HIGHWAY RESEARCH RECORD

Number 238

Transportation System Evaluation

12 Reports

Subject Area

15 Transportation Economics
55 Traffic Measurements
84 Urban Transportation Systems

HIGHWAY RESEARCH BOARD

DIVISION OF ENGINEERING NATIONAL RESEARCH COUNCIL
NATIONAL ACADEMY OF SCIENCES—NATIONAL ACADEMY OF ENGINEERING

Washington, D.C., 1968 Publication 1625

Price: $5.20

Available from

Highway Research Board
National Academy of Sciences
2101 Constitution Avenue
Washington, D.C. 20418

Foreword

This RECORD contains ten papers and two discussions that are generally characterized by their attention to techniques of systems analysis for comparing diverse technologies, means of evaluating alternative transportation plans, and the relationship of evaluation and implementation.

Morlok presents a framework within which the type of transport service for which a transport technology is inherently well suited can be readily identified. It is his hypothesis that a major weakness has been the lack of a complete characterization of system output, so that the capabilities of diverse technologies could not be compared.

Wang, Snell, and Funk discuss a technique for attacking the network investment problem, which determines the optional investment policy in the network and on this basis assigns a given trip demand to the improved network.

Harvey outlines a method to use in comparing alternative networks that requires only the interzonal volumes and the interzonal travel times, or other measures of separation for each network, and any difference in proposed construction costs.

Domencich, Kraft, and Valette discuss an urban transportation model that relates the number of zone-to-zone trips for a given purpose and mode simultaneously to socioeconomic variables and system characteristics in one analytical step. Another feature of the model is the function of time and costs of travel, which are disaggregated into line-haul and excess times and out-of-pocket and other operating costs and are incorporated as separate variables in an attempt to measure as explicitly as possible the effects of rating variables on both the total number of interzonal trips and their modal split.

Caswell discusses a proposed solution to the problem of computing expressway usage in terms of mean trip density, of certain high-speed facilities that are placed in a given region containing trip origins and destinations.

A series of papers discuss system evaluation studies and techniques used to develop transportation plans in Chicago, Seattle, Louisville, Milwaukee, and Minneapolis. These papers, along with two discussions, focus on the implications of various approaches in the evaluation of alternative transportation systems.

Contents

The Comparison of Transport Technologies

EDWARD K. MORLOK, Assistant Professor of Civil Engineering, The Technological Institute and the Transportation Center, Northwestern University

•OVER the past decade the transportation planning process has evolved into a rather sophisticated methodology for dealing with questions related to future investment in transport facilities—particularly highway facilities. Given any proposed change in the transportation network, it is possible to predict the consequences of this change that are associated directly with the transport system. However, at least two serious weaknesses in planning methodology remain. These are in the areas of developing alternative plans (1) and evaluating these alternatives in terms of their nontransport consequences (2).

At the present time there is very little basis upon which to develop transportation alternatives. There appear to be two reasons for this condition: (a) there exists only a weak understanding of the extent to which various types of transport services can assist in achieving various non-transport regional goals, and (b) there is little basis for making comparisons between transport technologies so as to enable rational choice of the technology mix that is to provide a set of services. The purpose of this research is to attempt to develop a framework that will permit a more comprehensive comparison of alternative technologies than has been possible in the past.

In the following section, a discussion and critique of earlier works in the comparison of transport technologies serves to identify the major strengths and shortcomings of methods employed in these studies and to sharpen our understanding of the requirements for a general comparison framework.

BACKGROUND

There have been a number of notable attempts to compare the characteristics of various existing and proposed transport modes. One might reasonably ask the question: Why do these not provide the framework necessary for the comprehensive comparison of various technologies? A review of these works would be helpful in answering this question and also in pointing out the direction of the research contained herein.

Perhaps the most ambitious attempt to compare the characteristics of various modes was the work of Meyer et al (3). This study was concerned with comparing the cost characteristics of the various existing (in 1955) modes. The basic measures used for comparative purposes were the long-run marginal cost per ton-mile for freight and per passenger-mile for the movement of persons. In the case of freight, cost was also considered a function of the size of the shipment and the length of the movement, and a distinction was made between three commodity classes (liquid, bulk, and manufactured) in the case of water carriers. For passenger traffic, a distinction was made between the costs for various types of accommodations on rail and air carriers.

One of the difficulties the authors faced was in dealing with differences among modes with respect to characteristics other than cost. This difference is most obvious in the time dimension, where the differences between, say, rail and truck freight, or air and bus passenger movement, are often orders of magnitude. Meyer et al tried to account for this in the case of freight by including in the cost of shipment a cost associated with the required in-transit inventory. No explicit treatment of this was found for person movement.

There were (and are), of course, many other differences between the services rendered by the modes considered by Meyer et al. Some of these are the location of access

Paper sponsored by Committee on Transportation System Evaluation and presented at the 47th Annual Meeting.

points and the effect on service and the need for feeder operations; variations in cost due to variations in location of origins and destinations relative to terminals; safety and comfort of persons; and damage to goods.

Using their comparison framework, the only means for dealing with these differences is qualitative. This is to directly associate certain characteristics of service with the institutional label of the mode. Perhaps this is adequate if one is solely interested in services of a type offered in the past. It is, of course, not adequate to deal with services and modes that depart from those of the past.

In a later work, which dealt with urban transportation, Meyer, Kain, and Wohl (4) developed a scheme for circumventing the difficulties of noncomparability found in the previous work. The scheme used involved the comparison of three modes (auto, bus, and rail) in an idealized environment. The services of each mode were made comparable—or as nearly comparable as technologically possible—and then the costs were compared. This enforced comparability took the form of requiring very short headways for common carriers so as to bring the waiting time close to that of the private automobile, and of requiring that a seat be given to each passenger, and so forth.

Having rendered the services as comparable as possible, it is assumed that users would be indifferent between the alternatives except insofar as their cost (or price) differed. Thus the criterion used to select one "best" mode from the alternatives is that of least total cost, for the desired level of output (capacity). In this manner the range of capacity for which each mode is "best" is identified.

The severe shortcomings of this type of comparison should be obvious. It fails to take into account the fact that different modes are inherently suited to provide different types of service. These differences relate not solely to capacity, but to many other service characteristics: frequency of departures, speed, the location of points of access, and price, to name a few.

A very commonplace situation can be used to illustrate this point. Consider two cities, which are connected by air and highway. Within a wide range of conditions, a plane trip would be faster and more expensive than an automobile trip (among many other differences). If you were to try to make these two modes comparable by reducing the speed of the aircraft and perhaps increasing auto speed to the upper limit of safety, you would be destroying an inherent characteristic of the air mode. It is doubtful that a cost comparison under these conditions would tell you very much about the types of services for which each mode is well suited.

There have been, of course, a number of studies of the economic and technological characteristics of individual modes. A notable example is the work of Land and Soberman (5). The authors present the technological or operating characteristics of urban transit and this leads into a discussion of cost functions. They do discuss many of the variables associated with the service rendered by this mode, but they do not develop any comprehensive cost or performance functions. Morlok extended their analysis to develop actual cost-output functions for a linear route (6). The measures of output considered were flow capacity, headway, station spacing, and speed.

These authors do not, however, discuss many variables that are associated with the spatial properties of the service. This is not a serious shortcoming in this instance, because rail lines are generally constructed in a radial pattern. However, if a rail network with highly interacting routes were under consideration, a means for dealing with properties of the network would be desirable.

Another example of this type of work is the book by Hay (7). Hay attempts to compare the technological, cost, and service characteristics of the major modes within a fairly comprehensive framework. This work is notable in that he does attempt to give a formal structure to the comparison, which brings to light many similarities and differences between modes which otherwise might be overlooked. The characteristics used as a basis of comparison are direct cost, transit or travel time, flexibility, reliability, damage, capacity, and price structure.

The major weakness of Hay's work is the absence of either qualitative or quantitative definitions of flexibility. This word denotes a set of characteristics that deal with the locational aspects of a service and with the extent to which a mode depends upon other modes for distribution services. These aspects of service are strongly related to most of the other more readily quantified characteristics.

A somewhat similar work on the subject of urban transportation technology has been written by Berry, Blomme, Shuldiner, and Jones under the sponsorship of the Transportation Center at Northwestern University (8). This book is notable in that the authors attempted to compare directly many distinct technologies, both conventional and novel, although their treatment of the automobile mode was somewhat brief. In marked contrast to some other works, these authors brought to light the inherent differences between modes with respect to the types and quality of service which could be offered. It is considerably more comprehensive in this sense than other studies of urban transportation technology.

From our standpoint, however, this book does not offer a usable framework for our comparison of intercity technologies. The reason is that the structure used in the comparisons is a rather loose one, and is not composed of rigorously defined elements, measures, and relationships. Here we are more concerned with making an advancement in methodology that will permit the making of strong, quantitatively based statements about different technologies.

We began our discussion with brief reference to transportation planning. It is appropriate to consider the methods used to characterize alternative technologies within the urban transportation studies. Even a cursory examination of these studies reveals that they consider only conventional technologies: automobiles, buses, and rail rapid transit (9). This is invariably justified by the assumption that no major technological advances in either private or public transportation will be made within the time horizon (usually 20 years) of the study.

This summary rejection of the possibility of significant technological breakthroughs is not wholly unrealistic, because of the extent of existing facilities and the (probable) high threshold cost of introducing a major technological change. However, the urban studies have done very little to advance our knowledge of the comparative characteristics of even those transport technologies that they consider. The major result of their studies in this area has been a set of cost-output quantity relationships defined on a linear route, that is, a route composed of a sequence of links (9, p. 4-5, 38-43, 81-98, 121-127). These have been helpful in deciding which of various roadway types (e.g., freeway, arterial street) would be most economical under various conditions. In some instances, similar curves have been developed for rail rapid transit.

There has been very little effort expended in attempting to deal with spatial and temporal aspects of the output of a system. Also, those cost functions that have been developed are invariably excessively simple, so that many variations of the spatial and temporal aspects of output could not be analyzed.

A concommitant characteristic of these studies is that often very few real alternatives are examined. Studies in large cities often examine three major alternative plans: one in which improvements are made only to the transit system, one in which only roads are improved, and one which involves some investment and operating changes in both public and private transport. The best of these three is obvious, because the first two can usually be rejected on intuitive grounds. Some variations within the third class of alternatives may be examined, in which case some "real" alternatives are considered and the conclusions of the study are not largely foregone. Partly because of this lack of concern for examining a wide range of alternatives, urban studies have done little to advance the characterization of transport technologies.

The foregoing comments have pointed out a need for the development of a framework for the comparison of various distinct transport technologies. This is becoming increasingly important, because the number and variety of new technologies as well as major changes in existing ones are increasing at a very rapid rate (10), and this situation is likely to continue with the entering of aerospace firms into the field. Perhaps the most concise statement of the need for research on the comparative evaluation of transport technologies is that by Lang (11) in the Foreword of a book on monorails:

> ...There has been little attempt made, for instance, to assess soberly the characteristics of our available transportation media and to compare them on their basic merits. It seems that partly as a result of this neglect we are not solving our problems as quickly as we should.

The need for this is also strongly implied in the work of Garrison and Marble (12, p.8) on the structure of transportation networks:

> One salient feature of the voluminous material on transportation is its heavy dependence on descriptive verbal expression and the lack of exact definition and generality in this expression.... Descriptive materials varying in completeness are available on most transportation systems, but the nature of these materials is such that they are not suitable for systematic study.... There are no good ways to compare information among systems or to put together pieces of information about an individual system and make statements about the system as a whole....

PURPOSE AND SCOPE

The purpose of this research study was to develop a framework of quantitative measures and relationships that will permit the direct comparison of the properties of diverse transport technologies. Our interest for comparative purposes focused on two areas: (a) the cost properties, and (b) the properties of the transport service provided. The former category includes such items as the investment in structures, land, and vehicles, as well as the cost incurred in operating and maintaining a system. The service provided by a system can be described only by a large number of variables, which relate to such properties as the location of access points, network configuration, flow capacity provided, frequency of departures, and many others. Part of this research is, therefore, concerned with the identification and quantification of measures of output capability and relating these to technological properties of the system.

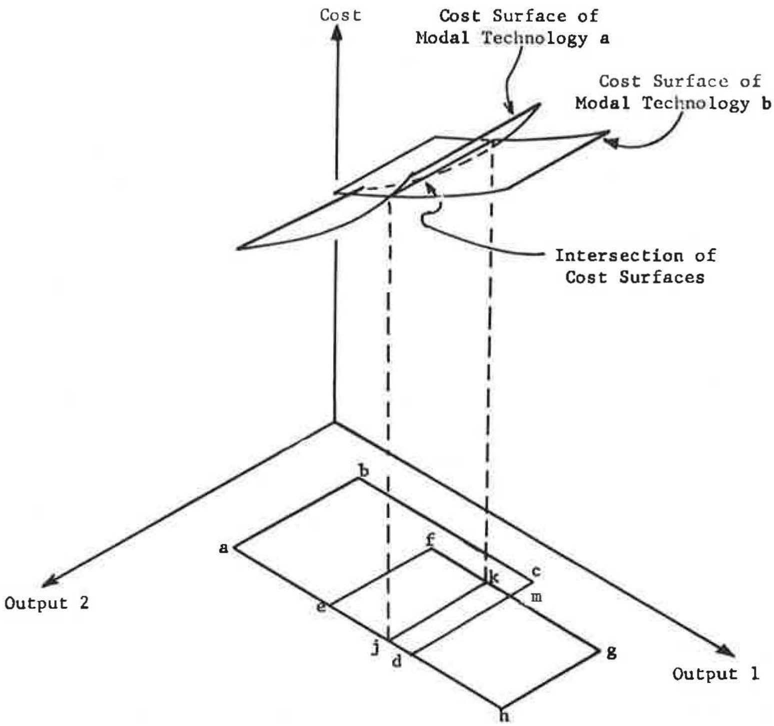
The results of this research can be viewed as consisting of two parts: (a) a vector of measures of the output capabilities of a transport system, which is related to the cost and technological properties of the elements of the system, and (b) a set of cost-output surfaces for three modes or technologies that were analyzed to test the efficacy of the approach. We are attempting to give a substantive structure to an area of study that has little basis at the present time.

In abstract terms, this research can be viewed in the following manner. We develop a vector of measures of the output capabilities of a transportation system, which describes a space which we call output space. By analyzing the physical performance properties of a transport technology, we find that portion of output space in which this technology can operate. An example of this is shown in Figure 1. Then we find the costs associated with each point in the output space and construct a cost surface on this space, after adding the appropriate cost dimensions. We can do this for any number of different technologies, and map the results onto the same space.

In order to find the types of service for which each technology is inherently suited, we need merely examine the cost-output surface for each technology. To illustrate the mechanism at work here, we refer again to Figure 1. For concreteness, we might consider the cost dimension to be total annual cost and the output dimensions to be average speed and flow capacity on a simple linear route, on which all other output variables are either of no interest or held constant. Technology a can operate over the output range abcd and technology b over the range of efgh. Although a can operate in the entire range abcd, it is not rational to use it in that range, for technology b costs less in the area djkm. Technology a should be used for outputs represented by abcmkj and technology b for outputs represented by jkgh.

Thus this scheme for analyzing transport capabilities for various technologies can result in information of the following types: (a) what the range of transportation outputs our technological capabilities enable us to produce is, and what levels of output we cannot now achieve, (b) how much it costs to produce any feasible level of output, and (c) which of the available technologies should be used at each point in output space.

Although we would like to develop this analytical methodology in such a manner that all modes of transportation could be included, this will be true only in a general sense.



Feasible Output Space of Technology a: abcd
Feasible Output Space of Technology b: efgh

Efficient Output Space of Technology a: abcmkj
Efficient Output Space of Technology b: jkgh

Figure 1. Simplified examples of cost-output space and representation of transport technologies.

The major elements of the framework and the relationships discussed will be sufficiently general and robust for this. However, the central focus of the research was on common carrier intercity passenger services. This serves to both cut down the magnitude of the task and to give it a specific orientation.

TRANSPORTATION SYSTEMS

Since this paper is conceptual in intent, a concern with definitions is a necessary prelude. Specifically, to be able to deal with the problem of comparing transport modes, we must suggest definitions for the terms transportation, transportation systems, and alternative transportation technologies.

Transportation

Our definition of transportation is drawn from the doctoral dissertation of Snell, in which he defines transportation as the translocation of objects, be they persons or goods, in physical space, in time, and in state (13, p. 52-57). State refers to such characteristics of an object as its monetary value and condition. Thus the product of the transportation of an object (in this case, a parcel of freight) can be represented in three-space as shown in Figure 2. The object is moved from location L_1 to location L_2, in a time interval $T_2 - T_1$; and the value of this good is increased from V_1 to V_2.

The transportation of many different objects, or of all objects—both goods and persons—could be represented within this framework. Of course, the number of dimensions used

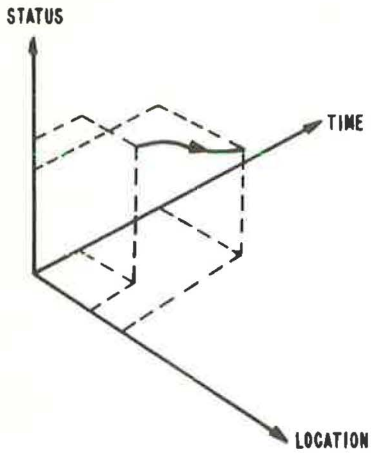


Figure 2. Transportation: The translocation of an object in location-time-status space.

to describe object state would have to be increased considerably, and many aspects of state for which we could not now give dimensions would have to be included in order to make the representation reasonably complete. In addition, a means of identifying each distinct object must be included, but this presents no conceptual problem. Thus we can define transportation as a change in the state of objects, in which each object is moved in time and physical space, and in which other attributes of each object are also likely to undergo a change.[1]

The closeness of this definition to that which is implicit in much of the regional science literature should be apparent. In the case of goods, the regional scientist usually describes the object (of transportation) by its location, the moment or period in time during which it is at that location, and its monetary value. Transportation will occur only when the increase in value resulting from the time-space translocation more than offsets the price paid for that transportation (15). Many of the conclusions of regional scientists regarding commodity flows and location decisions follow from this fundamental principle.

Transportation System

The definition of the transportation system rests in part upon our definition of transportation. We define the transportation system to be those physical objects and rules or procedures of operation which are engaged in the production of transportation. This definition is sufficiently general to include all of the current modes of transportation[2] and we feel will include any means of producing transportation likely to be available in the future.

As has been recognized in the transportation literature for almost 100 years, the various methods employed to produce transportation often display marked similarity. This is especially true among vehicular modes. The earliest published statement of and elucidation of this fact of which we are aware was written in 1870 (16); since that time there have been at least four others. Our conception of the elements of a transportation system differs somewhat from many of these, being based on what we feel are the basic functional elements of any system. These elements are (a) the way links, (b) the way interchanges (or intersections), (c) the terminals, (d) the vehicles, and (e) the control system. A definition (and example) of each of these, for the case of transportation of persons, is given in the following:

Boarding terminal—A facility which provides for the placing of travelers on the appropriate vehicle (or container), including processing of the traveler before boarding the vehicle. If the system is containerized, the transfer of containers between vehicle and the ground is included. (Airline terminal, including ticket sales, reservations, check-in and waiting areas, as well as restaurants, shops, etc.)

Alighting terminal—A facility which provides for the removal of travelers from a vehicle at their destination terminal or at vehicle transfer points. If the system is containerized, container transfer is included. (Same as above.)

---

[1]This concept of transportation is very similar to the general world view of many computer simulation languages, especially Simscript. In this language the analog of our object, an entity, undergoes changes in its status as a result of its passage through time and the operations performed upon it (14).

[2]We have deliberately ignored the question of defining the word "mode" and will not use it in this paper in any context in which a precise definition is needed. The present usage often refers to technology, ownership, and legal status, and defies any rationalization.

Way link—A path on which a vehicle can move in space. (Roadway of an expressway, between interchanges.)
Way interchange—A facility which permits vehicles to move from one way link to another. It can have many entrances and exits. (An expressway interchange.)
Vehicle—The device which is the interface between the object and the way system and which gives mobility to the object. (A railroad train.)
Control system—The set of devices, decision-makers, and associated rules which provide for the efficient and rational operation of the remainder of the system. (A traffic signal.)

