "A Disaggregate Behavioral Model of Auto Ownership", *Transportation Research Record 569,* Transportation Research Board, Washington, D.C., 1975.

30. Lerman, Steven R., and C. Manski, "Alternative Sampling Procedures for Disaggregate Choice Model Estimation", forthcoming in *Transportation Research Record,* Transportation Research Board, Washington, D.C., 1976.

31. Lisco, Thomas E., *The Value of Commuters' Travel Time: A Study in Urban Transportation,* Ph.D. Dissertation, Department of Economics, University of Chicago, 1967.

32. Luce, R. D. and P. Suppes, "Preference, Utility and Subjective Probability", in Luce, R. D., R. R. Bush, and E. Galanter (eds.), *Handbook of Mathematical Psychology,* Volume III, John Wiley, New York, 1965.

33. McFadden, Daniel, "Conditional Logit Analysis of Qualitative Choice Behavior", in Zarembka, Paul, (ed.), *Frontiers in Econometrics,* Academic Press, New York, 1974.

34. McFadden, Daniel, "Quantal Choice Analysis: A Survey", draft paper presented at the NSF-NBER Conference on Decision Rules and Uncertainty, March 22–23, 1974.

35. McGillivray, Robert G., "Binary Choice of Urban Transport Mode in the San Francisco Bay Region", *Econometrica, 40,* No. 5, September 1972.

36. Peat, Marwick, Mitchell and Company, *Implementation of the n-dimensional Logit Model,* prepared for The Comprehensive Planning Organization, San Diego County, California, May 1973.

37. Richards, Martin G., "Application of Disaggregate Modeling Techniques to Transportation Planning Studies", prepared for presentation at the PTRC Summer Annual Meeting at the University of Warwick, July 1975.

38. Richards, Martin G. and M. Ben-Akiva, *A Disaggregate Travel Demand Model,* Saxon House, D.C. Heath, Ltd., England, 1975.

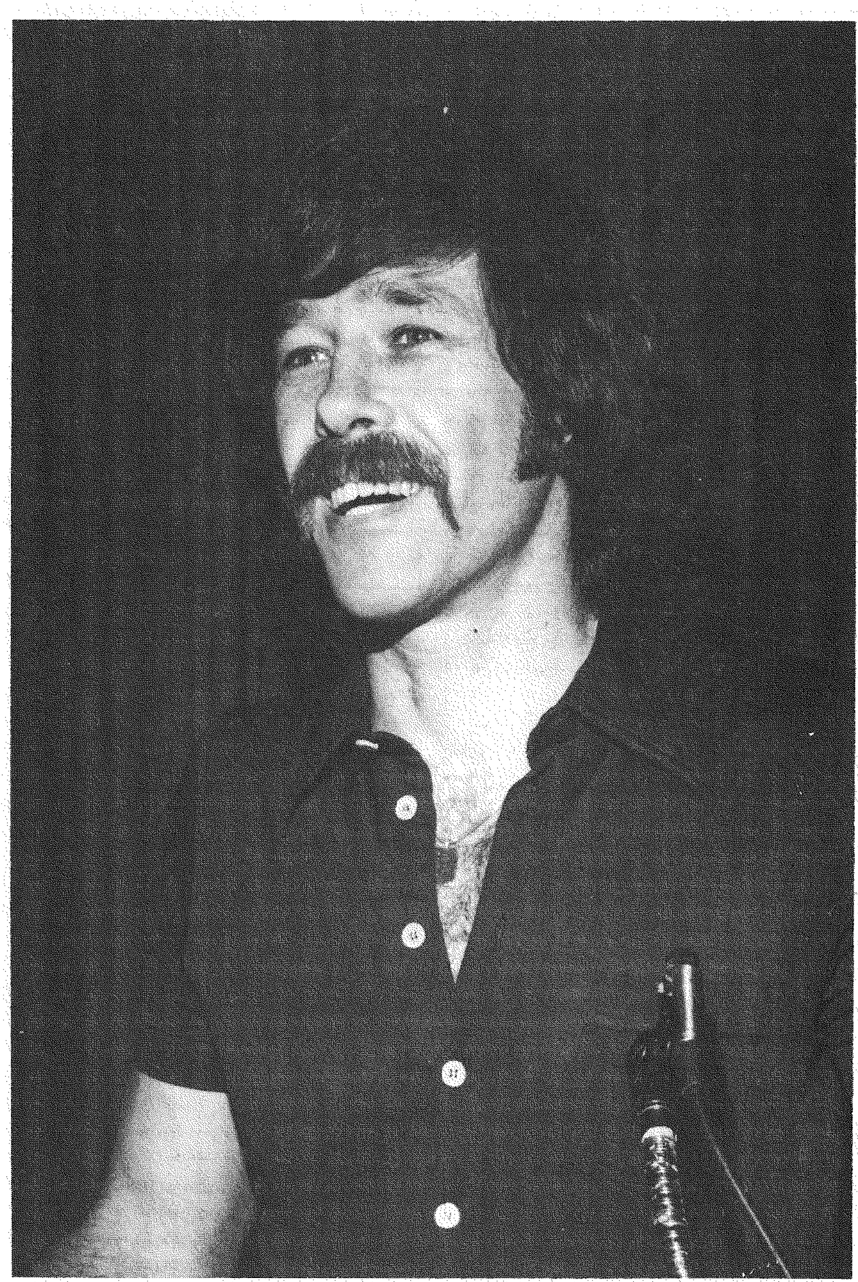
39. Terziev, Marc N., Moshe Ben-Akiva, and Paul O. Roberts, *Freight Demand Modelling: A Policy Sensitive Approach,* Massachusetts Institute of Technology, Center for Transportation Studies Report No. 75–6, April 1975.

40. Warner, S. L., *Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice,* Northwestern University Press, Evanston, Illinois, 1962.

41. Watson, P. and R. B. Westin, "Transferability of Disaggregate Mode Choice Models", *Regional Science and Urban Economics, 5,* 1975.

42. Worms, Vincent R., *The Implementation of a Joint Disaggregate Demand Model in An Urban Simulation,* M. S. Thesis, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, June 1976.

43. Young, Steven C. K., *The Application of Disaggregate Demand Models in Urban Travel Forecasting,* S.M. Thesis, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, May 1976.



Robert B. Dial, Director,
Office of Planning Methodology and Technical Support
Urban Mass Transportation Administration

Summary of Discussions: Transportation Demand Forecasting Techniques

Robert B. Dial
Director
Office of Planning Methodology and Technical Support
Urban Mass Transit Administration

Monday was certainly the most informative day for the participants of this seminar on "Emerging Transportation Planning Methods." This is not an idle boast. It is a fortunate fact. In more than any other area of planning, transportation demand forecasting techniques have enjoyed significant advances in these last few years. These advances sprang from the excellent work of small number of experts. The average practitioner, however, has not been privileged to the information that would enable him to utilize them in his work. This seminar brought together experts and practitioners, so as to improve the ability of the latter with the knowledge and experience of the former. And that's exactly what happened this fine, sunny, December Monday in Daytona.

Following Daniel McFadden's superb overview of disaggregate modeling in general, and the logit model in particular, a blue ribbon set of session leaders moderated panel discussions on the most important technical issues facing the transportation demand forecaster. They impressed all of us with the technical quality of sample solutions presented to these most vexing problems.

Peter Stopher led a panel discussion on Disaggregate Data and the Role of Nontransportation Variables. The single most attractive feature of behavioral models is their ability to deal with observations on the decisionmaking unit. This unit (e.g., the trip maker, his family, or his household) confronts conflicting desires and competes for scarce resources. Many of these desires and resources fall outside the traditional realm of transportation but are nonetheless crucial to the decision making process. This topic has been a specialty of Professor Stopher, and attendees of his session left with a much better understanding of these issues and how to cope with them.

Daniel McFadden led a panel discussion on Model Calibration and Statistical Analysis. This session (and indeed most of Monday's program) focused on the logit model. This emphasis was quite appropriate, because a large part of our improved ability to forecast travel demand stems from our improved understanding of and facility with this robust model. Professor McFadden's presence was also appropriate, as he is the key figure in the theoretical development of the logit model for transportation planning.

The first question anyone seems to ask about a model is, "How do you calibrate it?" One of Logit's nice features is its economy in calibration, thanks to some fairly modest data needs, some short-cut methods for the two-choice case, and some good software packages for the more common complex cases. Professor McFadden discussed logit calibration software like his Quail package and UTPS program ULOGIT. He illuminated the group on relevant statistical tests of "goodness of fit."

Antti Talvitie and Steven Lerman were discussion coleaders for a session on Supply Variables and Equilibrium. By "supply" variables, the transportation modeler means "transportation" variables, i.e., the variables that depict the quantity and quality of the transportation service. Of course, the performance of this service (i.e., the value of certain transportation variables) often depends on the quantity and nature of the persons, freight or vehicles transported. This "demand," however, is itself influenced by the performance offered. This circularity of service and usage has been liberally likened to the classical supply-demand equilibrium model of economics and is so named. No matter what it is called, it is a tough problem—particularly for the person who forecasts disaggregate demand for disaggregate "supply." The level of theoretical development is very low, and computer costs are very high.

Professor Talvitie has made interesting breakthroughs by depicting the transportation provided as a series of service equations, which are functions of demand levels and vice versa. In this simplified framework, powerful equilibration algorithms can be brought successfully to bear.

Professor Lehrman shared his rich experience in travel forecasting. Of particular interest was work in dial-a-ride where supply and demand interactions are probably more crucial and volatile than for any other mode. By apprising the attendees of how he met this and other formidable challenges, he impressed us all with his progress in this difficult area.

Frank Koppelman directed the panel discussion on Aggregate Prediction from Disaggregate Models. As was the case for all the panels, this one enjoyed the leadership of a superexpert. Professor Koppelman's Ph.D. dissertation covered this very topic, which is a thorn in the thumb of every disaggregate modeler. While it makes perfect sense to calibrate with a small disaggregate sample, forecasts have to be aggregations of a 100 percent "sample."

To date there are roughly three approaches to this problem. The simplistic approach of inserting average values into the independent variables of a disaggregate model is generally invalid and can render ridiculous results. Monte Carlo techniques, which randomly generate values for independent variables from their probability distributions, have theoretical appeal in their simplicity and consistency, but their computer costs are high. Perhaps the most promising approach is mathematical transformations of the disaggregate model formulation to its aggregate counterpart. Most often, however, these transformations require some simplifying assumptions; the implications of these assumptions regarding forecast consistency are quite unknown. Lest the reader be discouraged, the progress in this area has been significant, as Professor Koppelman made evident to his panel.

William Tye led the panel discussion that was the most fun. In it he described some quick and dirty Manual Forecasting and Sketch Planning Methods. Due to the creative efforts of Dr. Tye and others who have rediscovered the desk calculator, these back-of-the-envelope methods are valuable for at least two reasons. They are very educational, exemplifying important interactions and planning issues to students, planners and decisionmakers alike. Also, they are the best way to take a first look at any proposed transportation change. They can act as a cheap sieve, which separates innumerable alternatives into two groups, good and bad. The bad are eliminated from further consideration. The good are subjected to more rigorous (i.e., nonmanual) analyses. While these methods are naturally very popular for their simplicity, we were warned that they cannot by themselves yield explicit plans due to their simplistic assumptions.

Moshe Ben-Akiva led a panel discussion on Forecasting in Large Model Systems. At the opposite pole from Dr. Tye's, this session not only assumed we had a computer, it also assumed we had a very big and very fast computer. Professor Ben-Akiva (who in my opinion is the single most successful researcher in applying advanced techniques to problems of real-world scale and complexity) described a very comprehensive forecasting system he implemented. In it, he demonstrates how to deal with those hard problems of competition for the auto, trip chaining, and parking. Like the other sessions, and indeed the whole day, we all found it extremely enlightening and encouraging.

Paul Roberts summed up the day with a dinner speech that reflected the optimism we all felt. Now, we do in fact have a much better set of demand forecasting tools. Planners everywhere should look into this set and use the one that can satisfy their need for precision and accuracy within their cost and time constraints. These tools can't solve all of their problems, but they solve a lot more of them than was possible five years ago. Like this sunny day in December, this news was longed for and warmed us all.

II. TRANSPORTATION EVALUATION METHODS

Joseph L. Schofer, Professor of Transportation
Northwestern University

Emerging Methods in Transportation Evaluation

Joseph L. Schofer
Professor of Civil Engineering
Northwestern University

ACKNOWLEDGEMENT

Several people provided invaluable intellectual support in the preparation of this paper. David Gendell identified key elements in the literature; Edward Beimborn contributed a number of useful concepts regarding decisionmaking processes. Scott Rutherford offered special insights into the role of evaluation in specific planning concepts. David Schulz served as a colleague and critic in the development of some of the ideas in this paper. My father, August Schofer, applied the wisdom of his experience to a review of an early draft.

INTRODUCTION

This paper presents a strategic, tutorial review of the state of the art of evaluation in transportation planning. It is *strategic* in that it provides an overview of some of the macroscopic issues, rather than getting into mechanical details of particular techniques; it is *tutorial* in that it takes positions regarding what approaches are most desirable to follow in certain contexts.

In general, no uniform or standard technique is in use for transportation evaluation in the United States. While there is much similarity in what is done in support of specific decisions across the nation, it is common for evaluation studies to begin with another review of the state of the art, and another search for improved procedures. The federal government has offered only limited guidance in the selection of evaluation procedures, focusing principally on matters of detail (e.g., what factors must be considered) (1) and matters of process design (e.g., who must be involved, what reports must be prepared) (2). Only recently has the Department of Transportation moved in the direction of suggesting specific measures of effectiveness for use in transportation evaluation (3).

The absence of standard approaches reflects both the fluidity of the state of the art and the fact that no methods are available which solve even a majority of the difficult problems of evaluation. These premises suggest the need for practitioners to consider the merits of methods in use today and to initiate an organized effort to learn from each other, so that the process of evaluation will be generally advanced.

A review of a number of examples of current evaluation practice in transportation planning suggests that a wide range of approaches is being applied—at widely differing levels of quality. It also indicates that there are numerous opportunities for improvement in evaluation methodology. Some of the most significant improvements appear to be achievable through rather simple changes in philosophy and methodology. It is changes of this nature which are the principal focus of this paper.

DEFINITION AND MEASURES OF SUCCESS OF EVALUATION

Evaluation is a process of creating, analyzing and organizing information to support decisionmaking—the choice of an action from among a set of alternatives. It is *not* decisionmaking, nor is it usually a way to find the ultimate value of alternative courses of action—for that *is* decisionmaking. Evaluation is the technical process which links analysis, planning and design with decisionmaking. As such, evaluation is close to, must interact with, and should influence both planning and decisionmaking.

To evaluate *evaluation* itself—that is, to know what good evaluation is—we must find answers to questions such as the following:

1. Does the information provided to decisionmaking offer serious, constructive guidance to the choice process at a reasonable cost? That is, would the choice of a course of action be much different without evaluation?
2. Does the information provided to decisionmakers clearly reflect the true, salient characteristics of the alternatives?
3. Is the information provided sufficiently comprehensive to include all important, likely consequences of choosing a course of action?
4. Is the information presented sufficiently clear and succinct to ensure decisionmakers will understand and use the information provided?

5. Are the relationships between decisionmakers and planners sufficiently strong to assure that each knows the others' needs, abilities, and limitations, and thus each has confidence in, and finds credibility in, the actions of the others?

While this discussion will not answer these questions, the questions themselves suggest an approach to designing and assessing specific evaluation efforts. There is no successful formula for designing evaluation efforts which lead to good answers to these questions. But it is likely that an appreciation of the questions, and the pursuit of their answers, will lead to better evaluation.

EVALUATION AND DECISIONMAKING

Evaluation cannot be separated, intellectually or administratively, from planning; for virtually every step in the planning process should lead to the preparation of information to support decisionmaking. Thus, the approach to evaluation should govern, to a large extent, the nature and structure of the planning process itself, the data utilized, the tools applied, and the factors considered, predicted and tested. Where the evaluation process is designed as an afterthought, as a tail to be wagged at the end of a planning cycle, it is common to find that the "right"—most useful—information is not produced; that decisionmaking is less effectively supported by technical planning; and that, in the eyes of some planners, decisionmaking is "very political," i.e., "they didn't listen to us."

Similarly, good evaluation cannot be isolated from decisionmaking, as illustrated in the questions listed above. It is this relationship between the (technical) evaluation process and the (political) decisionmaking that most commonly escapes planners who design and conduct evaluation studies. Because the role of evaluation is to provide constructive informational support to decisionmakers, understanding the nature of decisionmaking and the needs and capabilities of those who are responsible for it will lead to better evaluation: this is simply a matter of understanding the nature of the market for the product.

It is common to find evaluation processes which are isolated from decisionmaking. Such processes often fail to identify and respond to the most important issues associated with a decision. As a result, decisions are sometimes made without regard to the products of the evaluation, suggesting, at a minimum, that planning resources have been wasted; in some cases, the results are poor decisions which can impose significant—but avoidable—social costs.

There are two extreme modes of isolation of evaluation processes. On the one hand, some planners/evaluators choose to ignore the decisionmaking process, perhaps because they do not understand it, but sometimes because they find it to be "too political, too anti-intellectual, too 'dirty'." Unfortunately, these attitudes can become self-fulfilling prophecies: ignoring that "too political" decision process can lead to a failure to provide the kind of enlightening informational support which might result in a *less* political decision process. Sometimes the result of this attitude toward decisionmaking is an excessively complicated evaluation procedure, involving intellectually high-powered tools which do not respond to the information needs and capabilities of decisionmakers. Such products can contribute to full employment for planners, but they often sit on the shelf while political decisions are being made elsewhere.

On the other hand, some planners assume that decisions are made prior to, or at least independent of, the planning and evaluation processes. Thus, the nature of evaluation becomes unimportant and, in some cases, the evaluation products reflect this attitude: evaluation is unstructured, information provided fails to describe the key aspects of alternatives, and decisionmakers find no guide at all in evaluation reports.

Many situations fall somewhere between these extremes. Decisionmaking is rarely examined systematically by the evaluator for the purpose of tuning the evaluation process to meet the needs of its intended market. Decisionmaking is assumed to be mysterious, political and unaffectable. While those accusations may, in some cases, be true, it is best for the planner to treat them only as hypotheses, and to work to refute them. Through a serious effort to understand the need of decisionmakers, we should attempt to show that decisionmaking is transparent, reasonably objective and consistent, and subject to guidance through the provision of good, clear information. It is probable that, through good evaluation, these, too, will become self-fulfilling prophecies.

UNDERSTANDING DECISIONMAKERS' NEEDS

Thus, both evaluation and decisions themselves may be improved if the planner can develop a better understanding of the information needs and processing capabilities of decisionmakers. Early and continuing interactions can productively include meetings to discuss major factors likely to influence future choices, the kinds of information and formats preferred, and the perspectives of decisionmakers on their roles. As the evaluation itself gets closer, it can be useful to review and discuss mockups of evaluation presentations; during the evaluation, decisionmakers must be encouraged to criticize the information presented, to make suggestions for improvement, and to request specific items not yet provided. Efficient cycling between evaluation and decision can sharpen the choice and establish it on strong foundations of fact.

A serious problem with this suggestion is that, in many cases, decisionmaking processes are highly implicit, and thus identifying and interacting with decisionmakers in advance of the choice may be impossible. This is common where the decision involves many jurisdictions, such as in the adoption of a regional plan or program. Frequently, choices are controlled by covert as well as overt decisionmakers—some of those having a major influence on the choice may not be visible to the public eye. The answer, in this case, must still lie in organized consideration of decisionmaking needs, even in the absence of the ability to sit down with decisionmakers. This might be accomplished by reviewing past decision processes and the people and factors affecting them. It might be aided by considering the objective, required procedures which must be followed to achieve implementation of a proposal; by tracing the agencies and individuals necessarily involved in those choices; and by seeking out the procedures, policies, and perspective of each of those factors.

Another problem in the interaction with decisionmakers is that it can be difficult to get their attention; most decisionmakers are already overloaded with information, and it is an added burden to ask them to relate to some of the more technical aspects of evaluation design. Indeed, in some cases it has proven difficult to get the attention of decisionmakers so that they might review and utilize evaluative information in the choice process itself. While there is no obvious solution to this problem, some promising approaches exist. For example, it is unwise to assume at the outset that decisionmakers are too busy or too disinterested to discuss their information needs as they influence evaluation design. It is worth attempting to establish such a dialogue. In addition, it may be useful to approach decisionmakers with the notion that evaluation (and indeed, planning) is intended to make their work easier and their performance better. To accomplish these ends, of course, it will be necessary to specify the needs more precisely. To the extent that the decisionmakers participate in the evaluation design, the products of the technical process should be more responsive to their needs.

Ultimately, the best way to attract and retain the attention of the decisionmaker is to provide the information he or she needs, in formats which are simple, clear and understandable. This general specification should be applied not only to the information presented to guide the choice of a transportation alternative, but also to information presented earlier on the planning process, evaluation, and how they can help decisionmaking.

The planner should recognize the possibility that, in some cases, decisionmakers themselves elect to operate in a vacuum, isolating themselves from the analytic contributions of planning, perhaps because of the political context, because they suffer from severe information overload, or because they fail to understand the process and products of planning. This is sometimes exemplified by cases where decisionmakers insist on adopting and holding to a course of action which has become clearly unattractive in an objective sense. Getting "backed into a corner" and being unwilling to change course is sometimes (not always) a sign of insufficient, inaccurate, or irrelevant information about alternative actions. While it is difficult for the planner to modify the political environment, he or she can probably adapt analyses and presentation formats to break the barriers of intellectual isolation with sharp information.

In some cases, planners *are* the decisionmakers. This can happen when a technical professional attains an elected or appointed office or where decisionmakers ask for specific action recommendations from planners. In such cases it is useful for the concerned individuals to understand clearly both of these roles, and to view them as distinct, but interrelated, steps. Otherwise, there is the danger of the technical process, with its limitations, having an undue influence on the choice. The decision can become a way of defending and justifying the technical products of planning and evaluation. Decisionmaking in such circumstances can sometimes "turn inward," ignoring the broader concerns of its true constituency.

GENERAL SPECIFICATIONS FOR DECISIONMAKER INFORMATION NEEDS

Evaluation will be improved to the extent that *each* planner gives serious, systematic consideration to the needs and capabilities of the decisionmakers with whom he or she works. Yet, it is possible to make some generalizations regarding the information needs of decisionmakers. For example, the level of sophistication of the information presented to decisionmakers should be appropriate to their ability to understand it. If decisionmakers fail to understand the information presentation, then something is wrong in the evaluation process. It is not impossible to develop a sense of the decisionmakers' understanding of evaluative data; one easy way is to ask about it.

Specifications on the technical sophistication of evaluative information should also be used to test the *processes* by which that information is prepared. Where decisionmakers cannot or do not understand the processes of analysis and synthesis of information, they may lack confidence in that information and be unwilling to rely upon it (4). Thus, using a highly sophisticated nonlinear aggregation scheme to collapse the dimensionality of evaluative information may be counterproductive if decisionmakers choose not to be involved in, or to understand, the complexities of the value-laden aggregation process. This suggests a need for balance in the sophistication of *each* aspect of the planning/decisionmaking process.

Decisionmakers should be provided with succinct, sharp descriptors of the most important factors likely to affect their choice. Determining those most important factors calls for close interaction with the decisionmakers themselves. Yet, this does not relieve the evaluator of the responsibility for pointing out important impacts which are not initially of interest to decisionmakers; the evaluator must function as a guide to the pitfalls and opportunities of the alternatives.

Focusing on the most important factors will often mean leaving out many interesting evaluative measures not likely to influence the choice. Yet, it must be remembered that parsimony is an essential virtue when it comes to retaining the attention of busy decisionmakers. Furthermore, what is an exciting measure to an analytic planner can often be obscure, uninteresting, and possibly even meaningless to a decisionmaker. What is most important depends heavily on one's perspective, and the perspective of the decisionmaker and his or her client is what's relevant here. Thus, regional measures alone may be of little interest to the representatives of one jurisdiction; similarly, while a transit manager may wish to think in terms of passengers per revenue mile, political decisionmakers may prefer to consider annual subsidy requirements and ridership on each route (5).

Information presentations must always balance brevity with comprehensiveness; achieving such a balance may be difficult. A truly comprehensive information set may be so cumbersome as to jeopardize the value of all of the information in it. Some information, which may be important to certain elements of society, must always be left out of the choice to make the information-processing task manageable. Schemes which aggregate some or all of the evaluative information into macroscopic indices or "supernumbers" attempt to achieve both comprehensiveness and brevity. Most fail because, in the aggregation process, they cover up salient facts which must be exposed if the decisions are to be responsive to the issues at hand (6).

Hierarchical or layered information structures appear to offer significant promise as ways to present comprehensive evaluative information without compromising the ability of the decisionmaker to digest it (7). Within such structures, the initial layers of information contain the key measures likely to be most significant in the choice. Successive levels or layers provide additional measures which add detail to the first group (e.g., qualitative descriptors supporting quantitative measures, indicators of distributional effects, etc,) and/or which extend the coverage of the measures set to make it more comprehensive. Table 1 illustrates a simple, hierarchical measures set. The hierarchical presentation may offer the decisionmaker all of the available layers of information, allowing him or her to select the level(s) at which he or she will function. Hierarchical structures can be increased in their responsiveness to decisionmakers' needs if supported by efficient query-response capabilities, wherein decisionmakers can request and receive richer information sets, clarifications on data provided, or measures from different perspectives.

TABLE 1. Hierarchical Measures Set

Brief Objective Statement	*Principal Regional Measures*	*Disaggregate (Subregional) Measures*	*Secondary (Supporting) Measures*
Save work trip travel time	Average work trip length (minutes)	Average work trip length by zone (minutes)	Average work trip speed
Provide transit system coverage	Percent of trip origins within 10 minute walk of transit entry	Percent of trip origins within 10 minute walk of transit entry, by zone	Transit coverage (route) map
Provide safe system	Total annual accidents	—	Total annual accidents per passenger mile
Reduce adverse noise impacts	Number of residents exposed to peak hour noise ≥70 dBA	Map of ≥70 dBA noise contour	—
Avoid relocating residents	Number of households relocated	Number of households relocated by income group Map of households relocated	Total monetary cost of relocation
Create jobs through construction, operation	Jobs created (person years)	Jobs created by jurisdiction	Total income due to added jobs

PUBLIC INVOLVEMENT IN PLANNING

Public involvement in the planning process will be critically important to the design and conduct of evaluation. The public has many potential roles, the most obvious and important of which is that of ultimate decisionmaker. While it is inappropriate and inefficient to take every transportation decision to the public—for that would undermine the functioning of representative government—elected and appointed officials eventually answer to the public, and thus views of the entire community are often important in evaluation and decisionmaking. Thus, when structuring information presentations for decisionmakers, the planner must appreciate the degree to which the public at large is the decisionmaker. Furthermore, the legal and ethical responsibilities for public accountability in both planning and decisionmaking underscore the need to make evaluation and its informational products not only open to public scrutiny, but also subject to that scrutiny by virtue of their transparency (8).

The evaluator should also consider the possibility that the public may sometimes serve as the best estimator of certain impacts of proposed courses of action, and thus may play a technical role as an information source, just as a simulation tool might serve as a predictive device (9). The public as expert is likely to have a well-defined, highly structured interactive role in the evaluation process (10).

But public involvement is a problem of the planning process as a whole, and not simply of evaluation; for the public must be concerned with goal formulation and the definition of alternatives if their involvement is to be meaningful (11). It is past time to be prescriptive about public involvement; the literature and the practice are sufficiently rich with ideas and methods to eliminate the excuse that we don't know how to do it (12, 13). What is done in a particular context is, and should be, determined by local practice and policy, as well as the nature of the decisions at hand. Because public involvement is viewed as a planning issue, not simply an evaluation issue, it will not be pursued further here. Yet it must be noted that most evaluation processes will need to consider the public as decisionmakers, and some will need to consider the public as evaluators as well.

ROLE OF THE DECISION IN STRUCTURING EVALUATIONS

It is important to establish a clear and explicit definition of the decision(s) to be made prior to the design of the evaluation process, for the nature of the decision will influence such elements as the measures of effectiveness, the time horizon, the scope of the analysis (i.e., factors to be included) and the scale of the analysis (i.e., geographic or social elements to be included). It is not unusual to find planning and evaluation processes which are conducted without a reasonable understanding of what decisions are to be made. For example, short-range decisions are sometimes made on the basis of information developed for long-range planning;

subregional decisions are often made using data prepared for regional studies. The decisions themselves might be better, and the process of arriving at them may be more efficient, if the information is developed expressly to support those choices.