It should be fairly clear from these definitions that any vehicular mode of transportation designed to carry passengers will have elements which perform the functions described. Those portions of a system designed to carry freight will have a corresponding set.

Statements as to what constitute the functional elements of a transportation system have appeared occasionally in the literature. The earliest we found was in an article by Potts rather pretentiously entitled "The Science of Transportation" (16). In this work, the elements were way facilities and vehicles. Way facilities included links, intersections of routes, and terminals. Since the author was mainly concerned with rail transportation, vehicles included locomotives, freight cars, etc. It is interesting to note that no mention was made of a control system, apparently reflecting the technology of the era. Although the discussion was clearly influenced by the then-current transport technology, Potts attempted to describe the effect on transportation service of certain technological advances, ranging from more powerful locomotives to flying machines and fluid cushion vehicles flowing through an almost frictionless medium in a tube.

In 1894 Cooley wrote on "The Theory of Transportation" (17), in which he divided a modal system into a number of elements—the way facilities, vehicles, and motive force. His discussion was essentially historical, not analytical in the sense of Potts' work.

More recent statements as to the components of a transportation system include those by Hay (7, p. 113) and Snell (13, p. 96). In both of these works the elements are essentially the same as ours. In a more recent paper by Manheim (18), the list is similar to ours except for the description of all way facilities as nodes and links. This is based on one (very common) means of abstractly representing the way facilities, which is derived from the edge (link) and vertex (node) concepts of graph theory. However, we choose to use the classification of elements based on function rather than the representation in most models, because other means of representation are possible.

Transport Technology

Although we have used the word technology frequently, we have not defined it, instead relying on the reader's familiarity with the common meaning of the term. The field of economics does have a very specific meaning for this word. A technology is a specific means of producing a good, including the capital equipment, labor, raw materials, and rules of operation (19). Thus the economist's definition of a technology is a very complete specification of the means by which an item is produced. Indeed, if one can employ two different means to fasten two parts of an otherwise identical machine, each of these technically represents a distinct technology.

This definition of technology is not entirely suited to our needs, for it forces us to look at the world in far more detail than we need or care to. Therefore we are taking a more pragmatic approach to the definitional question and will use a less precise, but hopefully more useful, definition. To distinguish it from the economist's definition we will call it a "technology class." We define a technology class as a set of technologies which are capable of producing transportation and which are sufficiently similar that the hardware and labor characteristics and operating procedures can be described by essentially the same variables and relationships. Thus we would consider a railroad operating between two cities at one capacity-train frequency level and essentially the same railroad operating at another capacity-train frequency level as the same technology class. Since different quantities of manpower, cars, locomotives, fuel, etc., would be required for these two different outputs, these are distinct technologies in the economist's sense. However, a vertical take-off and landing aircraft system and a railroad would be considered as two different technology classes, even if their products in terms of places served, capacity, frequency, etc., were essentially the same. Rail and air are considered different technology classes because, for example, the equations and variables used to describe the aerodynamic lift properties of an airplane are not used—and presumably need not be used—in the design or analysis of railroad locomotives and cars.

It is important to note that this distinction between technology classes has been decreasing in recent years and is likely to become much less important in the future. The reason for this is that researchers in transportation engineering are realizing that methods of analysis of one technology class have much in common with methods used for another. For example, the algorithms used to find the shortest path from one place to another in a network are as applicable to the routing of aircraft or freight trains as they are to the modeling of motorists' route choice behavior. We feel certain that the number and scope of problems that are recognized as common to many technology classes will increase very rapidly, and with this the need for the distinction will wane. This will be due in no small measure to the recognition and treatment of problems at a much higher level than those treated in the past. The nature of these problems is such that more than one technology class must be considered and that they must be treated in much the same terms. An example of this sort of analysis is the urban transportation study, in which highway and rail rapid transit must be considered.

MEASURING THE OUTPUT OF A SYSTEM

In the preceding section we defined the output or product of a transportation system in terms of the movement of a single object—presumably a person or a single shipment of freight. While this gives us a workable definition, we clearly cannot perform analyses of major changes in the system and still treat each object as a separate entity, because the information handling requirements would far exceed our computing machine capabilities as well as our comprehension. Thus more aggregate measures of output must be developed, using as a basis the macroscopic definition of transportation.

Macro Viewpoint

The first change in output description made necessary by the macro viewpoint is that the dimension of quantity must be introduced. This permits the measurement of various aspects of the capacity of the system, and also permits the description of trade-offs between the throughput of the system and qualities (such as time) of the product.

The second, and more difficult, change relates primarily to the time, location, and state dimensions. We are essentially concerned with measuring the general output capabilities of a system, not with the actual output in terms of the change in the state of particular sets of objects. Moreover, we would like to measure the output capabilities in terms of characteristics as independent of the particular set of objects that use the system as possible.

It is not possible at this time to demonstrate that this is in fact possible. However, we can suggest ways of handling the time, location, and state dimensions in such a manner that this independence is at least a very reasonable possibility. We shall treat each of the three dimensions separately.

In the case of location, the problem of aggregation has been partially solved in transportation planning studies by dealing with a finite number of zones, each with non-zero area, rather than continuous space, or a continuum of points. The tremendous mathematical complexities introduced by a continuous representation of space—even for relatively simplified problems—appears to be well recognized and accepted. To our knowledge, only two published research studies—one by Beckmann (20), the other from the staff of the Chicago Area Transportation Study (21)—have reasonably successfully dealt with problems in a continuous spatial context. However, when using the zonal representation, the size and boundaries of the analysis zones must be made so that the approximation is reasonable in relation to the purpose of the analysis.

TABLE 1

MEASURES OF PASSENGER TRANSPORTATION SYSTEM OUTPUT CAPABILITY

- Individual path between one origin and one destination:
  - Time vector
    - Total trip time
    - Waiting time (or departure frequency)
    - Vehicle(s) running time(s)
  - Cost vector
    - Fare or other out-of-pocket charges
    - Cost of other items (meals, lodging, etc.)
  - Comfort and convenience vector
    - Number of vehicle transfers
    - Availability of passenger services (meals, entertainment, etc.)
    - Ease of ticketing, reservations
    - Physical environment (temperature, pressure, cleanliness, etc.)
    - Psychological value (status, privacy, etc.)
    - Other values of trip (scenery, acquaintances, etc.)
    - Safety (probability of injury, death)
  - Quantity vector
    - Flow capacity
    - Storage capacity
- All reasonable paths between one origin and destination:
  - Ranges of values of elements of time, cost, quantity, and comfort and convenience vectors
  - Trade-offs between values of elements of time, cost, quantity, and comfort and convenience vectors
- Entire region under analysis:
  - Location vector
    - Density of access points
    - Density of routes
    - Network connectivity, redundancy
  - Time-space vector
    - Speed (or unit travel time)
    - Fraction of time spent waiting, moving, etc.
  - Cost-space vector
    - Cost per unit distance
    - Threshold cost
  - Comfort-space vector
    - Transfers
    - Fraction of access points with various types of passenger services
    - Amenities, scenery
  - Quantity-space vector
    - Flow capacity per unit area
    - Storage capacity per unit area
  - Ranges of values of elements of these vectors
  - Trade-offs between values of these vectors

Once the problem of measurement of location is solved, one can deal with the other dimensions. Since the zonal representation is most common, we shall limit our discussion to this type of spatial representation, although much is transferable to the other representation. The other dimensions of output are treated on a zone-to-zone basis, perhaps with distinction between classes of objects within each zone. In the case of time properties of the system, there are no major difficulties. Measures such as origin-to-destination travel time and the components thereof, frequency of departures (if service is scheduled), and time spent in various qualities of environment are readily constructed. Since we can measure time objectively, the measure is independent of the particular user.

It is much more difficult to achieve this independence with respect to other state change measures. Part of the reason lies with the fact that we do not have sufficient knowledge to state what these measures should reflect, particularly in the case of person movement. There is no operationally defined and measurable person analog to the market value of a good, for example.

We are thus forced to suggest measures which, based on our intuition, are probably significant causes of state changes on the part of users. Such items as the number of vehicle transfers, the availability of meal service and rest rooms, and the temperature and wind level of the area in which one must wait for a common carrier vehicle are measures which come to mind. Items such as those suggested can be measured objectively.

A list of specific measures that might be used to describe the output capabilities of a person transport system is given in Table 1. Each of these could be used to describe the capability between each pair of zones, or further aggregation could be made. Each measure potentially refers to a distribution of values, not necessarily a single number.

Our feeling that characteristics of a transportation system can be measured objectively is supported by some of the recent thinking of economists in the area of consumer demand behavior. Goods are no longer described solely by their name, e.g., an automobile, but by a collection of characteristics which the consumer purchases, e.g., speed, seating capacity, luggage capacity, and operating costs. As is described in a recent paper by Lancaster (22), this treatment of demand results in a much richer and more useful theory than the classical approach.

One central assumption of the reformation of demand theory is essentially ths same as our assumption that the output of a (transportation) system can be described objectively (22, p. 134):

> We shall assume that the structure which we have interposed between the goods themselves and the consumer's preference is, in principle, at least, of an objective kind. That is, the characteristics possessed by a good or a combination of goods are the same for all consumers and, given units of measurement, are in the same quantities, so that the personal element in consumer choice arises in the choice between collections of characteristics only, not in the allocation of characteristics to the goods.

Examples

Very similar in outlook is the recent work of Baumol and Quant in developing a model to predict person movement between areas of the Northeast Corridor via each of many possible present and future "modes" or alternative means of travel (23). Their work is an example of the application of Lancaster's approach to explaining consumer behavior. Baumol and Quant describe the product that a traveler purchases in making a trip by (a) the origin and destination, (b) total travel time, (c) the cost or price, (d) the frequency of departures, and (e) the relationship between alternatives with respect to these variables (23, p. 12-13). Although the empirical testing of their model is incomplete, the results appear encouraging (23, p. 19-25).

The measures used by Baumol and Quant are examples of one class of measures which can be derived from the definition of transportation. Again referring to Figure 2, the measures they use would be:

1. Origin and destination, L_1 and L_2;
2. Total travel time, $T_2 - T_1$;
3. The cost or price, which is measured on the status dimension as the change in the money the traveler has as a result of making the trip, $V_2 - V_1$ (in this particular case $V_2 < V_1$, contrary to the relationship in the figure);
4. The frequency of departures, which would be shown as the alternative times at which the traveler could move in certain portions of the space from L_1 to L_2; and
5. The relationship between alternative means of travel, which would be determined by comparing measures 1 through 4 for all alternatives.

All but the first of these measures involve a subtraction of measures associated with the object (passenger) at the origin and destination of the trip. These are simple status change measures, as is the first.

Other classes of measures are suggested by the model and are actually in use. The first of these is the rate measure, most commonly found as a measure of speed, e.g., $(L_2 - L_1)/(T_2 - T_1)$. Also, one could measure the rate of change of value of status (other than location) with respect to time. For example, the rate of deterioration of perishables could be measured as $(V_2 - V_1)/(T_2 - T_1)$. For person movement, time-cost trade-offs could be measured in this manner.

Furthermore, composite measures, such as passenger-miles, or passenger-miles per hour, can be derived from the model after the addition of the quantity dimension. If the superscript i were used to designate each object in the system, then the summation over all objects in a particular (spatial) area or in a particular time interval could be made. If $L^i(T_j)$ means the location of object i at time T_j, $f\left[L^i(T_j) - L^i(T_k)\right]$ denotes the distance moved in the interval T_j and T_k, and if Q^i is the quantity of i, then the passenger-miles per unit time of a system would be measured as

$$\frac{\sum_i Q^i \cdot f\left[L^i(T_2) - L^i(T_1)\right]}{T_2 - T_1}$$

The rate of unidirectional flow past a point is measured by $\sum_i Q^i/(T_2 - T_1)$, where i is summed over the appropriate objects. Thus there exists a very large number of possible output measures, ranging from simple status change measures to rates and complex measures.

Regardless of the exact form the dimensions and measures of the output of a transportation system take, the result is the specification of a vector by which the output of a system can be described. This vector defines a space, which we shall call transport system output space. The vector elements may correspond to those listed in Table 1, or they might differ according to the problem under analysis.

SAMPLE ANALYSIS

An example of the type of analysis proposed in previous sections is presented in order to illustrate its feasibility and potential use. While this example relates to high-speed intercity transportation, the applicability of the same methods to urban transportation will be pointed out and inferences about urban transportation analysis will be drawn in the concluding section.

The Problem

The specific problem considered is that of the comparison of three technology classes that are available for intercity transportation of persons—high-speed railroad system, vertical take-off and landing aircraft (VTOL) system, and a bus system in which the vehicles operate on conventional roads and freeways (24, p. 122-158). These were chosen in part because they represent some extremes in technology and in part because data on the cost and physical performance properties of each were available. Of course there are many other interesting and viable technologies, such as tracked air-cushion vehicle systems.

Definition of Output Space

The dimensions used to define the output space for the example problem reflect the purpose of the analysis: to provide a broad statement of the relative capabilities of each of the technology classes. The generality of this purpose dictated a very concise statement of output but also required that a very broad range of alternatives be capable of inclusion. The coarseness of the output space used here reflects these desired properties.

The most convenient means of exposition of the dimensions of this space is by considering a sequence of questions about the system. First, there is the question of where a person can travel, in the sense of what places are served. Then we could ask how, meaning via what routes or links, interchanges, and terminals. Third, the question of the amount and qualities of the transportation service offered between various places arises.

The first question is answered by the variable terminal density, measured by such units as terminals per square mile. As the value of this variable increases, more places are directly served and the ease of access to the system is increased. If reference is made to a particular region, the value basically determines which places are served.

The question of how these places are connected with one another is much more complicated, and really requires a group of measures in order to be answered. These measures are related to the concept of the connectivity of a graph, a graph being one means of abstractly representing a transport network. A graph is basically a set of vertices (points) and edges (lines), as shown in Figure 3. Connectivity is defined as the ratio of the number of edges of a graph divided by the maximum number which could exist (12, p. 22-24). If we limit ourselves to graphs in which only one edge can connect the same pair of vertices (simple graphs), and to those in which no edges can intersect except at vertices (planar graphs), this measure is

$$C = \frac{e}{3\,(v - 2)}$$

where C = connectivity, e = number of edges, and v = number of vertices.

The measure, connectivity, is very useful in distinguishing between various network configurations, as is shown in Figure 3. Its power in this connection is also demonstrated by the fact that the three classical transportation network patterns are distinguished as follows (24, p. 89-101):

Spinal, $\frac{1}{3} \le C \le \frac{1}{2}$, $v \ge 4$
Grid, $\frac{1}{2} \le C \le \frac{2}{3}$, $v \ge 4$
Delta, $\frac{2}{3} \le C \le 1$, $v \ge 3$

Thus this measure appears very useful as an indicator of network structure.

At first glance it would seem that terminals should correspond to vertices and the links between them to edges. This was not found to be the most useful representation, however, because most real world networks would then have connectivities less than two-thirds and much of the distinguishing power would be lost. Rather, interchanges—places where links intersect—were taken as vertices, and links between interchanges as edges. Terminals then could occur on a link.

This enabled one to change the scale of a network by simply changing the distance between interchanges, without necessarily changing the connectivity, the descriptor of the shape of the network. For example, with identical terminal locations, we could have a grid in which each terminal was the intersection of four (or two or three, in the case of terminals on the outer edges of the graph) links. Alternatively, the grid pattern

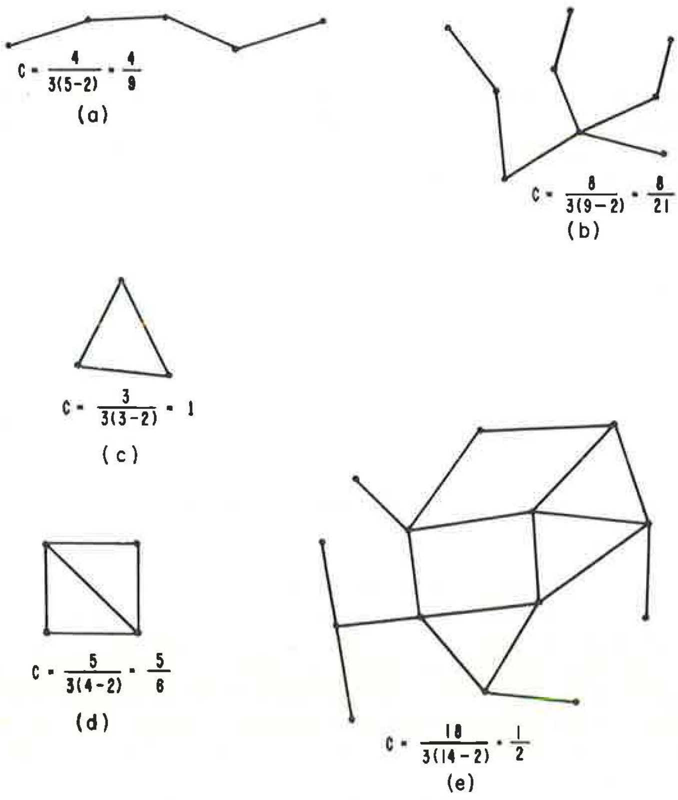


Figure 3. Examples of planar graphs and associated values of connectivities.

might be retained but only every other terminal might be such an intersection, with the intermediate terminals being on a continuous link. The connectivity measure defined on interchanges would be about the same for both, indicating the similarity of shape. Only the scale, as measured by interchange density, number of terminals per link, and link length, would differ. The portion of the system represented—consisting of the terminals, links, and interchanges—is conveniently described as the fixed network. This is in contrast to the flow or service network, which refers to the movements of vehicles on the fixed network.

The basic concept of the vehicle flow network is that of the vehicle service group. This is defined as a set of vehicles that follow the same path and make the same stops as one another, the only difference being in the time. An example of this is the set of trains that operates between New York and Washington, which follow the same route, shuttle back and forth, and make (almost) identical stops. Other examples are the express buses and air shuttle between the same cities.

Considering just a single service group, we can readily deal with two key dimensions: time and quantity. First, as for travel time between terminals, the time-distance curve is readily computed, given the route and vehicle characteristics. Another aspect of time is when movement can occur. Here we simplified reality by assuming a regular schedule, such as one departure per hour, except during two 2-hour peak periods, when departure rates are in integer multiples of the base rates.

Second, as for quantity, the common flow capacity measure of seats per hour was used.

Other important aspects of system output refer essentially to status change or perceived costs for which we now only have crude measures. One important class of these relates to the necessity of vehicle transfers on a trip. This was measured by the average of the number of other terminals to which one could travel from a single terminal without a transfer. Another aspect is the extent of express service, measured by the average of the number of other terminals to which one could travel from a single terminal without an intermediate stop.

The other measures used are availability of rest rooms, meal and beverage service, and volume of space per seat. As each of the modes was operated, these aspects of service were necessarily different, but alone probably would not have a great effect on patronage.

It is recognized that many other aspects of service are important from the user's viewpoint. However, the measures actually used in the analysis were limited to those given, in order to keep this first use of the comparison framework tractable.

Cost Functions

Cost functions were developed for each of the technologies—bus, rail, and VTOL—in which the output variables given were the independent variables. In some cases, bounds were placed on these due to limitations of the particular technology in question. Total annual cost—including annual capital cost and operating cost—was the dependent variable. Capital costs included the costs of vehicles, terminals, way facilities including right-of-way, maintenance equipment, and control facilities. Operating costs included those for labor, fuel, maintenance, terminal operation, and management. In the case of joint use of facilities (e.g., terminal) or internal services (e.g., management), with another "mode" (e.g., freight rail service), only the marginal costs are assigned to the modes under consideration. These relationships drew heavily from work done for the Northeast Corridor Transportation Project (24, p. 147-158) and the functional elements of the transportation system given earlier.