An important aspect of the decision itself is the perspective adopted by the decisionmaker: his "world view," or definition of that "world" which is relevant to the particular choice. In effect this means defining the boundaries of the system of concern for a particular choice (14). For example, the decision regarding implementation of a collector/distributor transit service linked to a rapid transit line would turn on different factors, and thus require different information, depending on whether it is viewed from the perspective of the transit system, the transportation system, the community in which the service is offered, the corridor, or the region as a whole. Table 2 illustrates the relationship between world views and decision factors included and excluded in an evaluation. Choice of a world view determines what factors are included in the decision, and what are left out. Such choices should be made with an explicit appreciation of their implications.

TABLE 2. World Views, Impacts Included, and Impacts Excluded in Evaluation: Bus Transit Service Improvement Example

World View	*Example of Impact Included*	*Example of Impact Excluded*
Bus service where improvement occurs	Ridership change on affected route	Ridership change on competing and complementary transit routes
Transit system	Net change in transit ridership	Change in congestion due to diversion of people from auto to transit
Transportation system	Change in overall travel costs	Effect of service improvements on CBD retail sales
Community	Effect of service improvement on CBD retail sales	Effect of subsidy requirements on State transport budget

Typically, decisionmakers will not voluntarily address the issue of selecting a perspective for the choice. Here there is a need for the planner to step in, offering guidance and, where necessary, making a choice of his or her own. Such a choice may not be difficult where the decisionmakers and the processes they follow are well known. Otherwise, interaction with the decisionmaker, perhaps using examples to suggest the significance of making a given choice of world view, is necessary. When in doubt, it is often useful to expand the boundaries of the system of concern to the *next* higher level, including *more* factors than may be necessary at the outset. It is almost always easier to narrow the focus of the analysis after it begins than to broaden it.

EVALUATION RELATED TO SPECIFIC DECISIONS

Long Range, Regional Plan Decisions

Selection of a long range plan is a decision commonly supported by formal evaluation processes. Such a choice is usually made by a group of decisionmakers representing various jurisdictions but attempting to operate from a regional perspective. Since decisions about long range plans are usually "reversible," they are sometimes easier to make and less political than short range, implementation decisions. The choice of a long range plan has been based principally on measures of long range and large scale impacts (15).

Where such choices have been easy it was often because they weren't very important; their reversibility meant that no real commitments were made. The aggregate, long range focus of the evaluation process emphasized the distance between the realities of today in each jurisdiction and the speculations about tomorrow for the region. The nature of the decision allowed evaluation to be abstract, and even academic, and the resulting evaluation products permitted decisions to be conceptual, unrelated to real changes in real systems. The consequences of the choices for specific groups and communities were rarely highlighted in the evaluations, a fact which facilitated decisions because the most important, and controversial, issues were ignored. A certain magic was associated with the concept of the regional plan, a magic which was based on the greater value of doing some large scale good for the region as a whole.

But no one lives in the region; each individual functions in one or more specific neighborhoods on a daily basis. When decisions about implementing particular components of long range plans have had to be made, we began to learn in specific terms about the impacts—the positive and negative—on various interest groups. It was at this level that controversies erupted; it was here that assaults were made on the magical concept of the regional plan (16). That plan, of course, was adopted in the absence of measures of specific, local impacts; the world view was different, and so was the decision at the stage of plan adoption. Thus decisions regarding the implementation of specific facilities or services often have been in conflict with choices about long methods for evaluating TSM options in the context of, and as contributions to, the long range plan.

The TSM requirement offers another opportunity for the planner to build stronger bridges to his or her clients, decisionmakers and the public, because of the immediacy and understandability of the options involved. Experience in meeting this requirement over the next few years should lead to planning and evaluation processes which are generally more responsive to the needs of the community.

Decisions about Existing Facilities and Services

The last kind of decision to be discussed here is that focusing on assessment of an existing facility or service. The relevant choice may be among abandonment, continuation, or improvement of the facility or service; it may also include extending or transferring the facility or service to another location. Such decisions call for monitoring evaluations or before/after studies. These evaluation processes are unique in that they depend little on predictive capabilities, but heavily on precise isolation and measurement of existing performance and impacts (23). Monitoring studies may appropriately rely heavily on measures of user (or nonuser) perceptions as indicators of true system impacts. It is not unusual for the professionals faced with the responsibility to do this kind of evaluation to have no background in evaluation (e.g., those working for an operating agency).

The importance of monitoring studies cannot be underestimated, for they not only support essential decisions about the development of the transportation system, but they also provide a basis for learning about how communities respond to transportation change, and thus enhance the predictive capabilities so essential to *a priori* evaluation (24, 25).

BASELINES FOR COMPARISON IN DECISIONMAKING

All types of evaluation focus on facilitating decisions regarding choice from among alternatives. A key element of the decision is the identification of the most appropriate base for comparison. The "base case" is often called the "do-nothing" alternative. Yet rarely is there the opportunity to do nothing. Frequently, under do-nothing, minor service changes are made of necessity (e.g., facility maintenance, fare increases, etc.). Furthermore, doing nothing in the face of increasing demand will probably influence the demand/supply equilibrium: the resultant increases in the price of travel are likely to limit demand to a level below that expected under any "do-something" option. Furthermore, facilities deteriorate over time and with use, so that doing nothing may lead to gradual price increases due to unmet maintenance needs.

Thus, do-nothing does not usually mean no change from present conditions, as sometimes assumed. Use of the terms "base case" or "minimum feasible action" may more appropriately reflect the complexity of what has been known as do-nothing. It will often be necessary to design the base case alternative just as do-something options are created.

The definition of the base case must be given consideration in the evaluation design because of its potentially critical influence on both the evaluation process and the decision itself. Indeed, clever selection of the base case can be used to drive the decision in a particular direction. For example, comparing a large investment in fixed guideway transit with the costs of carrying the same demand on an expanded urban freeway system may be unreasonable. Is the expanded freeway system the *only* other option? Would it be implemented? The minimum feasible action may involve TSM applied to the existing networks, resulting in reduced demand through higher transportation costs. This is a complicated "do-nothing" option, but it may reflect the realities of the alternatives more honestly.

Finally, some evaluation studies never openly recognize the possibility of doing nothing, or close to nothing. Instead, it is all too common to assume that something (major) *will* be done, and then to focus on the choice among the do-something options. Today, doing nothing has become a serious option in almost every transportation decision, and thus it merits explicit consideration in the design and conduct of the evaluation process (26).

STRATEGIC APPROACHES TO EVALUATION

Two strategic frameworks are in use for organizing information in transportation evaluation: efficiency analysis and effectiveness analysis. While some efforts have combined the best attributes of both of these approaches, in many cases only one has been applied in support of a particular decision.

Efficiency Analysis

Efficiency analysis is concerned with describing the relationship between all inputs required to implement an alternative and all outputs resulting from that alternative. Most typically it relates the gains or benefits produced by an action to the costs of that action. Benefit/cost analysis and related methods of economic evaluation (e.g., present worth, rate of return, equivalent annual net worth, etc.) are all forms of efficiency analysis; each attempts to characterize "how much bang for the buck" each alternative produces (27). Figure 1 illustrates the general procedure for benefit/cost analysis.

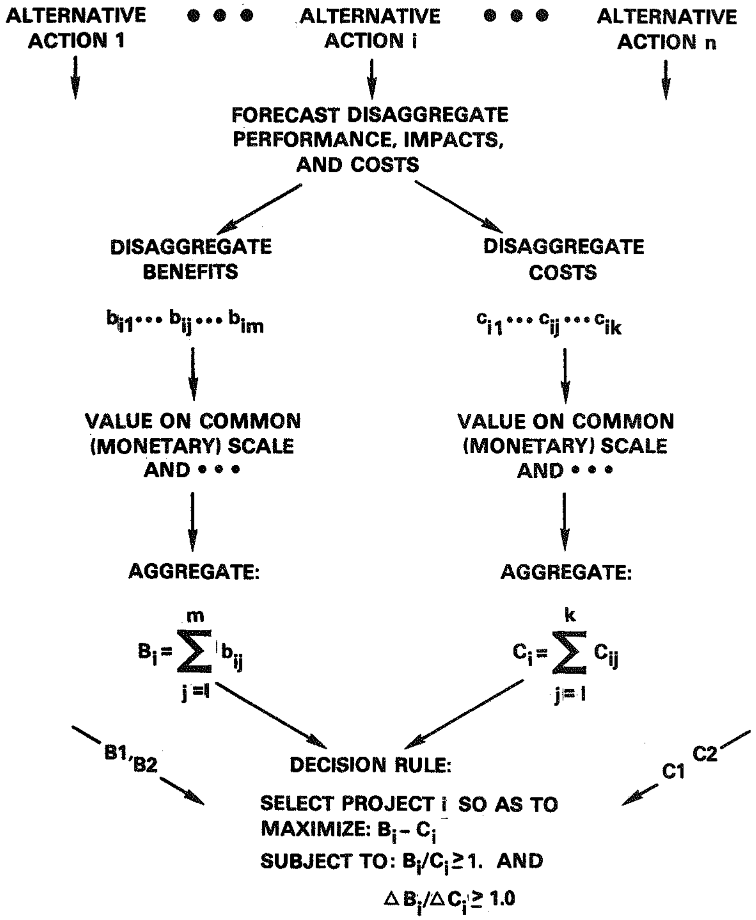


Figure 1.—General procedure for benefit/cost analysis.

Efficiency analysis is particularly attractive for use in evaluation because its product, usually a single number index, is intended to measure the relative worth of each alternative so as to provide strong and clear guidance to the decisionmaker. Indeed, most techniques actually make (or at least identify) the preferred choice; if the decisionmaker elects to accept this advice, his job is made quite simple. And the economic theory assures us that, properly conducted, the results of an efficiency analysis will lead us to the best choice of a course of action.

The limitation of efficiency analysis lies in the fact that it is impractical to conduct such a study according to the theory. In particular, benefit/cost or some other form of efficiency analysis requires that the planner identify, quantify, and value in correct monetary terms all of the benefits and costs associated with each alternative (28). Yet quantification and monetization in terms of prices reflecting the social value of resources are not universally possible. For example, the difficulties in establishing generally acceptable monetary values for such consequences as improving mobility for the handicapped, changing air quality, reducing ambient noise levels, separating socially integrated parts of a community, or even increasing the accessibility of a developing part of the region should be apparent. To be sure, it is possible to place values on each of these impacts; yet the question remains whether or not these are generally accepted and correct values (29).

To the extent that efficiency analysis leaves out major impact categories because of difficulties in measurement and monetary valuation, the results of the analysis will be biased; for example, we may tend to favor alternatives which have highly negative social and environmental impacts along with (or because of) low monetary costs.

This was the case a decade ago when community reactions to the negative consequences of urban highway investments led to policies emphasizing sensitivity to social and environmental impacts, and increased public involvement in planning (30). While it is possible to supplement efficiency analysis with nonmonetary or even nonquantitative information describing "other" impacts, the evaluation and decisionmaking processes still tend to be biased toward the monetary facts (31). This "harder" information appears to carry more weight in the minds of those who must process it, and thus it drives out more qualitative factors without regard to their importance.

Finally, efficiency analysis fails to deal directly with questions of social values: What do we really want to accomplish? What are the goals of society? How do the alternatives contribute to goal attainment? Goals are only reflected in efficiency analysis to the extent that the prices associated with various impacts reflect social values. Thus, efficiency analysis may be useful in finding alternatives which are good in a limited sense, but it may fail to lead us to the alternatives that we would prefer (32).

It should be noted that there are still important roles for efficiency analysis in evaluation. For example, in the face of major decisions in a community, it may be worthwhile to pursue a sophisticated and costly "social" benefit/cost analysis, which attempts to pursue aggressively the monetary value of every significant impact (33). In planning for TSM options, efficiency analysis may be entirely appropriate where the spectrum of benefits and costs is narrow, and where it is quite clear that benefits can be defined in terms of goal attainment. For example, benefit/cost analysis may be quite useful for prioritizing intersection improvement projects where the major benefits are savings in time and accident costs, and the major costs (negative impacts) are those for construction.

For the most part, however, it is unrealistic to apply traditional methods of efficiency analysis as the only approaches to evaluation. Indeed, it appears that few planning organizations are relying solely on such methods for evaluation. This methodological trend is supported by the tendency for decisionmakers to understand the dangers in making choices based on narrow and (to some) mysteriously complex analyses. Recent cases where decisions have gone against the more sophisticated forms of efficiency analysis suggest that the dominance of such techniques is an historical fact not likely to be repeated (34).

Effectiveness Analysis

Effectiveness analysis is concerned with measuring how each alternative contributes to attaining each of a set of prespecified objectives (35). Thus, it is objective (or goal) oriented, not input/output oriented. The objective attainment or effectiveness measures are chosen to best reflect the achievement of goals, and not necessarily to value alternatives in any single dimension (e.g., monetary units). Effectiveness analysis is something different than listing the costs, performances, and impacts of each alternative; there must be a structure imposed on such a list which responds to the values (goals, objectives) of the community (36). Thus, the list of consequences of alternatives is organized and prioritized to reflect preferred end states (37). Figure 2 illustrates a general procedure for effectiveness analysis.

While there is a high degree of rationality in effectiveness analysis, because it is founded on community objectives, its application depends on the availability of a comprehensive and meaningful set of objectives. Comprehensiveness means that the set of objectives should cover all impacts, positive and negative, which will be important in the decision. Meaningful objectives are those which suggest logical and unambiguous measures of their attainment; highly general, "motherhood and apple pie" objectives, agreed upon by all actors and met by all alternatives, do not contribute to evaluation. Similarly, objectives that

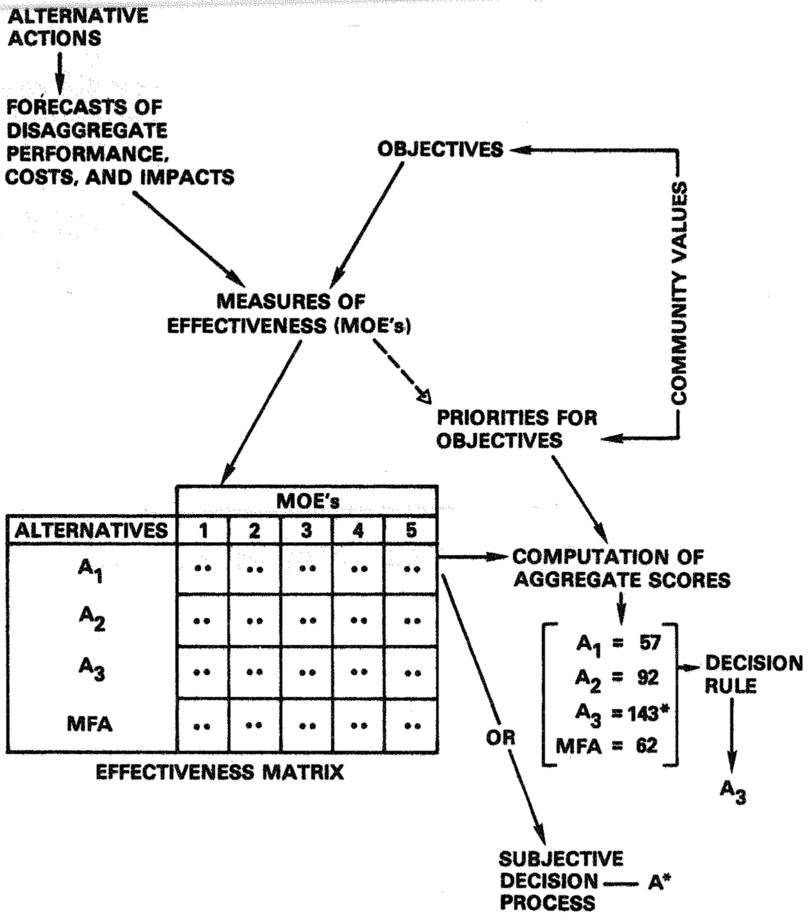


Figure 2.—General procedure for effectiveness analysis.

are equivalent to alternative courses of action should be avoided, since they determine the decision prior to evaluation.

Formulating Goals and Objectives

Organized efforts to formulate goals (more general) and objectives (quite specific) can be difficult, but are quite feasible, and numerous techniques have been applied with considerable success. Some of these include:

1. Large-scale citizen participation, simultaneously using numerous task forces, community forums, and/or open meetings, leading to citizen-defined and citizen-supported goals, e.g., "Goals for Dallas" (38);

2. One or more blue ribbon panels, made up of a representative group of citizens (not decisionmakers, and not merely the most influential actors), which structure draft goal statements (39);
3. Panels of decisionmakers representing the various jurisdictions in a region which attempt to specify goals (40); and
4. Analysis by the planner of formal and informal policy and goal statements made by decisionmakers and citizens of a region, for the purpose of synthesizing a representative goal set (41).

The least satisfactory goal formulation methods are those performed in secret by planners themselves and those which are not exposed to public review and critique. Thus, the last three methods described above should result in draft goal statements widely circulated to invite comments. The planner's goals are not generally those of his or her clients. On the other hand, in supplying methodological guidance to a goal-formulation process, the planner must work aggressively to point out the role of goals in evaluation, and thus strive for specificity, measurability, and alternative-abstractness in the goal set.

A more difficult job than specifying goals is deciding the priorities or relative values associated with attainment of each. Such decisions are highly subjective and generally must be made through public involvement and political choice. Some feel that it is impossible to establish a set of such priorities prior to the assessment of specific alternatives, for only when dealing with alternatives and their characteristics can decisionmakers define their priorities for performance, impact, and cost characteristics (42). Under such conditions, the decision regarding goal priorities is contiguous with the decision to select a particular alternative (43).

Still others are willing to accept the notion that priorities on objectives can be defined in advance of evaluation of the alternatives. Where such priorities can be defined in explicit, quantitative terms, effectiveness analysis can culminate in the application of a weighting/rating scheme which aggregates the measures of effectiveness for each alternative into a scalar score (44). Thus, the product of this approach is similar to the product of efficiency analysis. But this form of aggregation has serious disadvantages for decisionmaking, not simply because it requires *a priori* weights for objectives, but because it covers up the detailed characteristics of alternatives which may be critical to the choice (45). In general, effectiveness analysis requires neither explicit prioritization of objectives nor the development of aggregate effectiveness scores. Disaggregate effectiveness analysis, where decisionmakers are presented with a *list* of effectiveness measures for each alternative, is probably the most reasonable approach to contemporary problems of transportation evaluation.

Cost/Effectiveness Analysis

Cost/effectiveness analysis is a form of effectiveness analysis where the amount of effectiveness (objective attainment) is explicitly related to the monetary cost of each alternative. Where effectiveness is described through the use of aggregate indices, cost/effectiveness is reported as a single ratio for each alternative; where effectiveness can be defined in the same units as cost (i.e., monetary units) that ratio becomes (the inverse of) the benefit/cost ratio. Cost/effectiveness analysis was developed, however, to apply to cases where costs and effectiveness *cannot* be valued in the same dimensions (46).

Figure 3 is an example of a simple cost/effectiveness presentation. The diagram does not designate the preferred alternative. It does illustrate the *possible* trade-offs between levels of effectiveness and costs. These are called *phenomenological* trade-offs, and are determined by the attributes of the alternative (47). The decisionmakers' *willingness* to trade-off effectiveness and costs is their rate of *value* trade-off or utility function. The collection of undominated alternatives at a given cost level (e.g., $100) represents an efficient frontier (e.g., A_1, A_2, A_4); only alternatives on the efficient frontiers should be considered for implementation (48). Whether it is *worth* moving from a lower to a higher cost efficient frontier (e.g., $100 to $200 per day) depends on whether the increase in effectiveness is worth the cost increment to the decisionmakers.

A special form of cost/effectiveness analysis can be applied to alternatives where the effectiveness (or the costs) of each can be fixed at the same level. For example, consider the case of a corridor level alternatives analysis where it is possible to define several alternatives having the same or similar levels of effectiveness. Once equivalence of effectiveness is established, minimization of costs becomes a reasonable approach for identifying the preferred choice. Such a variant of cost/effectiveness is generally useful where a unitary "mission" can be defined for all alternatives. Where the missions of alternatives vary significantly, the decisionmaker is left with the burden of selecting that action which strikes the most reasonable balance between the arrays of costs and effectiveness (49).

EFFECTIVENESS, EFFICIENCY, AND DECISIONMAKING

Effectiveness analysis does not provide decisionmakers with the explicit guidance regarding choices offered by efficiency analysis. The latter would show, for example, the range of simple and incremental benefit/cost ratios for alternatives, and would thus identify those

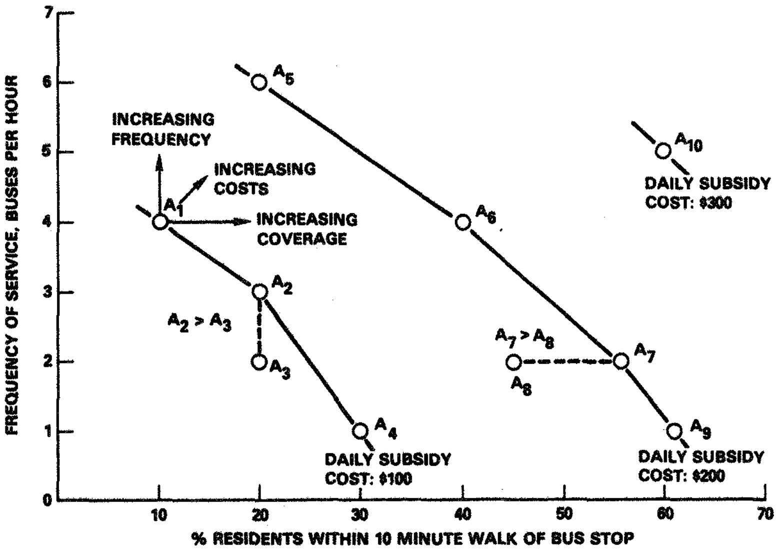


Figure 3.—Example of cost/effectiveness analysis—alternative bus transit services.

worth the investment and that which is most worth its cost (50). Effectiveness analysis often can do no more than show the array of effectiveness measures for the alternatives. This leaves the decisionmaker with the responsibility of ranking the alternatives and making a final choice. And while effectiveness analysis might lead the decisionmakers to the most effective alternative, it cannot assure them that the overall benefits justify the expenditure. This "go/no-go" question is only answered subjectively in the act of deciding on a course of action. Efficiency analysis, where all of its assumptions are met, does promise to determine whether it really is worth investing in a particular alternative. As stated above, however, those assumptions are rarely met in fact, and thus the credibility associated with investment choices supported by efficiency analysis is only apparent.

Effectiveness analysis can, of course, point directly to the most effective alternative if quantitative priorities for attaining the various objectives have been developed, permitting computation of scalar effectiveness scores; or if a *dominant* alternative is found, one that performs as well as or better than all others on every measure of effectiveness. Still there is no assurance that the preferred alternative is *worth* its cost.

A number of applications have tried to overcome the measurement and valuation limitations associated with efficiency analysis in order to take advantage of the clarity and definiteness of its products. The British, for example, have pursued a number of thorough benefit/cost studies wherein exhaustive efforts were made to capture, quantify, and monetize virtually all of the benefits and costs to society produced by the alternatives. The study of the Victoria Line Subway in London by Foster and Beesley and the site-location study for the Third London Airport by the Roskill Commission offer two good examples (51, 52). The former was an *ex post facto* research study, the latter an effort in support of a decision. Such studies underline two significant deficiencies in the analysis method. First, monetary values placed on many of the consequences are estimated using non-market methods, and thus are subject to serious question, suggesting that the choice of alternatives is not at all precisely defined. Second, and more important, the complexity of the analysis methods, and the aggregate nature of their product, limit decisionmaker confidence in the results and cover up details of the benefits and costs, including distributional effects, which are essential elements of the decision (53). Thus, the relevance of such studies to contemporary decisionmaking may be rather limited.

None of the arguments against efficiency analysis should discourage the use of valid economic tools to identify measure, and monetize those impacts of transportation change to which such methods are applicable. These certainly include changes in the cost of travel; one may also choose to use economic measures to evaluate travel time and accident savings. It is essential, however, to avoid assuming that those impacts measurable in economic terms are the most important consequences of transportation investments.

Finally, there appears to be some renewal of interest in efficiency analysis as an evaluation strategy for transportation decisionmaking. To the extent that such methods can be made more comprehensive and responsive to issues of value, they are worthy of pursuit. Still, the best benefit/cost analyses do not recognize goals and policies, and thus may not be sufficiently supportive of decisionmaking; they invariably leave out some consequences of alternatives because of measurement and valuation limitations. The result may be a divergence between evaluation products and decisionmakers' needs and interests; such a divergence can, over a period of time, undermine the credibility of the planning process as a whole.

DEFINITIONAL FRAMEWORK FOR MEASURING EFFECTIVENESS (54)

Effectiveness is the degree of objective attainment associated with an alternative action. An objective is a specific end state or target to be hit by the action to be selected. A good objective should be specific enough to suggest one or more ways to measure its attainment. As stated, objectives should not be defined as solutions to problems. Goals are more general than objectives; they are generalized end states or directions of desired movement. Goals typically are too general to suggest unambiguous measures of their attainment.

Criteria are ways to measure the degree of attainment of specific objectives by alternative actions. Thus, criteria are equivalent to measures of effectiveness. It is desirable to find quantitative criteria, for they are easier to manipulate, present, and (if necessary), aggregate. However, to ensure coverage of all relevant objectives by the set of criteria, it may be necessary to include nominal, qualitative, or judgmental criteria. Defining criteria is a technical task, but it warrants decisionmaker review to ensure that the measures selected are appropriate to the objectives. Defining goals and objectives, while calling for administrative and technical guidance from the planner, is a social and political task.

Standards are the minimum (or, as appropriate, maximum) acceptable levels of objective attainment; that is, they are the pre-specified limiting values of criteria. If criteria are test scores, standards are the minimum passing grades. Standards simplify many aspects of evaluation and decisionmaking, for they lead to simple, pass/fail choices. They are most appropriate where the phenomena to be evaluated are well known and understood, and thus where there are really no major decisions to be made. For example, it is reasonable to use standards to specify minimum transit-vehicle door widths or the proper superelevation for highway curves.

Standards are inappropriate, and sometimes dangerous, where there really are important decisions to be made, and where such choices cannot be taken prior to a careful assessment of the options. Thus, where significant benefit/cost trade-offs must be considered, adoption of arbitrary standards can assume away the essential issues. It is not necessary to establish and apply standards in evaluation. Where they are defined, it is important to ask why they are specified at a given level; only if the justification is objective, universal and irrefutable, is the standard reasonable.

Some recent evaluation studies have carefully pursued the development of standards based on what levels of effectiveness the community seems willing to accept, as well as on legally adopted standards which must be met (55). Such an approach can be quite useful where the list of criteria is long *and* where it is understood that the standards are not generally rigid. Still, using observed behavior for standard-setting may lock a community into a behavioral pattern which is unpopular but necessary due to constraining characteristics of the transportation system. This suggests the need for flexibility in the use of standards, and it calls for the serious exploration of trade-offs.

Goals, objectives and criteria are related through an hierarchical structure. At least one objective should be defined under each goal; at least one criterion should relate to each objective. There may, of course, be more; and a single criterion might well measure the attainment of more than one objective.

In a well-structured, "classical" effectiveness analysis, goals and objectives are defined first, then measures of effectiveness are selected. Many contemporary evaluation studies, however, begin with measures of effectiveness, principally because goal formulation is thought to be impossible or uninteresting (56). The notion that measures of effectiveness should be related to objectives can still be applied in this context as a test of the usefulness of each measure. Thus, the planner can attempt to synthesize the objective(s) for each criterion; to the extent that these objectives seem logical, appropriate to the regional value set, and comprehensive, the confidence in them can be increased.

DEFINING MEASURES OF EFFECTIVENESS

Good measures of effectiveness are critical to the success of evaluation. Unfortunately, selection of good measures is not yet a "handbook" problem, because the right measures depend heavily on the context (i.e., the decision and the decision process), the objectives, the world view, and the resources available for analysis and choice. On the positive side, selection of good measures of effectiveness is one of the most creative and influential elements of evaluation, in which the planner defines the overall nature of the evaluation and its products.

CRITERIA FOR GOOD MEASURES (57)

Although it is not possible to issue a set of good measures of effectiveness to every planning organization, it is useful to define and apply some criteria for the measures themselves. These criteria, or characteristics of good measures, can serve as a framework for defining and testing measures of effectiveness in operational evaluations. These criteria are as follows:

1. Each measure should be clearly and directly *linked to the objective* the attainment of which it is intended to measure. This link should be apparent to planners and decisionmakers.