Two types of analysis were performed using the cost functions and output space. These were analysis of trade-offs possible within each mode and a comparison of total costs. The purpose of these analyses is to give some indication of the range of choices as to output variables that exist for each mode and to compare them over the relevant range. In this manner an indication of the types of services for which each is suited is given. There are basically two comparisons, one corresponding to variables associated with the service network and the other corresponding to variables of the fixed network.

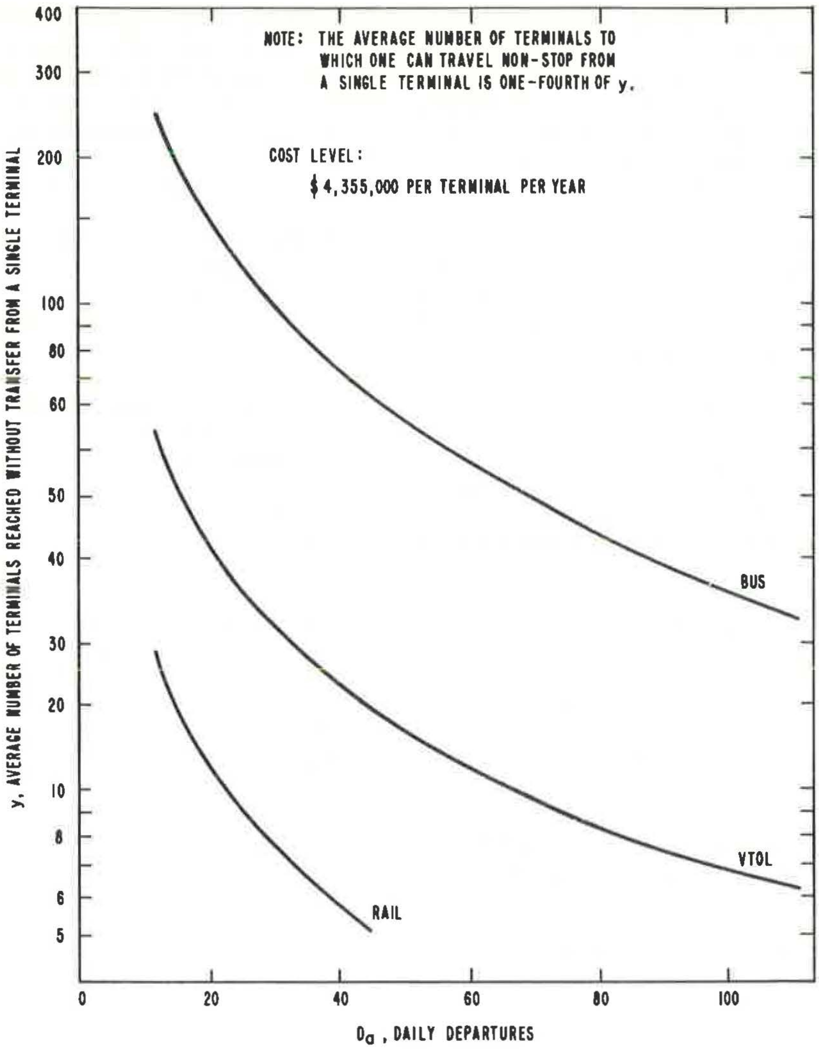


Figure 4. Trade-off between departure frequency and the number of terminals reached without vehicle transfer for rail, VTOL, and bus, at equal cost levels.

Trade-Off Analysis

The first comparison involves the relationship between the modes in terms of the two variables, (a) daily departures of each service (D_a) and (b) the average number of terminals to which one can go without transfer from a single terminal (y). The analysis considers only the cost of the service network, since there is no required change in the fixed network as these service variables change.

The results for the three modes are shown in Figure 4. The annual cost of each mode, per terminal, was held at $4,355,000. This number was chosen so that the rail curve would pass through the point ($y = 10, D_a = 23$).

The order of the modes in terms of the service variables y and D_a is unambiguous, with bus wholly dominating air and air wholly dominating rail. This ordering is precisely the inverse of that for speed, as shown in Figure 5. Thus, in terms of these three levels of service variables, no one mode dominates any other. Hence, there may be a market for each along the same route.

The reader may wonder why we do not compare the modal level of service trade-offs when the costs per unit of capacity are made equal. While this would be interesting, it is not, in general, possible. The costs of rail can only be reduced to those of bus when trains are exceedingly long—about 30 cars. Also, bus costs and air costs per unit of capacity cannot be made equal, as a moment's reflection will reveal. The ordering of service network costs per seat-mile from lowest to highest, is bus, rail, and air.

If one is only interested in departure frequency and the number of destinations reachable without transfer, and not in capacity per se, one can compare the costs of rail, air, and bus. A level of service trade-off curve coinciding with that of rail in Figure 4, costs \$388,300 per terminal with the bus technology and \$1,513,900 per terminal with VTOL technology. Thus bus costs are but 9 percent of rail and VTOL costs are only 35 percent of rail. However, there are such substantial differences in capacity that cost per seat-mile of bus is fully 71 percent of that for rail and the cost per seat-mile of air is 146 percent greater than that of rail. It should be borne in mind that there are other level-of-service differences, too, such as those with respect to speed and with respect to the very subjective area of comfort.

A second type of comparison refers to essentially spatial properties of their output, assuming reasonable levels of output with respect to vehicle flow properties. We are specifically concerned with the density of terminals (or the inverse, the tributary area) and the fixed network connectivity.

Here the terminal density, N′, is taken as a measure of the difficulty of gaining access to the intercity system. This difficulty increases with the area served by a terminal, and hence is inversely related to terminal density. Connectivity, C, is a measure of the ease of travel between terminals—assuming there is a vehicle flow on each way link, as is only reasonable. The greater the connectivity of the system, the more direct is travel between the points served. The tendency is toward reduced travel time, more places to which one can travel non-stop or without transfer, and greater capacity, ceteris paribus.

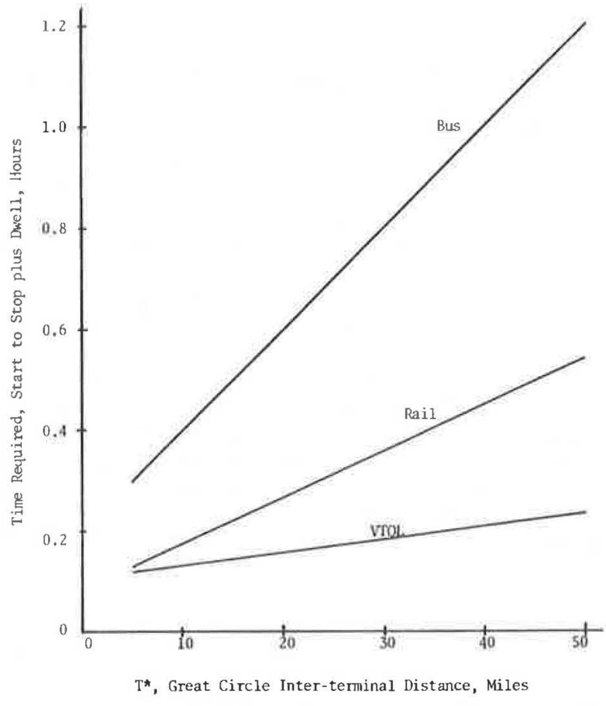


Figure 5. Modal time-distance curves including effect of circuity.

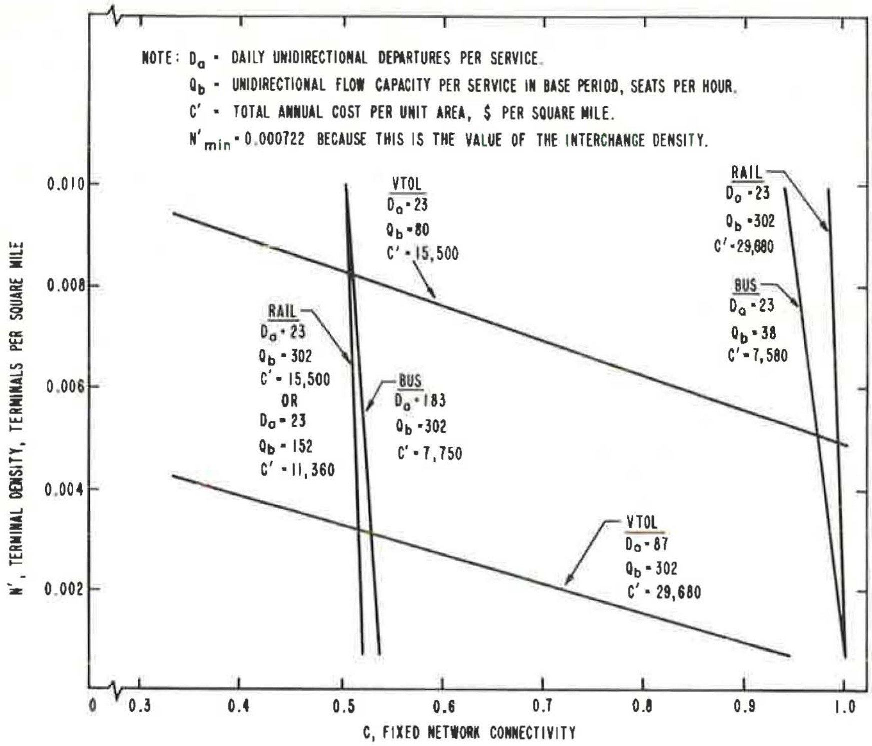


Figure 6. Trade-off between connectivity and terminal density for each mode.

The level of service assumed for this analysis is for each link to be used by two service groups. This corresponds to twice the minimum level of service required to just cover the network. On the average, a service group serves five terminals. The number of departures for all modes and the size of trains was varied so that comparisons at different levels of output could be performed. Link length was set at 40 miles and interchange density at 0.000722 per square mile, i.e., each interchange "serves" 1,385 square miles. This was based on interchanges having a hexagonal "tributary" area, which permits the widest range of network types.

The trade-off curves for the three modes under consideration are shown in Figure 6. The lower bound on terminal density is identical to the interchange density. The upper bound for analysis is somewhat arbitrary, but reasonable in that it corresponds to an inter-terminal spacing of about 5 miles. The limits on connectivity are self-explanatory.

The base level of flow capacity for this comparison was taken as that produced by rail with four-car trains, at a base period headway of one hour, for each service. This corresponds to a flow capacity of 604 seats per hour per link in each direction during the base period. Peak headways are one-half of base; hence, peak period capacity is 1,208 seats per hour per link. Since we have assumed a 60 percent load factor, these correspond to 364 passengers per hour (in each direction) during the base period and 725 passengers per hour (in each direction) in the peak periods. Because the carrying capacity of both buses and VTOL aircraft is fixed, it is necessary to operate 183 bus departures per day or 87 aircraft departures per day, per service, to produce a flow capacity equal to that of the trains. Other levels of vehicle flow (and hence capacity) were examined for all modes, also.

The trade-off between connectivity and terminal density is shown for selected values of daily departures in Figure 6. Even at the relatively high capacity of 604 passengers

TABLE 2
OPERATING SCHEMES FOR RAIL-VTOL AIRCRAFT COST COMPARISON

Scheme	Flow Capacity[a]		Assumed Usage[b]
	Per Service, Seats Per Hour	Per Way Link, Seats Per Hour	Per Way Link, Passengers Per Hour
1	151	302	181
2	226	452	271
3	302	604	362

Scheme	Daily Departures Per Service[c]		Train Length, Cars
	VTOL	Rail	
1	44	23	2
2	66	23	3
3	87	23	4

[a]Capacity figures refer to base period flow rates, unidirectional.
[b]Based on 60 percent load factor.
[c]Departures along each way link are twice these numbers.

per hour (base), the bus is cheaper than either rail or VTOL, at any particular level of connectivity and terminal density. This holds true for capacities many orders of magnitude larger than that quoted—values well beyond anything reasonable. At the stated flow capacity, and a connectivity within 5 percent of one-half, bus costs per unit area are only about one-quarter to three-tenths those of rail. As connectivity increases, the relationship is slightly more favorable to the bus. Bus costs are always less than one-half those of air, although the fraction varies considerably over the range of C and N′. However, as has been pointed out before, other values of the bus output vector are necessarily not equivalent to the rail or VTOL system. These include speed, availability of meal and bar service, and size of seat, etc.

The figure clearly illustrates the fact that bus and rail costs are essentially determined by the level of connectivity and are relatively independent of terminal density. This is indicated by the slope of the iso-cost lines, which show that a large reduction in terminal density will purchase only a very small increase in connectivity. This is not the case with VTOL. Here a reduction in terminal density from, say, 0.0004 to 0.0002 terminals per square mile, along an iso-cost curve, $29,680 per square mile, will purchase an increase in connectivity from 0.37 to 0.72.

This figure also indicates that VTOL is not always inherently better suited to highly connected networks than high-speed rail. Consider the example of desiring a flow capacity of about 300 seats per hour per link (or 180 passengers per hour per link) for which you were willing to spend about $11,350 per square mile. With rail technology this will buy a connectivity of about 0.5, regardless of terminal density. Using VTOL technology, there is a range of choice of connectivity of from 0.33 to 0.68 and of terminal density from 0.00072 to 0.0026 terminals per square mile. The rail and VTOL curves intersect at C = 0.52 and N′ = 0.0016. At values of connectivity less than 0.52 and terminal densities greater than 0.0016, rail has a lower cost than VTOL.

Total Cost Comparison

In order to find the combinations of connectivity and terminal density for which VTOL aircraft technology is less expensive than rail, and vice versa, the locus of points at which their respective cost surfaces coincide was found. This locus is dependent upon the level of flow capacity, or, more precisely, the number of departures of airplanes and trains, and the length of the trains. Loci were found for the combinations of capacity and daily departures shown in Table 2. These loci are shown in Figure 7. The locus for a capacity of 604 passengers per hour per link lies entirely below the minimum terminal density line. Rail is less expensive above this line, so VTOL is never cost optimal at this high level of capacity. This level of flow corresponds roughly to that on the main New Haven-New York-Washington rail route (25), but traffic between Philadelphia and New York is considerably heavier. This level of traffic on one mode may be rare or nonexistent elsewhere in the corridor.

The two other curves point out the strong influence of both connectivity and terminal density on the choice of cost-optimal mode. As connectivity increases, the range of terminal densities for which VTOL technology is least expensive increases. Also, as the desired level of flow capacity decreases, the larger the portion of the terminal density-connectivity space for which air is least expensive.

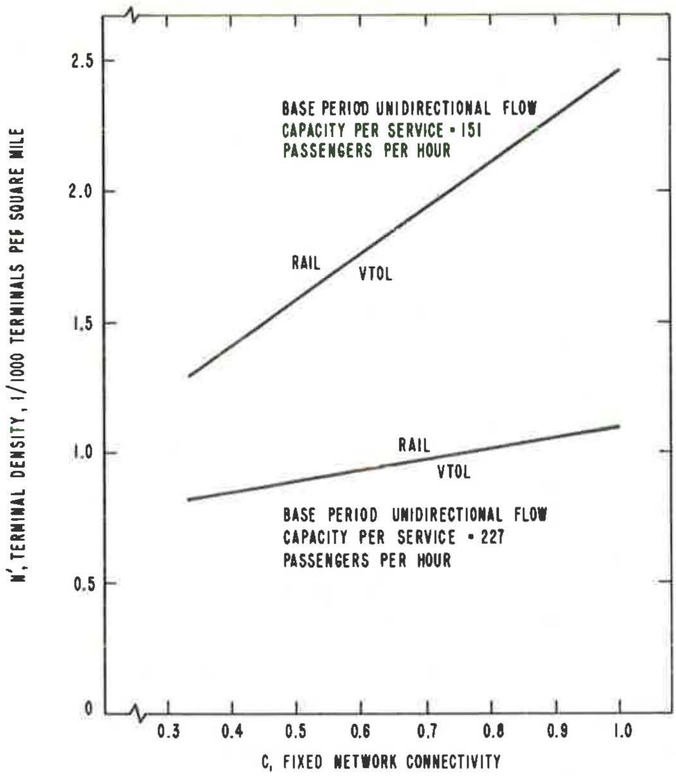


Figure 7. Rail-VTOL equi-cost frontiers. (Each line refers to a specific level of flow capacity. If the output desired lies above the line, rail is less expensive than VTOL. If the desired output lies below the line, VTOL is cheaper.)

Sensitivity of Results

Because of the speculative nature of many of the estimates of cost and technological performance parameters, particularly with respect to VTOL aircraft, it is important to discuss briefly the effect of changes in these parameters. As for the bus mode, it is clear from the cost values of Figure 6 and earlier figures that it is likely to retain its cost advantage regardless of any likely changes in parameter values. Similarly, it is likely to remain the slowest of the modes.

The loci of equal VTOL and rail costs, however, are very susceptible to change due to variations in parameters. If the cost associated with VTOL terminals were to be decreased, the slope of the iso-cost lines of Figure 6 would decrease; in other words, the N′ intercept would increase and the C intercept would remain fixed for the same total cost. Similarly, if the cost parameter associated with connectivity were to drop, more connectivity could be purchased for the same total cost. These same effects hold for both rail and bus, although with these technologies the uncertainty of the estimates is less.

The same changes will affect the equi-cost loci of Figure 7. As the cost associated with terminal density is reduced, the N′ intercept of the loci is increased, at both C = 0 and C = 1, and the entire curve is shifted upward. As the cost parameter of connectivity is decreased, the N′ intercept at C = 1 is increased, but that at C = 0 is unchanged.

A large-scale sensitivity analysis of the results was not undertaken as part of this study. This was not considered appropriate for two reasons. First, the purpose of this research was to develop a methodology or framework for the quantitative comparison of the cost-output properties of transport technologies, not to obtain better estimates of the parameters that describe each element of a transport technology. Second, there appears to be no basis in the literature for making estimates of the likely range of values of parameters, since only point estimates were given for most of these. It seems extremely premature to conduct sensitivity analyses at this stage in the development of the comparison methodology. It is appropriate that we return to this methodology, and in particular to the discussion of Figures 6 and 7.

It is likely that link flows will be below 400 passengers per hour (unidirectional) on high-speed common carrier links in most of the corridor. Also, we strongly suspect that terminal tributary areas of 300 to 1000 square miles are likely to be of greatest interest. If these suppositions are true, then the choice between VTOL and a high-speed rail system (resulting from fixed facility and vehicle improvements to existing rail lines) is not obvious. Both may be substantially in evidence, each serving its own territory. Hopefully, rational planning would provide for close coordination, so that these two technologies could operate in what the traveler would view as one high-speed system.

CONCLUSIONS

This research has been an attempt to develop a framework within which the type of transport service for which a transport technology is inherently well-suited can be readily identified. In our review of previous work in this area we found that a major weakness has been the lack of a complete characterization of system output, so that the capabilities of diverse technologies could not be compared. Therefore, much of our effort was directed toward identifying the dimensions and developing operational measures of system output. The measures developed drew heavily from the abstract notions of graph theory, as well as the more concrete concepts of location, time, and quantity. These dimensions form transport system output space.

Each transport technology can operate within a certain portion of this output space, which we call the feasible output space for that technology. Since the feasible output spaces of two or more technologies often overlap, we are interested in their relative resource use in producing similar levels of output. Also, we are interested in knowing what it costs to produce the level of output associated with each feasible point in the output space. To enable us to estimate cost functions of various technologies, we developed a generalized transport cost model, in which we associated fixed and marginal costs with each of the functional elements of a vehicular transport system. In this manner, we have developed a general cost-output space within which various transport technologies can be compared, and within which the region of output space for which each technology is inherently suited can be identified.

To test the efficiency of this theoretical development, we compared three technologies with it: vertical take-off and landing aircraft, high-speed rail, and bus. The ease with which modal cost and performance parameters were transformed into those of the functional model indicated the soundness of the representation. We then used this representation to construct actual cost-output functions for these technologies, using the dimensions of our output space. The costs and values of the components of the output vector corresponding to each of the technologies were compared, and many statements regarding the type of transportation service for which each is suited were made.

Briefly, the conclusions were that, for the Northeast Corridor situations examined, bus system costs tend to be lower than rail or air system costs for reasonable levels of output. However, the difference between the values of many variables of the output space makes the bus mode inferior with respect to some level of service properties. The feasible output space for rail and VTOL aircraft systems intersect, except for the speed dimension (where VTOL is superior), so that these technologies are very close substitutes in some regions of output space. In this output region that is feasible for both technologies (ignoring speed), the cost surfaces also intersect. Rail is preferred at low values of connectivity, high values of terminal density, and high levels of link flow capacity. In terms of output variables associated with only the service network, such as frequency of departures and the opportunities for non-stop and non-transfer travel, there exist very substantial trade-offs with cost constant in all three of the modes.

From these comments, it is clear that a good or an efficient transportation system for a region as diverse in activity patterns and density as the Northeast Corridor is likely to be one that includes many different technologies. Each of these would perform where it is relatively best suited. Hopefully, intermodal coordination will be such that the traveler can easily make use of the entire system and not be artificially restricted in intermodal transfers.

A number of inferences can be made about the analysis of urban transportation systems from this study of intercity systems. The first is that the comparison methodology presented here probably could be applied to urban technologies and situations. Of course, the specific measures of output used probably would require change. This type of analysis would assist, however, in the determination of what types of service each existing or new technology is suited for. This in turn should aid the transportation planner in developing better plans.

This study points out the existence of two distinct types of transportation alternatives, whether they be for an urban area or a megalopolitan region. On the one hand we have technological alternatives—represented by the automobile, rail rapid transit, bus rapid transit, the Starrcar, etc., and we have transportation system alternatives—represented by different points in a system output space. Most of the debate on urban transportation alternatives seems to be about technological alternatives. Yet there exists a very large number of significantly different alternatives which employ the same mode (or modes) but are distinguishable in terms of travel time, capacity, and other system output properties, as well as price and usage controls. This latter class of alternatives is just as important as the former—perhaps more so. It deserves as much attention, for it reflects the issues of the general qualities and quantities of transportation capability to be provided in a region, not just essentially how the transport capability will be produced.

The use of a comparison framework such as that presented here would aid in the identification of the range of alternatives that exist for an area. It would help to identify the bounds on systems and to indicate the direct costs associated with the various alternatives included within the extremes. Hopefully, this would uncover alternatives not now considered, and this would potentially improve the programs of transportation investments suggested by the studies. If the resulting plans drew from alternatives that would not have been considered in the absence of the formal comparison methodology, the program is presumably a better one as a result. Even if the plan does not draw from newly uncovered alternatives, however, the certainty with which the planners feel the recommended plan is best in terms of their criteria is increased, for more alternatives were considered and rejected. In either case, the planning process benefits.

The relationship between what we have attempted to do and the planning process is succinctly stated by Heymann in "Transport Technology and the 'Real World ' " (26):

> ... What the planner can do here is not to select a single best transport system, but to marshall the data on the costs of alternative transport systems that will achieve different combinations of objectives.
>
> ...the transportation planner should attempt to present to the community a series of feasible, efficient alternatives....

We hope that we have made a contribution to this effort and to the emerging science of transport systems analysis.

ACKNOWLEDGMENT

The critical comments and helpful suggestions of Dr. Donald S. Berry and Dr. Abraham Charnes of the Technological Institute, Northwestern University, and of Dr. Paul W. Shuldiner of the U. S. Department of Transportation, and Mr. Henry W. Bruck, formerly of the U. S. Department of Transportation, are gratefully acknowledged. The assistance of the staff of the Technical Analysis Division of the Nation Bureau of Standards in computer programming and execution of the graphics is appreciated.

REFERENCES

1. Ferguson, George A. Development of Transportation System Alternatives. Highway Research Record 148, p. 1-8, 1966.
2. Irwin, Neal A. Criteria for Evaluating Alternative Transportation Systems. Highway Research Record 148, p.9-19, 1966.
3. Meyer, J. R., Peck, M. J., Stenason, J., and Zwick, C. The Economics of Competition in the Transportation Industries. Harvard University Press, Cambridge, 1960.
4. Meyer, J. R., Kain, J. F., and Wohl, M. The Urban Transportation Problem. Harvard University Press, Cambridge, p. 171-306, 1965.
5. Lang, A. Scheffer, and Soberman, Richard M. Urban Rail Transit: Its Economics and Technology. The MIT Press, Cambridge, 1964.
6. Morlok, Edward K. Rail Transit Costs. In Transportation in the New Cities, Chapter VI. The Transportation Center at Northwestern University, Evanston, 1965.
7. Hay, William W. An Introduction to Transportation Engineering. John Wiley and Sons, New York, 1961.
8. Berry, D. S., Blomme, G. W., Shuldiner, P. W., and Jones, J. H. The Technology of Urban Transportation. The Transportation Center at Northwestern University, Evanston, 1965.
9. Morlok, Edward K. Chicago Area Transportation Study, Final Report, Volume III, 1962.
10. Massachusetts Institute of Technology. Survey of Technology for High Speed Ground Transport. Report No. PB 168-648 U. S. Department of Commerce, p. IV-1—IV-90, 1965.
11. Botzow, Herrmann S. D., Jr. Monorails. Simmons-Boardman Publishing Corp., New York, 1960: Foreword by A. S. Lang, p. III.
12. Garrison, W. L., and Marble, D. F. The Structure of Transportation Networks. The Transportation Center at Northwestern University, Evanston, 1962.
13. Snell, Jack E. A Framework for the Analysis of Urban Transportation Research. Unpublished PhD dissertation, Department of Civil Engineering, Northwestern University, 1966.
14. Markowitz, Harry M., Hausner, Bernard, and Karr, Herbert W. SIMSCRIPT: A Simulation Programming Language. Prentice-Hall, Englewood Cliffs, N. J., 1963, p. 3-10.
15. Moses, Leon N. Location and the Theory of Production. Quarterly Journal of Economics, May 1958, p. 259-272.
16. Potts, Joseph D. The Science of Transportation. Journal of Social Science, No. 2, p. 115-128, 1870.
17. Cooley, Charles H. The Theory of Transportation. Publications of the American Economic Association, Vol. IX, No. 3, May 1894.
18. Manheim, Marvin L. Principles of Transport Systems Analysis. Highway Research Record 180, p. 11-20, 1967.
19. Soberman, Richard M. Economic Analysis of Highway Design in Developing Countries. Highway Research Record 115, p. 44-63, 1966.
20. Beckmann, Martin J. A Continuous Model of Transportation. Econometrica, Vol. 20, No. 4, Oct. 1952, p. 643-660.
21. Creighton, R. L., Hoch, I. Schneider, M. and Joseph, H. Estimating Efficient Spacing for Arterials and Expressways. HRB Bull. 253, p. 1-43, 1960.
22. Lancaster, Kelvin J. A New Approach to Consumer Theory. Journal of Political Economy, Vol. 74, April 1966, p. 132-157.
23. Baumol, William J., and Quandt, Richard E. The Abstract Mode Model: Theory and Measurement. Northeast Corridor Transportation Project Technical Paper No. 4, U. S. Department of Commerce, 1966.

24. Morlok, Edward K. An Analysis of Transport Technology and Network Structure. Unpublished PhD dissertation, Department of Civil Engineering, Northwestern University, 1967.
25. Pennsylvania Railroad—Passengers on Board, New Haven Railroad—Passengers on Board, and Total—October 1966. U. S. Dept. of Commerce, Office of High Speed Ground Transportation, unpublished.
26. Heymann, Hans, Jr. Transport Technology and the "Real World". Paper No. P-2765, The RAND Corp., Santa Monica, Calif., 1963, p. 8, 9.

Toward a Solution for the Optimal Allocation of Investment in Urban Transportation Networks

JIN-JERG WANG, ROBERT R. SNELL, and MONROE L. FUNK
Respectively, Graduate Research Assistant, Associate Professor, and Assistant Professor, Civil Engineering Department, Kansas State University

This paper demonstrates the application of a discrete version of the maximum principle to the problem of optimal investment in a transportation network. Network investment problems that include nonlinear relationships between travel time, traffic volume, and investment are considered. The technique determines the optimal investment policy in the network and on this basis assigns a given trip demand to the improved network. The objective is to provide an investment policy that will cost least to construct and operate.

•THE economic analysis of a transportation network provides valuable guidance in developing a comprehensive, long-range transportation plan which, as concluded by Zettel and Carll (1), is the basic objective of a transportation study. Being part of the public services and competing for the use of limited resources, the transportation system should be built and operated economically, while at the same time it should meet the standards and goals of the community in order to promote growth and meet the needs of the economic activities. Specifically, the objectives of a transportation system have been summarized (2) as:

1. Provide a means for moving people and goods safely, freely and economically;
2. Provide a choice of mode of travel;
3. Make the city a more attractive place to live; and
4. Provide the means for fulfilling the travel needs and desires of the urban population within their ability to pay.

Theoretically, an optimal transportation system which best fits the economic and social objectives would be based on criteria that reflect these objectives. However, this evaluation would be very difficult if it were to be done quantitatively.

Certain aspects of transportation system evaluation are subject to quantitative analysis, such as the addition of capacity, improvement of level of service, and optimal allocation of funds for these purposes. The problem of adding capacity and improving the level of service of an urban street network has been dealt with by several researchers. In 1958, Garrison and Marble (3) presented a linear programming formulation for the analysis of network improvement. Travel cost for each link was assumed to be constant and the investment was assumed to increase the capacity linearly. The objective was to minimize the sum of the investment and travel cost subject to constraints such as flow balance, budget, and capacity limits. The simplex algorithm was employed to seek the optimal solution.

Carter and Stowers (4), in 1963, again utilized linear programming to formulate a model for funds allocation for urban highway system capacity improvement. The basic

Paper sponsored by Committee on Transportation System Evaluation and presented at the 47th Annual Meeting.

formulation was the same as that of Garrison except that each link was represented by two arcs, one with free flow capacity and normal operating cost, the other with higher operating cost (due to congestion) and a capacity equal to the difference between possible and practical capacity. The ratio of the capacities of these two arcs was kept constant as the capacities were improved.

In 1964, Roberts and Funk (5) developed a linear programming model for the problem of adding links to a transportation system. The locations of possible additional links in the system were first decided. In seeking the optimum, the additional link was either completely built or not built at all. If the link were added, the cost was included in the objective function. If it were not added, flow on the link was blocked. In this formulation an integer programming technique was used. The paper also suggested a possible application of dynamic programming in treating the stage-wise construction problem. As a result, in 1966, Roberts et al (6) combined the use of linear programming and dynamic programming techniques to solve a stage-wise link addition problem.

Hay, Morlok, and Charnes (7) developed a model for optimal planning of a two-mode urban transportation system. A two-mode system, auto transport and public transit, was to be built in an urban corridor. The auto roadway capital cost was linearly related to capacity and speed. Transit speed was fixed with the capacity linearly related to capital cost. The length of the transit route was also assumed fixed. The choice of mode was linearly related to the travel time ratio between road and transit. Again, the linear programming technique was used to formulate the problem and seek the optimum. In this formulation, the travel time was excluded from operating cost and was treated as a constraint to reflect the minimum level of service desired and the maximum speed obtainable. For a true optimum, it was necessary to change the length of the transit route and run the program several times.

Distinct from the linear programming type models, Ridley (8), in 1965, developed a method for seeking the optimum investment policy to reduce the travel time in a transportation network. The unit travel time was assumed to be decreasing linearly with the investment. It was also assumed that the flow was far below the link capacity. The objective was to minimize the total travel time. Because the travel time was a function of both investment and traffic volume, the objective function was nonlinear in nature. For some special cases, such as no budget limit, fixed traffic volume, fixed investment, and single origin-destination, the formulation can be simplified into a linear programming model. For the general case having budget and travel time constraints, the bounded subset method was utilized to search for the optimum.

HYPOTHETICAL PROBLEM NETWORK

To demonstrate the model to be formulated in this study a hypothetical urban network was created. The network shown in Figure 1 gives the peak hour trip distribution pattern. The network shown can be considered to be one quadrant of a city with the CBD at node (4, 4). Thus all trips originating at the various nodes in the quadrant are destined to the CBD.

The streets shown in Figure 1 that make up the hypothetical urban street network are assumed to have various characteristics. Each street link in the network can have a different free flow travel time and a different coefficient of investment to reflect different construction and right-of-way costs.

The network was divided into two parts by a diagonal line passing through nodes (1, 4) and (4, 1). The lower part, which is adjacent to the CBD, was assumed to be more densely developed than the upper part. Thus it was assumed that the maximum speed attainable would be 60 mph in the densely developed lower half and 70 mph in the less densely developed upper half. These are the maximum speeds used to derive the model constants in the Appendix although they may not be possible from a practical viewpoint.

THE OBJECTIVE FUNCTION AND TRAVEL TIME EQUATION

The objective of the model formulated in this study was to minimize the sum of investment cost and travel time cost. Investment was considered as an independent variable and it was assumed that it could be expressed in terms of dollars per mile. How-

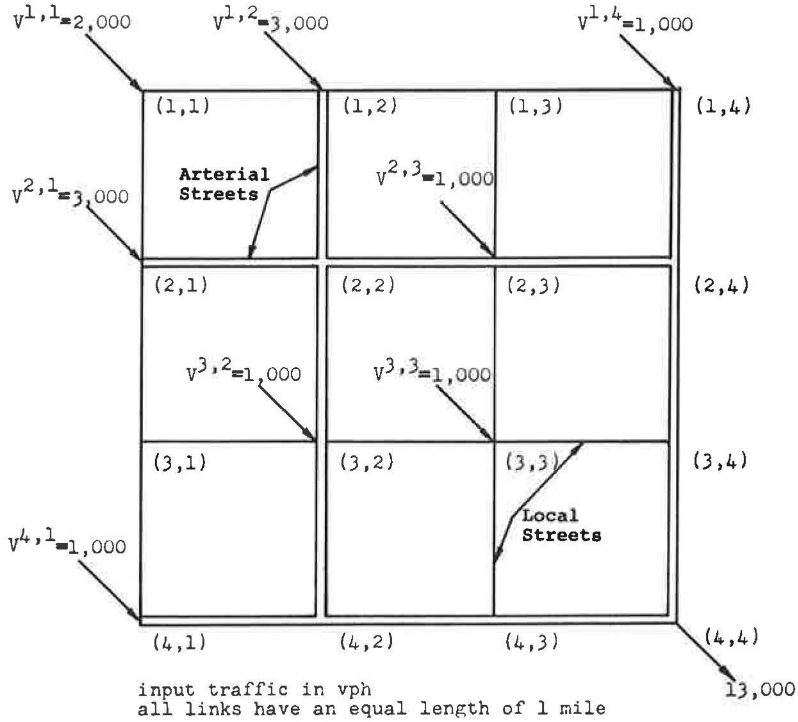


Figure 1. Hypothetical network and peak-hour traffic distribution.

ever, unit travel time was, in general, dependent on traffic volume and roadway conditions.

To express unit travel time as a function of traffic volume and investment, some basic relationships were observed:

1. Unit travel time increased as the traffic volume increased;
2. Unit travel time decreased as the investment increased;
3. Unit travel time had a lower limit (free flow travel time); and
4. With constant travel time, capacity increased as investment increased.

Keeping the basic relationships in mind and further assuming that the free flow travel time is constant for each link, and that traffic volume served is proportional to investment for a constant travel time, an equation of the following form may be hypothesized:

$$t = K_1 + \frac{K_2}{\theta} V \tag{1}$$

where

- t = unit travel time (hr/mi/veh);
- K_1 = free flow travel time (hr/mi/veh)—the magnitude depends on the maximum speed obtainable or allowed by regulation;
- K_2 = coefficient of improvement (dollar-hr/mi²/veh²)—its magnitude depends on link location and reflects the difficulty of improvement;
- θ = equivalent hourly investment per unit length (dollar/mi/hr); and
- V = traffic volume per unit time (veh/hr).

In the case where old facilities exist, the investment should be expressed as:

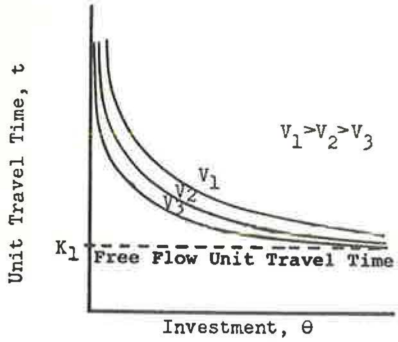


Figure 2. Travel time-investment curve with fixed volume.

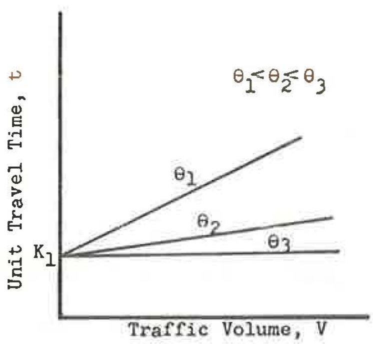


Figure 3. Travel time-volume curve with fixed investment.

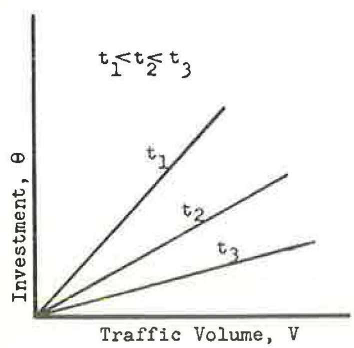


Figure 4. Volume-Investment curve with fixed travel time.

$$\theta = K_3 + \theta' \tag{2}$$

where K_3 (dollar/mi/hr) represents the existing investment and θ' (dollar/mi/hr) is the additional investment

The general form of the unit travel time equation then becomes

$$t = K_1 + \frac{K_2}{K_3 + \theta'} V \tag{3}$$

The characteristics of this equation are demonstrated in Figures 2, 3, and 4.

Letting L be the length of the link and C_t the cost of time, the objective function then becomes

$$S = \text{link investment} + \text{travel time costs}$$

$$S = \theta' L + \left(K_1 V + \frac{K_2}{K_3 + \theta'} V^2\right) L\, C_t \tag{4}$$

Since the objective function to be optimized is nonlinear, a technique designed to handle this type of function must be employed. Although several such techniques are available, this study utilized a discrete version of the maximum principle (9). The use of the maximum principle to assign trips to an urban street network has previously been demonstrated (10, 11).

FORMULATION OF NETWORK INVESTMENT PROBLEMS

In this paper we shall consider the problem of optimal network investment under various conditions. This requires several model formulations, one of which will be presented here. The other formulations (12) will not be presented here in the interest of brevity; they will, however, be demonstrated by example problems.

Seven investment conditions are to be considered. They can be grouped in two categories as follows:

1. Investment allocation in a network with no existing facilities:
   a. No budget constraint.
   b. Fixed overall system budget.
   c. Fixed budget at each node.
2. Investment allocations to improve a network with existing facilities:
   a. System improvement budget = 0.
   b. No budget constraint.
   c. Upper and lower limit on individual link improvements.
   d. Fixed overall system improvement budget.

In this section the model formulation for the condition of no budget constraint will be derived, which will take care of conditions 1a, 2a, 2b, and 2c listed.

For each of the investment conditions, future traffic demand patterns for the area are assumed to be known. Also assumed as given is the geometric configuration of the network. As stated earlier, the objective is to build (or improve) a transportation network that will accommodate the assumed demand and will cost least to construct and operate. This will be accomplished, in this paper, by searching for the optimal sequence of decision variables.

To facilitate formulation of the model, various assumptions were necessary:

1. No turn penalties;
2. Zone centroids coincide with the nodes;
3. Traffic directions are preassigned;
4. Traffic distribution is fixed;
5. Transportation network can be represented by a rectangularly arranged combination of links;
6. Travel time is the only factor that influences the traffic assignment; and
7. Unit travel time on each link can be expressed as

$$t_j^{n,m} = K_{j1}^{n,m} + \frac{K_{j2}^{n,m}}{\theta_j^{n,m} + K_{j3}^{n,m}} X_j^{n,m} \tag{5}$$

where j = 1, for horizontal links; j = 2, for vertical links; and (n, m) designates the particular node in the network.

Figure 5 shows a basic N × M rectangular network with node (N, M) as the destination and all other nodes as origins. With the input trips at each node assumed to be given, the problem is to find an investment policy under each investment condition such that the total cost is a minimum.