2. All important *objectives should be supported by one (or more) measures*; that is, the set of measures should be comprehensive.

3. Measures should be *"sharp"*; that is, each should measure that aspect of objective attainment which is its intended target. For example, the objective "provided transit coverage of base-period shopping trips" is not measured by the cost/revenue ratio of the service provided, nor by the ratio of peak to base riders; it *is* measured by route maps showing shopping and residential areas, or by the number of bus trips per hour arriving at shopping areas. In general, if an objective relates to service offered, a sharp criterion will measure service delivered. Where the objective is concerned with service consumed, the measure will be similarly focused.

4. Measures should be appropriately *sensitive to real changes* in objective attainment. This sensitivity is often related to sharpness, as defined above, and to the scale of measurement. Thus a disaggregate (e.g., trip-based) measure is a more sensitive measure of system change than corridor or network-wide measures (e.g., point-to-point running time vs. vehicle-hours of travel). Aggregate measures which don't vary much over alternatives should be avoided.

5. Measures should be *clearly understandable* to decisionmakers and the public; overly complicated measures tend to be uninformative and self-defeating. For example, cumbersome indices which attempt to combine several measures of effectiveness can be unintelligible (e.g., weighted, regional-scale averages of community-level service quality). A good test is to ask, "What does this measure *really* mean? Is an increase in its value obviously desirable or undesirable?" The absence of good answers suggests a need for measurement work.

6. To the extent possible, measures should be *objective*—i.e., they should not be arbitrary. Thus, overall performance or impact indices which weight and aggregate various quantities are arbitrary if the weights are not rationally based. On the other hand, subjective (judgmental) measures are not necessarily arbitrary if they can be linked to facts. Thus, asking judges to rate alternatives in terms of their aesthetic characteristics as good, fair, or poor, can be relatively objective if we define "good, fair, and poor" explicitly prior to the judgmental process. Arbitrariness in judgmental evaluation can also be reduced by using a panel of judges because the reliability of aggregated judgments by panels is generally higher than

reliability of individual estimates (58). In monitoring evaluations, it is especially appropriate to employ service users as judges to measure certain components of effectiveness.

7. Each measure should be considered for its *cost-effectiveness*; that is, each measure should provide a sufficient increment of guidance to the decisionmaker to justify its costs of data collection, analysis, and forecasting.
8. Measures should be *unbiased* in terms of the range of alternatives considered; that is, when pursuing multimodal evaluations, measures which apply to, or favor, only one of the modes should be avoided. This will tend to lead to measures that are at a higher level of abstraction than mode-specific measures. For example, vehicle-miles of travel can be a useful indicator of the amount of travel in a corridor, but it doesn't tell much about transit travel. Many of the traditional travel-forecasting models tend to produce measures which focus on specific modes to the exclusion of others.
9. At least some of the measures should *reflect the scale at which the critical issues lie.* Thus, if the community is concerned with the quality of service to be offered in specific neighborhoods, aggregate regional or corridor measures may be insufficient. Disaggregate indicators, including trip scenarios (e.g., walk 5 minutes to bus; wait 10 minutes, board bus with seated load, travel 28 minutes to cover 5 miles, etc.) should be considered to supplement the aggregate measures. This may be a challenging problem to solve because of limitations in models and planning resources in most agencies. Yet a scale mismatch can result in the development of information unresponsive to the key issues and, ultimately, in misunderstandings and conflicts. The careful use of judgmental methods to disaggregate measures should be considered where alternative techniques are infeasible.
10. While comprehensiveness is an important attribute of the set of measures, it is also important that, to facilitate understanding and choice, the *measures set be as small as possible.* There is no substitute for good taste in the selection of the minimum useful set of measures of effectiveness. Unnecessary redundancy of measures should be avoided.

Where and What to Measure

In defining measures of effectiveness, the planner often faces choices regarding at what point in the overall system effectiveness is to be measured. For example, we can measure service produced (e.g., buses/hour), service utilized (e.g., passengers/hour), client perceptions (e.g., user satisfaction measured through attitudinal studies), or environmental changes (e.g., land-value shifts, land-use changes, air-quality changes, etc.). This range of choice of measurement points is illustrated in Figure 4. The choice of the point of measurement ought to be explicit and well considered, for it affects not only the evaluation work load, but also the outcome of the decision process. Further, some "downstream" measures are dependent on "upstream" measures, and thus using both amounts to multiple counting.

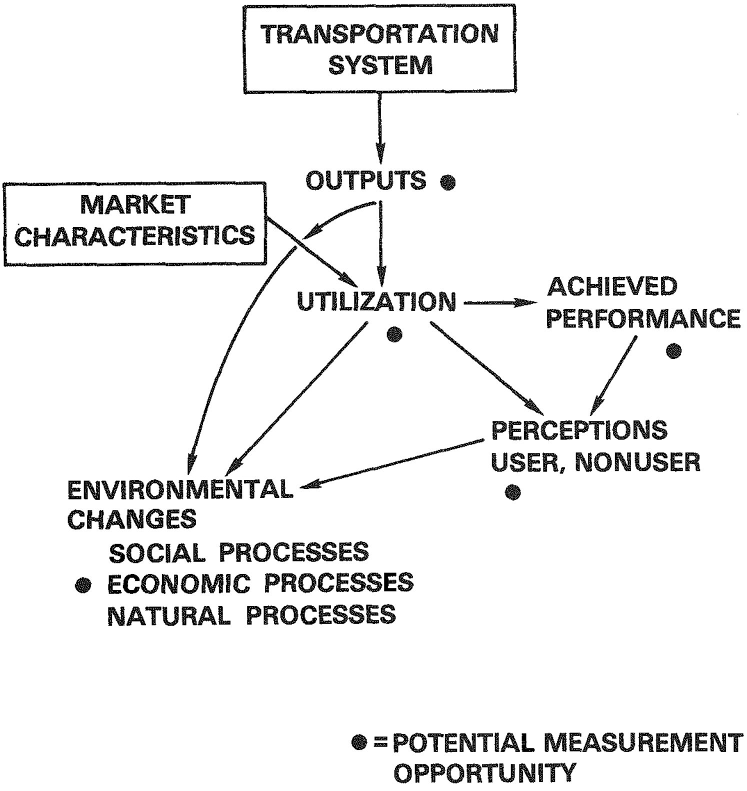


Figure 4.—Opportunities for measuring the effectiveness of transportation systems.

Thus, measuring land-use change, land-value change, *and* travel cost savings can amount to counting the same consequences three times at three different points in the system. This problem relates to the issue of world view. The evaluator needs to define a boundary around the system of concern such that each major consequence is measured only once; that boundary may be at the level of system outputs, utilization, or environmental change—but it should be consistently specified.

Measuring system or service outputs is considerably easier than measuring utilization, for the former are less dependent upon demand-forecasting tools. Where objectives focus on services offered, or where the link between services offered and utilization is easily judged by decisionmakers, working with outputs is an attractive approach. Measuring environmental changes (e.g., land use, air quality) is much more complex and unreliable due to the stepwise forecasting processes required, and the uncertainty in the intervening variables. Yet, it may be most appropriate to specify objectives in terms of environmental change, thus placing a special challenge on the evaluation process.

Careful consideration must be given to the cascading relationship between measures at different levels: estimates of changes in air quality or energy consumption are critically dependent upon both measures of service delivered and forecasts of utilization. Any errors in upstream (e.g., output or utilization) forecasts will be compounded in downstream (e.g., environmental) measures. The range of error in the measures, and the compounding of errors across measures, should be included in the evaluation and presentation of information to decisionmakers.

JUDGMENTAL MEASUREMENT METHODS (59)

Judgmental methods for measuring effectiveness are appropriate not only for treating factors subject only to subjective assessment (e.g., aesthetics, vehicle cleanliness, comfort), but also for more general evaluation efforts wherein low cost and limited data collection are preferred over objectivity and precision. Judgmental evaluation can be quite accurate, reliable (i.e., reproducible), and credible (because it is easily understood) if properly designed. Among the attributes of a good judgment design are the following:

1. A panel or panels of judges should be used to avoid, through aggregation, the biases and inconsistencies of the estimates of an individual;
2. Judges should be carefully selected to reflect breadth of viewpoints, knowledge of the subject area(s) to be evaluated, and knowledge of the community environment;
3. Judges should be well-prepared for their tasks through careful briefings regarding the process which they are to experience, the assumptions and theories involved, and the products they are expected to produce;

4. All judges should be provided with a consistent and clear set of objective information on which to base their judgments. Thus, if the issue is to judge likely user satisfaction with a new transit service, all judges need to review comprehensive, user-scale measures of the service to be offered;
5. The judgmental process itself should be carefully structured. For example, the nature of the judgment to be made at each step should be specified, and evaluators should be given adjectivally-tied, nominal scales for response to promote inter-judge consistency; and
6. It is desirable to provide judges an opportunity to interact, both before and after their judgment, to promote consistency, to share information, and to resolve significant conflicts. The judgment process itself, however, should be a private one. Iterative, interactive, judgmental evaluation can produce rich, reliable, and comprehensive information at low cost.

MEASURING TIME-STREAMS OF CONSEQUENCES

The consequences of transportation change occur over extended time periods. It is desirable to include the time-streams of performance, impacts, and costs in the evaluative information for decisionmaking. Although most public officials focus their concerns on the near term, when transportation improvements have long-term consequences, both decisionmakers and the society as a whole need to be so informed. There are two major issues associated with the treatment of time-streams of consequences: prediction of the time-streams themselves, and presentation of the information to decisionmakers.

Because prediction of consequences of transportation change tends to be complex and costly, it is rare to find planning studies in which true time-streams are forecast; instead, most studies predict performance and impacts for only one point in the future and assume that the developmental trajectories of those consequences are "well-behaved" (e.g., linear) between now and then. But because systems and services are often implemented in stages, and because the growth of demand cannot be expected to be so well-behaved, such approaches to time-stream estimation are subject to considerable error. It would be desirable, in planning studies having long term foci, to identify and predict consequences for several key turning points in system behavior over the time horizon considered in the evaluation. In this way some useful information about the *paths* to be traveled by the transportation system, rather than simply estimates of future "end states," can be developed.

Presentation of time-stream information complicates decisionmaker interaction considerably. Use of discounting to collapse the dimensionality of monetary impact measures remains a valid technique (60). The problem arises when we are faced with an increasing number of consequences not subject to monetary valuation. One of the more useful options in such cases is to find ways to present to decisionmakers descriptors of the time-streams themselves. These might be in the form of plots of measures of effectiveness versus time; but, more likely, this would be accomplished by offering scenario or tableau descriptions of system state at a few points in future time, as shown in Figure 5. The level of detail—and of certainty—would go down as the information set, shifted ahead in years, and this should be demonstrated explicitly in the presentation.

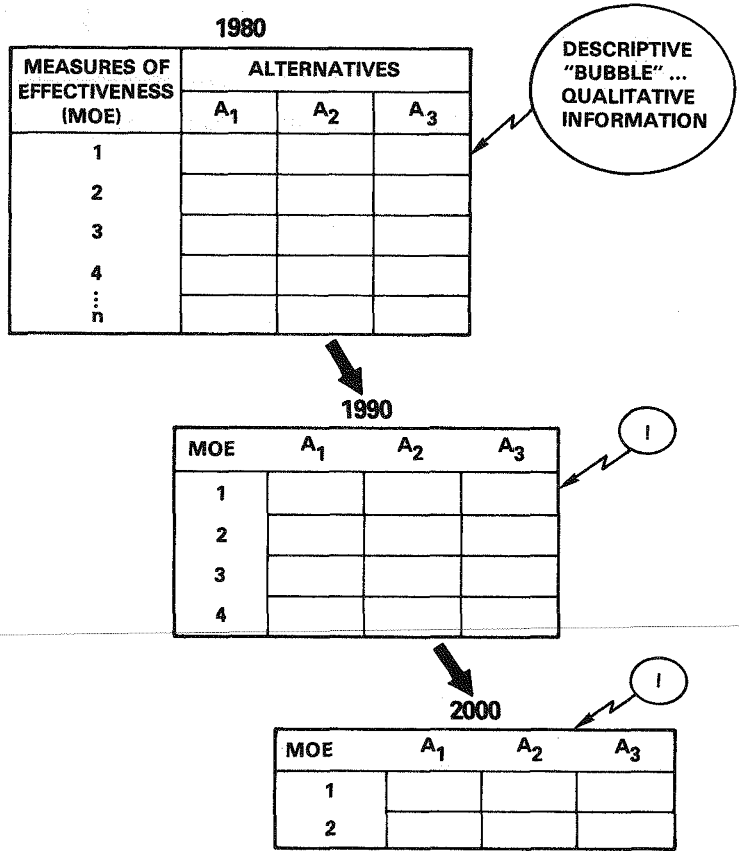


Figure 5.—Sequence of tableaus for illustrating time streams of effectiveness measures.

Shifting the level of detail and certainty over the time horizon offers a backdoor answer to the question of what time horizon to use in an evaluation. We can usually evaluate alternatives quite far into the future, but we cannot provide much (reliable) information about the far term. Providing descriptions at three points in time (0–1 year (implementation period), 5 years, and 15 years) should push forecasting tools to their limits, and should suggest the future trends rather clearly.

MEASURING DISTRIBUTIONAL EFFECTS

Transportation decisionmaking is the process of allocating services, impacts and costs to various people and places. Virtually all transportation investments produce distributional biases in the allocation of these costs and benefits. Responsive decisionmaking requires consideration of these effects. Furthermore, a number of contemporary transportation programs have as their focus the delivery of services with particular distributional biases (e.g., service to the elderly and handicapped). Thus, transportation evaluation must consider the distributional effects of alternative actions.

Accomplishing this is not easy, for it means adding still another dimension to the already complex information set presented to decisionmakers. Yet, we should be encouraged by the possibility that the political decisionmaker will be more interested in distributional effects (*who* benefits, *who* pays) than in many of the other important issues in evaluation. Indeed, political resolution has been the principal mechanism for treating distributional questions in public-sector planning and evaluation. The result has typically been an allocation of costs and benefits which is most responsive to the majority, although the political process also tends to protect the minority from extremes of misallocation.

It is likely that the political process will produce a more equitable and needs-oriented allocation of benefits and costs if better evaluative information on distributional effects is provided. This generally requires estimating some of the measures of effectiveness for each of several incidence groups; or, where such groups can be geographically defined, for various locations. Geographically-specific measures are best presented in maps. This complicates the information set, but most decisionmakers are able to comprehend maps, particularly if the scales are large enough to identify places and groups to which they are responsible.

Public involvement offers a feasible approach to both the estimation and evaluation of distributional effects of transportation actions. In this approach, community groups might be presented with objective measures of system outputs (e.g., service delivered, properties acquired, etc.), and invited to estimate community impacts and provide comments and criticism. The results could be used not only for evaluation, but also for redesign of alternative actions. Public involvement at this scale is likely to be relatively effective because it can focus on specific characteristics of alternatives as they are likely to influence well-defined groups.

In general, the increasing concern for equity in the allocation of transportation services is the strongest argument in favor of the use of disaggregate evaluation techniques which identify, measure and preserve descriptions of the benefits and costs of actions as they affect various components of the community.

TREATMENT OF UNCERTAINTY

Uncertainty is always associated with estimates of future performance, impacts and costs of transportation actions. This stems both from factors outside of the transportation system which affect it, such as energy availability/price, state of the economy, spatial patterns of activities, and social values; and from factors inside the system, such as the characteristics of new technology. Traditional approaches to planning and evaluation have ignored this uncertainty, operating from the assumption that all relevant factors could be predicted precisely and without error. Recent trends suggest that this assumption is not only incorrect, but that it can lead to poor decisions, commitments to unattractive developmental paths, high costs, and community dissatisfaction.

Recent federal regulations tend to re-focus the planning process toward the short term, and toward options with lower costs, impacts and implementation requirements, so that reversing previous decisions in the face of unexpected outcomes becomes more feasible (61). Working principally with the near term, and with flexible investments, is a useful way to deal with uncertainty. But we cannot ignore longer-term planning, for there is need for a framework within which to do short-range planning (62). Furthermore, studies of the distant future are essential if we are to be prepared with radically different technologies to apply in the face of large shifts in resource availability or in values.

Improved predictive tools can help overcome problems of uncertainty in several ways. Models which better account for the causal variables influencing a system can respond more logically to expected shifts in external parameters. Models capturing more elements of system behavior can assist in evaluating options under a range of possible futures. Still, better models will not solve many of the problems of uncertainty in planning, because the external factors which bring about that uncertainty tend to be outside the realm of both transportation and technical model-building.

Uncertainty in evaluation should be faced squarely and, within the limits of available tools, the *set* of possible levels of system performance, impacts and costs should be estimated. One way to do this is through sensitivity analyses to determine the impact on the evaluative measures of variations in those key, independent variables which are most uncertain. In a sensitivity analysis, each key variable is shifted incrementally, and estimates of the effectiveness measures are produced for each shift. A serious combinatorial problem arises when there are several such variables and a wide range for each. For example, if there are 4 variables on which the ranging is to take place (e.g., population, gasoline price, manufacturing employment, and level of transit subsidy), and if 3 levels are to be tested for each, 81 different sets of evaluative measures must be produced. Organizing and presenting this information to decisionmakers may be an impossible task.

An alternative to this simple form of sensitivity analysis is to establish a small number of scenarios, or descriptions of possible futures, in terms of all of the key independent variables. Predictive techniques can then be applied independently to each of these to produce one set of effectiveness measures for each scenario. By keeping the number of scenarios small (e.g., 3 or 4), and by defining the futures themselves judiciously, it is possible to provide decisionmakers with a reasonable picture of possible transportation futures.

Two approaches to defining the scenarios themselves have been proposed. In one, contingency analysis, futures are selected to represent the most serious challenges to the transportation system. They would reflect the most difficult situations in which transportation must perform, and the resulting effectiveness measures should indicate how well alternative actions will perform under conditions of adversity.

The other approach to scenario definition is to specify several alternative futures which together reflect the range of possibilities, from optimistic to pessimistic (63). It is not necessary, and indeed, it is counter-productive, to anticipate which of these futures is most likely to occur, for such an attempt tries to eliminate uncertainty, which is not possible; furthermore, it may reduce the decisionmakers' concern for uncertainty, which is undesirable. The result of this approach to scenario definition is a set of contexts which should suggest how alternative transportation actions will respond to the most likely range of future conditions. Unlike sensitivity analysis, however, only a few alternative futures (described in terms of all of the relevant variables) would be considered.

In both of the scenario approaches, the difficult challenge—aside from securing the resources to pursue additional evaluations for each future—is in the definition of the scenarios themselves. A promising approach to this problem is through a multidisciplinary conference setting, relying on the judgment of experts rather than reverting to more complex forecasting models (64). Such conferences can be used to define alternative futures in general, to identify the key variables associated with them, and to select approximate levels for those variables. The technical planning and evaluation process must then take over to refine variable estimates, and to predict values for the measures of effectiveness.

Derivatives of the scenario format for considering uncertainty can be applied even to simple, small-scale evaluation problems. For example, uncertainty might be characterized in terms of 2 or 3 different levels for a single key variable, and measures of effectiveness estimated accordingly. Thus, the results might show two values for each measure, one for high fuel costs and one for low.

The decisionmakers' concern under all of these approaches is to seek out the most robust alternatives—those which do best under *all* of the possible conditions—rather than the most attractive combination of future and alternative. Since it is not generally feasible to predict *which* future will obtain, actions responding in an acceptable fashion to *each* possibility need to be identified. This new strategy of decisions must be accepted by both planners and decisionmakers if this process is to be successful.

MANAGING COMPLEX INFORMATION SETS

The highly complex information set required to pursue a comprehensive evaluation as defined in this discussion presents a significant challenge not only to the planner who must develop it, but to the decisionmaker who must attempt to comprehend and use it. There are four components in this information set: the measures of effectiveness, descriptions of the incidence (distributional effects), time streams, and uncertainty associated with those measures.

It is impractical to work with all these dimensions in every evaluation. Indeed, only the largest and most complex projects may warrant detailed analyses of all of these dimensions. In specific cases, however, the planner—working with decisionmakers—must determine which of these components are important, and thus which merit primary consideration. An open-minded approach to evaluation calls for consideration of each of these dimensions at the beginning of each study. Of course, the measures of effectiveness will be the most important elements of the evaluation. Next, concerns for distributional effects can be expected to have high priority. The issues of time streams and uncertainty will be more important in longer-term planning where significantly new actions and resource constraints are possible.

In any case, the task of presenting information to decisionmakers and the public will be a difficult one, requiring a high level of creativity for success. One option which has been pursued for many years is that of aggregation—application of methods for collapsing the dimensionality of the evaluative information set to clarify and simplify the choice process. Benefit/cost and other economic efficiency analysis techniques aggregate information by arraying all impacts on a common, monetary scale.

Linear weighting/rating schemes were developed to reflect the fact that all important impacts can not logically be measured in terms of money. In such schemes, the effectiveness of an alternative in attaining a particular objective is rated, analytically or subjectively; scores across all objectives are aggregated by weighting the ratings according to an *a priori* estimate of the utility associated with attaining each objective (65). A number of non-linear weighting/rating schemes have been designed in response to the fact the utilities associated with various levels of attainment of one objective, and with trade-offs among objectives, are typically non-linear (66). For example, a decisionmaker may be willing to spend money to reduce work-trip travel time in a community; but the amount he or she is willing to pay for successive increments of time saved will decline. Further developments in the field of weighting/rating methods include tools which accommodate different objective weights for different incidence groups, and which aggregate scores over all incidence groups (67).

A major requirement of all of the weighting/rating techniques is that the utilities or weights associated with objective attainment be derived by decisionmakers and/or the public, since they should reflect community values. Yet getting the attention of these groups, educating them regarding the evaluation tools, and assisting them in estimating stable, consistent, *a priori* policy weights is a most difficult—some say impossible—task (68). To be sure, a number of such methods are in use with some success. There are examples, however, of cases where it was impossible to secure the attention of decisionmakers, where the junior staff estimated utilities, and where those who held the ultimate responsibility for choice had no confidence in the evaluation results. Generally, such approaches to aggregating evaluative information are not attractive, except where the decisionmaking process is relatively clear, and where the relationship between planners and decisionmakers is one of high confidence and cordiality, thus ensuring successful interactions for weighting tasks prior to the final decision. Where those conditions are met, it is probable that aggregation techniques may not even be needed.

Aggregation schemes try to make the decisionmakers' job easier; yet they may well make it more difficult by covering up the richness of detail which could be most influential in the choice process. Whatever simplification is achieved in decisionmaking through such methods may only be apparent. Pro-aggregationists argue that all the busy decisionmaker can cope with is "the bottom line," a single number score for each alternative. Yet decisions about actions with major impacts can rarely be made in such a simplistic way. Furthermore, a major concern of decisionmakers today is maintaining the ability to *justify* a choice—to explain in a rational manner why a certain action was preferred. Aggregate information is often insufficient for such explanations.

Practical alternatives to aggregation methods lie in creative formats for displaying rich, disaggregate evaluative information. Recent concepts in matrix displays, relating alternative actions to measures of effectiveness, offer useful ways to highlight some of the complex tradeoffs between alternatives and objective attainment. The addition of subregional displays with similar structures provides an approach to evaluating distributional consequence (69). The incidence of benefits and costs is also easily captured in maps, which offer a natural format for interaction with the public. The opportunities offered by computer-assisted information systems, and particularly interactive computer graphics, may open new paths to the presentation of disaggregate, hiearchical information sets to decisionmakers (70). The use of scenarios, rich, verbal and pictorial descriptions of community life and travel experiences under various transportation options, can convey the attributes of options in meaningful, personalized formats most likely to encourage interest among both decisionmakers and citizens. In general, disaggregate presentation formats need to be designed to meet decisionmakers' information needs by providing the data necessary to support the choice process.

The appropriate level and method for aggregation in a given evaluation study is still a topic of considerable debate. It is likely that tensions between the desire to provide a rich (disaggregate) information set and the need to make the choice process feasible will continue indefinitely. A natural, and desirable, result of these tensions is the creation of new, innovative ways of presenting evaluative information sets to decisionmakers.

CLOSURE

This discussion has offered no miracle cures for evaluation problems. Instead, it has highlighted a number of issues thought to be of importance, and it has proposed some promising strategic approaches. The tactics are left to the practitioner, who surely knows his or her problems best, and who is in the most advantageous position for applying creativity to their solution. Let this paper be a reminder that the strategic issues must not be overlooked, but must be dealt with directly to develop a strong foundation for specific, successful evaluation tasks. Figure 6 summarizes some of the key steps which should lead the planner to better evaluation.

- DESIGN EVALUATION PROCESS AND PRODUCTS EXPLICITLY AND EARLY:

  DEFINE –
  - *DECISIONS*
  - DECISION-MAKERS
  - DECISION-MAKING PROCESS
  - WORLD VIEW
  - TIME FRAME

  IDENTIFY –
  - BASELINE
  - GENERAL RANGE OF ALTERNATIVES
  - NATURE OF UNCERTAINTY
  - KEY IMPACT GROUPS
- FORMULATE OBJECTIVES, DEFINE MEASURE OF EFFECTIVENESS, CREATE PRESENTATION FORMATS
- OPERATIONALIZE MEASURES, ANALYSE AND INTERPRET RESULTS
- PRESENT RESULTS TO DECISION-MAKERS AND THE PUBLIC, INTERACT, AND RECYCLE
- KEEP THE PROCESS OPEN TO DECISION-MAKERS, THE PUBLIC, AND TECHNICAL PROFESSIONALS
- AVOID PREMATURE COMMITMENT TO A PREFERRED ACTION TO RETAIN OBJECTIVITY OF THE PROCESS
- FOCUS THE EFFORT ON THE PRODUCTS: DECISIONS AND IMPROVEMENTS IN TRANSPORTATION AND COMMUNITIES

Figure 6.—Key steps to better evaluation.

REFERENCES

1. "Public Hearings and Location Approval," Policy and Procedure Memorandum 20–8, issued by the Federal Highway Administration, U.S. Department of Transportation, January 14, 1969.
2. "Guidelines for Implementing Section 102(2) of the National Environmental Policy Act of 1969, Section 1653(f) of 49 U.S.C., Section 470f of U.S.C. and Section 309 of the Clean Air Act of 1970," Policy and Procedure Memorandum 90–1 issued by the Federal Highway Administration, U.S. Department of Transportation, August 24, 1971.
3. "Major Urban Mass Transportation Investments," proposed policy, Urban Mass Transportation Administration, U.S. Department of Transportation, *Federal Register,* August 1, 1975, p. 32546.

4. C. W. N. Thompson and R. G. Weiss, "Implementation Analysis for Experimental Social Systems," presented at the Joint ORSA/TIMS conference, Chicago, 1975.
5. B. M. Hudson, M. Wachs, and J. L. Schofer, "Local Impact Evaluation in the Design of Large Scale Urban Systems," *Journal of the American Institute of Planners,* Vol. 40, No. 4, July 1974.
6. E. A. Beimborn, "Structural Approach to the Evaluation and Comparison of Alternative Transportation Plans," presented at the 55th Annual Meeting of the Transportation Research Board, Washington, January 1976.
7. J. L. Schofer, "Evaluative Measures for the Year 2000 Transportation System Plan," Chicago Area Transportation Study, May 1976.
8. "Process Guidelines (Social, Economic, and Environmental Effects on Highway Projects)," Policy and Procedure Memorandum 90–4, Federal Highway Administration, U.S. Department of Transportation, June 1, 1973.
9. G. Fellman, "Sociological Field Work Is Essential in Studying Community Values," *Highway Research Record 305,* Highway Research Board, 1970.
10. J. L. Schofer, R. S. Gemmell, and G. L. Peterson, "Social Impact Assessment: A Sourcebook for Highway Planners, Volume VII—The Judgmental Impact Matrix Approach: A Framework for Evaluating the Social and Environmental Impacts of Transportation Alternatives," Federal Highway Administration, June 1975.
11. "Urban Panel, Project Design and Methodology," North Central Texas Council of Governments, Dallas, 1976.
12. M. L. Manheim, *et al., Transportation Decision-Making—A Guide to Social and Environmental Considerations,* National Cooperative Highway Research Program *Report 156,* Transportation Research Board, 1975.
13. R. Yukubovsky, "Citizen Participation in Transportation Planning—A Selected Bibliography," Planning Division, New York State Department of Transportation, Albany, February 1973.
14. E. N. Thomas and J. L. Schofer, *Strategies for the Evaluation of Alternative Transportation Plans,* National Cooperative Highway Research Program *Report 96,* Highway Research Board, 1970.
15. M. Wachs, B. M. Hudson, and J. L. Schofer, "Integrating Localized and Systemwide Objectives in Transportation Planning," *Traffic Quarterly,* April 1974.
16. A. Lupo, F. Colcord, and E. P. Fowler, *Rites of Way: The Politics of Transportation in Boston and the U.S. City,* Little, Brown, and Co., Boston, 1971.
17. M. Wachs, B. M. Hudson, and J. L. Schofer, *Op. Cit.*
18. "Transportation Improvement Program," Federal Highway Administration and Urban Mass Transportation Administration, U.S. Department of Transportation, *Federal Register,* September 17, 1975, p. 42976.
19. *Ibid.*
20. "Draft Annotations Pursuant to Policy on Major Urban Mass Transportation Investment," Urban Mass Transportation Administration, U.S. Department of Transportation, March 1976.

21. "Measures of Effectiveness for Multimodal Urban Traffic Management," Federal Highway Administration Request for Proposal 466, 1975.
22. "Transportation Systems Management Review," Federal Highway Administration and Urban Mass Transportation Administration, August 1976.
23. D. T. Campbell and J. C. Stanley, *Experimental and Quasi-Experimental Designs for Research*, Rand-McNally, Chicago, 1966.
24. J. L. Schofer, "Controlling Socio-Technical Systems: Needs, Problems, and a Response," Discussion Paper 13, Center for Urban Studies, University of Illinois at Chicago Circle, February 1970.
25. D. T. Campbell, "Reforms as Experiments," *American Psychologist*, Volume 24, April 1969.
26. "Guidelines for Implementing Section 102(2) of the National Environmental Policy Act of 1969 . . . ," *Op. Cit.* (note 2).
27. P. R. Stopher and A. H. Meyburg, *Transportation Systems Evaluation*, Lexington Books, 1976.
28. E. N. Thomas and J. L. Schofer, *Op. Cit.*
29. C. D. Foster and M. E. Beesley, "Estimating the Benefits of Constructing an Underground Railway in London," *Journal of the Royal Statistical Society*, Volume 126, Part I, 1963.
30. J. E. Seley, "Development of a Sophisticated Opposition: The Lower Manhattan Expressway Issue," Discussion Paper VII, Research on Conflict in Locational Decisions, Regional Science Department, University of Pennsylvania, Philadelphia, September 1970.
31. J. C. T. Mao, "Efficiency in Public Urban Renewal Expenditures through Benefit/Cost Analysis," *Journal of the American Institute of Planners*, March 1966.
32. C. Foster, *et al.*, *Lessons of Maplin: Is the Machinery for Governmental Decision-making at Fault?*, The Institute of Economic Affairs, London, May 1974.
33. Commission on the Third London Airport *Report*, Her Majesty's Stationery Office, London, 1971.
34. C. Foster, *et al.*, *Op. Cit.*
35. E. N. Thomas and J. L. Schofer, *Op. Cit.*
36. "Boise Metropolitan Area Public Transportation Plan," a report to the Ada Council of Governments, Boise, Idaho, prepared by De Leuw, Cather, and Company, August 1975.
37. J. L. Schofer and D. G. Stuart, "Evaluating Regional Plans and Community Impacts," *Journal of the Urban Planning and Development Division*, American Society of Civil Engineers, Volume 100, Number UP 1, March 1974.
38. "Goals for Dallas," Newsletter, September 1976; and, Dallas *Times Herald*, July 4, 1976, p. A-7.
39. M. Skutsch and J. L. Schofer, "Goals Delphis for Urban Planning," *Socio-Economic Planning Sciences*, Volume 7, 1973.
40. "Urban Panel Project Design and Methodology," *Op. Cit.*

41. "A Preliminary Statement of Goals for the Michigan State Railroad Plan," prepared for the Michigan Department of Highways and Transportation, by CONSAD Research Corporation, July 1974.
42. *Evaluation: 1995 Highway-Public Transportation Networks,* prepared by the Regional Transportation Planning Board, Chicago, undated.
43. N. Lichfield, P. Kettle, and M. Whitbread, *Evaluation in the Planning Process,* Pergamon Press, Oxford, 1975.
44. C. C. Schimpeler and W. L. Grecco, "Systems Evaluation: An Approach Based on Community Structure and Values," Highway Research *Record 238,* Highway Research Board, 1968.
45. E. A. Beimborn, *Op. Cit.*
46. E. N. Thomas and J. L. Schofer, *Op. Cit.*
47. R. DeNeufville, "Toward a Comprehensive Systems Analysis for Transportation Planning: An Urban Example," *Papers,* 8th Annual Meeting of the Transportation Research Forum, Montreal, 1967.
48. S. K. Agarwal and J. L. Schofer, "A Transportation Network Design Model for Regional Development Planning," *ITCC* (International Technical Cooperation Centre) *Review,* Volume IV, Number 3, July 1975.
49. D. Novick (Editor), *Program Budgeting,* Harvard University Press, Cambridge, 1965.
50. A. R. Prest and R. Turvey, "Cost-Benefit Analysis: A Survey," *Economic Journal,* Volume 75, Number 300, December 1965.
51. C. D. Foster and M. E. Beesley, *Op. Cit.*
52. Commission on the Third London Airport, *Op. Cit.*
53. N. Lichfield, P. Kettle, and M. Whitbread, *Op. Cit.*
54. M. Wachs and J. L. Schofer, *Op. Cit.*
55. *West Yorkshire Transportation Studies,* "Issues, Objectives, Measures and Standards," prepared by WYTCONSULT (Freeman Fox and A. M. Voorhees), February 1976.
56. J. L. Schofer, "Evaluative Measures for the Year 2000 Transportation System Plan," *Op. Cit.*
57. *Proceedings of the ORSA/TSS Urban Transportation Workshop,* Operations Research Society of America, August 1971.
58. J. L. Schofer, R. S. Gemmell, and G. L. Peterson, *Op. Cit.*
59. *Ibid.*
60. P. R. Stopher and A. H. Meyburg, *Op. Cit.*
61. "Transportation Improvement Program," *Op. Cit.*
62. D. F. Schulz, J. L. Schofer, and N. J. Pedersen, "New Approaches to Long-Range Transportation Planning," presented at the 56th Annual Meeting of the Transportation Research Board, Washington, January 1977.
63. *Proceedings, Year 2000 Alternative Transportation Futures Conference,* Chicago Area Transportation Study and Northwestern University Transportation Center, March 1976.
64. *Ibid.*
65. C. C. Schimpeler and W. L. Grecco, *Op. Cit.*

66. "Evaluation of Alternative Transportation Plans for the Indianapolis Region—Supporting Documentation," prepared for the City of Indianapolis by Barton-Aschman Associates, Evanston, Illinois, February 1977.
67. W. F. Keller, "Method for Development of a Mass Transit Evaluation Model Based on Social System Values," Highway Research *Record 427,* Highway Research Board. 1973.
68 M. L. Manheim, *et al., Op. Cit.*
69. J. L. Schofer and D. G. Stuart, *Op. Cit.*
70. *Application of Interactive Graphics in Citizen Participation,* Transportation Research Record 553, Transportation Research Board, 1975.

Thomas B. Deen, Chairman of the Board
Alan M. Voorhees & Associates, McLean, Virginia

Practical Considerations in Transportation Decisionmaking

Thomas B. Deen
Chairman of the Board
Alan M. Voorhees & Associates, Inc.
McLean, Virginia

When Dave Gendell asked me to prepare this paper, he suggested that I not attempt prescriptions with respect to new planning methods or additional formulations concerning travel forecasting, evaluation, or land use interaction, since these were being well taken care of elsewhere. Rather, he wanted this particular paper to be more of a change of pace, and to view the entire subject from a different perspective. I have interpreted it as license to look at the transport planning profession today, evaluate our limited successes, and take stock of the adversary, which suppresses our efforts toward real improvement.

I would suppose that if we were to ask most transport planners "why" we plan, we would get answers such as "to maximize benefits," or "to improve the distribution of benefits for transportation facilities," or "to reduce adverse environmental impacts and improve energy conservation," or "minimize costs," or "generally to improve the quality of decisionmaking about transport actions in local areas." While we may agree that all of these express our objectives, we must question the extent to which we are succeeding. Indeed we might ask "Are we having any impact at all?" If so, "Is the impact we are having to improve decisionmaking and even if we think we improve decisionmaking, does the public agree?" I would suggest that at present the answer to this is at best a qualified, tentative, and hesitant "yes" and it might well be "no." The list of rejected transportation plans and projects is sufficiently long and so well known that I need not elaborate on this point. In the past I think the technician and expert dealing with urban problems was seen in a much more favorable light than today and it is reasonable to ask, why this deterioration in our public image?

I would submit that it is not so much us as the nature of the problem, which has changed. In the past the problems were simpler than they are now in two specific aspects:

1. There was a consensus on the nature of the problem. In the past U.S. cities were much like those we see in some of the richer developing countries today where, because of rapid growth and increasing wealth, you can't get a phone installed and when you do, the other phones always seem to ring busy; water pressure is low and sometimes polluted; half or more of the area may be unsewered; power blackouts are common; many roads are unpaved; the schools are crowded; the housing inadequate; the traffic hopelessly congested; and buses overloaded. The urban professional's job is seen very clearly by both himself and the public in such situations. Namely, do something about it—find solutions to these obviously acute problems.
2. The main test of potential solutions was the test of efficiency. That is, the technician was expected not only to solve the problem but to do it most efficiently given the limited resources available. This means that he would not build a required road with too much capacity since that would waste resources, or too little capacity since that would not solve the problem, but rather he would size it correctly and he would build it with the least cut and fill and the least demolition—in short, the minimum cost solution. The problem and the guidelines were clearly understood and within such a framework the public and officials were largely willing to allow the planning and design of solutions to the technician.

Now, however, the nature of the problem has changed. No longer are we dealing with these primary physical problems, but rather secondary and sometimes tertiary problems. After all, our water flows, our phones work, our streets are paved, our lights burn, our housing exists and is usually structurally sound, our traffic moves. In short, we could well pat ourselves on the back as being party to a rather incredible accomplishment. Edward Banfield, in his controversial book, "The Unheavenly City Revisited" (1) suggests that there really isn't an urban crisis and things are probably better than they have ever been; it's just that our standards have changed as has our perception of the problem.

Our transportation problems in our cities today are part of a class of social welfare problems that professors Webber and Rittel (2) identified in a very perceptive paper a few years ago, as being "wicked" problems as distinguished from other problems which are "tame or benign." In the past our transportation problems were of the tame variety, whereas today they are in the wicked category. Tame problems are those in which (1) the objective is clear, and (2) you can tell whether the problem has been solved. For example, solving a mathematical equation is of the tame variety. It's very clear the objective is to find the value of x and having once found the value, there are ways to prove whether your answer is correct or not. Similarly, in trying to perform an analysis of a chemical compound, it's very clear as to the objectives and it's clear as to whether the right solution has been found.

On the other hand, wicked problems, which I prefer to call devilish problems, are of a different class, have different characteristics and because they are devilish they make our task difficult and perhaps even hopeless if we still have illusions that rigorous or "correct" solutions can still be found. It's perhaps useful to know our adversary which causes us all of these problems and therefore I would like to explore some of the characteristics of these devilish problems of which urban transportation is clearly a group.

1. A devilish problem is one which is interlinked to all others. For example, we no longer perceive our task as simply the solution to the transport problems of a single link or corridor, or mode, or of even a multi-modal, multi-link network problem. But we see it encompassing the problems of regulation and legal aspects as well, and are expected to solve not only these problems in the short term, but also simultaneously to examine the implications in the long term. Furthermore, we see all of this as but one link tied to a still larger network affecting local land development, housing, taxation, equity, distribution of benefits, and opportunities, pollution problems, energy conservation, and on and on. All links are tied to all other links. We are quite sensitive and no longer surprised to find that the solution to one link causes great problems in still another link and so we continually expand the boundaries of the system within which we are dealing until we know less and less about more and more and indeed may be coming close to knowing nothing about everything.

2. A devilish problem is one in which the solutions don't flow from an understanding of the problem. One of the first tasks that I had as a young professional was to direct one of the early transportation studies along the lines of the procedure manuals which had been developed by the National Committee on Urban Transportation (NCUT). One of the guiding principles of this book was that if you collected enough data and analyzed it properly you would gain insights into the nature of the transport problem and, in fact, solutions would flow from such a procedure. As part of my activities, I had developed a Coordinating Committee made up of the various jurisdictions and

constituencies involved in the area and had a massive data collection program underway. During one of the meetings of this Steering Committee, the Finance Director of the City raised his hand and asked—"Tom, what are you going to do with all this data when you get it?" I answered dutifully that it would be obvious what we were going to do with the data once we had collected and analyzed it. The answer was adequate for the moment, but I'm no longer as optimistic about such a procedure as I once was and, in fact, am convinced that the data collection effort in a transport problem must be designed at the outset in accordance with the way you perceive the problem. It's simply not possible to get data on all aspects. If you were doing a transit study in which you perceived the problem to be routes and schedules, the information on routes and schedules won't help much if in fact the problem turns out to be how to deal with large financial deficits, nor will the financial data help if the solution turns out to be auto restraints and priority lanes. In effect, we must make a guess as to the nature of the possible solution and aim our data collection effort at that; and in so doing, we will have defined the problem.

Consider the poverty problem. One may perceive it as a problem of low income, in which case data collection and analysis will be aimed at the problem of national and regional economics and ways to upgrade them; or the problem may be seen as a lack of skills in the labor force, in which case data collection and analysis will be aimed at the educational system; or we may see it as a problem of deficient physical and mental health, in which case the health care system is explored. Also the problem may be seen as the problem of spatial dislocation, in which case transportation may be involved; or it could be caused from family deprivation, or it could be deep-seated self-identity problems, etc., etc. Since we must work in a world of limited resources in terms of money, time, and even intelligence, we must pick the area in which we believe the solution falls, and in so doing, we will have defined the problem.

3. Devilish problems are those which have no end. I think perhaps I have never concluded a significant transportation study project with the feeling that I had really explored it all the way to its end. Rather it always ends as a result of running out of resources, either time or money or other resources, and we end up satisfying ourselves with expressions like "it's good enough," "it's the best we can do," or "to go further is beyond the scope of the study." Tame problems in contrast have endings and you know when you have come to them.

4. Devilish problems have no true or false solutions, just better or worse solutions. In other fields, you can get a qualified person to come in and check your solutions. For example, if you are working out a structural formula for the size of column or a beam, another qualified engineer can come and check your calculations and verify whether you have or have not satisfied the load requirements. A similar kind of an expert can be found in case the problem is a chemical compound. In social planning, however, where devilish problems are involved, there are many parties equally equipped to judge your solution; but none have the power to set formal decision rules to judge the correctness of the solution.

5. A devilish problem is one of which there is no immediate or ultimate test of the solution to the problem. Even after implementation of a transportation solution, there is no way really to tell whether you have "solved" the problem or whether you have really made it worse. In most professional areas you are able to judge the solution. A lawyer either wins or loses the case; with the doctor, the patient either lives or dies; with the structural engineer, the bridge either stands and carries the load or it fails. For the munitions expert, the bomb explodes; and the laser cuts; and the computer computes. Transportation solutions, when implemented, are not so simple because in fact what happens is that the solution generates wave after wave of consequences over a long time span which is extended and unbounded both functionally, geographically, and temporally, and we have no way of checking out all of the impacts to judge whether we have found a good solution.

6. Any transportation problem can be considered as but a symptom of a higher level problem. For example, traffic congestion at 8th and Broadway can be seen at its most basic level as a lack of capacity at the intersection which is solvable by channelizing the intersection, altering the signal system, or putting in a pedestrian bridge, etc. At one level up on the hierarchy, the problem is seen as the fact that the freeway is dumping traffic three blocks away at 5th and Broadway, in which case the solution might well be an alteration of freeway ramps. At still a third level, the problem may be viewed as a lack of a rapid transit system for the corridor. At level four that problem is seen as, in fact, a lack of an equitable Federal financing policy for all modes of transportation in local areas. At still a fifth level, the whole problem can be viewed as another symptom of the irrational love affair that Americans have for the use of private cars.

Similarly, street crime can be seen at its most basic level as a lack of police on the street; at a second level, it can be seen as a lack of state and Federal financial aid for cities; and, at still a third level, it can be viewed as another symptom of the moral decay or permissiveness of our national attitudes, or to too much poverty or of too much wealth, or a lack of proper distribution of the same.

At what level the analyst attacks the problem is somewhat arbitrary depending upon the size of the analyst's ego, the amount of money that he has available at the moment, or the extent to which his sponsoring agency has him fully under control.

7. A devilish problem is one in which the affected parties, and thus the ultimate decisionmakers, are very pluralistic. We, as a nation, are becoming a nation of diverse groups with divergent interests and ever increasing pluralism. We no longer are surprised to find that solutions satisfying one group turn out to be abhorrent to the other. Ten points of benefit derived from a transportation solution for Group A turns out to be 10 points of disbenefit for Group B and thus we have a "zero sum" solution. Our problem as transportation planners is compounded by the fact that it's becoming increasingly difficult to find non-zero sum solutions.

Given these obvious complexities and clear perspectives, the transportation planning problems are in fact of the devilish class, and given the reduced success likelihood that all of this implies, one must ask why we continue to have large efforts and investments going into transportation planning. We continue to have planning, I believe, for several reasons that are somewhat unrelated to the amount of success that we are having. These may be subsumed into four groups:

1. The public is still asking ever larger questions and asking us to evaluate ever larger impacts, and many of the American constituency still believe firmly that problems can be solved by expenditure of more funds. Thus, we continue to get more money for planning.

2. Planning is often a method for putting off the requirement to make unpopular decisions. Politicians and officials recognize that many of the possible solutions are "zero sum" and in order not to offend those groups who perceive themselves to be disbenefitted from various solutions, they prefer not to make the decision at all. Thus, more planning studies.

3. Anthony Downs in his interesting book "Inside Bureaucracy" (3) suggested a list of laws that seem to apply to virtually all human institutions or bureaucracies. One of these laws may be paraphrased as follows—"That those bureaucracies which do not charge their clients for their services must find non-monetary means of rationing their services to clients." Several Federal and State agencies which are providing assistance to local areas for planning and implementing transportation facilities do so on a discretionary grant basis which means they are giving away money to those candidates they find to be worthy. According to Tony Downs' law, such agencies (e.g., the Urban Mass Transportation Administration) must find non-monetary ways to ration this product (money). This is most easily done (and perhaps most rationally as well) by subjecting such candidates to a series of planning efforts in which each series is succeeded by a new set of questions which can only be addressed by additional studies.
4. Planning has its own constituency now and succeeds most admirably in reducing middle income unemployment (among planners). This may well be a part of the reason we continue to have a great many planning efforts.

However correct the previous answers may be, they are not complete answers. There still are transport problems to be solved and still decisions to be made concerning transportation. Many of these involve technical issues at least in part and elected officials must ultimately make these decisions and *someone* must advise them. The *someone* is likely to be a transportation planner who must do planning studies in order to render appropriate advice.

This leads me toward my conclusions which consist of some rather modest suggestions about how transportation planners may conduct themselves in this uncertain and complex environment in which they know in their hearts that they probably cannot satisfy all the demands of their constituency.

1. That transportation planners shed our pompous ideas and illusions about finding the "best"solutions and instead approach our jobs with a humility more fitting a profession which isn't quite sure what it's about.
2. That we continue to improve upon our methodologies as best we can. While it may not be possible to find the "best" solutions, it's certainly possible with our present methods to distinguish those that are hopelessly bad, and improved planning techniques which assist us in making those judgments should be encouraged.

3. That we invest at least as heavily in trying to improve our communications skills with officials and with citizens as we do in trying to improve technical methodologies. This would presumably include training in communications skills and the use of graphics and other media in our presentations and in writing and speaking, etc., as mandatory inputs to the transportation planner's continuing education.

   In this regard, we must really make an effort to shed our jargon. We are perhaps among the worst of all professions in both the use of archaic and unintelligible jargon and the essentiality that we not use it. The statement "thus using a highly sophisticated non-linear aggregation scheme to collapse the dimensionality of evaluative information may be counterproductive if decisionmakers choose not to be involved in or to understand the complexities of the evaluation and aggregation process" are not statements that you should use with your Mayor or Council, or the general public.

4. Transportation planners should not assume themselves more pure than others. I find there seems to be a terrible righteous pride among many planners who tend to hold the notion that because we are "professional" this somehow means we have a key to the truth and are truly objective, whereas the motives of others must be held as suspect. My experience suggests that most officials want to do "right": meaning that everything else being equal they will opt for the best technical solutions.

   However, other things often are not equal and the politician must often consider a broader array of objectives than the transportation planner. For example, he certainly has to balance spending priorities and be concerned about a balanced budget at some point. He also has his own constituency which may not be the same group that the transportation planner sees as the proper constituency. For example, a mayor may be elected only from the Central City, whereas the transportation planner's concern is with a more regional outlook.

   Furthermore, the politician must see proposals as not only being "right" but also they must "seem" right. He has to recognize that they will make sense to the electorate.

   And finally he has to be concerned about his own survival—a concern to which transportation planners, when the chips are down, are also very much subject. When one considers these factors, he is likely to see the actions of the politicians and officials as much more rational than if he neglects these items.

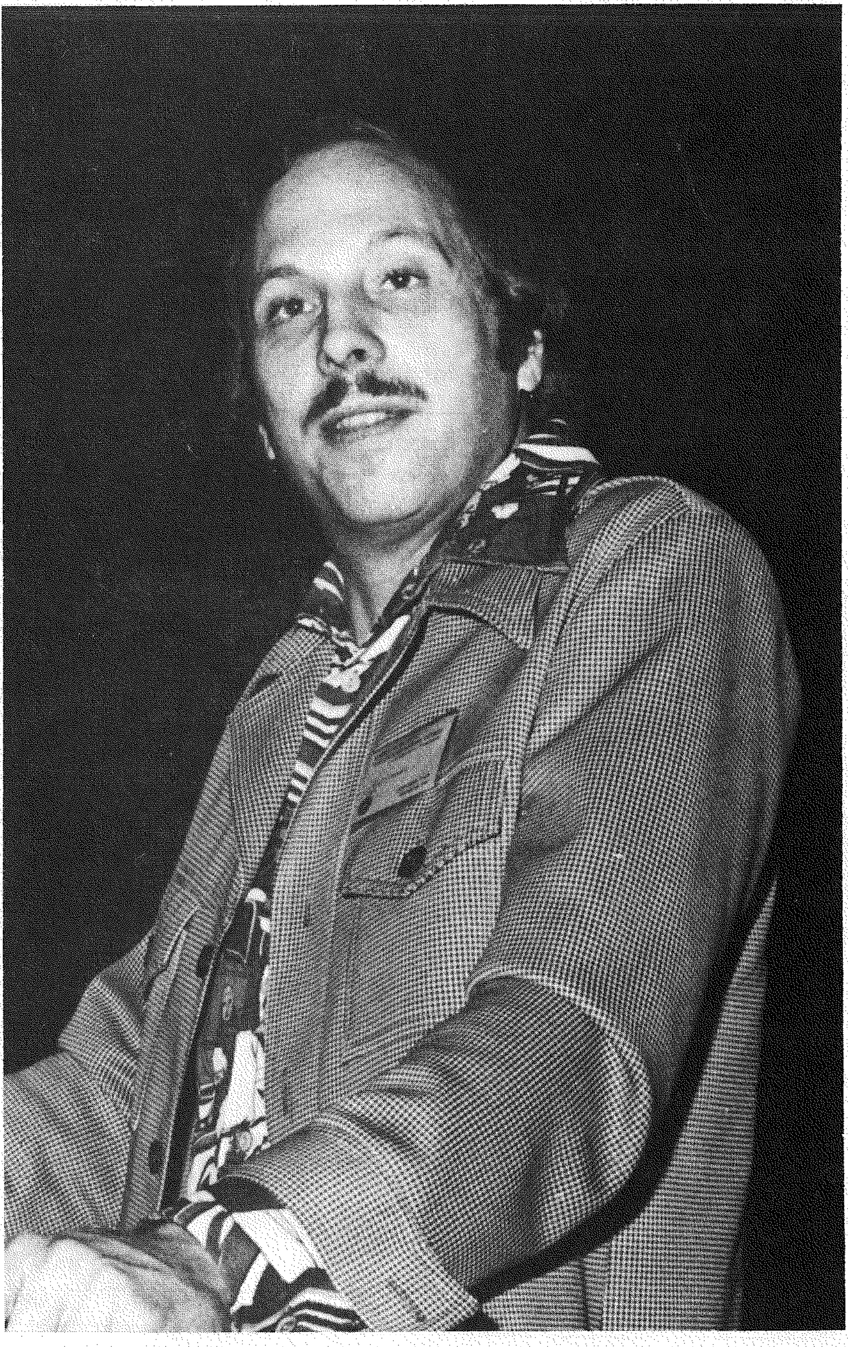
5. That transportation planners don't resist strongly the temptation to invest too much of the resources of their planning

efforts into forecasting and number analysis and too little of both time and treasure to interaction with their public and decisionmakers.

In conclusion, despite the rather pessimistic assessments made earlier, I believe transportation planners still have a role to fill and they can have an impact and be a useful element in society. But despite the importance of developing ever better planning techniques real success is more dependent on the degree to which they establish credibility with their sponsors and with officials which they serve.

REFERENCES

1. Edward C. Banfield, "The Unheavenly City Revisited," Little, Brown and Company, Boston, 1970.
2. Horst, W. J. Rittel, and Melvin M. Webber, "Dilemmas In A General Theory Of Planning," University of California, Berkeley, California, November 1972.
3. Anthony Downs, "Inside Bureaucracy," Little, Brown and Company, Boston, 1967.



David S. Gendell
Federal Highway Administration

Summary of Discussion: Transportation Evaluation Methods

David S. Gendell

Chief
Transportation Studies Division
Federal Highway Administration

Workshop on Treating Distributional Effects and Uncertainty, Robert E. Paaswell, Chairman. The discussion in this workshop centered around the identification of benefits and the relationships between identifiable benefits, community objectives, and the nature of the planning process. It was suggested that the issues of equity, who benefits from and who pays for transportation investments, is going to be one of the major transportation issues during the next 4 or 5 years. Fundamental questions include how equity should be defined and how the distribution of benefits and costs associated with alternative investments can be measured. While planners might focus their concern on the *measurement* of such distributional consequences, there is increasing pressure to ensure that public sector *investment programs* provide an equitable distribution of dollars per person spent on transportation to avoid the increasing possibility of legal action. The prospect of expanded litigation on the issue of distributional effects was foreseen; the desirability of resolving questions of equity on their technical merits rather than allowing decisions to be made in the courts was emphasized.

The difficulty of basing evaluation and decisionmaking on explicitly specified community objectives was pointed out. In particular, because of reduced Federal spending, and increased competition for funds, there was growing pressure to get the most Federal dollars into communities through every possible program, without serious consideration of other objectives in some cases.

It was suggested that attempts to specify the objectives of the planning and evaluation process at the outset, and thereafter holding firmly to those objectives, deny the reality that objectives change as the planning effort is pursued; objectives are constantly being renegotiated between decisionmakers and the community. They should be expected to change throughout the planning process and in many cases are greatly influenced by the availability of Federal funds. It was suggested that careful and explicit consideration of objectives remains a good idea, as long as one recognizes that they will change during the planning process.

Some expressed the opinion that too large a proportion of planning resources is spent on formulating goals and objectives, and that sometimes these activities are pursued to the exclusion of serious planning and decisionmaking. Often it is only possible to reach agreement on goals and objectives at an abstract level which is not particularly useful in alternatives evaluation and choice. The need to carefully limit resources allocated to formulating goals and objectives was suggested, so that an appropriate allocation of resources for the overall planning process can be achieved. Still, an attempt to understand the structure of the basic value system within which planning and decisionmaking are to occur should be sought to provide reasonable support for the evaluation process.

The workshop raised the possibility that it might be useful to consider the notion of opportunity costs more explicitly in public sector transportation evaluation and decisionmaking. Opportunity costs reflect what the community must give up in order to invest in a particular alternative. Thus, opportunity costs reflect high level tradeoffs which must be made in order to implement a particular action. For example, transportation expenditures must be weighed against investments in other fields such as health care, education, social services, etc. A dilemma raised by this workshop was how to incorporate the idea of opportunity costs and the distributional impacts of a transportation plan into the evaluation process.