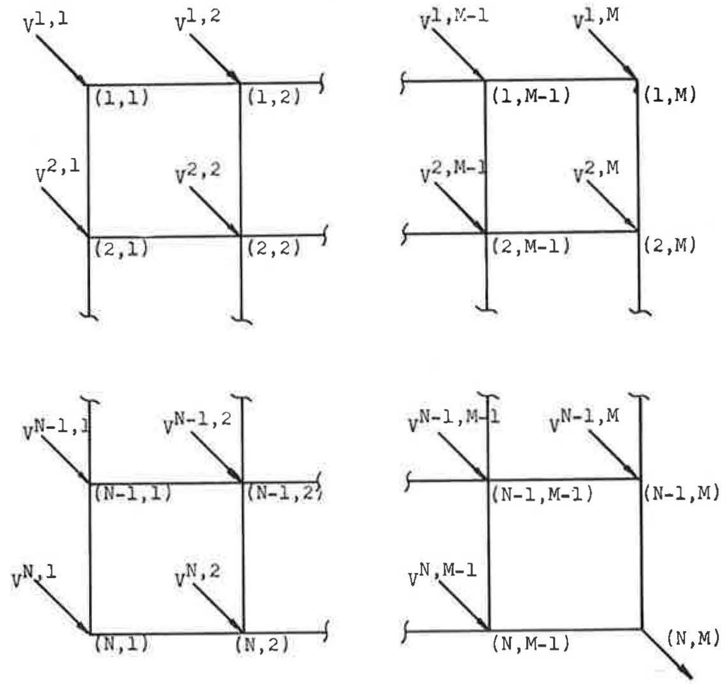


Figure 5. Basic N × M network.

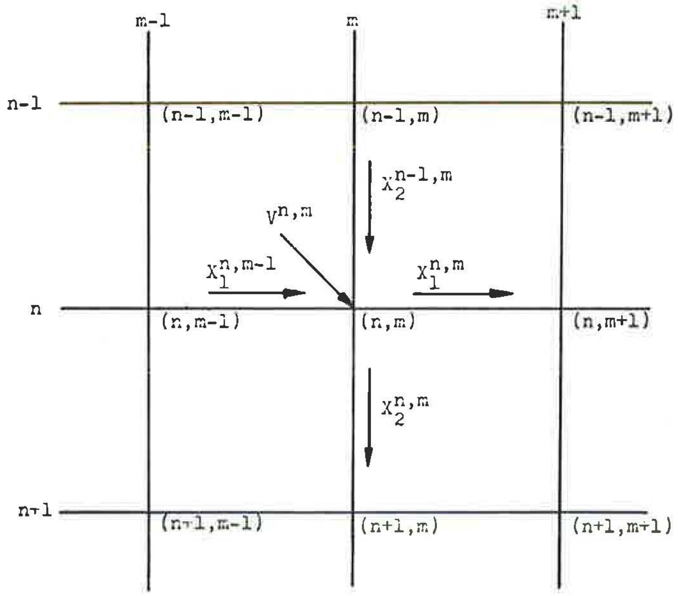


Figure 6. Typical interior node of a rectangular network.

The performance equations for a typical interior node, as shown in Figure 6, were developed as follows:

$$X_1^{n,m} = \left(X_1^{n,m-1} + X_2^{n-1,m} + V^{n,m}\right)\theta_3^{n,m} = AI^{n,m}\theta_3^{n,m} \quad (6)$$

$$X_2^{n,m} = \left(X_1^{n,m-1} + X_2^{n-1,m} + V^{n,m}\right)\left(1-\theta_3^{n,m}\right) = AI^{n,m}\left(1-\theta_3^{n,m}\right) \quad (7)$$

$$X_3^{n,m} = \theta_1^{n,m} L_1^{n,m} + X_3^{n,m-1},\ \theta_1^{n,m} \geq 0 \quad (8)$$

$$X_4^{n,m} = \theta_2^{n,m} L_2^{n,m} + X_4^{n-1,m},\ \theta_2^{n,m} \geq 0 \quad (9)$$

$$X_5^{n,m} = K_{11}^{n,m} X_1^{n,m} L_1^{n,m} C_t + \frac{K_{12}^{n,m} L_1^{n,m} C_t}{\theta_1^{n,m} + K_{13}^{n,m}}\left(X_1^{n,m}\right)^2 + X_5^{n,m-1}$$

$$= K_{11}^{n,m} AI^{n,m}\theta_3^{n,m} L_1^{n,m} C_t$$

$$+ \frac{K_{12}^{n,m} L_1^{n,m} C_t}{\theta_1^{n,m} + K_{13}^{n,m}}\left(AI^{n,m}\theta_3^{n,m}\right)^2 + X_5^{n,m-1} \quad (10)$$

$$X_6^{n,m} = K_{21}^{n,m} X_2^{n,m} L_2^{n,m} C_t + \frac{K_{22}^{n,m} L_2^{n,m} C_t}{\theta_2^{n,m} + K_{23}^{n,m}}\left(X_2^{n,m}\right)^2 + X_6^{n,m-1} =$$

$$= K_{21}^{n,m} AI^{n,m} \left(1 - \theta_3^{n,m}\right) L_2^{n,m} C_t$$

$$+ \frac{K_{22}^{n,m} L_2^{n,m} C_t}{\theta_2^{n,m} + K_{23}^{n,m}} \left[AI^{n,m} \left(1 - \theta_3^{n,m}\right)\right]^2 + X_6^{n,m-1} \tag{11}$$

where

$$AI^{n,m} = X_1^{n,m-1} + X_2^{n-1,m} + V^{n,m} \tag{12}$$

and

$X_j^{n,m}$ = state variables representing flows from node (n, m), j = 1, 2;

$\theta_j^{n,m}$ = decision variables representing investments on links leaving node (n, m), $\theta_j^{n,m} \geq 0$, j = 1, 2;

$K_{j1}^{n,m}$ = free flow travel time on links leaving node (n, m), j = 1, 2;

$K_{j2}^{n,m}$ = coefficient of investment on links leaving node (n, m), j = 1, 2;

$K_{j3}^{n,m}$ = existing investment on links leaving node (n, m), j = 1, 2;

$L_j^{n,m}$ = link length on links leaving node (n, m), j = 1, 2 where j = 1 for horizontal links and j = 2 for vertical links;

$X_3^{n,m}$ = state variable representing the total investment on horizontal links from node (1, 1) through node (n, m);

$X_4^{n,m}$ = state variable representing the total investment on vertical links from node (1, 1) through node (n, m);

$X_5^{n,m}$ = state variable representing the total travel time cost on horizontal links from node (1, 1) through node (n, m);

$X_6^{n,m}$ = state variable representing the total travel time cost on vertical links from node (1, 1) through node (n, m);

$X_7^{n,m}$ = state variable representing the total investment on both links from node (1, 1) through node (n, m);

$\theta_3^{n,m}$ = decision variable representing the fraction of the vehicles departing node (n, m) on the horizontal link, $0 \leq \theta^{n,m} \leq 1$;

C_t = time cost; and

$V^{n,m}$ = input trips at node (n, m).

The Hamiltonian function at this node is defined as

$$H^{n,m} = Z_1^{n,m} X_1^{n,m} + Z_2^{n,m} X_2^{n,m} + Z_3^{n,m} X_3^{n,m} + Z_4^{n,m} X_4^{n,m}$$

$$+ Z_5^{n,m} X_5^{n,m} + Z_6^{n,m} X_6^{n,m} \tag{13}$$

Substituting Eqs. 6 to 11 into Eq. 13 and taking derivatives with respect to state variables, the adjoint variables are obtained as follows:

$$Z_1^{n,m-1} = \frac{\partial H^{n,m}}{\partial X_1^{n,m-1}} = Z_1^{n,m} \theta_3^{n,m} + Z_2^{n,m} \left(1 - \theta_3^{n,m}\right)$$

$$+ Z_5^{n,m} K_{11}^{n,m} \theta_3^{n,m} L_1^{n,m} C_t + Z_6^{n,m} K_{21}^{n,m} \left(1 - \theta^{n,m}\right) L_2^{n,m} C_t +$$

$$+ 2Z_5^{n,m} \frac{K_{12}^{n,m} L_1^{n,m} C_t}{\theta_1^{n,m} + K_{13}^{n,m}} AI^{n,m} \left(\theta_3^{n,m}\right)^2$$

$$+ 2Z_6^{n,m} \frac{K_{22}^{n,m} L_2^{n,m} C_t}{\theta_2^{n,m} + K_{23}^{n,m}} AI^{n,m} \left(1 - \theta_3^{n,m}\right)^2 \tag{14}$$

$$Z_2^{n-1,m} = \frac{\partial H^{n,m}}{\partial X_2^{n-1,m}} = Z_1^{n,m} \theta_3^{n,m} + Z_2^{n,m}\left(1 - \theta_3^{n,m}\right) + Z_5^{n,m} K_{11}^{n,m} \theta_3^{n,m} L_1^{n,m} C_t$$

$$+ Z_6^{n,m} K_{21}^{n,m}\left(1 - \theta_3^{n,m}\right) L_2^{n,m} C_t$$

$$+ 2Z_5^{n,m} \frac{K_{12}^{n,m} L_1^{n,m} C_t}{\theta_1^{n,m} + K_{13}^{n,m}} AI^{n,m} \left(\theta_3^{n,m}\right)^2$$

$$+ 2Z_6^{n,m} \frac{K_{22}^{n,m} L_2^{n,m} C_t}{\theta_2^{n,m} + K_{23}^{n,m}} AI^{n,m} \left(1 - \theta_3^{n,m}\right)^2 \tag{15}$$

$$Z_3^{n,m-1} = \frac{\partial H^{n,m}}{\partial X_3^{n,m-1}} = Z_3^{n,m} \tag{16}$$

$$Z_4^{n-1,m} = \frac{\partial H^{n,m}}{\partial X^{n-1,m}} = Z_4^{n,m} \tag{17}$$

$$Z_5^{n,m-1} = \frac{\partial H^{n,m}}{\partial X_5^{n,m-1}} = Z_5^{n,m} \tag{18}$$

$$Z_6^{n-1,m} = \frac{\partial H^{n,m}}{\partial X_6^{n-1,m}} = Z_6^{n,m} \tag{19}$$

The original conditions for the state variables are given as

$$X_1^{0,0} = X_2^{0,0} = X_3^{0,0} = X_4^{0,0} = X_5^{0,0} = X_6^{0,0} = 0 \tag{20}$$

The objective function is

$$S = X_3^{N,M} + X_4^{N,M} + X_5^{N,M} + X_6^{N,M} \tag{21}$$

Therefore, by definition, the boundary conditions for the adjoint variables are

$$Z_1^{N,M} = Z_2^{N,M} = 0 \tag{22}$$

$$Z_3^{N,M} = Z_4^{N,M} = Z_5^{N,M} = Z_6^{N,M} = 1 \tag{23}$$

Substituting Eq. 23 into Eqs. 16 to 19, the following equation is derived:

$$Z_3^{n,m} = Z_4^{n,m} = Z_5^{n,m} = Z_6^{n,m} = 1 \quad \text{for all } (n, m) \tag{24}$$

The Hamiltonian function then becomes

$$H^{n,m} = Z_1^{n,m} X_1^{n,m} + Z_2^{n,m} X_2^{n,m} + X_3^{n,m} + X_4^{n,m} + X_5^{n,m} + X_6^{n,m} \tag{25}$$

In order to have S a minimum, the following conditions are necessary:

$$\frac{\partial H^{n,m}}{\partial \theta_1^{n,m}} = 0, \ \theta_1^{n,m} > 0$$

$$\frac{\partial H^{n,m}}{\partial \theta_2^{n,m}} = 0, \ \theta_2^{n,m} > 0$$

$$\frac{\partial H^{n,m}}{\partial \theta_3^{n,m}} = 0, \ 0 < \theta_3^{n,m} < 1$$

when $(\theta_1^{n,m}, \theta_2^{n,m}, \theta_3^{n,m})$ is an interior point, or $H^{n,m}$ = minimum with respect to those $\theta_j^{n,m}$ which are at a boundary point of the constraints.

Substituting Eqs. 6 to 11 into Eq. 25 and taking derivatives with respect to the various decision variables, the following equations are obtained:

$$\frac{\partial H^{n,m}}{\partial \theta_1^{n,m}} = L_1^{n,m} - \frac{K_{12}^{n,m} L_1^{n,m} C_t}{\left(\theta_1^{n,m} + K_{13}^{n,m}\right)^2} \left(AI^{n,m} \theta_3^{n,m}\right)^2 \tag{26}$$

$$\frac{\partial H^{n,m}}{\partial \theta_2^{n,m}} = L_2^{n,m} - \frac{K_{22}^{n,m} L_2^{n,m} C_t}{\left(\theta_2^{n,m} + K_{23}^{n,m}\right)^2} \left[AI^{n,m}\left(1 - \theta_3^{n,m}\right)\right]^2 \tag{27}$$

$$\begin{aligned}\frac{\partial H^{n,m}}{\partial \theta_3^{n,m}} = {} & \left(Z_1^{n,m} - Z_2^{n,m}\right) AI^{n,m} + \left(K_{11}^{n,m} L_1^{n,m} - K_{21}^{n,m} L_2^{n,m}\right) AI^{n,m} C_t \\ & + 2 \frac{K_{12}^{n,m} L_1^{n,m} C_t}{\theta_1^{n,m} + K_{13}^{n,m}} \left(AI^{n,m}\right)^2 \theta_3^{n,m} \\ & - 2 \frac{K_{22}^{n,m} L_2^{n,m} C_t}{\theta_2^{n,m} + K_{23}^{n,m}} \left(AI^{n,m}\right)^2 \left(1 - \theta_3^{n,m}\right)\end{aligned} \tag{28}$$

Taking the derivative of Eq. 28 with respect to $\theta_3^{n,m}$, the following equation is obtained:

$$\frac{\partial^2 H^{n,m}}{\left(\partial \theta_3^{n,m}\right)^2} = 2\, \frac{K_{12}^{n,m}\, L_1^{n,m}\, C_t}{\theta_1^{n,m} + K_{13}^{n,m}} \left(AI^{n,m}\right)^2 + 2\, \frac{K_{22}^{n,m}\, L_2^{n,m}\, C_t}{\theta_2^{n,m} + K_{23}^{n,m}} \left(AI^{n,m}\right)^2 \qquad (29)$$

Setting Eqs. 26 and 27 equal to zero and applying the boundary conditions of the decision variables, the values of $\theta_1^{n,m}$ and $\theta_2^{n,m}$ can be obtained from the following equations:

$$\theta_1^{n,m} = \sqrt{K_{12}^{n,m}\, C_t}\; AI^{n,m}\, \theta_3^{n,m} - K_{13}^{n,m} \text{ when } \theta_1^{n,m} > 0 \qquad (30)$$

or

$$\theta_1^{n,m} = 0 \text{ when } \sqrt{K_{12}^{n,m}\, C_t}\; AI^{n,m}\, \theta_3^{n,m} - K_{13}^{n,m} \le 0 \qquad (31)$$

$$\theta_2^{n,m} = \sqrt{K_{22}^{n,m}\, C_t}\; AI^{n,m} \left(1 - \theta_3^{n,m}\right) - K_{23}^{n,m} \text{ when } \theta_2^{n,m} > 0 \qquad (32)$$

or

$$\theta_2^{n,m} = 0 \text{ when } \sqrt{K_{22}^{n,m}\, C_t}\; AI^{n,m} \left(1 - \theta_3^{n,m}\right) - K_{23}^{n,m} \le 0 \qquad (33)$$

When both $\theta_1^{n,m}$ and $\theta_2^{n,m}$ are greater than zero, Eqs. 30 and 32 can be substituted into Eq. 28 to obtain

$$\begin{aligned}
\frac{\partial H^{n,m}}{\partial \theta_3^{n,m}} &= \left(Z_1^{n,m} - Z_2^{n,m}\right) AI^{n,m} \\
&+ \left(K_{11}^{n,m}\, L_1^{n,m} - K_{21}^{n,m}\, L_2^{n,m}\right) AI^{n,m}\, C_t \\
&+ 2\sqrt{K_{12}^{n,m}\, C_t}\; L_1^{n,m}\, AI^{n,m} - 2\sqrt{K_{22}^{n,m}\, C_t}\; L_2^{n,m}\, AI^{n,m} \\
&= AI^{n,m}\left[\left(Z_1^{n,m} - Z_2^{n,m}\right) + \left(K_{11}^{n,m}\, L_1^{n,m} - K_{21}^{n,m}\, L_2^{n,m}\right) C_t \right. \\
&\left. + 2\left(\sqrt{K_{12}^{n,m}\, C_t}\; L_1^{n,m} - \sqrt{K_{22}^{n,m}\, C_t}\; L_2^{n,m}\right)\right] \qquad (34)
\end{aligned}$$

$\theta_3^{n,m}$ is eliminated by the substitution and the value of $\frac{\partial H^{n,m}}{\partial \theta_3^{n,m}}$ becomes independent of $\theta_3^{n,m}$ as shown in Eq. 34. This implies that the value of $H^{n,m}$ is linearly related to $\theta_3^{n,m}$ and the extreme of $H^{n,m}$ with respect to $\theta_3^{n,m}$ occurs at a boundary. In this case, to obtain the minimum value of $H^{n,m}$, $\theta_3^{n,m} = 0$ when $\frac{\partial H^{n,m}}{\partial \theta_3^{n,m}} > 0$ or $\theta^{n,m} = 1$ when $\frac{\partial H^{n,m}}{\partial \theta_3^{n,m}} < 0$.

If $\frac{H^{n,m}}{\theta_3^{n,m}} = 0$, $\theta_3^{n,m}$ can be any value between 0 and 1 because the value of $H^{n,m}$ is independent of $\theta_3^{n,m}$.

When either $\theta_1^{n,m}$ or $\theta_2^{n,m}$ is equal to zero, or when both are equal to zero, Eq. 34 is no longer valid; Eq. 28 is then set equal to zero and solved for the optimal value of $\theta_3^{n,m}$.

Special Case

In an urban area, the space available for street or freeway construction is often limited. For example, a freeway with more than 8 lanes may be difficult to construct near the CBD. At the same time it may be the policy of the area to provide at least a minimum level of service in all parts of the urban area. Thus, it may be desirable to place upper and lower total investment limits on various links in the network. This is the condition 2c stated earlier.

Mathematically, this investment criterion can be expressed as follows:

$$\theta_{1\,min}^{n,m} \le K_{13}^{n,m} + \theta_1^{n,m} \le \theta_{1\,max}^{n,m} \tag{35}$$

$$\theta_{2\,min}^{n,m} \le K_{23}^{n,m} + \theta_2^{n,m} \le \theta_{2\,max}^{n,m} \tag{36}$$

This formulation provides the equations for searching the optimum sequence of the decision variables, $\theta_1^{n,m}$, $\theta_2^{n,m}$, and $\theta_3^{n,m}$ and the associated values of the state variables.

The optimum seeking procedure developed for this problem is as follows:

1. Assume a set of decision variables, $\theta_3^{n,m}$.

2. Calculate $X_1^{n,m}$, $X_2^{n,m}$ and $AI^{n,m}$ by Eqs. 6, 7, and 12, starting at n = m = 1 and proceeding to n = N and m = M.

3. Calculate decision variables $\theta_1^{n,m}$ and $\theta_2^{n,m}$ by Eqs. 30 and 32 and check the boundary conditions for each special case.

4. Calculate the values of $X_i^{n,m}$, i = 3, 4, 5, 6, by Eqs. 8 to 11, starting at n = m = 1 and proceeding to n = N and m = M.

5. Calculate the adjoint vectors, $Z_i^{n,m}$, i = 1, 2, with the above $X_i^{n,m}$ values, by Eqs. 14, 15, and 22, starting at n = N, m = M and proceeding backward to n = m = 1.

6. Using the above values of $X_i^{n,m}$ and $Z_i^{n,m}$, calculate $\frac{\partial H^{n,m}}{\partial \theta_3^{n,m}}$ and $\frac{\partial^2 H^{n,m}}{\left(\partial \theta_3^{n,m}\right)^2}$ by Eqs. 28 and 29.