The desirability of using both efficiency and effectiveness analysis for evaluation in certain circumstances was also suggested. It is desirable to achieve a balanced application of the two methods to satisfy the needs of the community, the decisionmakers, and external funding authorities. It was suggested that efficiency and effectiveness analysis are brought together at the point where cost effectiveness studies are undertaken.

Workshop on the Evaluation of Project Plans—G. Scott Rutherford, Chairman. One of the concerns of this workshop was the current emphasis on short-range, smaller scale projects (e.g., transportation system management projects), and the resulting consequences for traditional, long-range planning activities. In particular, the need to explore and evaluate the longer range implications of short range projects was emphasized. It was suggested that the relevance of long-range planning is greater for areas which are experiencing, or are likely to experience, significant growth. In other areas, the focus on shorter range management of existing transportation system seems more reasonable, but not to the point of ignoring long range implications of actions.

This workshop considered methods for evaluating transportation system management options. While some suggested that simplified evaluation techniques, focusing primarily on efficiency analysis, might be appropriate for simpler TSM options, the workshop concluded that it would be dangerous to narrow evaluation efforts too far. The use of a broader, cost effectiveness framework similar to that used in long-range planning would be desirable to assure that potentially significant (and longer range) impacts were not overlooked. Of course, it was recognized that in some cases it may not be appropriate to go into as much depth as would be used in a larger scale and longer range evaluation process; still, the use of a broader framework is to be preferred. More basically, the need to do serious evaluation of TSM alternatives was generally agreed upon.

Expressing a dissenting view, some suggested that the application of long range, regional scale evaluation methods of TSM is not likely to be successful. TSM evaluation must function at something less than the regional scale, adopting a world view that focuses on the impacts which are most important to people. In response to this point, it was suggested that there is a need to look at the *same categories* of measures for TSM evaluation as would be used for long range, regional scale evaluation. The measures, however, may not be precisely the same, and they certainly won't be at the same scale; measures need to be operative at the scale of influence of the improvement itself. At some point, however, it may be desirable to evaluate the larger scale implications of collections of even rather localized TSM actions.

Workshop on Evaluation of Regional Plans and Programs—Garrison P. Smith, Chairman. This workshop explicitly considered the nature of decisions and decisionmaking, particularly at the regional scale. It was suggested that it is useful to recognize a hierarchy of decisions and decisionmakers; decisionmakers are not only elected officials, but in general may be considered as implementers. Evaluation must meet the needs of all of these individuals involved in the process.

It was suggested that planners should recognize that many decisions come up for consideration repeatedly; decisions are often remade based on new information and new contextual situations. The decision that is the outcome of a planning and evaluation process is not likely to be fixed. Evaluation can and should reflect the fact that, in all probability, it will be necessary to re-evaluate and rethink some of the decisions in the future. Planning and evaluation are continuing processes, and consideration of alternatives probably stops only at the point where facilities are actually built and opened to traffic. It is useful to plan for the contingency that all of the expected future decisions will not be made in expected ways. That is, components of systems and facilities may not be built, and those that are implemented need to be assessed as to their likelihood of functioning reasonably by themselves (i.e., contingency analysis).

The desirability of looking at time streams of projects was pointed out. No systems come into being as complete units, but are implemented as components, and the full sequence of implementation decisions cannot be anticipated in advance. Thus, there is a need to examine the performance and impact of partially-built systems and facilities and to evaluate staging plans to assure that intermediate development levels produce reasonable outcomes.

This workshop also considered the problem of communicating the results of evaluation to decisionmakers and the public. The relatively small proportion of planning resources devoted to communications, especially when compared with the resources consumed in model development and application, was a matter of concern. The importance of preparing and summarizing information for presentation to decisionmakers must receive consideration in the study design process to assure that agency resources are allocated and managed properly to produce the desired products.

Workshop on Preparation and Interpretation of Evaluation Results—Edward A. Beimborn, Chairman. This workshop also considered the need for improved communications between planners and decisionmakers and the community at large. The need was expressed for better methods for presenting and displaying complex information sets which describe the distribution of costs and benefits, the characteristics of contingencies and uncertainties, and the time streams of impacts. Despite the complexity of the information sets to be presented, the techniques needed are those which will simplify presentations so that they will be understandable to the decisionmaker who is generally a layman. When evaluative information is overly complex, the credibility of the planner and planning process is threatened because those who must understand the results are unable to do so. Thus, there is a need for simpler presentation mechanisms, better graphics and maps, and more effective use of small group interaction techniques.

It was suggested that it might be desirable to reduce the level of information presented to only those measures which showed significant differences between alternatives, and to those trade-offs between alternatives which may be clearly identifiable. The process of interaction between planner and decisionmaker should be sufficiently open and responsive as to allow a more detailed examination and discussion of evaluation results where necessary. The use of "supernumbers," aggregate "indices" which attempt to present an overall score for alternatives, was judged to be an unwise approach to evaluation by participants in this workshop, because such indices cover up more information than they provide.

Members of this workshop suggested an additional role which the evaluation process can play, and that is the building of decisionmaker confidence in a particular choice. This is a way to help develop closure in the planning process, such that a strong decision is produced and implementation of an appropriate action can take place.

A major issue considered in this workshop was the trade-off between local and regional concerns in planning and evaluation, particularly as they relate to the efficacy of the citizen involvement process. Citizen involvement almost invariably focuses on very localized concerns; people live in the short run and behave in response to short run, localized pressures and concerns. To the extent that the planning process can be more responsive to producing evaluative information which is salient at the local level of analysis, citizen participation programs may be made more effective. Because the most vocal concerns which are expressed in citizen participation activities are those related to localized negative impacts, some participants in this workshop felt that one role for the planner was to be an advocate for regional scale benefits produced by major transportation improvements, since few interest groups were willing to speak up for this point of view.

The most appropriate role for the planner was a major topic of discussion in this workshop. It was suggested that there is an inherent conflict of interest between the planner as an objective individual who helps to evaluate alternatives from a balanced perspective, and the planner who is necessarily involved in implementation and who is seeking the approval of a variety of actors for projects on his agenda. Most planners, at some point in their activities, must become advocates of plans and the conflict with the objectivity which they must maintain at other times is a difficult one to deal with.

Workshop on Strategic Approaches to Evaluation—Richard D. Worrall, Chairman. Participants in this workshop were also concerned about the ability of the planner to remain objective throughout the evaluation process. It was concluded that it will be difficult, if not impossible, for an individual performing an effective planning function to fail to become biased towards one or more of the alternatives with which he is working. It will be important to make sure that the planner learns to recognize these biases within himself and to deal with them explicitly.

This workshop concluded that no single abstract, strategic model was appropriate for all evaluation tasks. It is necessary to tune the approach to the particular project to be evaluated, the context or environment of that project, and the actors involved in the decisionmaking process. It is important to attempt to design the evaluation process at the start of planning activity, so that evaluation becomes a fundamental part of planning. It was suggested that the evaluation function might be viewed as the ultimate outcome or product of planning itself. Thus, evaluation should in some respects determine the planning process design.

In the design of the evaluation it is useful to put oneself in the position of the various decisionmakers, so that the informational products can be structures to meet their needs directly. For example, it was suggested that mock-ups of evaluation presentations be prepared and presented to teams of planners for the purpose of identifying problems which can be resolved prior to the ultimate presentation to decisionmakers and the public.

It was suggested that it is important to avoid inundating decisionmakers with more information than they can use effectively. A two-stage information presentation, as suggested by the Workshop on Preparation and Interpretation of Evaluation Results, was also discussed by this workshop. The most important evaluative information, including that information felt by the planner to be most useful in decisionmaking (based, perhaps, on prior discussions with decisionmakers) should be presented first in a clear and succinct package. Supporting this presentation should be a comprehensive package of information available to answer questions and provide detail where it is needed.

Finally, this workshop recommended that interactions with decisionmakers should not be viewed as a single activity to be pursued at the end of a planning process, but should be an ongoing involvement focusing not only on a few, obvious decisionmakers, but broadly oriented toward the community for which the planning exercise is being performed.

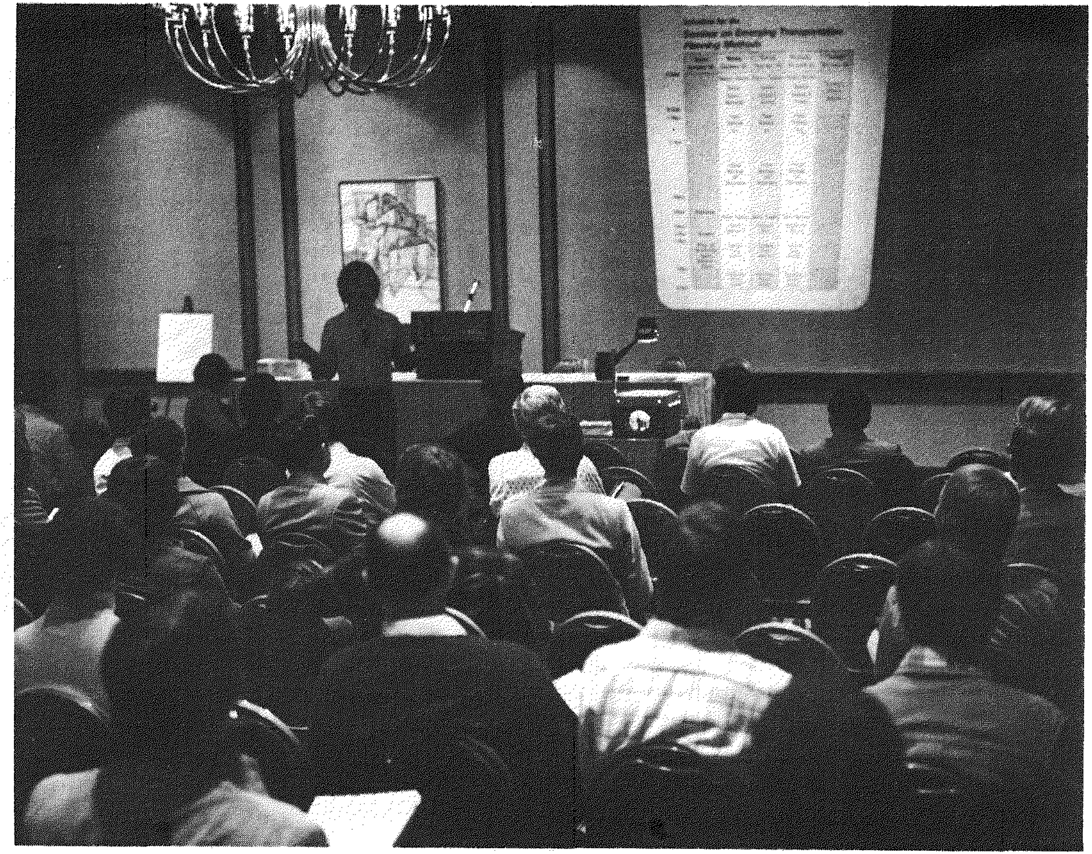
Workshop on Evaluation and Decisionmaking—Charles C. Schimpeler, Chairman. Participants in this workshop also agreed that no single approach to evaluation will be appropriate for all problems. This is particularly true because of operative political forces, and the wide variation in roles of technical professionals in contributing to and supporting decisionmaking processes.

The role of public involvement in evaluation and decisionmaking was discussed; it was felt that a broad and open public involvement process contributes significantly to the credibility of planning and decisionmaking while providing an opportunity for the advocacy of important, special viewpoints. It was also recognized that public involvement can be difficult and expensive, requiring a substantial commitment of resources, effort and skills. It is important to specify openly and early the role of public involvement activities, including the authority and responsibility of advisory committees. This will ensure that the focus of decisionmaking remains in the hands of responsible public officials, while reducing the likelihood that expectations of public participants will be unduly raised.

Finally, it was suggested that some interest groups concerned about particular decisions may not function well in the context of a broad and open public involvement process in support of decisionmaking. In such cases, planners might anticipate the formation of secondary response channels which are external to the public involvement program. It would be desirable for the design of the evaluation and public involvement programs to identify the likelihood of such responses in advance, and to attempt to build them explicitly into the overall process.


Panel D

III. TRANSPORTATION/LAND USE INTERACTIONS



Introductory Remarks by Edward Weiner

The relationship between transportation and land use is one of the least understood and most ignored subjects in urban transporation planning. Discussion of land use-transportation interactions can be found in the transportation planning literature as early as the midfifties and in urban planning sources even earlier. The trend-setting Chicago Area Transportation Study used a simple land use forecasting model to develop inputs to their travel forecasting models.

During the 1960's, researchers and a number of urban transportation planning studies developed land use models. Each succeeding model appeared more complex and data hungry than its predecessor, as researchers attempted to encompass more elements of the urban development process. By the late 1960's, land use models had become too complex, required too much data, were too costly to construct and operate, and many still did not produce useable results. Land use modelling efforts fell into a period of dormancy, and funding for such work all but disappeared.

However, planners were still concerned about their lack of ability to treat land use-transportation interactions in the planning process. In the 1970's, there was renewed interest in the development of simple, useable, policy-oriented land use models to assist in evaluating the transportation planning process. As this process broadened to encompass a wider range of social, environmental and, most recently, energy consumption issues, concern about the effects on land use distribution became increasingly important. Even the emphasis on transportation system management required an understanding of the impact of such measures as pricing and regulatory changes on land use.

The presentation that follows discusses the integrated forecasting of land use and transportation. The model that is described is an evolutionary outgrowth of the model developed by Ira Lowry in 1964 and the Projective Land Use Model (PLUM) developed by William Goldner for the Bay Area Transportation Study. The current version of this newer model is ready for application in urban transportation planning; operating experience with this model should provide useful information for further refinement.

Stephen H. Putman,
Associate Professor of City and Regional Planning
University of Pennsylvania

The Integrated Forecasting of Transportation and Land Use

Stephen H. Putman

Associate Professor of City and Regional Planning
University of Pennsylvania

INTRODUCTION

Serious efforts to construct urban land use simulation models began in the early 1960's. In all of these efforts the relationships between transportation and land use were only partially represented. In each case, a matrix of zone-to-zone impedances was unput to a land use model. These impedances were measures of the difficulty of interaction between activities located in different zones. Depending upon the availability of data, travel times, travel costs, or airline distances (or various combinations thereof) were used as impedances. The impedances were the only link between transportation and land use in these modelling efforts.

During this same period the modelling of transportation (especially highway) networks progressed to the point of rather routine application of trip generation, distribution and assignment program packages. The connection from the land use distributions to these transportation models was simply the input of the activity distributions and land use data to the network model package.

As part of its response to an increasing concern with the premature obsolescence (over-congestion) of urban highways, in March 1971 the U.S. Department of Transportation* issued a request for a proposal to do a research study entitled "Interrelationships of Transportation Development and Land Development." The contract for the study was awarded to Professor Stephen Putman at the University of Pennsylvania, with the work beginning in November of 1971. It was evident from the start of the work that the premature obsolescence question was but one factor of the more general question of urban spatial configurations.

---

* Federal Highway Administration, Office of Highway Planning, Urban Planning Division.

The proposed method of analysis for the general problem was an integrated set of land use and transportation models. As the project emphasis was on the integration of these models and their consequent capability for analyzing land use-transportation interactions, it was decided to use modified versions of existing models rather than constructing wholly new ones. The land use model portion of the package was a much modified version of PLUM, called IPLUM. The transportation part of the package was an "in house" network package incorporating capacity restrained, incremental tree-by-tree assignment. In early tests of the model package, in order to assure eventual compatibility with FHWA software, a trial linkage was constructed and several test runs were also made using the FHWA-UTPS network programs.

This project, completed in the summer of 1973, demonstrated the first successful integration of a land use model with a transportation model, incorporating feedbacks.[1] Trip generation was accomplished by extracting and converting the implicit trip matrices from the land use model activity allocations. Congested network times resulting from these trips were then fed back into the land use model. The entire process was run, in an iterative manner, to equilibrium.[2]

Though this first project did not involve extensive policy testing, there were some interesting policy conclusions. In particular, the simulation results demonstrated that attempts to deal with highway congestion by means of new highway facility construction will, in the absence of stringent land development controls, lead to even greater highway congestion in future years. In addition, it was shown that a "do nothing" policy of allowing existing facilities to congest might lead, in the long run, to lower levels of congestion. As part of these conclusions it was clearly demonstrated that new highway facilities, in the absence of land use controls, lead to increased urban sprawl, while congested facilities tend to inhibit sprawl. Finally, these test runs indicated relationships throughout large metropolitan regions; e.g., the effects of land use controls imposed in one county on development in adjacent counties, that would probably pass unnoticed with more traditional analysis techniques.

In May 1973, a grant was received from the National Science Foundation for a study titled Laboratory Testing of Predictive Models. The purpose of this work was to develop a method for testing and

---

[1] Putman, S. H. (1973) "The Interrelationships of Transportation Development and Land Development" Urban Planning Division, FHWA, Department of Transportation, Washington, D.C. (in 2 Volumes, revised and reprinted in 1976).

[2] Putman, S. H. (1974) "Preliminary Results From an Integrated Transportation and Land Use Model Package" *Transportation* Vol. 3, pp. 193–224.

evaluating predictive land use models, and to apply this method to existing land use models. After a thorough review of existing land use models, the efforts of the project were devoted to work with the EMPIRIC model and a version of IPLUM (which resulted from the DOT work described above) as a representative of Lowry derivative models.

The research strategy in this new work was to first attempt to estimate both models' parameters for a common data set, and then to do a series of sensitivity tests of both models. No difficulties were encountered in estimating parameters for EMPIRIC. When it came to estimating the parameters for IPLUM it was discovered that there was no formal technique in U.S. modelling practice for estimating the parameters of Lowry derivative models. In dealing with this dilemma, the research eventually led to consideration of the entropy-maximizing approach of Alan Wilson.[3] Through this approach IPLUM was reformulated and subsequently calibrated (i.e., its parameters were properly estimated). The resulting model was sufficiently different from IPLUM to be named Disaggregated Residential Allocation Model—DRAM.[4]

The two models, DRAM and EMPIRIC, were then put through a series of tests of their responses (i.e., changes in outputs) to different types of change in inputs. The results of these tests showed DRAM to be much more responsive to changes in inputs than was EMPIRIC. This, along with DRAM's better articulated theoretical structure, suggested a clear superiority over EMPIRIC for use in planning policy tests and evaluations. This work was completed with the preparation of a draft final report in the summer of 1975.[5]

June 1975 saw the beginning of a second round of work on the Integrated Transportation and Land Use Package. While some further work had been done in the preceding two years,[6] a National Science Foundation grant titled Development of an Improved Integrated Transportation and Land Use Model provided the funds necessary to carry the development of ITLUP to its present state. The chief goal of this effort was to upgrade ITLUP to include DRAM, and to add several

---

[3] Wilson, A.G. (1970) *Entropy in Urban and Regional Modelling,* Pion, London.

[4] Putman, S. H. (1977) "Calibrating a Disaggregated Residential Allocation Model—DRAM" *London Papers in Regional Science* Vol. 7, pp. 108–124.

[5] Putman, S. H. (1976) "Laboratory Testing of Predictive Land-Use Models; Some Comparisons" Office of Transportation Systems Analysis and Information, DOT, Washington, D.C.

[6] Putman, S. H. (1976) "Further Results from the Integrated Transportation and Land Use Model Package (ITLUP)," pp. 165–173.

additional models to the package. Also to be explored were a basic employment model, a modal split procedure, and an air pollution model. The whole package, including parameter estimation procedures, was to be prepared for dissemination to experienced users in the field.

Additional funds from the Office of the Secretary, DOT were included in this grant to cover the preparation of a number of policy tests using the improved ITLUP package on an actual data set. These policy tests are currently under way (as of December 1976).

As part of the work of the current grant, many calibrations of DRAM were performed, including some for a 19th century data set,[7] with highly encouraging results confirming the general form of the model. Similarly, a simple model of basic employment location has been developed and yields rather promising results. All of this work is nearing completion and will be written up in the near future.

The importance of this work, quite apart from the various interesting theoretical-empirical findings, lies in the preparation of advanced planning technology for distribution and use to practitioners in the field. Use of these methods can reasonably be expected to improve planning analyses, while the results of such uses can be used to direct future research efforts towards the development of even more accurate and useful methods.

These new techniques, along with numerical examples of their derivation and application, are presented in the following pages.

INTEGRATING TRANSPORTATION AND LAND USE MODELS

It is common knowledge that in the past two decades there has been enormous public investment in highway construction. This same period has also witnessed large scale shifts in population from rural areas to urban areas; and within the urban area, from center-city to suburb. Even without attempting here to define the explicit causal relationships between these two developments, it is possible to describe a related phenomenon which is observed, not infrequently, when the construction of a new section of roadway is followed, all too soon, by very heavy usage and subsequent congestion. Specifically, the nature of this process seems to be that: (a) Due to the inadequacy of existing facilities a decision is taken in a metropolitan area to improve transport facilities (usually by road building) for a particular part of the area.

---

[7] Putman, S. H. (1977) "Calibrating a Residential Location Model for Nineteenth-Century Philadelphia" *Environment and Planning A* Vol. 9.

Then (b) assuming that the decision is approved, in anticipation of the new roadway, land developers and/or speculators become involved with properties near the proposed route. As construction of the facility begins, some homeowners and businesses may consider and even act upon location decisions. Upon completion of the facility, additional location decisions are made. Finally (c) a relatively short time after completion of the facility, the demand for its use greatly exceeds the demand which existed prior to the decision to construct it. Consequently the design capacities are soon reached and often exceeded, resulting in congestion and premature obsolescence of the facility. While this is a rather generalized description of a very complex process, it is reasonably accurate. The question is whether it is feasible, via integrated transportation planning and land use planning, to avoid or ameliorate the occurrence of this particular phenomenon in the future. Further it is necessary to analyze this process and to determine whether (a) balanced development of transportation facilities and land use is feasible, and (b) if it is feasible, what means are available to accomplish it.

There is considerable evidence indicating that the demand for highway travel is rather sensitive to changes in highway capacity. This sensitivity, as described above, frequently results in heavy utilization and congestion of new facilities soon after their construction. It has been argued that the solution to this problem is to construct more facilities. It is possible that at some point, if this policy were followed, an equilibrium situation would indeed be reached. This conclusion can be supported by asserting that the elasticity of demand for highway travel is finite. However, if the "population" generating that demand continually increases at the same time, it is not clear that the total demand can easily be satisfied in this manner. The limit of this strategy, at the extreme, would be reached when so much land is converted to roads that land development for other purposes is restricted, with a consequent limit on trip generation and road use. It is obvious that this "equilibrium" is not the desired solution and is hopefully not feasible in any case. It will therefore be necessary to analyze possible "intermediate" solutions.

A balance between transportation facility development and land development implies a market equilibrium of the demand for use of the transportation facility and its supply, i.e., its speed and capacity characteristics. There are two basic alternatives available to the planner who wishes to modify an existing transportation and land use situation sufficiently to achieve such a balance, though it is likely that the best strategies will be mixtures of the alternatives. The first alternative is to allow demand for transportation to fluctuate freely, with no interference, and attempt to cope with it. This would be accomplished, as in the past, by new highway construction or by implementation of mass transportation systems. The second alternative is to attempt to restrict the demand for transportation so that facilities do not become overloaded. This could be accomplished in three ways: (1) the existing transportation facility could be made more costly to use, e.g., by the imposition of tolls or by allowing congestion to develop, thus imposing a time penalty on users; (2) land development controls could be imposed, thus reducing (or slowing the growth of) trip generating activities; and (3) a mixture of these two actions could be implemented. Finally, a mixture of the two basic alternatives could be attempted, where an attempt would be made to cope with a certain amount of transportation demand (by improving transportation facilities) at the same time that an attempt was being made to control its increase. To summarize, the essence of this problem is controlling (i.e., altering, directing, and modifying) the spatial organization of the metropolis. In particular, the concern is with its spatial expansion and with the feasibility of balanced development of land use and transportation facilities.

The inherent complexity of a comprehensive analysis of both transportation and land use, along with the requirement that the analysis method be capable of providing a self-consistent procedure for testing the sensitivity and response to integrated transportation and land use control policies, strongly suggests the development of integrated transportation and land use models. Serious efforts to construct urban land use simulation models began in the early 1960's. In all of these efforts the relationships between transportation and land use were only partially represented. In each case, a matrix of zone-to-zone impedances was input to a land use model. These impedances were measures of the difficulty of interaction between activities located in different zones. Depending upon the availability of data, travel times, travel costs, or airline distances (or various combinations thereof) were used as impedances. The impedances were the only link between transportation and land use in these modelling efforts.

During this same period the modelling of transportation (especially highway) networks progressed to the point of rather routine application of trip generation, distribution and assignment program packages. The connection from the land use distributions to these transportation models was simply the input of the land use data to the network model package.

In the transportation and land use literature of the 1960's and early 1970's one will find occasional references to the desirability of more fully integrating land use and transportation models (not to mention planning). This integration of models would attempt to eliminate the principal failing of contemporary land use or transportation studies, by explicitly including feedback loops. Typically, a transportation study assumes a future land use pattern as given, and designs a transportation system to cope with it. This procedure ignores the redistributive effects which are produced by the construction of the system. Transportation systems obviously do not just suddenly appear but are constructed in stages with consequent redistribution of activities all during the period of construction. The typical land use study accepts a transportation system as given, and then estimates the consequent distribution on the network. An integrated model package would attempt to capture the interrelatedness of the transportation system and the distribution of activities.

The integration of a transportation network model with a land use model is, in principle, a rather straightforward matter. Consider the problem of forecasting the future distribution (e.g., spatial pattern) of population and employment in an urban area. We are given a description of the transportation network which will exist at that future time, along with regionwide projections of population and employment. Further we know, for a specified base year, the population and employment distributions and the trip patterns.

We begin the forecasting process by loading the current trip pattern on the projection year network using a capacity restrained assignment procedure. This yields an estimate of travel times on the future network in the unlikely event that both the region's activity levels and distributions remained unchanged from the base year. These travel times are used as input (in the form of a zone-to-zone impedance matrix) to the land use model, which produces the first estimate of the forecast year activity distributions. The activity distribution is, in turn, used to make the first estimate of the forecast year trip pattern (origin-destination matrix).

The second iteration begins with the assigning of the first estimate of the forecast year trip matrix to the forecast year network (the base year trips having been previously removed). This yields a second, and more likely, estimate of travel times on the forecast year network. These travel times are then input to the land use model which produces a second, revised, estimate of the forecast year activity distributions. The activity distributions are then used to generate the second, revised estimate of the forecast year trip pattern. The entire process is then repeated, until an equilibrium is reached.

It may be helpful to consider this process with the aid of the following block diagrams. First, in Figure 1 we see simplified representations of a typical transportation planning process and a typical land use planning process.

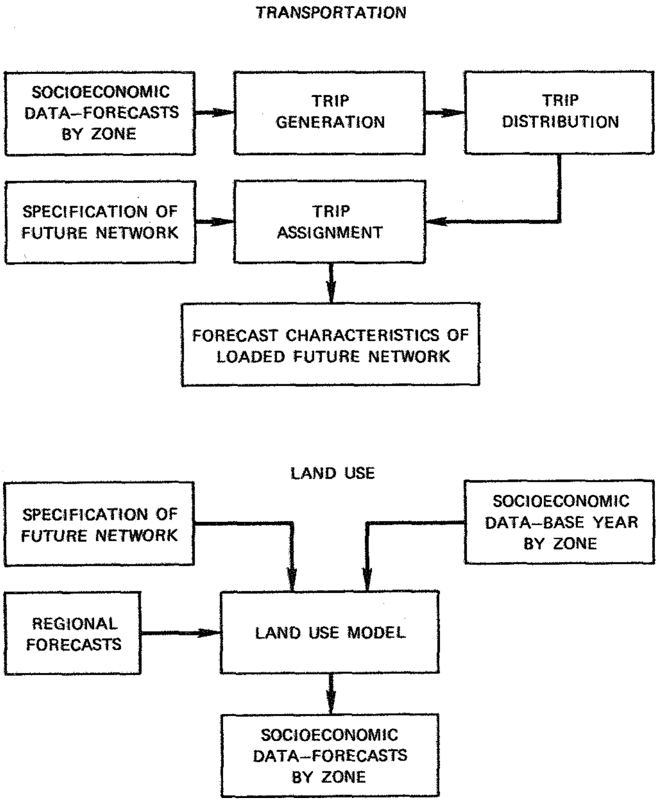


Figure 1.—Transportation and land use planning—two separate activities.

In Figure 2 we see how these two processes may be linked, first by using the outputs of the transportation planning process as input to the land use planning process, and second, by using the outputs of the land use planning process as input to the transportation planning process. Having shown this connection, it becomes clear as to why it may be necessary to iterate (i.e., repeat) the steps in the process several times before coming to an equilibrium solution.

In Figure 3 we see a translation of the integrated transportation planning and land use planning processes into a set of model procedures. First a set of base year data on both land use and transportation is used to estimate a base year O.D. trip matrix. These trips are then loaded on (assigned to) the future year transportation network. The time and cost characteristics of this partially congested network are used, along with regional economic forecasts, as input to forecasts

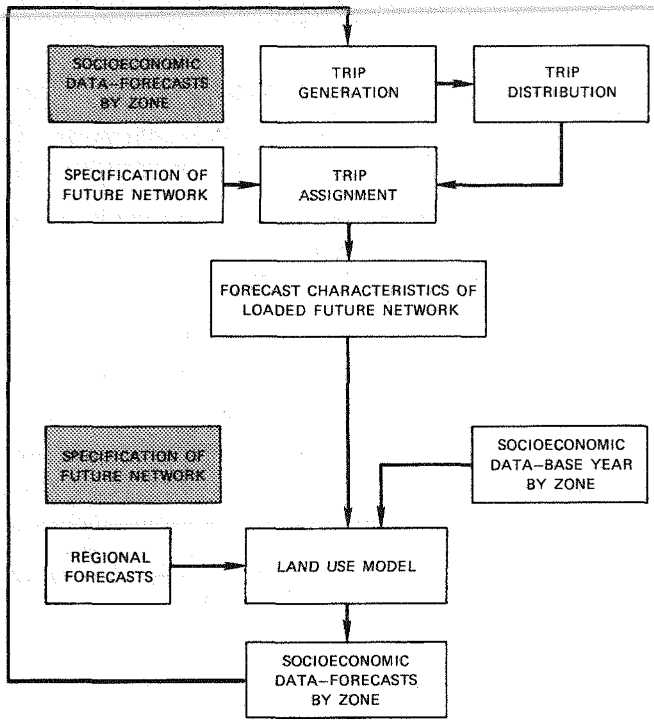


Figure 2.—Integration of land use and transportation planning.

of the future year land use pattern. From these future year land use forecasts, we obtain a future year O.D. trip matrix. These estimates of future year trips are then assigned to the future year network. It is certain that at least two cycles through this process will be required. After these first two cycles the results (forecasts) are checked for equilibrium, after which the process is terminated.

The first operational version of ITLUP followed the above procedure rather closely. In that version of the package, the network model used an incremental tree-by-tree, capacity restrained assignment algorithm. The basic employment distribution was an exogenous input. The population and non-basic employment distributions were estimated with a Lowry-derivative model (a modified version of the incremental PLUM model).

Policy and sensitivity tests were run with this version of ITLUP. The most important methodological conclusion from this work was a clear demonstration of the need for an integrated package to make accurate

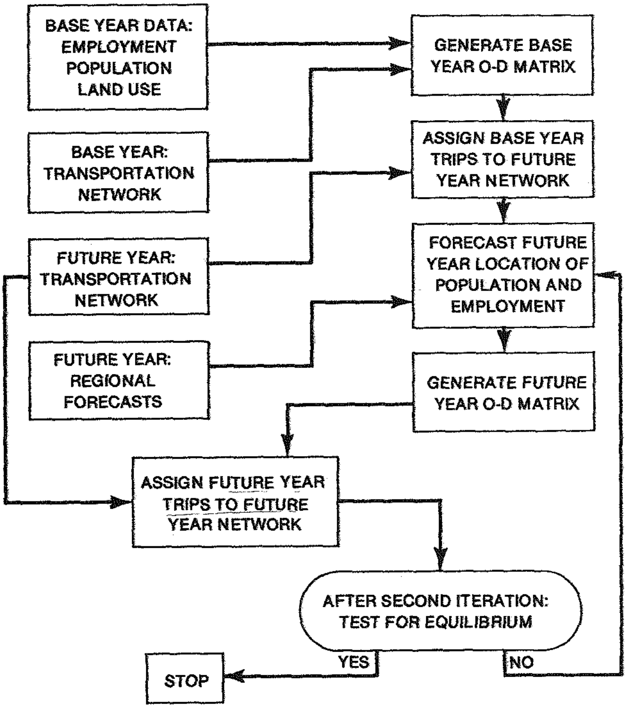


Figure 3.—Model procedure for integrated land use and transportation planning.

forecasts. Use of the same land use model in an unintegrated noniterative forecasting procedure yielded less accurate forecasts than those obtained from ITLUP. The most important substantive conclusion from this work was that improvements in a region's transportation system will, in the absence of coordinated land use controls, usually lead to further urban sprawl and congestion along with continual demand for new transportation improvements which will serve to repeat the cycle. The reverse was also found to be true; allowing the development of transportation congestion reduced tendencies to urban sprawl and led to the development of a regional pattern of clusters of denser development.[8]

---

[8] Putman, S. H. (1973) "The Interrelationships of Transportation Development and Land Development" op. cit.;

Putman, S. H. (1974) "Preliminary Results from an Integrated Transportation and Land Use Models Package." op. cit.

Since the publication of the initial results from ITLUP, considerable additional work has been done.[9] Subsequently other researchers have published findings which corroborate the ITLUP results.[10]

In 1973 another project was begun, under N.S.F. sponsorship, with the intent of comparing various operational land use forecasting models.[11] During this project the residential submodel of ITLUP was reformulated to allow for its proper calibration.[12] This reformulation was based on the Wilson maximum entropy approach.[13] Further work with this form of the model, called DRAM (Disaggregated Residential Allocation Model) has produced important results not only from the standpoint of parameter estimating and forecasting, but for the more direct integration of transportation and land use models.

There appears to be a great deal of misunderstanding of the maximum entropy derivation of these new transportation and land use models. Further, it seems that some of their more important implications are quite overlooked in U.S. planning practice. Consequently, the next section of this paper will present a simple and hopefully lucid discussion of the underlying basis for these models while the following section will present an illustration of one such model.

SPATIAL INTERACTION MODELS AND MAXIMUM ENTROPY

It is well known that the initial development of transportation and land use models was largely based on the observed fit of these phenomena to the Newtonian gravity model. While the descriptive validity of these gravity models was reasonably good, a number of persistent doubts remained. One of the most serious of these, from a theoretical point of view, was the question of why human spatial interactions should resemble the interactions of planetary bodies.

The entropy maximizing derivation of spatial interaction models completely obviates the use of the gravity model analogy. The derivation thus depends not on a rather implausible analogy to a physical

---

[9] Putman, S. H. (1976) "Further Results from the Integrated Transportation and Land Use Model Package (ITLUP)."

[10] Berechman, J. (1976) "Interfacing the Urban Land-Use Activity System and the Transportation System," *Journal of Regional Science,* Vol. 16, No. 2, pp. 183–194;

Peskin, R. L. and J. L. Schofer (1976) "Assessment of Transportation Energy Conservation Strategies through Simulation" presented at Fourth Inter-society Conference on Transportation, Los Angeles, California, July.

[11] Putman, S. H. (1976) "Laboratory Testing of Predictive Land-Use Models: Some Comparisons."

[12] Putman, S. H. (1976) "Calibrating a Disaggregated Residential Model—DRAM."

[13] Wilson, A. G. (1970) *Entropy in Urban and Regional Modelling.*

phenomenon, but on an analogy with statistical mechanics. Further, it has been shown in the literature that several models which derive from concepts of micro-economic behavior, e.g., travel cost minimizing and to a certain extent market-clearing behaviors, are compatible with the maximum entropy approach.[14]

The principles of the maximum entropy approach may be shown with the following simple example.[15] Imagine six employed persons living in a residence zone i and commuting to various work zones j. Let there be one residence zone, i = 1 and three work zones corresponding to j = 1, 2, 3. Suppose that the six workers are named A,B,C,D,E,F. We may now specify the origins and destinations of the worktrips of each worker. Each possible, fully described, system of (a) origin, (b) three destinations, and (c) six worktrips with their origins and destinations may be called a *microstate* of the system. Six of these possible microstates are shown in Figure 4. There are obviously very many microstates of even this simple system.

Let us now consider Microstate 1 where the trips between i and j = 1 are 3; between i and j = 2 there are 2 trips; and the total trips between i and j = 3 are 1. Microstate 6 may also be seen to have this same distribution of trips, from i to j = 1 there are 3, i→2 = 2, i→3 = 1. Clearly there are many microstates which could be drawn and which would have this same arrangement of total trips. This particular arrangement of zone-to-zone trips, if described independently of *which* worker is making which trip, may be called a mesostate of the system. Four mesostates of the system are shown in Figure 2.

Comparing Figures 4 and 5, it can be seen that Microstates 1 and 6 are possible microstates of Mesostate A. Microstate 2 is a possible microstate of Mesostate B. Microstate 5 is *not* a possible microstate of Mesostate D. Thus each mesostate describes a set of possible microstates.

If we now consider that there might be several residence zones in addition to the one which has been used in this example, then a more aggregate level of description of the system would be the total trips leaving each origin and the total trips arriving at each destination. Let us assume that two workers live in zone i = 2 and four workers live in

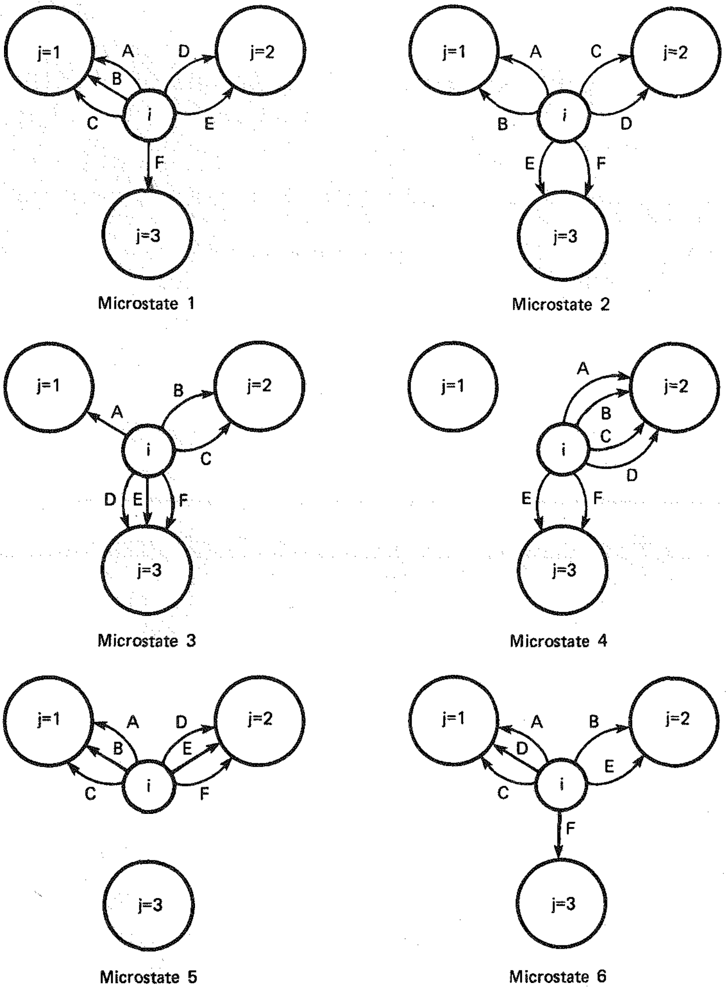
---

[14] Senior, M. S. (1973) "Approaches to Residential Location Modelling 1: Urban Ecological and Spatial Interaction Models (A Review)" *Environment and Planning* Vol. 5, pp. 165–197;

Nijkamp, P. (1975) "Reflections on Gravity and Entropy Models" *Regional Science and Urban Economics,* Vol. 5, pp. 203–225;

Choukroun, J. M. (1975) "A General Framework for the Development of Gravity-Type Trip Distribution Models" *Regional Science and Urban Economics* Vol. 5, pp. 177–202.

[15] Drawn in part from Senior (1973), op. cit.



Microstate 1

Microstate 2

Microstate 3

Microstate 4

Microstate 5

Microstate 6

Figure 4.—System Microstates.

zone $i = 3$ in addition to the six already defined as living in $i = 1$. Further assume that these additional workers are named G,H,I,J,K,M.

A microstate of this newly expanded system would be a list of the origins and destinations of the worktrips of each of the twelve workers. A mesostate of this system would be a list of the total number of worktrips from each origin zone to each destination zone. Finally, a *macrostate* of this expanded system is a list of the total trips leaving each origin and the total trips arriving at each destination. Figure 6 shows four macrostates of the expanded system.

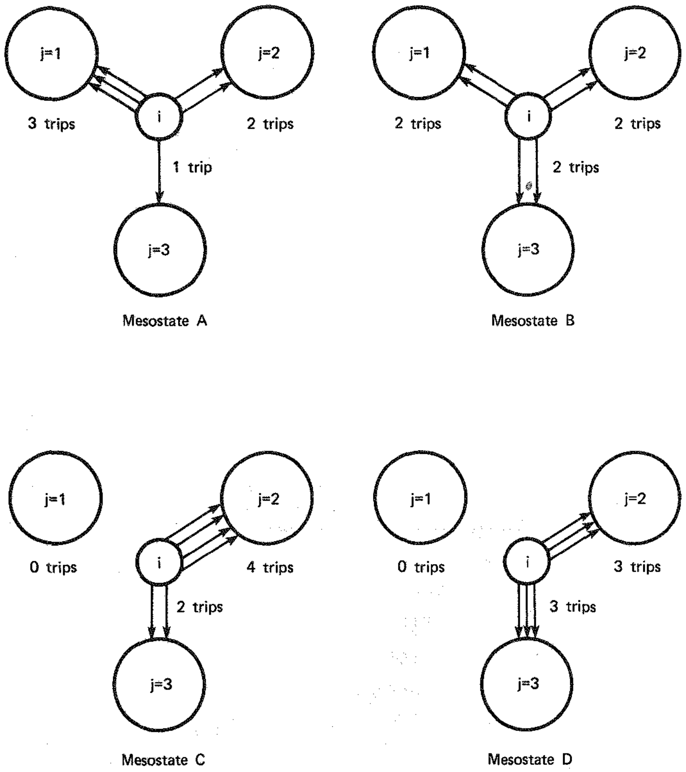
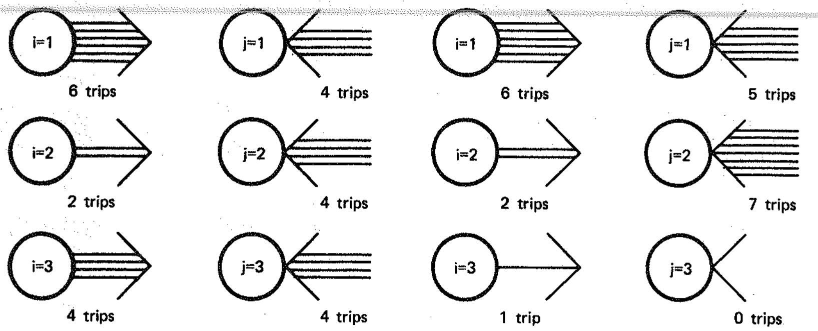


Figure 5.—System Mesostates.

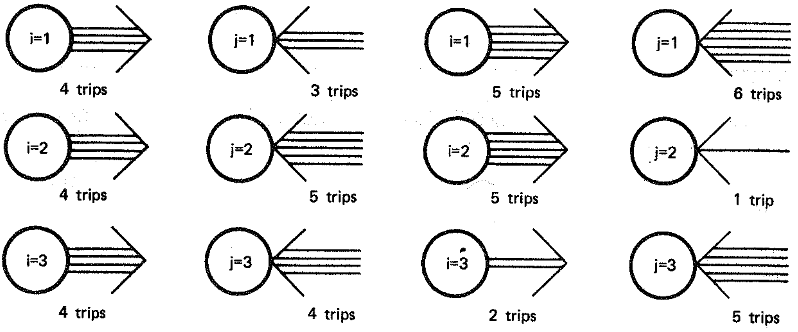
Referring to Figure 6, Macrostates 1 and 2 with six trips leaving $i = 1$ contain all the previous examples of microstates and mesostates. Macrostates 2 and 3, with the trips leaving $i = 1$ not equal to six, correspond to other system states which do not include the microstates and mesostates given as examples. We should also note in passing that one could have defined a macrostate for the example of a single origin used at the start of this discussion. This would have been, in a sense, a degenerate case, as the trips leaving the single origin would always have been equal to six.

In operational urban simulation models virtually no attempt is made to simulate at the level of microstates, i.e., individual behavior. Most of the models operate at the mesostate and/or macrostate level. The entropy approach deals with these, and requires two key assumptions. First, all microstates are assumed to be equally probable. Second, the most likely, mesostate or macrostate is assumed to be the one with the greatest number of possible microstates.



Macrostate 1

Macrostate 2



Macrostate 3

Macrostate 4

Figure 6.—System Macrostates.

We may now easily derive a spatial interaction model for the mesostate level of detail.[16] First, defining the variables.

T_{ij} = the number of workers living in i and working in j (this is the variable to be estimated)

O_i = the number of workers living in i (a given)

D_j = the number of jobs in j (given)

C_{ij} = the cost of travel from i to j (given)

C = the total travel expenditure of the system (given)

In addition there are several constraints on the system. First, the sum over all j destination zones of all trips arriving from zone i is equal to the total trips leaving i. That is

---

[16] Wilson, A. G. (1970) *Entropy in Urban and Regional Modelling* Pion Limited, London, pp. 1–14.

$$\sum_j T_{ij} = O_i$$

Second, the sum over all i origin zones of all trips going to j is equal to the total trips arriving in j. That is

$$\sum_i T_{ij} = D_j$$

Finally, the total travel cost is equal to the sum of the trips between each i-j pair times the cost of travelling between that pair.

That is

$$\sum_i \sum_j T_{ij} c_{ij} = C$$

Now, the most probable mesostate is the one with the maximum number of possible microstates, subject to constraints (1), (2), and (3). Consequently, what is desired is the description (i.e., equation) of the matrix $\{T_{ij}\}$ which has the greatest number of microstates, i.e., the greatest number of ways, $W(\{T_{ij}\})$ of getting $\{T_{ij}\}$ given the constraints. The number of ways of getting $\{T_{ij}\}$ is a problem of combinations of individual workers to given origin-destination pairs. The microstates in Figure 1 represent six of many possible combinations of one origin, three destinations, and six workers. If T is the total number of workers, i.e.

$$T = \sum_i O_i = \sum_j D_j = \sum_i \sum_j T_{ij}$$

then by combinational theory

$$W(\{T_{ij}\}) = \frac{T!}{\prod_{ij} T_{ij}!}$$

For the Figure 4 example:

$$i=1 \quad j=3 \quad T=6$$

Then[17]

$$W(\{T_{ij}\}) = \frac{6!}{(T_{11}!)\ (T_{12}!)\ (T_{13}!)}$$

---

[17] Note 6! means 6 factorial, which equals 6x5x4x3x2x1 = 720. Also note that 0! is defined to be 1. Further note that Π is the product operator, as Σ is the sum operator.

And, if we wish to know the number of microstates for Mesostate A in Figure 5.

$$W(\{T_{ij}\}) = \frac{6!}{3!\ 2!\ 1!} = \frac{720}{6(2)} = 60$$

The number of microstates for Mesostate B in Figure 5.

$$W(\{T_{ij}\}) = \frac{6!}{2!\ 2!\ 2!} = \frac{720}{8} = 90$$

By trial and error, one may substitute values in the denominator of this equation and discover that the minimum value of the denominator (subject to the constraint that the sum of all the trips equals six) is at 2!2!2! Thus the maximum value of W is 90, which suggests that in the absence of any further information about the system of our example, the most probable mesostate is when the six trips are evenly distributed to the three destinations.

We recall however that it is the equation for the T_{ij} which we are trying to determine. Consequently we must maximize $W(\{T_{ij}\})$ as given in equation (4) subject to the constraints (1), (2), and (3). This is mathematically possible, and involves the use of Lagrangian multipliers and Stirling's approximation to the numerical value of large factorials.[18] The result gives

$$T_{ij} = A_i B_j O_i D_j \exp(-\beta c_{ij})$$

where

$$A_i = \frac{1}{\sum_j B_j D_j \exp(-\beta c_{ij})}$$

$$B_j = \frac{1}{\sum_i A_i O_i \exp(-\beta c_{ij})}$$

where β is an empirically derived parameter, or if C in equation (3) were known, β could be solved for numerically.

Before proceeding, we should note that the name maximum entropy for this derivation stems from the fact that equation (4) is defined, in statistical mechanics, to be the entropy of the system. It is, of course, equation (4) and therefore the system entropy which we maximize to derive equations (5), (6) and (7). At this same point, it must be mentioned that there are alternative interpretations of entropy which may

---

[18] Stirling's approximation for large values of x yields; $\ln(x!) = x\ln(x) - x$

be used. From a practical point of view the results lead in the same direction, which is that a spatial interaction model whose equations resemble the traditional gravity model may be derived from assumptions which do not include any reference to Newtonian gravitational phenomena, but which do refer to probability statements and "most probable" distributions.

Having shown the derivation of a spatial interaction model via the maximum entropy approach, we may also show its relationship to a transportation cost minimizing approach of the sort which economists assert should constitute the underlying behavior of spatial interaction. A simple cost minimizing approach may be constructed as follows. Suppose that there are a given number of trips O_i, originating, and of trips D_j, terminating in all zones. Suppose that each trip T_{ij} has a cost c_{ij} of travelling between i and j. The problem may then be stated as:

Minimize: $$\sum_i \sum_j T_{ij} c_{ij}$$

Subject to: $$\sum_i T_{ij} = D_j$$

$$\sum_j T_{ij} = O_i$$

This is known as the "transportation problem" of linear programming and often has additional constraints as part of the problem formulation.

Now, consider the entropy maximizing approach as presented above. It may be shown algebraically that maximizing W is equivalent to maximizing ln W and by further manipulation that

Maximizing: $$W(\{T_{ij}\}) = \frac{T!}{\Pi_{ij} T_{ij}!}$$

is equivalent to

Maximizing: $$-\sum_i \sum_j T_{ij} \ln T_{ij}$$

Further we note that in general, maximizing a function is the same as minimizing its negative. That is

Max: $f(x) =$ Min: $-f(x)$

Consequently, the entropy maximizing approach may be rewritten as

Minimize: $$\sum_i \sum_j T_{ij} \ln T_{ij}$$

Subject to: $$\sum_i T_{ij} = D_j$$

$$\sum_j T_{ij} = O_i$$

$$\sum_i \sum_j T_{ij} c_{ij} = C$$

Suppose now that for a particular problem situation there is a minimum numerical value M, of entropy for this problem. Then we may consider the relationship between the cost minimizing formulation and the entropy minimizing formulation in the following problem statement where we simply add one more constraint to the cost minimizing problem.

Minimize: $\sum_i \sum_j T_{ij} c_{ij}$

Subject to: $\sum_i T_{ij} = D_j$

$$\sum_j T_{ij} = O_i$$

$$\sum \sum T_{ij} \underline{\ln}\ T_{ij} \geq M$$

As this last constraint is non-linear, the solution of the problem in this form is quite difficult. However, its statement in this form shows that the entropy constraint simply adds a "noise" level into the cost minimizing problem. Obviously, if M=O, there is no "noise" and thus the expected spatial interactions will exactly equal the cost minimizing solution.

As the level of detail appropriate to land use and transportation modelling, say 100 to 200 zones for a metropolitan area of several counties, the entropy derivation seems relatively satisfactory. Thought of in terms of viewing the metropolitan area from about a mile above it, the statistical mechanics, most probable distribution, is a very reasonable way to describe the observable phenomena.

In the next section, a residential allocation model based on these principles will be developed. This will be followed by a discussion and numerical example of its proper calibration.

A SIMPLE RESIDENTIAL LOCATION MODEL

In developing equations (5), (6), and (7) it was assumed that both O_i (number of workers living in area i) and D_j (number of workers employed in area j) were known. In a residential location model we are trying to estimate O_i. We therefore replace O_i in those equations with W_i, a measure of the residential attractiveness of area i. When we do this we eliminate the need for the origins balancing factor A_i. This gives the following equations:

where

$$T_{ij} = B_j W_i D_j \exp(-\beta c_{ij})$$

$$B_j = \frac{1}{\sum_i W_i \exp(-\beta c_{ij})}$$

But, T_{ij} is the number of workers living in area i and working in area j. We wish N_i, the number of workers living in area i, to be the dependent variable, so we sum over all workplaces. This gives

$$N_i = \sum_j T_{ij} = \sum_j B_j W_i D_j \exp(-\beta c_{ij})$$

If we now substitute equation (9) into equation (10) we get

$$N_i = \sum_j \left[\frac{W_i D_j \exp(-\beta c_{ij})}{\sum_i W_i \exp(-\beta c_{ij})} \right]$$

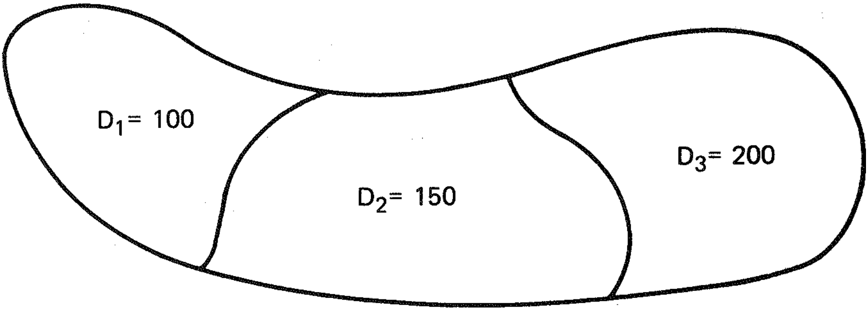
and, simplifying this equation, we get

$$N_i = \sum_j D_j \left[\frac{W_i \exp(-\beta c_{ij})}{\sum_i W_i \exp(-\beta c_{ij})} \right]$$

It may be helpful to note that the term (fraction) within the brackets in equation (12) is simply the relative attractiveness-accessibility of an individual area i compared to all other areas in the region being modelled. Thus equation (12) is merely stating that the number of workers living in i is the sum over all j of the workers employed in each j times the probability that workers employed in that j will live in i.

An obvious question here is how to measure the attractiveness W_i of each area i. Many such measures have been proposed. Current work with a version of this model indicates that a description of the prior (base year) household composition of area i along with a measure of the available land and its degree of development provide an adequate description of the area's attractiveness.[19] At this point a numerical example of the model may be helpful to understanding its operation.

Assume a region divided into three zones. In each of these zones there is a known quantity of employment.



Suppose further that the cost of travel between zones is given by the following matrix.

Zone	1	2	3
1	1.5	2.5	4.0
2	2.5	2.0	3.5
3	3.0	3.5	1.5

Next, we assume that some measure of residential attractiveness has been defined for each zone, based perhaps on its population composition and, say, density. So, $W_1 = 3$, $W_2 = 4$, $W_3 = 5$.

Finally, without discussing how we found its value, let us assume that the value of β is 2.0.

We may first calculate the matrix of $\exp(-2.0c_{ij})$. For example when c_{ij} equals 1.5 then $\exp(-2.0\text{x}1.5) = \exp(-3.0) = 0.0498$. So, filling in the new matrix of $\exp(-2.0c_{ij})$

Zone	1	2	3
1	0.0498	0.0067	0.0003
2	0.0067	0.0183	0.0009
3	0.0025	0.0009	0.0498

---

[19] Putman, S. H. (1976) "Calibrating . . .," op. cit.