7. Adjust the values of $\theta_3^{n,m}$ by adding an amount equal to Δ, where

$$\Delta = - \frac{\dfrac{\partial H^{n,m}}{\partial \theta_3^{n,m}}}{\dfrac{\partial^2 H^{n,m}}{\left(\partial \theta_3^{n,m}\right)^2}}$$

and check the boundary condition.

8. With the new values of $\theta_i^{n,m}$, return to step 2 and repeat the procedure until the value of the objective function, Eq. 21, is sufficiently close to the previous value to indicate adequate convergence.

SOLUTION TO NETWORK INVESTMENT CONDITIONS

Input data for the various investment conditions are summarized in Table 1. Since construction cost and right-of-way costs will not be the same throughout the network area, two values of K_{13} were assigned to the links even though these links represent the same type of facility. For the same reason, in the link constraint condition 2c, links have different values for maximum and minimum investment levels. The derivation of these data is discussed in the Appendix. Time cost of travel was assumed to be $1.55 per hour per vehicle as suggested by AASHO (13).

Example 1: Investment in a network with no existing facilities.
Suppose we are planning a network for a given set of trips where no facilities presently

TABLE 1

INPUT DATA FOR EXAMPLE PROBLEMS

Nodes (n,m)	Links (i)	$K_{i1}^{n,m}$	$K_{i2}^{n,m}$	$K_{i3}^{n,m}$	$\theta_{imax}^{n,m}$	$\theta_{imin}^{n,m}$	$V^{n,m}$	$SI^{n,m}$
1,1	1	0.0143	0.00003	$ 8.0	$ 80	$10	2,000	$40
	2	0.0143	0.00004	8.0	80	10		
1,2	1	0.0143	0.00006	8.0	80	10	3,000	40
	2	0.0143	0.00005	10.0	80	10		
1,3	1	0.0143	0.00008	8.0	80	15	0	60
	2	0.0143	0.00006	8.0	80	10		
1,4	1	--	--	--	--	--	1,000	40
	2	0.0167	0.00010	15.0	100	15		
2,1	1	0.0143	0.00005	10.0	80	10	3,000	40
	2	0.0143	0.00005	8.0	80	10		
2,2	1	0.0143	0.00006	10.0	80	10	0	60
	2	0.0143	0.00005	10.0	80	10		
2,3	1	0.0167	0.00010	15.0	100	15	1,000	80
	2	0.0167	0.00008	12.0	100	15		
2,4	1	--	--	--	--	--	0	50
	2	0.0167	0.00015	15.0	100	15		
3,1	1	0.0143	0.00006	8.0	80	10	0	60
	2	0.0143	0.00006	8.0	80	10		
3,2	1	0.0167	0.00008	12.0	100	15	1,000	80
	2	0.0167	0.00010	15.0	100	15		
3,3	1	0.0167	0.00015	12.0	100	15	1,000	100
	2	0.0167	0.00015	12.0	100	15		
3,4	1	--	--	--	--	--	0	60
	2	0.0167	0.00025	15.0	100	15		
4,1	1	0.0167	0.00008	15.0	100	15	1,000	40
	2	--	--	--	--	--		
4,2	1	0.0167	0.00015	15.0	100	15	0	50
	2	--	--	--	--	--		
4,3	1	0.0167	0.00020	15.0	100	15	0	60
	2	--	--	--	--	--		
4,4	1	--	--	--	--	--	-13,000	0
	2	--	--	--	--	--		
Total				$272.0				$860

i = 1 for horizontal links
i = 2 for vertical links

GI = $300.00
C_t = $1.55/hour

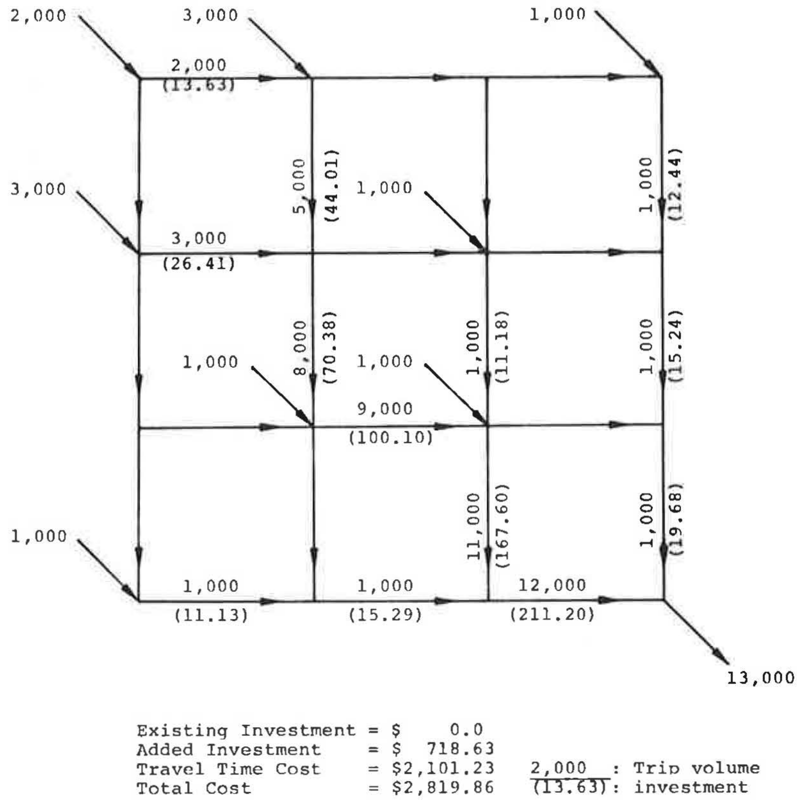


Figure 7. Investment allocation in a network with no existing facilities and no budget constraint—Example 1a.

exist $\left(K_{13}^{n,m} = 0\right)$. This might be the situation in a completely undeveloped area. It is desired to provide facilities for the area subject to three possible investment conditions.

Example problem 1a develops a network where no constraint is placed on the funds that can be spent for network facilities. This can be considered the theoretical optimal system since funds will be expended until the decrease in total travel costs are equal to the additional investment and the system is not burdened by sunk investments.

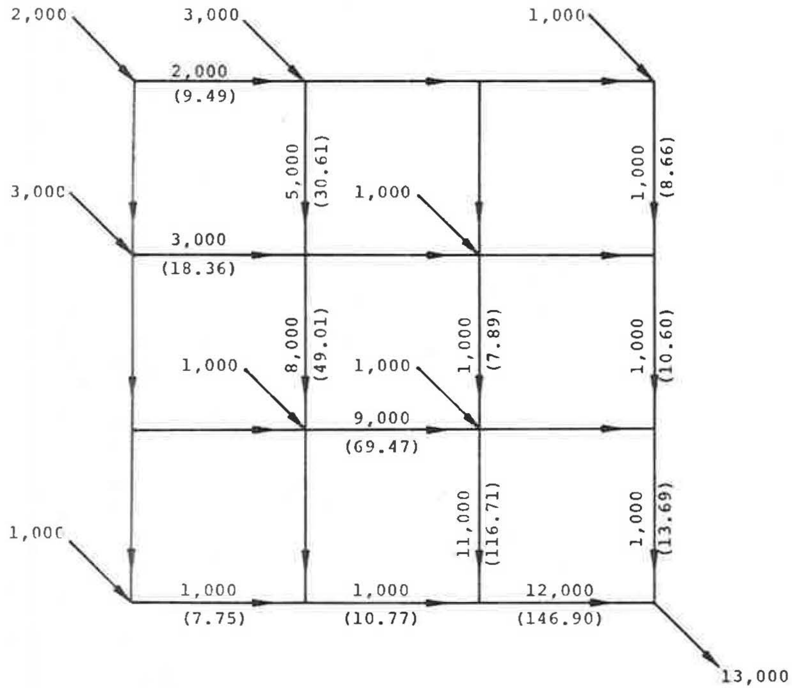
The resulting system is shown in Figure 7. Note that the system developed forms the minimum path tree in which only one route is built for each origin-destination pair and all trips are assigned to this route.

Example problem 1b again develops a network where no facilities presently exist but where a budget limitation that is less than the theoretical optimal budget determined in example problem 1a is placed on total investment expenditures. The optimal solution for this condition is shown in Figure 8. Again the system developed forms a minimum path tree as in example 1a. A link-by-link comparison shows that less funds are expended on each link resulting in increased travel costs. The increase in travel costs is greater than the decrease in investment costs.

Example problem 1c develops an undeveloped network and places limitation on funds that can be spent on the horizontal and vertical links leaving each node (n, m). The node budgets, $SI^{n,m}$, are tabulated in Table 1.

This type of situation might be encountered where the individual regions, represented by the nodes, are allowed to expend budgeted funds in their region only.

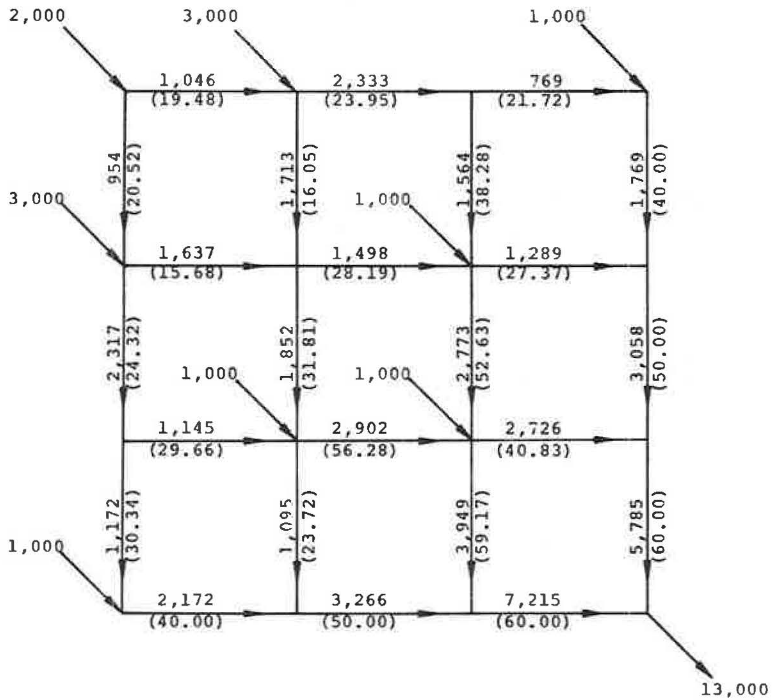
The model is formulated in such a manner (Fig. 9) that all budgeted funds must be spent, resulting in a total network investment of $860. This is a greater expenditure than invested under no system budget limitations in example 1a. Further, the added expenditure failed to reduce total trip travel costs below the level achieved in example 1a.



Existing Investment = $ 0.00
Added Investment = $ 500.00
Travel Time Cost = $2,415.86
Total Cost = $2,915.86

2,000 : trip volume
(9.49) : added investment

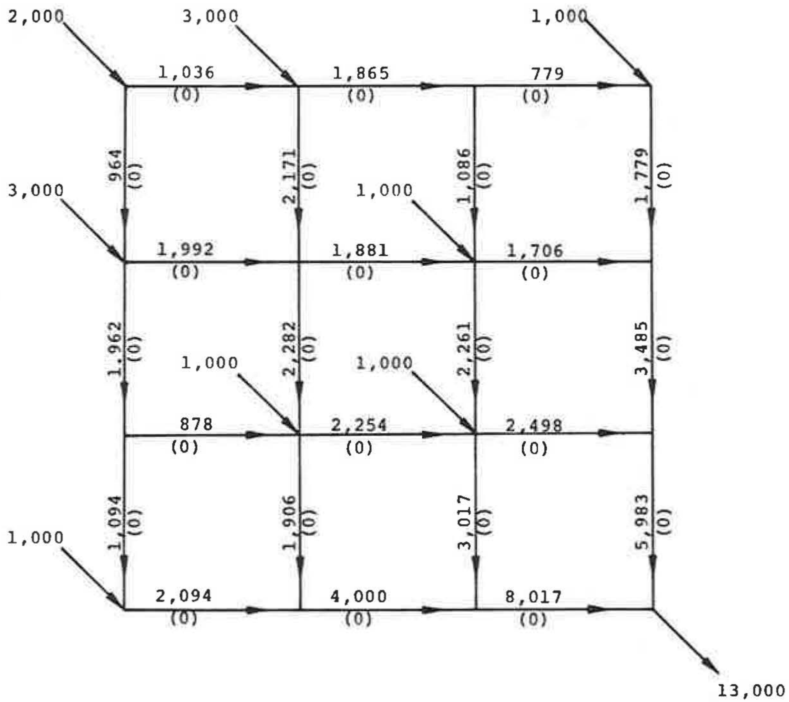
Figure 8. Investment allocation in a network with no existing facilities and a fixed overall system budget—Example 1b.



Existing Investment = $ 0.00
Added Investment = $ 860.00
Travel Time Cost = $2,252.91
Total Cost = $3,112.91

1J46 : trip volume
(19.48) : investment

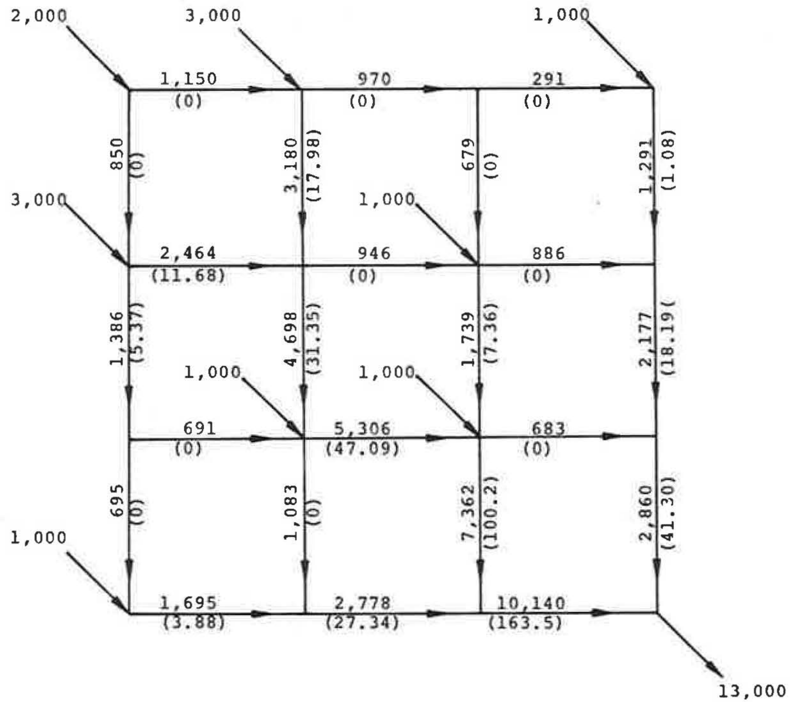
Figure 9. Investment allocation in a network with no existing facilities and a fixed budget at each node—Example 1c.



Existing Investment = $ 272.00
Added Investment = $ 0.00
Travel Time Cost = $4,585.94
Total Cost = $4,857.94

1036: trip volume
(0) : added investment

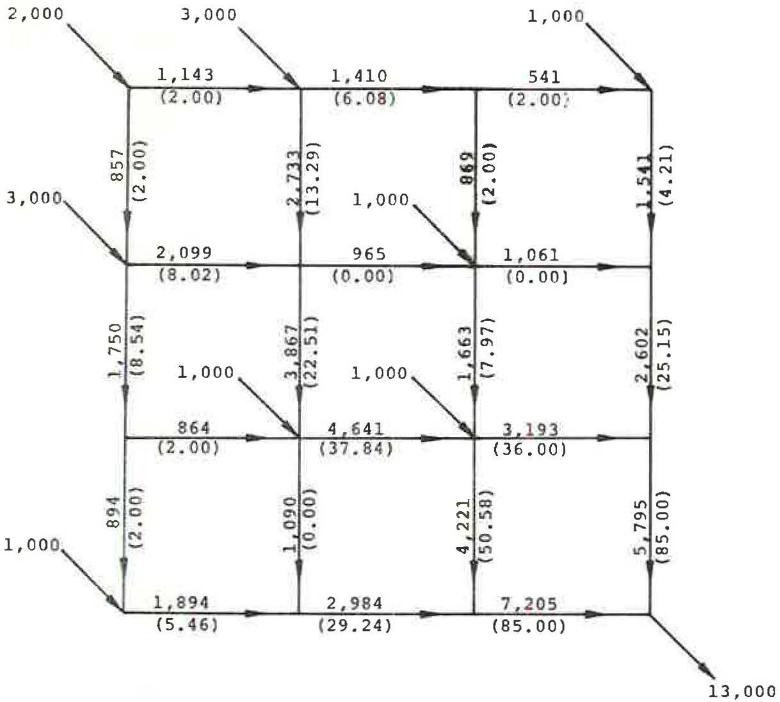
Figure 10. Travel time cost on network with existing facilities before additional investment—Example 2a.



Existing Investment = $ 272.00
Added Investment = $ 477.96
Travel Time Cost = $2,100.96
Total Cost = $2,850.92

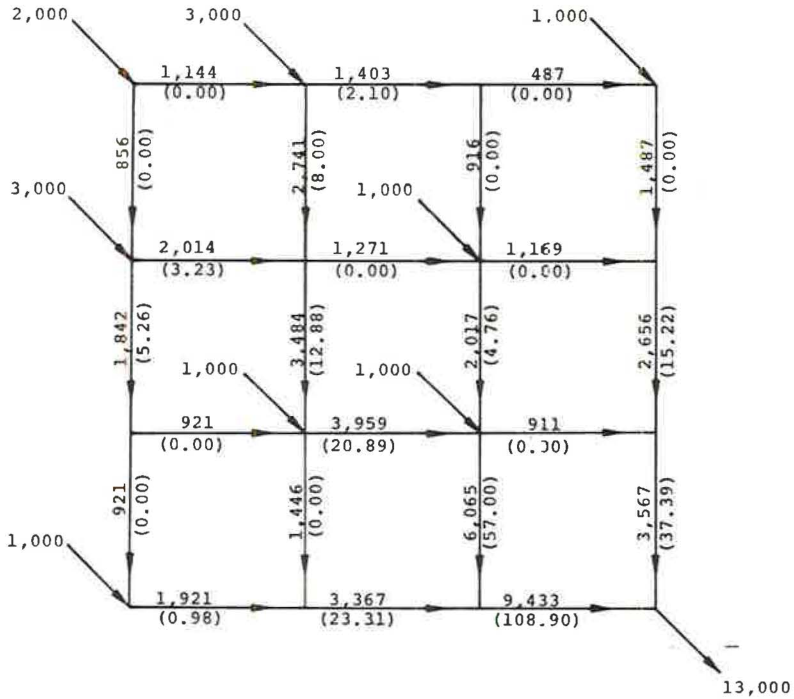
1150: trip volume
(0) : added investment

Figure 11. Investment allocation on a network with existing facilities and no budget constraint—Example 2b.




```
Existing Investment = $     0.00
Added Investment    = $   445.04     1,143 : trip volume
Travel Time Cost    = $2,158.95     (2.00): added investment
Total Cost          = $2,875.99
```

Figure 12. Investment allocation in a network with existing facilities and upper and lower limits on link investment level—Example 2c.




```
Existing Investment = $   272.00
Added Investment    = $   300.00     1,144 : trip volume
Travel Time Cost    = $2,339.38     (0.00): added investment
Total Cost          = $2,911.38
```

Figure 13. Investment allocation in a network with existing facilities and a fixed overall system budget—Example 2d.

In other words, the area development scheme, as simulated by node budget constraints, resulted in increased trip travel cost even though the total investment was somewhat greater than the budget in example 1a, indicating that this can be an uneconomic method of system development.

Example 2: Investment in a network with existing facilities.
In the next four example problems it was assumed that existing facilities do exist. The magnitude of the existing investment in each link is given by $K_{i3}^{n,m}$ in Table 1. The total existing investment is equal to $272 for the entire network. The objective in these example problems is to improve an existing network, subject to specified constraints, in an optimal fashion.