Starting with the first zone, we may now calculate its employed residents.

$$N_1 = D_1\left[\frac{W_1\exp(-2.0c_{11})}{W_1\exp(-2.0c_{11}) + W_2\exp(-2.0c_{21}) + W_3\exp(-2.0c_{31})}\right] \cdot +$$

$$D_2\left[\frac{W_1\exp(-2.0c_{12})}{W_1\exp(-2.0c_{12}) + W_2\exp(-2.0c_{22}) + W_3\exp(-2.0c_{32})}\right] +$$

$$D_3\left[\frac{W_1\exp(-2.0c_{13})}{W_1\exp(-2.0c_{13}) + W_2\exp(-2.0c_{23}) + W_3\exp(-2.0c_{33})}\right]$$

Numerically this is

$$N_1 = 100\left[\frac{3(0.0498)}{3(0.0498) + 4(0.0067) + 5(0.0025)}\right] +$$

$$150\left[\frac{3(0.0067)}{3(0.00.67) + 4(0.01.83) + 5(0.0009)}\right] +$$

$$200\left[\frac{3(0.0003)}{3(0.0003) + 4(0.0009) + 5(0.0498)}\right]$$

$$N_1 = 100\left[\frac{0.1494}{0.1887}\right] +150\left[\frac{0.0201}{0.0978}\right] +200\left[\frac{0.0009}{0.2535}\right]$$

$$N_1 = 79.17 + 21.58 + 0.71 = 110.7$$

Similarly we may calculate

$$N_2 = 129.3 \qquad N_3 = 210.0$$

Thus we have 111 employed residents in zone 1, 129 in zone 2, and 210 in zone 3. Note that without using a regional control total for employed residents, their sum is 450, the number of persons employed in our closed region. To get total population from employed residents regional or zone-specific multipliers are used.

Policy tests may be made with this model by changing its inputs. Suppose, for example, that a transportation improvement between zones 1 and 2 produces a 0.5 reduction in travel cost between them, yielding the following cost matrix

Zone	1	2	3
1	1.5	2.0	4.0
2	2.0	2.0	3.5
3	3.0	3.5	1.5

Recalculating the employed residents yields 123 in zone 1, 117 in zone 2, and 210 remaining unchanged in zone 3.

Another policy might result in 50 jobs moving from zone 3 to zone 2. This gives $D_1 = 100$, $D_2 = 200$, and $D_3 = 150$.

If, using the original travel costs, this new employment distribution was used as input the resulting distribution of employed residents would be 121 in zone 1, 166 in zone 2, and 162 in zone 3.

CALIBRATION OF THE SIMPLE RESIDENTIAL MODEL

In the numerical example above, the value of β in equation (12) was assumed to equal 2.0. In model calibration the value of β, plus other parameters which may appear in more complex forms of this model, is precisely what needs to be estimated.

Consider a situation where the employed residents, residential attractiveness, employment and zone-to-zone travel cost are known for a given set of zones at a given point of time. We are interested in finding the value of β. Due to the nonlinearity of equation (12) it is not possible to use a standard regression procedure to find β. What is required is some form of non-linear search procedure which finds a best value of β in an efficient fashion.

Briefly stated, let $\hat{N}_i$ be an estimate of N_i as per,

$$\hat{N}_i = \sum_j D_j \left[\frac{W_i \exp(\hat{\beta} c_{ij})}{\sum_i W_i \exp(\hat{\beta} c_{ij})} \right]$$

Then we must define a criterion to describe how closely N_i is matched by $\hat{N}_i$, and then find the value of $\hat{\beta}$ which produces the best possible value of the criterion.

There are several different criterion functions which might be chosen, such as R^2, maximum likelihood, or minimum relative error squared. There are arguments for and against different criteria in different problems. We will continue the discussion based on R^2 as a criterion. This may be defined as:

$$R^2 = 1 - \left[\frac{\sum_i (\hat{N}_i - N_i)^2}{\sum_i (N_i - \bar{N}_i)^2} \right]$$

where, again

N_i = observed employed residents in zone i

$\hat{N}_i$ = estimated employed residents in zone i

$\bar{N}_i$ = mean of N_i

Given this function one may find a β which maximizes R^2, by a trial and error procedure. In actual practice this would be a tiresome procedure. A more efficient alternative would be to use gradient search.

The gradient of any function may be calculated from the partial derivatives of that function with respect to each of its variables. For example, for a given function, say $\mathcal{F}$ (x,y,z), the gradient $\nabla \mathcal{F}$ is given by

$$\nabla \mathcal{F} = i \frac{\partial \mathcal{F}}{\partial x} + j \frac{\partial \mathcal{F}}{\partial y} + k \frac{\partial \mathcal{F}}{\partial z}$$

Note that each term in the above vector has been multiplied by a constant (the directional unit vector) in each coordinate direction. Now, suppose we let

$$w = \mathcal{F}(x,y,z) = x^2 + y^2 - z$$

then

$$\frac{\partial w}{\partial x} = 2x; \frac{\partial w}{\partial y} = 2y; \frac{\partial w}{\partial z} = -1$$

If the directional unit vectors in each coordinate direction are defined as i, j, and k respectively, then

$$\nabla w = i2x + j2y - k$$

Thus, the gradient is a vector orthogonal (at right angles, or perpendicular) to a particular mathematical surface. If the gradient is projected back on the mathematical surface, it points in the direction of steepest ascent. More simply, gradient search is a sophisticated "hill climbing" procedure.

In doing a gradient search the gradient, ∇ , of the criterion, with respect to the parameter(s), must be found. For our case,

$$\nabla R^2 = \frac{\partial R^2}{\partial \hat{\beta}} = \frac{\partial R^2}{\partial \hat{N}_i}\left(\frac{\partial \hat{N}_i}{\partial \hat{\beta}}\right) = -2\left[\frac{\sum\limits_i \left[(\hat{N}_i - N_i)\left(\frac{\partial \hat{N}_i}{\partial \beta}\right)\right]}{\sum\limits_i (N_i - \bar{N}_i)^2}\right]$$

To continue, we refer to equation (13) and to simplify the notation, let

$$W_i \exp\,(\hat{\beta} c_{ij}) \rightarrow L_{ij}$$

$$\sum_i W_i \exp\,(\hat{\beta} c_{ij}) \rightarrow M_j$$

Now, substituting equations (16) and (17) into (13),

$$\hat{N}_i = \sum_j D_j \left[\frac{L_{ij}}{M_j}\right] = \sum_j D_j L_{ij} M_j^{-1}$$

Then, taking the partial derivative

$$\frac{\partial \hat{N}_i}{\partial \beta} = \sum_j D_j \left[M_j^{-1}\left(\frac{\partial L_{ij}}{\partial \hat{\beta}}\right) + L_{ij}\left(\frac{\partial (M_j^{-1})}{\partial \hat{\beta}}\right)\right]$$

Again, taking the partial derivative

$$\frac{\partial L_{ij}}{\partial \hat{\beta}} = W_i c_{ij} \exp(-\hat{\beta} c_{ij})$$

and substituting again from equation (16)

$$\frac{\partial L_{ij}}{\partial \hat{\beta}} = c_{ij} L_{ij}$$

Again, taking the partial derivative

$$\frac{\partial (M_j^{-1})}{\partial \hat{\beta}} = -\,M_j^{-2} \sum_k W_k c_{kj} \exp(-\hat{\beta} c_{kj})$$

and substituting again from equation (16)

$$\frac{\partial(M_j^{-1})}{\partial\beta} = -M_j^{-2}\sum_k c_{kj}L_{kj}$$

Now, substituting equations (20) and (22) into (18),

$$\frac{\partial\hat{N}_i}{\partial\beta} = \sum_j D_j\left[M_j^{-1}c_{ij}L_{ij} + L_{ij}\left(-M_j^{-2}\sum_k c_{kj}L_{kj}\right)\right]$$

$$= \sum_j D_jL_{ij}M_j^{-1}\left[c_{ij} - M_j^{-1}\sum_k c_{kj}L_{kj}\right]$$

This result can then be substituted back into equation (15) to yield

$$\nabla R^2 = -2\left[\frac{\sum\limits_i \left\{(\hat{N}_i - N_i)\left[\sum\limits_j\left(\frac{D_jL_{ij}}{M_j}\right)\left(c_{ij} - \frac{\sum\limits_k c_{kj}L_{kj}}{M_j}\right)\right]\right\}}{\sum\limits_i (N_i - \bar{N}_i)^2}\right]$$

Now, a numerical example may serve to clarify the purpose of these machinations.

Let us begin with the following "observed" data.

$D_1=100$ $W_1=3$ $N_1=111$
$D_2=150$ $W_2=4$ $N_2=129$
$D_3=200$ $W_3=5$ $N_3=210$

and

$C_{11}=1.5$ $C_{21}=2.5$ $C_{31}=4.0$
$C_{21}=2.5$ $C_{22}=2.0$ $C_{23}=3.5$
$C_{31}=3.0$ $C_{23}=3.5$ $C_{33}=1.5$

This data set represents the data which would normally be collected for use in calibrating an actual land use model. The question in such a case would be to determine a value of β that would give the best fit between the models estimates of $\hat{N}_1$ and the observed N_1 values. In this example, we know that a value of -2.00 for β would give a perfect fit. (Because, the observed N_1 in this example were obtained in the previous pages by actually working through the model with a value of -2.00 for β.) We may begin with a trial value of -1.00 for β. With this we get:

$$\hat{N}_i = 101 \qquad \hat{N}_2 = 132 \qquad \hat{N}_3 = 217$$

We may then calculate the R^2 between these estimated $\hat{N}_1$ and the N_1 observed. When we do, we get $R^2 = 0.97455$ and, if we also calculate ∇R^2, we get -0.01882 as the value of the gradient of R^2. This tells us to move the value of β in a negative direction in order to improve the R^2. Suppose, then we try a new value of -1.52 for β. With this we get

$$\hat{N}_i = 107 \qquad \hat{N}_2 = 129 \qquad \hat{N}_3 = 214$$

If we then calculate the R^2 between this second set of estimated N_1 and the N_1 observed, we get $R^2 = 0.99216$ and we get a value of -0.3306 for the gradient of R^2. This tells us that the R^2 has improved, but that it can be further improved by moving β further in the negative direction. Suppose that we now take a new value of -2.52 for β. Then we get

$$\hat{N}_1 = 113 \qquad \hat{N}_2 = 131 \qquad \hat{N}_3 = 206$$

We then calculate $R^2 = 0.99486$ and 0.01556 for ∇R_2. Again, we have an improvement in R^2, but now the gradient tells us to make a positive increment in β in order to get a better R^2. At a value of -2.12 for β, we get

$$\hat{N}_1 = 112 \qquad \hat{N}_2 = 130 \qquad \hat{N}_3 = 208$$

We calculate $R^2 = 0.99963$ and $\nabla R^2 = 0.00597$. So, we take another positive step. At a value of -1.96 for β, we get

$$\hat{N}_1 = 111 \qquad \hat{N}_2 = 129 \qquad \hat{N}_3 = 211$$

And, we calculate $R^2 = 0.99995$ and $\nabla R^2 - 0.00234$.
Finally, a small negative step to -2.00 for β yields

$$\hat{N}_1 = 111 \qquad \hat{N}_2 = 129 \qquad \hat{N}_3 = 210$$

And, $R^2 = 1.00000$ and $\nabla R^2 = -0.00000$. Thus, we are at the optimum or maximum value of R^2, as a function of β.

There are, of course, a number of ways to make this procedure more efficient, but the principle remains the same. There are also choices to be made as to search strategy. For example we could search to maximize the value of R^2, or we could search to minimize the value of ∇R^2. We should note also that for actual data sets we do not expect to get an $R^2 = 1.00$; rather we continue the search procedure until we get a $\nabla R^2 \approx 0.0$ and assume that this is the best set of parameters which can be gotten for the given model equation and data set. In practice, the use of these search programs is no more difficult than the proper use of standard multiple regression packages.

INTEGRATING TRANSPORTATION AND LAND USE FORECASTING

The simple residential model derived above illustrates the general structure and an appropriate calibration procedure for more sophisticated operational models of the same form. One such model is DRAM, mentioned at the beginning of this paper. Developed from the general principles described above and calibrated as shown, DRAM has yielded consistent sets of parameters and reasonably good data fits for five different U.S. metropolitan areas in the 1960 s or 1970 s plus a most interesting (and still consistent) set of results for nineteenth century Philadelphia.[20] It is this model which is now included in the Integrated Transportation and Land Use Package—ITLUP.

The direct connection of the land use and transportation models should now be easily seen. Any Lowry type residential model allocates employees to place of residence. In so doing it generates an implicit matrix of work trips. If such a model is not described via entropy-maximizing it is nevertheless possible to extract this implicit work trip matrix, apply conversion factors to put it in terms of vehicle trips, and use these trips to load the network. (We note that shopping trips may be derived in a similar fashion.) This is what was done in the earlier versions of ITLUP where a modified PLUM model was used for residential locations. If an entropy derived model is used for residential location, then a similar trip matrix is an explicit part of the model formulation and operation (see equation (8), for example). In such models this trip matrix, again modified by conversion factors from person-trips to vehicle trips, may be taken directly from the model run and loaded on (assigned to) the transportation network.

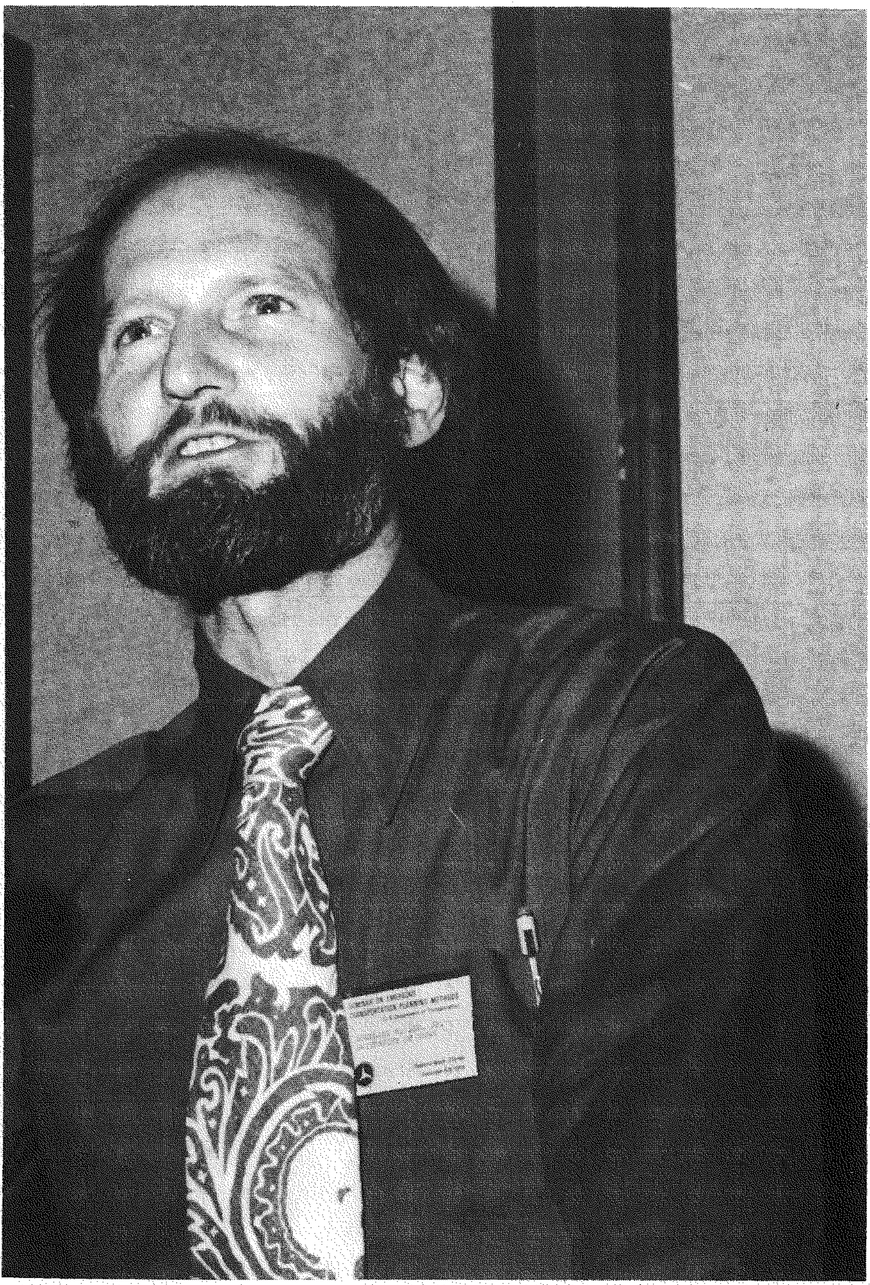
The simplicity of these statements is, in a sense, anticlimactic. This is the way to link land use and transportation models. Once the structure of these land use models is understood, the linkage to transportation is both obvious and inescapable. While the addition of land use models along with the use of capacity restrained networks and feedback from congested networks may seem to be an awesome increase in model complexity, our working experience with the model package suggests otherwise. We anticipate the eventual inclusion of similar land use

---

[20] Putman, S. H. (1977) "Calibrating a Residential Location Model for Nineteenth Century Philadelphia" *Environment and Planning,* forthcoming.

models in standard transportation planning program packages such as the UTPS. Once the normal feelings of uncertainty about new techniques are overcome, their use will be seen to be quite straightforward.

The ability of such integrated model packages to properly represent important metropolitan level phenomena normally overlooked by separate land use and transportation models suggests they will play an important role in metropolitan regional planning in the coming decade.



Douglass B. Lee, Jr., Associate Professor
University of Iowa

Improving Communications Among Researchers and Planners in the Transportation and Land Use Field

Douglass B. Lee, Jr.
Associate Professor, University of Iowa

INTRODUCTION

The problem of how to plan transportation facilities in conjunction with land use activities, recognizing the mutual interdependencies and the large but subtle impacts of each upon the other, has remained challenging but awkward to work on and, as yet, unsolved. Several major schools of thought have emerged on transportation and land use planning methodology—notably the comprehensive plan-making process and the large-scale transportation and land use models—but most of these seem to suffer from a conceptualization of the solution that is somewhat remote from the reality of the political decision process.

To address this issue, a conference was held in November 1976 at the University of Pennsylvania in Philadelphia. The conference organizers chose not to discuss and evaluate the efficacy of various approaches to, and methodologies for, addressing transportation and land use planning issues. Instead, the following strategem was devised: a small number of case studies were prepared illustrating how political decisionmaking takes place in typical settings, and these formed the basis for discussion by professionals and researchers in the field. Another feature of the case studies was that actual participants—including politicians—were involved in the discussion. This strategem was implemented through the Forecasting and Evaluation Division of the Office of the Secretary of the U.S. DOT, and relied heavily upon the Committee on Urban Activity Systems of the Transportation Research Board. Carl Swerdloff is Chief of the former agency and Dave Boyce is Chairman of the committee.

This paper is the result of the effort. Excluded from the present version are the two case studies themselves, contained in the final report to the U.S. DOT.

In the past, transportation planning and land use planning have been undertaken separately, by separate agencies, largely at separate levels of government, and transportation planning itself was largely undertaken by mode. More recently, the need to integrate these activities has been recognized, but the means for doing this—from institutional structures to analytic techniques—are still subject to considerable debate. As experience is gained, as questions are resolved and new questions emerge, as tastes and values and perceptions change, the need for effective communication increases.

At one level, communication is something akin to technology transfer—the dissemination of concepts and techniques that improve the substantive quality of technical efforts. At a more general level, communication concerns the interactions between and among professionals and researchers, with regard to how the state-of-the-art can be better utilized and advanced. At the highest level, communication between the technical and political sides of the decision process in the transportation and land use planning field is the subject of paramount interest. The scope of this report is primarily within the last two levels, although the conclusions and recommendations have clear implications for the kinds of technology that are needed and usable.

Researchers, Professionals, and Policymakers

For discussion, persons in the land use and transportation planning field are grouped into three categories:

1. *Researchers.* Academically-based non-profit institutions and individuals, with a variety of disciplinary backgrounds.
2. *Professionals.* Publicly employed land use planners, transportation planners, consultants, and federal administrators. Administrators of research funds (e.g., NSF, DOT), consultants who engage in contract research, and non-profit research institutions are included in the professional category primarily because of organizational characteristics rather than research capabilities.
3. *Policymakers.* Politicians, appointed officials, citizens, trade and professional groups, and others who exert influence on decisions through the political process and who are expected to represent the interests of a particular constituency.

Some of the lines are, of course, hard to draw, but doing so helps simplify the discussion. Figure 1 illustrates the relationships.

Subobjectives

Within the overall topic, three subtopics emerged in the discussion at

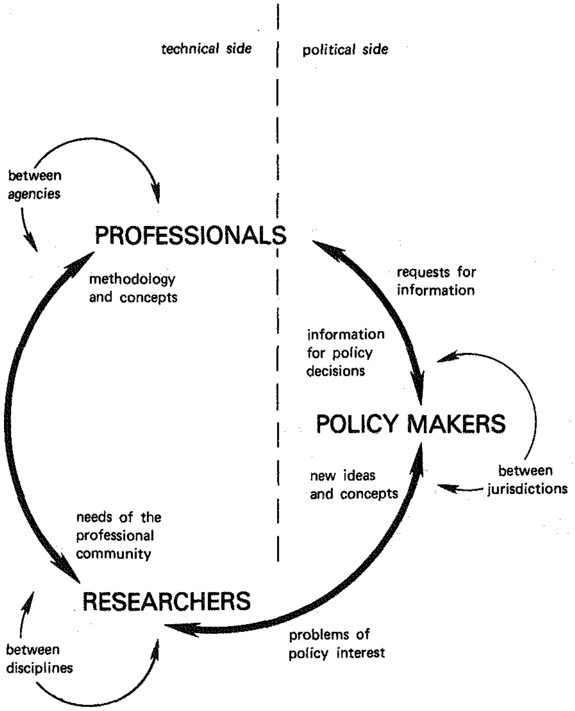


Figure 1.—Relationships among researchers, professionals, and policymakers.

the workshop:

1. Obtaining better utilization of research results (including professional experience) already available, both among professionals and among policymakers.
2. Improving the professional response to policy questions.
3. Informing the research community of the technical needs of the policymaking process, and encouraging researchers to work on these topics.

Conclusions and recommendations from the workshop discussion are presented first in the report, followed by the case studies themselves. Some readers may wish to read or skim the case studies—to gain an idea about the context for the discussion—before plunging into the somewhat more abstract conclusions.

GENERALIZATIONS AND CONCLUSIONS

The main virtue of case studies as a format for discussion is that they demonstrate the decisionmaking context into which technical information may or may not be incorporated, and by implication, the kinds of technical analyses which might potentially be useful in planning decisions. One possible undesirable feature is the difficulty of separating, in this instance, aspects of the transportation and land use planning process from the changing environment in which it is currently imbedded. Familiarity with both is undoubtedly of help in this regard.

The Current Transportation Planning Process

An overall hypothesis presented to the workshop was the observation that the nature and style of the transportation and land use planning process has shifted drastically over the past decade or less, and that the technical side of this process has not kept pace with the needs of the political side. The technical process, as it has evolved, seems to exhibit a number of drawbacks from a policy standpoint:

1. It has a high technical closure, which is to say that the process is predicated on the assumption that the determination of the best plan or project is capable of being solved entirely on a technical basis; political considerations are then seen as a "compromise" to the ideal.
2. Despite the inclusion of feedback loops, "social and community-value factors," and the three C's, the technical planning process is inherently linear, proceeding from certain givens through a predetermined sequence of analyses to a proposed plan. Transportation planners seem to want only a batch of inventories and a land use forecast in order to produce a long-range plan.
3. As in the comprehensive plan paradigm, the transportation planning process attempts to make a complete set of decisions in conjunction with each other and simultaneously, rather than adapting a sequence of decisions to the needs of the political process.

Discussion. It seems true that the current planning process is excessively concerned with plan making, overburdened with cumbersome complex models and techniques that are more hindrance than help, and politically naive in many ways. Under federal regulations and court suits, the process has often taken on a strong case of "compliance planning": the fulfillment of the letter of procedure requirements while abandoning the intent of planning. Highway departments—as in the I–66 case—sometimes seem to feel that time is on their side, and eventually the highways will get built once the regulations are satisfied. The notion that the nature of the decision process has changed, irretrievably, has eluded them. Typically, the 3–C process has not involved the relevant decisionmakers nor has it adequately informed them on policy issues.

The new planning process involves a significant citizen interaction component, and professionals, as well as politicians must incorporate this interaction into the process. Citizen participation is now a critically important element, but it also brings new difficulties:

1. Counterplanning or advocacy groups may be effective in special circumstamces, but their representation may be too narrow or their resources too thin to offer anything more than simply another political bloc.
2. Citizen or political groups may abuse the intent of regulations such as NEPA, by blocking, on purely procedural grounds, projects which they oppose.
3. Citizens are commonly unwilling or uninterested in digesting alternatives at the time the decision is being made; only when immediately threatened do they react, making it almost impossible to get meaningful and timely discussion, especially about regional issues.
4. Citizen groups and political jurisdictions frequently change their minds, which makes it hard to develop long-range programs.
5. Citizens cannot evaluate choices based on transportation inputs, (freeways, buses), which is what they are asked to do now; rather, they can only choose among outcomes, e.g., lifestyles.
6. Determining who the client is and what alternatives are to be evaluated becomes a more difficult problem in a pluralistic setting.

Because the tough issues in transportation (e.g., land use, environment) are not purely technical, there is a need to incorporate citizen and political interaction into the planning process, yet it makes much harder the problem of getting things done. Initially, the administrative structure for highway design and construction was intended to centralize authority and implementation for the sake of management efficiency; now that the previous goals are not universally accepted, a different process is required.

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

Institutional Setting for Planning and Policy Decisions

Political jurisdictions participating in the I–66/METRO decision included two local counties, a handful of cities and villages, regional and subregional agencies, two states and the District of Columbia, and half a dozen federal agencies along with the Secretary of Transportation. A similar list could be constructed for the Mt. Hood controversy, although the federal involvement was not as intense and the directly-affected jurisdictions fell entirely within one state.

Without a doubt, the most striking difference between the two cases—and the characteristic which makes the Mt. Hood situation unusual—is the institutional structure used to overcome the fragmentation of authority and achieve a consensus among the various interest groups. The Governor of Oregon and the Mayor of Portland, with the concurrence of the appropriate powers at the state level, established a Governor's Task Force with adequate political and other resources to seek and articulate an alternative to the freeway in the Mt. Hood corridor. This task force was appended to and drawn from the regional council of governments, but was structured to be able to move more effectively. In contrast, the Governor of Virginia retained the maximum leverage at the state level and refused to participate in any negotiations among the affected jurisdictions.

Many persons conclude that such fragmentation of decisionmaking authority precludes rational decisions; their remedy is stronger regional government. This approach, however, is not consistent with decentralized democratic ideals and, moreover, is naive in believing that the lack of a solution is due to a lack of centralized power. What is lacking is a forum for debate and negotiation, in which the various entities with something at stake come together and work out a balanced resolution of the issues. Regional COGs are beginning to serve as this kind of forum, but the additional element of the specially constituted issue-oriented task force (as undertaken initially in Portland and belatedly for I–66) seems to offer the desirable flexibility.

Discussion. Unquestionably, the institutional framework for decisionmaking is of essential importance in the successful resolution of policy issues. The forum must be representative and have the resources to acquire the necessary (but not too much) information.

Although the Secretary of Transportation was (and still is) involved in the I–66 decision, his participation has had the effect of leaving the decision to the local jurisdictions; in Portland, the Governor was instrumental in establishing a new direction for transportation, but the decisions themselves were left to the locals to resolve. Oregon has explicitly recognized the local focus of transportation decisionmaking by requiring local concurrence on all state DOT projects. While this has the effect of diminishing the relative share of the federal impact on such decisions, the federal role is probably larger than ever in terms of absolute activity. There are some persons who are apprehensive about turning over decisionmaking to the local level, but this shift is feasible (even I–66 illustrates that) and is occurring. Other people are of the opinion that the federal role is both too large and too constraining; attempts to specify what is professional practice and what constitutes planning create much more of an overburden of unnecessary work than the benefits would justify. While there is an obvious conflict between different levels here, the workshop seemed to accept the center of gravity being roughly at the metropolitan level as a starting point.

A number of problems in intergovernmental relations—between different levels of government and between jurisdictions at the same level in the same geographic area—need to be solved if a task force approach is to succeed. At the federal level, there are still artificial constraints on financing ("red" dollars and "green" dollars) which divert funds into modes rather than problems. There is some debate as to the extent to which these barriers have been eliminated, but it was agreed that, although considerable progress has been made in introducing flexibility, strong biases still exist.

An inevitable question that arises in considering the relationship between the technical and political components of the decision process is the relative importance of technical analysis. While there are those who believe, at least implicitly, that "correct" solutions can be arrived at on a purely technical basis (implying that the political process is "wrong" if it does not adopt technical recommendations), there are also those who express the belief that whatever the political process decides must be correct.

Neither of these views provides much useful guidance to the professional or researcher, but there is a parallel dimension that has more serious implications for planning. At one extreme are those who feel that the political process is pretty well locked up or unavailable for technical inputs, and that analysis can have little or no impact; these may be called the technical pessimists. At the other extreme are those who think that the political process should—and can—recognize "good" ideas and act on them without additional effort. In between are those who see the professional role as one of serving to support and strengthen (e.g., avoid pitfalls) a political consensus already undertaken, and those who see the primary technical role as one of clarifying issues, introducing desirable alternatives not being considered, and actively seeking to achieve a political consensus. These in-between strategies are all possible, depending upon the political context and the skills of the professionals.

Another way of regarding the professional planning problem is that of attempting to obtain "weights" or values from the political side on various alternative outcomes. Success in the planning process depends on several factors, including but not limited to the skills and personalities of the individuals involved, both professional and political; good results, as things stand now, come from a combination of good work and fortuitous circumstances. The case studies provided many examples of the political uses of information (controlling the opposition, providing ammunition, agreement on concepts prior to agreement on outcomes, credibility, drawing out passive actors), which emphasize the creative ability and marketing skills required of the professional planner.

Environmental Impact Statements

Some people are of the opinion that EISs are becoming both ridiculously expensive (into the millions in some cases; the one for I–66 was on the order of half a million) and substantively inconsequential. While they served their purpose of causing the political system to review some decisions more carefully before they are made, the EIS process may have some significant flaws:

1. Courts sanctioned the expansion of scope of the term "environment" to include everything, with the result that a concept which had some technical validity when applied to the natural environment is no longer workable on the global scale.
2. The basis upon which EISs were justified was that the public should have full information about the consequences of a project, and then make up its own mind; the courts could not require good decisions, only adequacy of information. But information collected for purposes of evaluation requires a

framework for evaluation, and impact statements do not have this. Gobs of data are collected and generated, arrayed in matrices, but no normative basis exists for using it. At best, it provides ammunition.

3. While the impact statement procedure is more compatible with the political process than is the long-range plan, it is still mistaken in implicitly assuming that a major public investment decision is a one-shot choice among a finite list of alternatives.

Although certainly a step forward in many ways, the EIS process needs to be gradually dismantled and incorporated into a new planning process. Some of the ways in which this can be accomplished include:

1. A series of mini-impact statements to match the sequence of issues and public debate described above under the transportation planning process. Ceteris paribus assumptions would be much more plausible in this context, and evaluation of tradeoffs would be both clearer and more valid.
2. The structuring of information in the impact statement should be in accordance with the normative basis for the planning process (e.g., social costs, social benefits, and distributional equity).

Discussion. While the EIS process has produced some obvious examples of poorly directed effort along with many excellent studies, the effect of NEPA has been very positive overall and—most emphatically—should not be dismantled. There are however, a number of ways in which productiveness could be increased:

1. If planning agencies are in continuous close communication with various community groups and agencies which will be affected by a project being considered, many of the objections can be handled in the design stages without the need for extremely elaborate and detailed EISs. It is the controversial projects that are forced to produce the foot-thick volumes, and this is as much a political requirement as anything else; projects for which a consensus has already been reached before formal public approval are less likely to be taken to court, even if the EIS is fairly modest in scope.
2. Some guidelines could be developed to indicate when an EIS is adequate enough; supported by a little case law, most of the extreme examples would not arise. Exhaustiveness in these instances has become a legal fetish rather than good policy analysis.
3. Concerns normally addressed in an EIS can be studied throughout the proposal, planning, and design phases,

resulting in a more effective and more streamlined EIS at the time it is formally required.

4. Environmental factors should be folded into a general framework of costs and benefits.

An EIS can be used to prove almost anything, especially if the long-range context is absent as it is in the I–66 case, but even in this instance the impact statement demonstrated that the technical justification for the Virginia Highway Department's preferred solution was weak or nonexistent. The EIS process can also incorporate citizen participation, which the 3–C process does not do very effectively.

RECOMMENDATIONS

Ideas for improving communication that were presented to, or emerged from, the workshop can be organized according to the three objectives listed at the end of the Introduction; namely, improving utilization of existing knowledge, improving professional response to the political process, and areas of research needs. While the specific recommendations often spill across several of these objectives, the primary intent is usually narrower.

Improving Utilization: Professional Conduct

Professional planners often appear to feel that politicians and citizens are unwilling to accept the recommendations of the experts, and at the same time are unwilling to do their homework well enough to understand how the recommendations were reached. To them, the political side exhibits little tolerance for new ideas, cannot digest much in the way of complexity, only takes an interest when an immediate self-interest is threatened, asks questions that are either simplistic or rhetorical or designed to reinforce an already-determined position, and makes decisions largely on political grounds rather than merit.

On the other hand, politicians and citizens perceive that technical professionals are secretive and devious, never answer questions directly, couch their answers in meaningless technical jargon, rigidly refuse to consider another viewpoint or incorporate factors not already accepted, are obtuse either by preference or by training, fail to comprehend political realities and take them into account, and do all their plans in private before venturing to discuss anything with the public. Requests by politicians for more information from agencies supposedly working for them are met with responses that either duck the question asked, take too long to be of any use, or simply recite dogma already presented.

Both of the above images contain a modicum of truth and the communication which results seldom provides much guidance for public policymaking. While improving this situation requires a joint effort, there is much that professional transportation planners can do in the short run:

1. Structure the planning process to introduce concepts and receive feedback on them on a continuous basis, not at the end of a plan-making effort.
2. Present information with a minimum amount of extra baggage, providing detail and backup only if it is asked for.
3. Refer generally to logical ideas that can be presented with clarity, as opposed to technical procedures, rules of thumb, and other essentially artificial (at least from the standpoint of the layman) patterns of reasoning.
4. Respond to questions quickly, with a reformulated question if that is necessary and will serve to sharpen debate.
5. Consider the impacts of alternative policies on various interest groups and constituencies, and anticipate their concerns.

Discussion. With a few exceptions, workshop participants agreed that the stereotype of the politician applied to very few people they had encountered, while the professional stereotype applied to all too many. Most politicians are conscientious and make the best of the resources available to them, and professionals could go a lot farther than they currently do in meeting the needs of politicians.

The specific recommendations were thoroughly endorsed, and the workshop provided numerous personal examples of each of the proposed ways of improving communication with the political side of the planning process. Participants also added several more to the list of recommendations:

6. Aggressively seek to bridge or cross organizational lines in order to facilitate agreement on issues. Both of the case studies—Portland, in particular—demonstrate that the most effective professionals move freely between various agencies and levels of government in dealing with an issue, and also interact directly with the political side on an informal basis. Bureaucracies, by their nature, create obstacles to communication which planners should work to overcome.
7. Change roles occasionally, even if only on a temporary or part-time basis. Both formal education and formal or informal political activity can provide the benefits of alternative viewpoints, as can a change of agencies.

8. Anticipate questions before they are asked, and be prepared to answer them.
9. Because of the shifting content and style of the field, many professionals are poorly educated for their current roles; workshops, continuing education, and variable means of technology transfer could help ameliorate this problem.
10. Professionals should recognize that they are, in fact, part of a political process, and that the line between political and technical is precise only in the abstract. Consultants are more used to seeing the world this way than are agency professionals.

In reference to the full list of recommendations, a number of comments are noteworthy: more information is not an unmitigated good, and may serve instead to muddle debate or distract from the main issue (a planner is not a vacuum cleaner spewing out data); simplification does not mean that the planner is substituting for political values, but planners must learn to be balanced without being exhaustive; professionals should know their jobs and do them, while insisting that politicians do the same; providing quick answers frequently requires techniques that have not been developed or are not readily available; making decisions in the open is hard on politicians, and inadequate or poorly presented information is extremely frustrating.

Improving Utilization: Removing Institutional Barriers to Policy Research

At the extreme (between, say, professional land use planners and academic economists), communication between academics and professionals is nonexistent; at best (probably between policy-oriented academics and contract research consultants), the interchange is minuscule in comparison to what would be desirable. Professionals view academics as unconcerned and ignorant about professional problems, and unresponsive to their needs; academics sometimes regard professionals as hostile and unreceptive to new ideas.

Academics depend upon peer evaluation for advancement, and policy-relevant research or policy applications are worth few credits while being, at the same time, more difficult. Universities are conservative institutions and prefer academic acclaim (which is safe) to policy relevance (which is controversial and risky). These biases can be overcome by individual researchers, but it would help if external incentives (awards, prestige, visibility) could be favorably enhanced.

Professionals have a hard time escaping the day-to-day pressures on the job (particularly if they are productive professionals), and they cannot take the time to sort through and digest the huge pile of publications that might be relevant. Ideas and methodologies intended for professionals should be thoroughly thought out from the professional's perspective and effectively presented; if a professional is expected to spend time and energy wrestling with new ideas, he or she should be able to feel confident that the expenditure is worthwhile.

There has long been a small share of academics who talk directly to policymakers, sometimes productively and sometimes not. When academics can make their ideas understandable and can incorporate political considerations into their analyses, they are generally well received. For policy-related fields, this communication is valuable and ought to be encouraged. Unfortunately, when academics attempt to communicate what they have learned about important policy problems to the professional community, the reception is often less favorable.

Discussion. The desirablility of breaking down barriers to the conduct of applied research and the utilization of those results is evident, but it is hard to specify how to do it since "good" policy-oriented research is a very judgmental thing. Policy relevance lies in the eye of the beholder. A few suggestions, however, can be made:

1. Ask academics to help formulate the problem. Even if some academics may not have a highly developed sense of the nuances of public policy, they often have good ideas, and the debate generated is worthwhile in itself.
2. Encourage academics to adopt professional roles on a temporary or part-time basis; again, the experience is valuable for all in the long run.
3. Provide some rewards—acclaim, citations, research funding—for good policy research, to offset the inherent university bias against it.

Many good examples of these kinds of activities can be found, but they could be undertaken on a larger scale and more systematically.

Professional Responsiveness: Short Range Decision Focus

Conflicting but strong evidence exists that the Federal DOT is moving away from the large-scale comprehensive plan-making approach developed in the 1950's and epitomized by PPM 50–9. Some of this evidence is reflected in short-range transportation improvement programs (TIP), transportation systems management plans (TSM), revised guidelines and procedures (PPM 90.4 and Federal Register, September 17, 1975), "alternatives analysis" policy, simplified hand-held methodologies, pricing and other demonstration programs which attempt to derive the most from existing facilities, low capital requirement strategies, and greater emphasis on policy rather than plans. Slowly, but eventually, this redirection will become spelled out in more detail, and state, regional and local agencies will follow suit.

While the new emphasis is certainly an improvement, so fundamental a shift cannot be carried out within the existing transportation and land use planning paradigm. What is needed are mechanisms for incorporating long-run implications and concerns into a short-run, sequential, incremental decision process.

Discussion. A strong love-hate conflict seems to surround this topic. While almost everyone acknowledges the improvement in policy responsiveness that comes from greater emphasis on immediate concerns, there are many who are very uncomfortable with what appears to be a purposefully and even perversely myopic view. Many professionals who appear to be perfectly happy operating without a plan still want to produce one when given the chance. Always there is a search for something fixed, something that can be taken as given, which most commonly turns out to be a long-range land use plan. If only we had some agreement on where we are going, it is said, operating in a short-range framework would present no problems.

The Washington-Northern Virginia circumstances of I-66/METRO clearly lacked such an agreement on regional land use and transportation policies, and planning suffered from that lack. Major de facto policies seem to be in direct conflict: rail rapid transit and expanded freeway capacity in the same corridor, both serving the District; environment, growth management, neighborhood protection, and transit concerns in concert with an insistence on low density, unregulated, auto-oriented development and life styles. In Portland, the major policy direction was provided by the agreement of the dominant political forces to pursue a pro-transit, reduced auto strategy, without knowing in detail what the land use consequences would actually be.

Obviously it is not possible to keep all options open indefinitely, and if some are to be closed off, it ought to be done self-consciously rather than by default. Making marginal changes to an existing system which is generally accepted as correct presents few problems, but reaching agreement on new direction and following through on it is another matter. The mess we are in is certainly a result of the failure of long-range planning, and pretending we don't need it won't make the problem go away.

Research Areas: Simpler Methodologies

Increasing emphasis is being placed on simplified manually operated procedures, quick response analytic techniques, sketch planning methods, and approaches that could be characterized more as policy analysis rather than plan making. A technical process that is capable of grappling with complex and long-range issues in a short-range context must offer several important features:

1. The ability to break complex problems down into a series of smaller and simpler problems that can be addressed in an ordered sequence and will still result in a resolution of the overall issues. Obviously, this ordering must be tailored to the particular problem and political context
2. The ability to focus on issues with selective vagueness, i.e., leaving most of the problem implicit while one portion is dealt with. This could be thought of as directing a powerful light on the trouble spot from the right angle, leaving the rest in darkness.
3. The ability to clarify—even oversimplify—the immediate issue so that the political process will be guided by the most useful concepts.

The dangers of not taking a comprehensive, holistic approach are greatly exaggerated; it is actually easier to consider alternative future scenarios and various contingencies from a highly focused perspective than from one which seeks to know everything in great detail.

Discussion. Proceeding along these lines is a desirable direction to go, and a great deal of progress has already been made, but there is a much longer way yet to go than has been covered so far. Professionals in agencies and consulting firms have made important contributions, as well as academics, but much time and effort are still wasted in excessively cumbersome and overly sophisticated analytic procedures. An example in the land use and transportation area is the land use models which attempt to integrate both kinds of relationships—transportation impacts on land use and vice versa—into a single optimizing algorithm; these models seem to be narrowly limited for policy purposes.

Multiple simple methodologies that look at the same problem from different angles are preferable to single elaborate methodologies, and fewer data are required. Cycles of quick and dirty analyses that iterate (in conjunction with the political process) toward a solution are preferable to one-shot solutions. The process may become messier and harder to manage, but the elimination of unnecessary effort should lead to cost savings while providing improved feedback. Resolution of the Mt. Hood freeway controversy in Portland has managed to get by without elaborate complex techniques, and much of the technical work was accomplished in a short time under conditions of continuous political interaction.

Research Areas: The Research Model

There is a tendency on the part of research funding or contracting agencies as well as some academic fields to apply a model derived from the physical and biological sciences to policy-oriented research, the epitome of this model being the controlled experiment. When experiments of a relatively modest scale can be actually undertaken, then the results can and ought to be studied in an experimental design format. For most issues in the transportation and land use planning field, the model is misleading: transportation impacts on land use (in contemporary settings) are small relative to a variety of other influences, and many of these other factors are subject to direct or indirect policy control. A good example is the study of the land use impacts of BART: a system that large cannot possibly be forced into an experimental design format, and the impacts themselves are at least as much a matter of other public policies as they are a consequence of BART. The inevitable conclusion is that, holding everything else (roughly) constant, BART has little effect on land use.

In this type of situation, social sciences attempt to replicate the controlled experiment by means of statistical procedures (econometrics, multiple regression, factor analysis, multi-dimensional nonmetric scaling, etc.), but because the variation to be controlled massively exceeds that to be measured, these methods always produce weak results. What does it mean to hold land use plans constant, on the average?

A policy oriented research strategy would start with the policy question (how should land around rail transit stations be developed?), structure the alternatives (low, medium, and high densities with environmental and adjacent neighborhood protection), and evaluate the results (incremental costs versus benefits, with amenity level held constant). The research content of such a study is to develop the methodology for carrying out studies of the type, since no integrated methodology currently exists. One of the advantages of a policy oriented strategy is that empirical questions can be resolved to the level of precision necessary to distinguish alternative policies, rather than an external disciplinary standard such as confidence level.

Discussion. Presentation of this recommendation to the workshop was not clear, and it received little discussion.The general feeling was that it related to the kinds of questions that were asked of researchers, rather than the type of methodology used; depending upon what questions are asked, researchers can be expected to apply the most suitable methodology. An example of the kind of question is "having approved METRO, how do we make the best use of it?" Certainly the answer is neither high density development at every station nor no development, but some balance; finding this balance requires research because the theory and the methodology have not been developed.

Research Areas: Substantive Research Needs

The overriding needs in the transportation and land use research field are substantive; the other recommendations offered above are for the purposes of clearing the path for the productive undertaking and effective utilization of this research. Within the substantive needs, the most important are synthetic, in the sense of constructing a new framework for evaluation of transportation planning decisions rather than extending an existing framework at the margins. Analysis is needed of the type which can aid in problem formulation and in breaking larger issues into smaller ones.

At a slightly more specific—although still very general—level, three substantive areas deserve greatly increased attention. Again, it is in the synthesis that the greatest challenges lie, for many of the pieces are already available and operational.

1. *Social Benefits.* These may accrue to individuals (travel, travel time reductions) or to a community (national defense evacuation); they may be evaluated in economic markets (indicating a willingness-to-pay) and/or in political arenas (reflecting a slightly different willingness-to-pay). Social benefits may be quantified (dollars, hours, pounds, etc.) or described in qualitative terms. Methodologies are needed for quantifying or evaluating benefits to the extent that is appropriate, and for accumulating benefits without redundancy or doublecounting; this is especially true for secondary or indirect benefits, which commonly provide the major rationale for public expenditures.
2. *Social Costs.* Essentially the same thing needs to be done for the cost side as well. There is a pervasive myth that, as professionals, we know the costs quite well and it is the benefits that are hard to pin down; but the cost side is, in fact, equally weak. Again, not all costs are readily priced (air pollution) or easily quantified (noise pollution), and many of the gaps are both conceptual and empirical. Simplified cost functions (supply functions) for producing alternative service levels on alternative modes are needed, as well as sound techniques for allocating costs to user classes.

3. *Horizontal and Vertical Equity.* Those who use the terms frequently have no idea what they mean. Apparently the most popular form of equity is vertical or distributional equity, which is the distribution of net benefits by income class, but many others are of equal importance: every interest group (including political jurisdictions, business lobbies, clean air groups, the federal government, etc.) should be able to compare its net benefits with those of other groups. Planners must be able to construct these accounts as a matter of public service and as a check on particular groups who claim their ox is being gored. Having this kind of information would also allow planners to anticipate political reactions to proposals and adjust the proposals or prepare for the objections in advance. It provides a natural linkage to the political process, as has already been encouraged.

Discussion. The social costs-social benefits-equity framework certainly can provide a wealth of useful information, and the cost-benefit framework appears to be becoming more widely used as resources in the transportation field become tighter in relation to demands upon them. It must be stressed that acceptance of the concept of social benefits does not mean that everything should be priced or even quantified—much must be left in a qualitative form, and there may be many items about which it is difficult to say whether they are benefits, disbenefits, or simply transfers. There also may be alternatives to going in this direction. Overall, however, further work in these three major areas is bound to have some positive effects on planning practice. Do not expect to see many cost-benefit ratios or the like; the "bottom line" must be supplied by the political process, and an analytic framework is nothing more than a means of organizing information for that process.

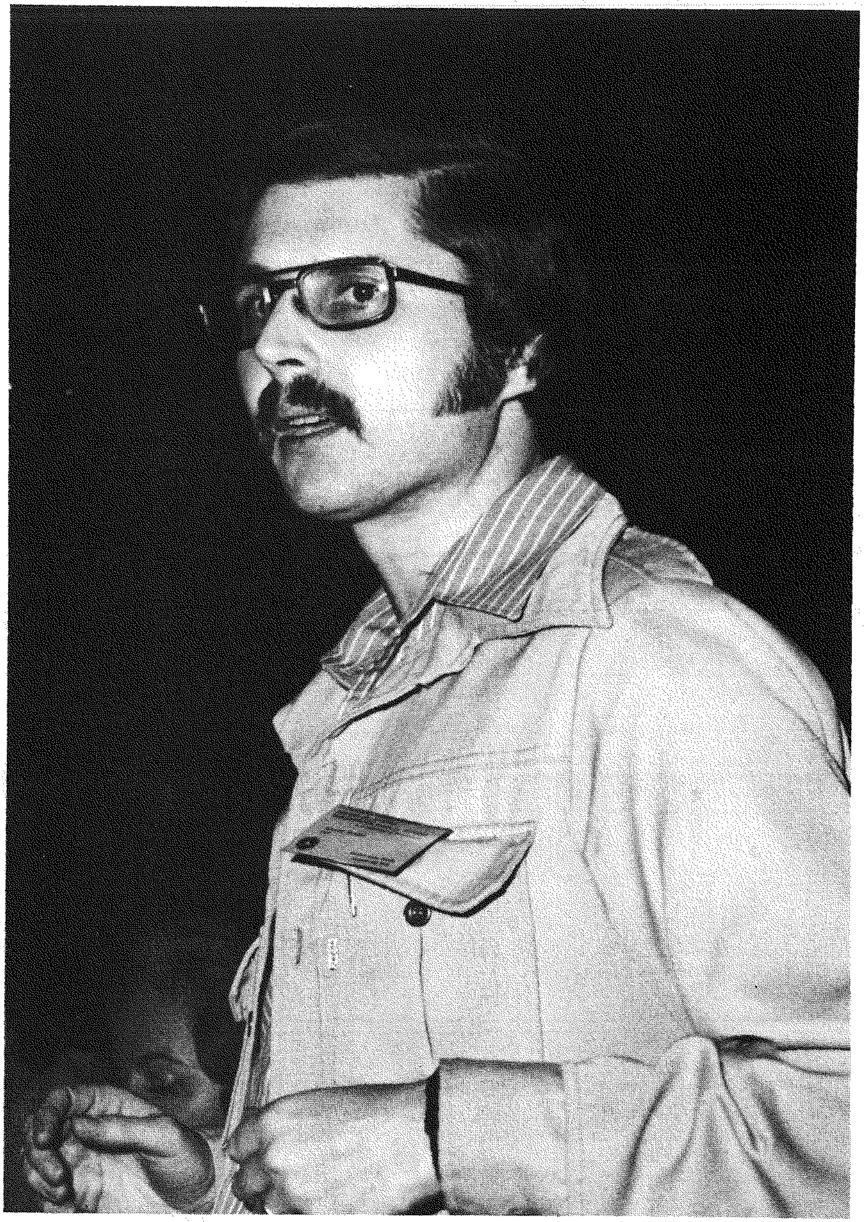
A number of more specific research areas and topics emerged in the discussion throughout the workshop:

1. *Financing.* Far too little attention is given to the questions of who pays for something and how it is paid for.
2. *Documenting Good Planning Practice.* Effort should be directed at communicating examples of the best planning available to other professionals; monitoring, evaluation, and knowledge transfer on a voluntary basis can be more effective than attempting to regulate or specify professional practice.
3. *Transportation as Policy Instrument.* How can transportation be used to help produce changes in other conditions, what are the tradeoffs between building transportation systems and locating jobs, how is transportation related to downtown

development, what are the effects of pedestrian malls on retail sales?

4. *Real Costs of Travel.* What are the full, hidden, indirect, external, etc., costs of alternative transportation forms, on a comparable basis?
5. *Level of Mobility.* Once a minimum adequate level of mobility has been achieved by a society, more mobility becomes a tradeoff between one good and many others, and we have no basis for establishing what is a good or satisfactory level of mobility.
6. *Low Capital Alternatives.* As resources become scarcer, more concern is expressed for how the existing system is being utilized. Research topics include the effect of tolls on congestion, the effect of parking limitations or charges on traffic and mode choice, and impacts of similar administrative policies.
7. *Non-passenger Priorities.* What is the proper balance between urban and rural, intra- and inter-city, passenger and freight, air versus ground, etc.?

Items on the above list are not intended to be highly refined or specific research problems, but they arose in the course of discussing the uses of information in the planning process and reflect the kinds of questions that would be of interest to policymakers if useful answers could be provided.



Edward Weiner, Manager, Urban Analysis Program
Office of Transportation Systems
Analysis and Information
Office of the Secretary of Transportation

Summary of Discussion: Transportation/Land Use Interactions

Edward Weiner

Manager
Office of Transportation Systems
Analysis and Information
Office of the Secretary of Transportation

Six workshop sessions were held. The names of these sessions were:

A. Land Use Modelling, Decisionmaking and Politics
B. Land Use/Transportation Modelling Structures
C. Application of Land Use Models in Comprehensive Planning
D. Short Range Forecasting and Land Use Impacts
E. Land Use/Transportation Forecasting at the Community Scale
F. Details of the Integrated Transportation Land Use Package

During the evening plenary session there was further discussion of the reports presented by the workshop chairmen. From this, the following major points emerged.

1. There is considerable interest on the part of planning professionals in the use of large scale urban simulation models for policy analysis.
2. Further clarification of the models' underlying assumptions is needed to provide assurance to planning professionals and decisionmakers.
3. Questions as to data requirements remain to be clarified.
4. There is a general concern as to the degree to which successful use of these models requires technical artistry.
5. Prompted by the current problems of many metropolitan areas, there is the general question as to the usefulness of these models in declining regions.
6. Apart from these technical questions, but equally important, is the general problem of involving persons other than planners and technicians in the forecasting processes.
7. Finally, there is a clear need for the development of forecasting techniques which will be applicable at other levels of detail; in particular at the community or micro scale of application.

Each of these points in discussed below in more detail in order to give a better impression of the concerns of the discussions.

1. Large scale urban simulation models have a somewhat chequered history. Many missed deadlines and broken promises in the mid 1960's led to planners taking a cynical attitude towards models in the early 1970's. This attitude seems now to be giving way to one of cautious optimism. The obvious complexity of urban problems leaves no practical alternative to large scale urban simulation models for assessing the impacts of policies on urban spatial patterns. This continuing requirement, coupled with the promise of the readily applicable and transferable models now emerging from current research, implies greater model usage in future policy analyses. Continued reporting and eventual dissemination of these models to the planning profession seem to be a much desired goal.
2. Another difficulty with large scale urban simulation models has been the obscurity of their workings and the general cloud of confusion surrounding their application. This is partly due to the models themselves, as one would expect a moderate degree of complexity to be required to simulate the very complex system of interrelationships present in urban areas.

   But the confusion is also due to the model builders. Both from a desire to protect proprietary rights and from a fair degree of confusion as to the true workings of the phenomena being modelled, the results of their work have often been more than a little bit obscure.

   While there is still confusion as to the causes of urban spatial phenomena, much of the proliferation of models has ended. Research now focuses on variants of only a few model types, with general agreement often being found on questions both of theory and application. In such an atmosphere it is far more likely that the models' workings will be more clearly described. This endeavor is of great importance, as it is only by explaining the models to their potential users that proper applications can be encouraged.
3. In past model applications, further difficulties have been caused by their considerable appetite for input data. This appetite resulted, to a great extent, from lack of knowledge of urban spatial phenomena. The tidal wave of information systems and data processing techniques which swept the social sciences with the advent of computing technology also swept hundreds of variables per zone into models whose sole working hypothesis was that, if all possible variables were included in the model, it was bound to produce good forecasts. It is now

clear that this was not a very clever approach. There were difficulties with the statistical validity of such enormous data sets, not to mention the difficulty of updating such files for input to forecasts beyond the base period. That is, if one were forecasting from 1975 to 1985 and then from 1985 to 1995 one would need to update all the 1975 input variables (by use of the model or otherwise) to 1985 in order to make the 1985–1995 forecast. Coupling these problems into the long strings of variables estimated for use as input to estimating other variables makes it quite clear why it is sensible to seek to reduce the number of model variables to the greatest extent possible. Recent research suggests that quite adequate results may be achieved with only seven to ten variables per zone. At that level, current models use no more, perhaps less, data than a planner would require for any non-model analysis procedure.

4. There is no question that proper use of large scale urban simulation models requires a certain degree of technical artistry. But, this is also true for most technical devices, including shop machinery and typewriters. The difficulty in the past has been that there was no clear division between the builders and designers of models and their users. This is customarily the situation for emerging techniques. Land use models have now reached the same state as the models in the early BPR network packages, and are soon to be disseminated by similar means. As the BPR package evolved to the current UTPS, so do we expect large scale land use simulation to evolve to a point of regular routine usage. Thus we will see a shift from requiring the artistry of a model builder to requiring the artistry of a model user.

5. The usefulness of these models in declining metropolitan areas is partly dependent on the definition of declining area. For a region which is growing overall, but which has a declining central city area, the models are quite adequate. For metropolitan areas that are declining overall, the answer is less clear as there are no current data sets on which models can be tested for this circumstance. This question is currently under investigation.

6&7. These are clearly problems requiring the further attention of model builders in the case of the small area requirements, and both builders and users in the case of decisionmaker and citizen involvement. The current types of large scale simulation model were never intended to model micro level problems and areas. Their use should not be attempted at those levels. Model builders and users have long been remiss in involving their "consumer" in the process and need to expend much more energy at this task.

These are the general comments and conclusions reached at the plenary session for the third day of the session. The general feeling seemed to be one of cautious optimism towards the future application of these techniques.

☆ U. S. GOVERNMENT PRINTING OFFICE : 1980 6[illegible]/6058