Example problem 2a might be considered the "benchmark" condition since it represents travel costs on the network before any improvement investment takes place. The trip assignment pattern developed under these conditions is shown in Figure 10. Note that each link in the network is being utilized.

In example problem 2b the network is improved with no limit placed on the magnitude of the investment. The solution to this condition is shown in Figure 11. An investment of $477.96 resulted in a reduction of travel costs of $2,484.98 as determined in example 2a.

Figure 12 shows the solution to example problem 2c when upper and lower total investment limits are placed on each link. These limits, $\theta_{i}^{n,m}{}_{max}$ and $\theta_{i}^{n,m}{}_{min}$, are given in Table 1. In only two locations were the limits in effect. The links leaving node (1, 1) are bounded by the lower limit. In other words, the traffic using these links does not fully utilize the minimum level of total investment required. The links entering node (4, 4), the CBD, are bounded by the upper investment limit of $100 on each link. Thus these links are carrying traffic in excess of their economic limit. If additional investment were possible on these two links total travel time cost would be reduced. It should be noted that since existing investments do exist on every link, every link is used to accommodate trips.

Example problem 2d demonstrates the effect of a network budget limitation (Fig. 13). The model formulation required that the budget, set at $300, was to be completely spent on the network. The budget of $300 plus the existing investment of $272 results in a total investment of $572, which is $146.63 less than the optimal investment of $718.63 determined in example 1a. While the investment cost is $146.63 less, the total travel cost was $238.15 greater than the optimal solution 1a. It might be noted that on several links no additional investments were required since existing investment was sufficient to handle trip demands.

CONCLUSIONS

A new technique for the analysis of transportation system investment problems has been presented. Considering each node of a rectangular urban network as a stage, a discrete version of the maximum principle was utilized to formulate a transportation system model. An investment model was derived for the condition when no budget limitation was present. Other investment models illustrating different conditions were presented through the use of example problems.

As opposed to linear programming models, the maximum principle is capable of attacking transportation system investment problems that include nonlinear relationships between travel time, traffic volume, and investment cost.

Although the models presented were applied to only single copy networks, no difficulty should be experienced in extending the technique to more complex networks.

Although this paper marks only a first step in an attempt to apply a new technique to the complex problem of optimal network development, some generalized statements and conclusions are in order:

1. In Table 2 are given the solutions to the various investment conditions. Here it is noted that the least constrained system, example problem 1a, produces the least-cost solution. As soon as constraints, in the form of budgets and/or sunk investments, were placed on the system, total costs increased, producing non-optimal solutions.

TABLE 2

RECAP OF VARIOUS SYSTEM INVESTMENT CONDITIONS AND COSTS*

(1) Example Problem	(2) Network Investment Conditions	(3) Existing Investment Cost	(4) Added Investment Cost	(5) Total System Travel Cost	(6) Total Cost (3)+(4)+(5)
1a	No Budget Constraint $K_{i3}^{n,m} = 0$	0.0	$718.63	$2,101.23	$2,819.86
1b	Fixed System Budget $K_{i3}^{n,m} = 0$	0.0	500.0	2,415.86	2,915.86
1c	Fixed Budget at each Node $SI^{n,m} = \theta_1^{n,m}+\theta_2^{n,m}$ $K_{i3}^{n,m} = 0$	0.0	860.00	2,252.91	3,112,91
2a	System Budget = 0 $K_{i3}^{n,m}$ = Table 1	$272.00	0.00	4,585.94	4,857.94
2b	No Improvement Budget Constraint $K_{i3}^{n,m}$ = Table 1	272.00	477.96	2,100.96	2,850.92
2c	Upper & Lower Limit on Link Investment $\theta_{imin}^{n,m} \le K_{i3}^{n,m}+\theta_i^{n,m} \le \theta_{imax}^{n,m}$ Table 1 $K_{i3}^{n,m}$ = Table 1	272.00	445.04	2,158.95	2,875.99
2d	Fixed System Budget $K_{i3}^{n,m}$ = Table 1	272.00	300.00	2,339.38	2,911.38

*All costs assumed to be equivalent hourly costs.

2. The models were so formulated that all budgeted funds had to be expended. This is in keeping with government policies at almost all levels. When a budget was allocated in a non-optimal fashion, as in the case of example problem 1c, a non-optimal overall solution resulted.

3. It is felt that the models are realistic since added investments produced reduced total travel costs.

4. A systems effect is necessary to achieve a true optimum. That is to say that all system benefits must be compared to total system costs to determine the optimal solution. When this is not allowed, as in example problem 1c where each node has budget constraint, a non-optimal solution occurs.

5. Sunk investment in the form of existing facilities can act as a constraint and produce a non-optimal solution when compared to the theoretically optimal condition 1a.

6. With the exception of condition 2a, the optimal solutions to the various situations all fall within a 10 percent range. It thus appears that no matter what the condition of investment might be the solution to this condition may not be too far from the true theoretical optimum determined in 1a. It may be that other considerations may be more important than construction and operating costs in determining the optimal network development policy.

Although the functional relationships derived in this paper could be improved, the research did demonstrate the ability of the discrete maximum principle to solve nonlinear optimization problems. Improved data and additional research into the relationship between travel time, capacity, traffic flow, and investment are needed to make the models more realistic and useful.

The next logical step in this research should be aimed at the multi-copy problem, the problem of mixed modes and the problem of optimal staging in a dynamic situation. Finally the relationship between land-use and transportation needs to be formalized and brought into the optimal investment problem.

REFERENCES

1. Zettel, Richard M., and Carll, Richard R. Summary Review of Major Metropolitan Area Transportation Studies in the United States. ITTE Special Report, University of California, Berkeley, Nov. 1962.
2. Committee on Urban Transportation of the Highway Division. Urban Transportation Problem. Jour. Highway Div., ASCE, Vol. 83, No. HW3, Oct. 1958, Proc. Paper 1801.
3. Garrison, W. L., and Marble, D. F. Analysis of Highway Networks: A linear Programming Formulation. HRB Proc., Vol. 37, p. 1-17, 1958.
4. Carter, E. C., and Stowers, J. R. Model for Funds Allocation for Urban Highway Systems Capacity Improvements. Highway Research Record 20, p. 84-102, 1963.
5. Funk, J. L., and Roberts, P. O. Toward Optimum Methods of Link Addition in Networks. Report, Department of Civil Engineering, Northwestern University, Sept. 1964.
6. Roberts, P. O., Burns, R. E., and Tabora, P. N. A Model for the Formulation of Transportation System Alternatives. Preprint 316, ASCE Structural Engineering Conference, Miami, Florida, Jan. 1966.
7. Hay, G. A., Morlock, E. K., and Charnes, A. Toward Optimal Planning of a Two-Mode Urban Transportation System: A Linear Programming Formulation. Highway Research Record 148, p. 20-38, 1966.
8. Ridley, T. M. An Investment Policy to Reduce the Travel Time in a Transportation Network. Operations Research Center, University of California, Berkeley, Dec. 1965.
9. Fan, Liang-tseng, and Wang, Chiu-sen. The Discrete Maximum Principle—a Study of Multistage Optimization. John Wiley, New York, 1964.
10. Yang, T. C., and Snell, R. R. Traffic Assignment by the Maximum Principle. Jour. Highway Div., ASCE, Vol. 92, No. HW2, Oct. 1966, p. 1-14.
11. Funk, M. L., Snell, R. R., and Blackburn, J. B. Optimal Allocation of Trips to a Street Network. Jour. Highway Div., ASCE, Vol. 93, No. HW2, Nov. 1967, p. 95-114.
12. Wang, Jin-Jerg. Toward a Solution for the Optimal Allocation of Investment in Transportation Network Development. Masters Thesis, Department of Civil Engineering, Kansas State University, 1967.
13. Road User Benefit Analysis for Highway Improvements. American Association of State Highway Officials, Washington, 1960.
14. Aitkin, P. W. Cost of Constructing Transportation Facilities in the Chicago Standard Metropolitan Area. CATS Research News, Vol. 1, No. 14, p. 8-11, Aug. 2, 1967.
15. Moskowitz, K. A Note on Prorating Highway Costs. Unpublished Report, June 1964.
16. McElrath, D., Rothenberg, T., Shuldiner, P., and Funk, M. L. Automobile Transportation Systems: Cost Characteristics. The Mid-Twentieth Century City Study, Preliminary Report No. 3, The Transportation Center, Northwestern University.
17. Highway Capacity Manual, HRB Spec. Rept. 87, 1965.

Appendix

DERIVATION OF CONSTANTS IN UNIT TRAVEL TIME EQUATION

Unit travel time has been expressed as

$$t = K_1 + \frac{K_2}{\theta + K_3} V \qquad \text{(A-1)}$$

where

- t = unit travel time (hr/mi/veh);
- K_1 = free flow travel time (hr/mi/veh)—the magnitude depends on the maximum speed obtainable or regulated;
- K_2 = coefficient of improvement (dollar-hr/mi^2/veh^2)—its magnitude depends on link location and reflects the difficulty of improvement;
- K_3 = existing investment (dollar/mi/hr);
- θ = equivalent hourly investment per unit length (dollar/mi/hr); and
- V = traffic volume per unit time (veh/hr).

In this section, a set of K values is derived from data reported by other researchers. The purpose of this section is twofold: (a) to justify the fitness of the equation, and (b) to obtain a set of K values for the example problems.

Values of K_1

The K_1 value is equal to the reciprocal of the maximum speed obtainable or regulated in each area. Several common values are as follows:

Maximum Speed (mph)	K_1 Values (hr/mile)
70	0.0143
60	0.0167
50	0.0200

For the example problems, maximum speeds were assumed to be 70 mph in less densely developed areas and 60 mph in densely developed areas. The K_1 values are therefore 0.0143 and 0.0167 hours per mile respectively.

Values of K_2

1. Near CBD Area:

The average cost of an 8-lane freeway near the CBD, as estimated by Aitken (14), is \$15,500,000 per mile. Assuming 30-year life and 6 percent interest, annual cost is equal to \$1,130,000 per mile. If we further assume peak hour traffic is 10 percent of daily traffic, the equivalent peak hour cost becomes

$$\$1,130,000 \times \frac{1}{360} \times \frac{1}{10} = \$314 \text{ per mile per hour}$$

This freeway can handle 1100 vph per lane at unit travel time of 0.02 hr/mile. Assuming K_1 = 0.0143 hr/mi/veh (70 mph speed), K_2 is derived as follows:

$$0.0143 + \frac{K_2}{314}(1100 \times 8) = 0.020$$

or

$$K_2 = 0.00207 \text{ dollar-hr/mi}^2\text{/veh}^2 \qquad \text{(A-2)}$$

TABLE A-1

COST CHARACTERISTICS OF URBAN HIGHWAYS

Characteristic	Local Street	Arterial	Freeway
Practical capacity (vph/lane)	500	700	1800
Average speed (mph)	20	25-40	45-65
Right-of-way cost ($/mile)	250,000	450,000	4-8 million
Construction cost ($/mile)	300,000	500,000	4-6 million
Total cost ($/mile)	550,000	950,000	8-14 million

Using Haikalis' data and adjusting for the downtown area, Hay et al (7) used an arterial street with 2000 vph volume at unit travel time of 0.0333 hour per mile costs $3,400,000 per mile or $250,000 per mile annually. Equivalent peak hour cost becomes:

$$\$250{,}000 \times \frac{1}{360} \times \frac{1}{10} = \$69.5 \text{ per mile per hour}$$

Assuming $K_1 = 0.025$ hr/mi/veh (40 mph speed), K_2 is derived as follows:

$$0.025 + \frac{K_2}{69.5}\, 2{,}000 = 0.0333$$

$$K_2 = 0.000288 \text{ dollar-hr/mi}^2\text{/veh}^2 \qquad \text{(A-3)}$$

2. Average Urban Area:

The overall average cost for an 8-lane urban freeway is $5,000,000 per mile as estimated by Moskowitz (15). Assuming 30-year life and 6 percent interest, equivalent peak hour cost becomes:

$$\$5{,}000{,}000 \times 0.0726 \times \frac{1}{360} \times \frac{1}{10} = \$101 \text{ per mile per hour}$$

Using Figure 3.38 in the Highway Capacity Manual (17), a typical freeway with 70-mph average highway speed can handle 1800 vph per lane at a speed of 45 mph. The K_2 value is derived as follows:

$$K_1 = 0.0143 \text{ hr/mi/veh}$$

$$0.0143 + \frac{K_2}{101}(1800 \times 8) = 0.0222 \qquad \text{(A-4)}$$

$$K_2 = 0.0000553 \text{ dollar-hr/mi}^2\text{/veh}^2$$

TABLE A-2

COST AND TRAVEL TIME OF URBAN HIGHWAYS

Characteristic	Local Street	Arterial	Freeway
Number of lanes	2	4	6
Total volume (vph)	1,000	2,800	10,800
Average speed (mph)	20	32.5	55
Total cost ($/mile)	550,000	950,000	8-14 million
Equivalent peak hour cost ($/mile)	11.1	19.2	161-282
Assumed maximum speed (mph)	35	40	70
Minimum unit travel time (hr/mile)	0.0286	0.025	0.0143
Average travel time (hr/mile)	0.05	0.0308	0.0182

As summarized from "Automobile Transportation Systems: Cost Characteristics" (16), Table A-1 shows relationships among volume, average speed, and cost for three types of urban roads. Using these values and the assumed maximum speeds and average lanes, Table A-2 is obtained. The K_2 values are, then, derived as follows:

Local street:

$$0.0286 + \frac{K_2}{11.1} \times 1000 = 0.05$$

$$K_2 = 0.000227 \text{ dollar-hr/mi}^2\text{/veh}^2 \qquad \text{(A-5)}$$

Arterial street:

$$0.025 + \frac{K_2}{19.2} \times 2800 = 0.308$$

$$K_2 = 0.0000398 \text{ dollar-hr/mi}^2\text{/veh}^2 \qquad \text{(A-6)}$$

Freeway:

$$0.0143 + \frac{K_2}{161}\, 10800 = 0.0182$$

$$K_2 = 0.0000582 \text{ dollar-hr/mi}^2\text{/veh}^2 \qquad \text{(A-7)}$$

$$0.0413 + \frac{K_2}{282}\, 10800 = 0.0182$$

$$K_2 = 0.000102 \text{ dollar-hr/mi}^2\text{/veh}^2 \qquad \text{(A-8)}$$

3. Rural Area:

Cost data for rural highways are not generally available. However, the cost of a rural freeway may be assumed as equal to the lowest cost of a freeway in an urban area.

On this basis an 8-lane freeway will cost about \$3,000,000 per mile (16). Using Figure 3.38 in the Highway Capacity Manual (17), a typical freeway with 70-mph average highway speed can handle 1800 vph per lane at 45 mph. Equivalent peak hour cost becomes

$$\$3{,}000{,}000 \times 0.0726 \times \frac{1}{360} \times \frac{1}{10} = \$60.5 \text{ per mile per hour}$$

The K_2 value is derived as follows:

$$0.0143 + \frac{K_2}{60.5}(1800 \times 8) = 0.0222$$

$$K_2 = 0.00003322 \text{ dollar-hr/mi}^2\text{/veh}^2 \qquad \text{(A-9)}$$

Excluding Eq. A-5, K_2 values are summarized as follows:

Type of Area	Range of K_2 Value
CBD	0.000207-0.000288
Average urban area	0.0000398-0.000102
Rural	0.0000332

The wide range of K_2 values in an average urban area is caused by the wide variance of urban freeway costs as shown in Table A-1. In general, K_2 value is fairly consistent in each area. This indicates a fairly good correlation between the equation and the real-world situation.

Values of K_3

The K_3 value represents the existing facilities in terms of cost per mile per hour. Equivalent peak hour cost, for each type of road, derived in the previous sections gives the average values of K_3.

The K values used in the example problems are summarized in Table 1.

A Method of Network Evaluation Using the Output of the Traffic Assignment Process

THOMAS N. HARVEY, Department of Civil Engineering, Massachusetts Institute of Technology

•THE objective of the research underlying the method described in this paper has been to discover and develop a means of network evaluation that will help to determine just who benefits or loses from a particular network change (1). A further motivation has been to determine quantitatively how much each group gains or loses. One particular difficulty with existing methods of evaluation is the need to assume that the level of demand remains fixed between alternative networks, that is, that demand is unaffected by changes in the transportation network or in the operation of it (2, 3). Consequently, the research has entailed looking for a method that will help determine the value of trips apparently diverted, generated, or eliminated by alterations to the network or to the means of operating it.

What has been found is a way of looking at network flows that will permit the determination of the net gain or loss accruing from them, no matter how demand differs between networks. In this paper, the method will first be presented with the type of demand model to which it appears to be best suited. Then, possible ways of using the method with existing means of estimating travel demand will be shown.

Although the method can be used to compute the net user benefit accruing from flows over a given network, it is most useful in comparing alternative networks on the basis of the difference in net user benefits accruing from them. It requires interzonal traffic volumes and interzonal separations from each alternative network. In some cases, as will be shown, it needs interzonal travel demand as a function of interzonal separation-interzonal travel demand functions.

In most cases, the interzonal separations should be those which it is assumed are experienced by users of the network. The variable used to describe interzonal separation can be travel time, travel cost including value of time, or some other such measure. It is important to note, however, that the resulting evaluation will be only as comprehensive as this measure of interzonal separation is.

It should also be recognized that the user benefits computed using the method are intended to be one dimension in a larger framework of network evaluation. Predicted user benefits must be combined with estimates of construction and maintenance cost in network comparisons. Individual link volumes need to be assessed for congestion, overcapacity, and traffic noise. Other data on other network attributes must also be examined. Still, user benefits are vital to any network evaluation, if not its most important element.

UNDERLYING ASSUMPTIONS AND THE BASIC ECONOMIC RATIONALE

It is first assumed that the markets for travel between zones in urban areas can be expressed in terms of demand and supply curves. Figure 1 shows a demand and a supply curve, each with its normally assumed shape. The demand curve is a plot of the prices of interzonal travel vs the number of trips that would be taken between the zones at each price. It is thus a plot of what people are willing to pay to travel between zones.

The supply curve defines what the producer would have to be offered in terms of a unit price in order to be induced to output a given number of units. Since the producer here is just the impersonal transportation network, the supply curve is a plot of what

Paper sponsored by Committee on Transportation System Analysis and presented at the 47th Annual Meeting.

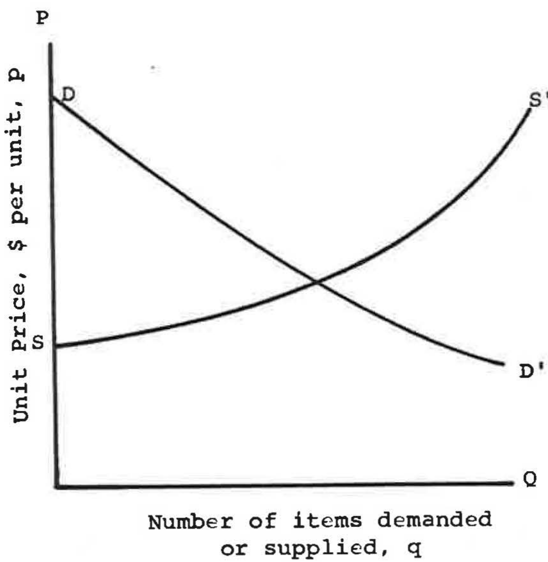
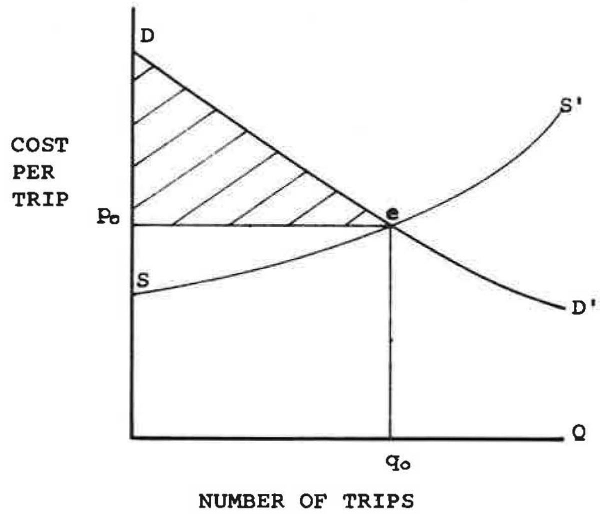


Figure 1. Supply and demand curves.



TRAVEL FROM ZONE 1 TO ZONE 2

D-D' = plot of what potential travelers are willing to pay

S-S' = cost per trip as function of number of trips undertaken

p_o = equilibrium cost = what travelers are required to pay; q_o = equilibrium number of trips

area D-e-p_o = difference between what travelers are willing to pay and what they have to pay = traveler's or consumer's surplus

Figure 2. Equilibrium of supply and demand for interzonal travel and determination of user benefit from interzonal travel.

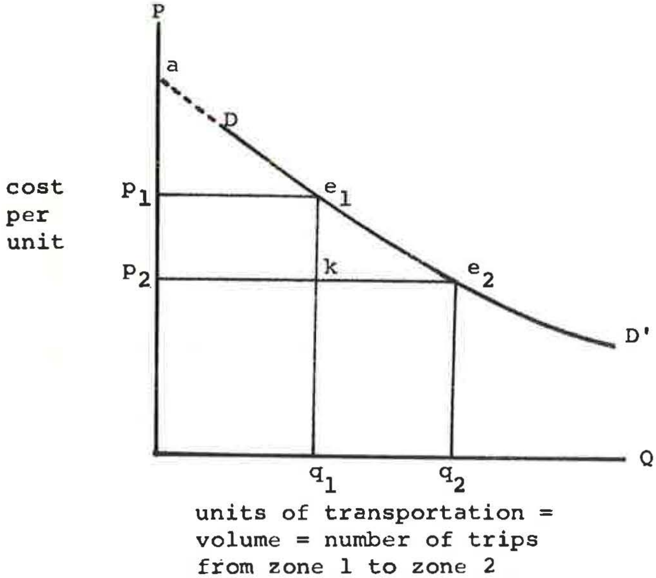


Figure 3. User benefit accruing to travel between two zones from an improvement to the transportation system.

people have to pay to travel between two zones as a function of the number of trips taken between the zones. The terms "price," "cost," "accessibility," and "separation" as used in this paper to define the cost of travel between two zones are intended to be synonymous. Thus, the supply curve is directly analogous to a volume-travel time or a volume-travel cost capacity restraint function but for travel between zones rather than over a single link.

An equilibrium of supply and demand is defined by the intersection of the curves of supply and demand for travel between two given zones. The equilibrium thus occurs when the cost (price, separation, etc.) of travel between the two zones just equals what the last traveler (as determined by the demand curve) is willing to pay to make the trip. Therefore, the interzonal volume and the interzonal separation are the coordinates of a supply-demand equilibrium point. Figure 2 shows just such an interzonal equilibrium. Note that the separation axis is in terms of the unit cost of travel. The unit cost is defined as the sum of the out-of-pocket expenses and the value of the time consumed in making the trip. Although interzonal separation could be expressed in terms of travel time or some other measure, the foregoing definition will be assumed in the remainder of this paper unless otherwise specified.

The Measure of Benefit-Consumer Surplus

Once the equilibrium point is determined, the amount of user benefit accruing from travel between the two zones can be computed. It should be recalled that the demand curve is a plot of what potential travelers are willing to pay to travel between the zones. The equilibrium price or separation is what the travelers are compelled to pay to make the trip. The difference between what these travelers are willing to pay, the demand curve, and what they have to pay, the equilibrium price, is defined as the benefit that accrues to these travelers as a result of their taking the trip. This measure of user benefit is thus synonymous with consumer surplus. Figure 2 graphically displays this concept of user benefit.

Consumer surplus was introduced as a measure of value by the British economist, Alfred Marshall, in the late nineteenth century (4). Since that time, controversies have arisen between economists as to the conditions and validity of its use, but these controversies do not invalidate its use in the present case and are too varied and involved to be summarized here (5). Consumer surplus is probably most clearly presented as a measure of the amount of user benefit accruing from improvements in the transportation system by Mohring and Harwitz (6). Figure 3 shows the composition of user benefit in this context. The initial cost and volume of travel, e.g., between two zones, are p_1 and q_1, respectively. An improvement is made to the system that results in a lower cost of travel between the two zones, p_2, and a consequent increase in travel to volume q_2. The benefit to the initial users is the lessening of the total cost of q_1 units of travel from p_1q_1 to p_2q_2, or the area $p_1e_1kp_2$. The net benefit accruing to the additional trips taken after the improvement, q_2-q_1, is defined by the area e_1e_2k. It is the difference between what the new travelers are willing to pay, as defined by the demand curve, D-D', and what they have to pay, p_2. (The new trips are shown by Mohring and Harwitz to be due to the substitution of transportation-intensive goods and services for other goods and services.) The total gain is thus equal to area $p_1e_1e_2p_2$. This gain is merely the difference in consumer surplus before and after the improvement, that is, area ae_2p_2 minus area ae_1p_1. Mohring and Harwitz further show that this measure of gain gives a close estimate of the total economic benefit from an improvement in the transportation facilities, provided that (a) transportation comprises a reasonably small proportion of the average consumer's budget and (b) business firms are reasonably "competitive" in the strict sense defined by economic theory (6). (A competitive firm in this sense is one which has enough competitors that it alone cannot arbitrarily set the prices of the goods it produces.)

Thus, the total benefit from transportation improvements may be closely approximated by user benefit and user benefit is defined by the resultant change in consumer surplus.

Supply Curves for Travel Over Networks

In order that supply-demand concepts may be applied to estimation of the benefit accruing from transportation improvements, it is necessary to examine the supply function for travel over a network. The supply curve is normally used to indicate what the producer of the item being supplied would have to be offered in terms of a unit price to induce him to produce another unit of the item. Since it is also normally assumed that he wishes to recover just whatever marginal increase occurs in his total costs due to producing the extra unit, his supply or offer curve is assumed to be his marginal cost curve.

In the case of the interzonal travel market, the basic trip-producing unit is assumed to be the link. The supply curve is a plot of the average unit cost of travel (as perceived by the traveler) over the link as a function of the number of vehicles using the link in a given time period. Marginal cost is not used here because it is average cost which the user perceives. In other words, it is assumed that the user does not take into account the changes in the costs to other users caused by his presence on the link; he is only interested in the cost to himself. The following formulas may help to clarify the difference between marginal and average costs by showing their respective relationships to total cost (7):

$$AC = \frac{TC}{q}$$

$$MC = \frac{dTC}{dq} = \text{slope of total cost curve at } q$$
$$= \text{change in total cost caused by addition of one more vehicle}$$

where

TC = total cost of q trips over a given link, \$;
AC = average cost per trip, \$/trip; and
MC = marginal cost of a given trip, \$/trip.

The shape of the supply curve for trips over a link thus depends on the physical characteristics of the link and the interactions between the vehicles traveling over it. A supply curve for an entire interzonal trip over a given route may be derived merely by adding up the supply curves for each of the links comprising the route, such as the minimum path between the zones, as shown in Figure 4.

The supply curve for travel between two zones is actually made up of portions of the supply curves for each of several alternative routes, as shown in Figure 5. Figure 5 shows that as congestion builds up on any one route, a more circuitous route may become the minimum path. Also, the supply curve as shown implicitly assumes a certain level of network loading, that is, trips between other zone pairs over the same links. In the example of Figure 5, only link 4-2 carries traffic between another zone pair, zones 3 and 2. The contribution of this traffic to the curve for path 1-4-2 and hence to the supply curve for travel between zones 1 and 2 is shown by the dashed lines.

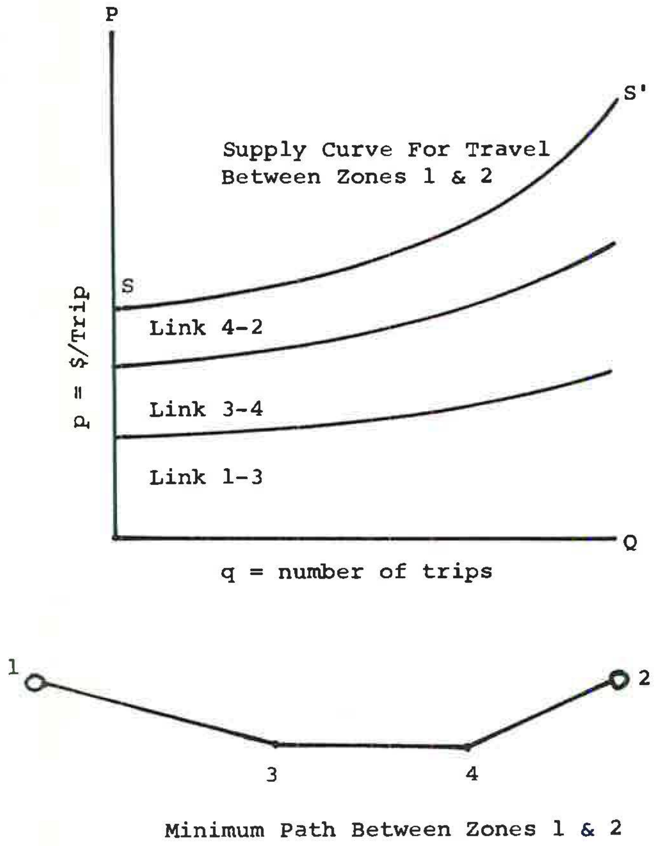


Figure 4. Obtaining interzonal supply curves.

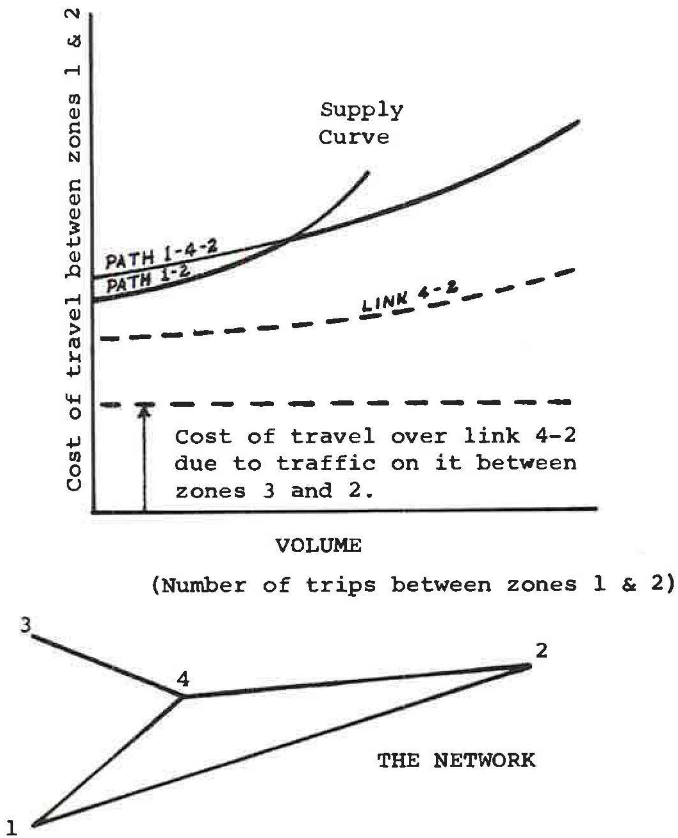


Figure 5. Make-up of interzonal supply curves.

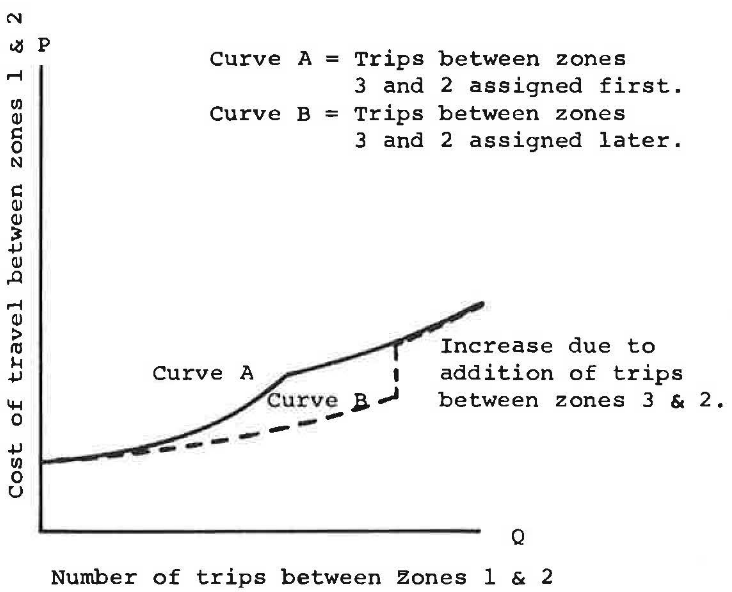


Figure 6. Supply curve differences due to different loading sequences.

Thus the loading of traffic between zone pairs in different sequences may produce different supply curves for travel between any one given pair of zones. Figure 6 illustrates two such possibilities for the travel between zones 1 and 2. Note that as soon as travel between zones 3 and 2 is added to the network, that is, the network is more completely loaded, the alternative supply curves get closer together (in this case, they coincide). This implies that the equilibrium interzonal travel times are consistent on a network with all trips assigned to it no matter what the loading sequence happened to be. It is important to note this consistency because it implies that the evaluations produced by the method outlined in this paper are also consistent no matter what the loading sequence.

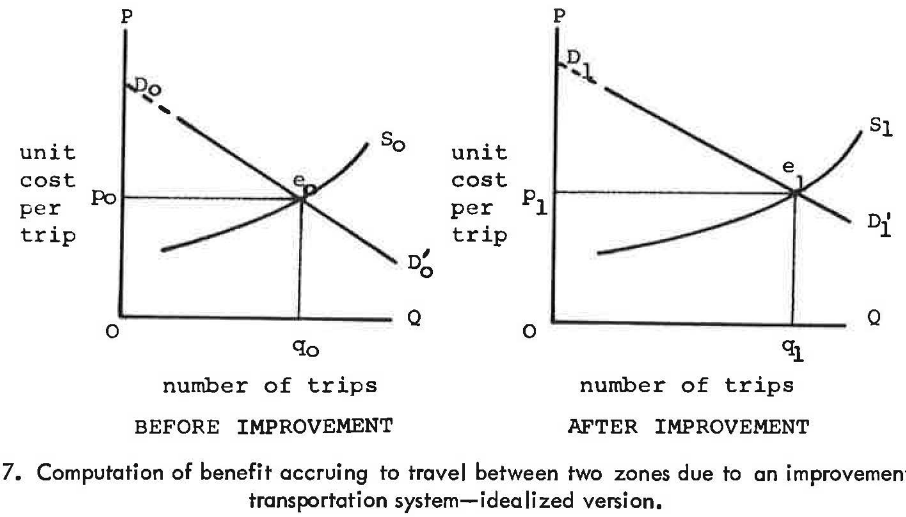
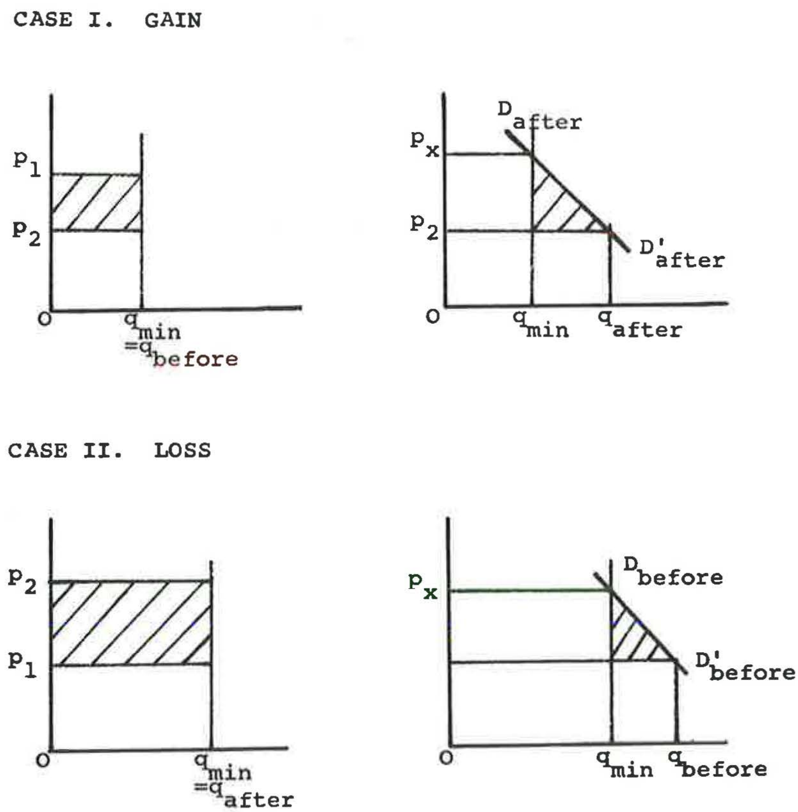


Figure 7. Computation of benefit accruing to travel between two zones due to an improvement in the transportation system—idealized version.

THE METHOD—AN IDEALIZED VERSION

The method of evaluating network changes consists of merely summing the difference in consumer surplus accruing from interzonal travel, as outlined in the previous section, over all zone pairs. The idealized version outlined in this section requires a demand function for travel between each pair of zones over each network. Compuation of the difference in benefit between networks would proceed as shown in Figure 7. Curves $D_0 - D_0'$ and $D_1 - D_1'$ are the two required demand curves for travel between zones 1 and 2 over the initial and improved networks, respectively. Hence, the benefit over



p_1 = interzonal price before change in network

p_2 = interzonal price after change in network

q_{min} = number of interzonal trips taken both before and after the network change = $\min(q_{before}, q_{after})$

q_{before} = number of trips taken between zones concerned before change in network

q_{after} = number of trips taken between zones concerned after change in network

p_x = price coordinate of point on demand curve at q_{min}

Figure 8. A more practical form of the idealized version.

each network is defined by areas $D_0e_0p_0$ and $D_1e_1p_1$. The change in benefit due to the improvement is the difference between the two areas, $D_1e_1p_1$ minus $D_0e_0p_0$. This quantity, summed over all zone pairs, would give the total benefit due to the improvement.

It would also be necessary to sum the evaluation over all time periods for which the demand pattern is assumed to be significantly different. The division of time periods depends on the time period on which the traffic assignment is based. For example, if the traffic assignment is an hourly one, the relevant time periods between which the pattern of demand changes might be the morning peak, evening peak, and off-peak; if the traffic assignment is performed on an ADT basis, the time periods of differing demand patterns might be the average annual weekday and the average annual weekend day.

The idealized version of the method in the form described has the disadvantage of requiring the computation of areas under the demand curve near the vertical axis. In many cases the demand curve is asymptotic to the vertical axis or is undefined in this region. The area under such a curve is therefore infinite or undefined. It appears most reasonable, therefore, to compute the change in benefit, in consumer surplus, in the manner illustrated in Figure 8.

For those interzonal trips, q_{min}, taken both before and after the network change being evaluated, the change in benefit is equal to the change in price, $p_2 - p_1$; the sign of the quantity $p_2 - p_1$ will determine whether a gain or a loss has been incurred for these trips. For those trips added or deleted (or diverted) by the network change, the change in benefit is equal to the area under the demand curve and above the lower equilibrium trip price. When a gain—a reduction in user costs—has been incurred, the area concerned is that under the interzonal demand curve for travel after the network change and above the interzonal trip cost or price on the network after the network change and above the interzonal trip costs or price on the network after the change. Conversely, when a loss—or an increase in trip price—is incurred, the area concerned is that under the interzonal demand curve for travel before the change and above the equilibrium interzonal trip price on the network before the change. For each case, gain or loss, the total change in user benefit accruing to the trips between the zones concerned is the sum of the benefit for q_{min} and for q_{after} or q_{before} minus q_{min}—the sum of the cross-hatched areas in the left and right figures for each case.

It should be noted that p_x is the trip price determined by the intersection of the demand curve with a vertical line drawn at q_{min}. Because the interzonal demand curve may shift—change position and shape—due to the network change, p_x is not necessarily equal to p_1 in Case I nor to p_2 in Case II.

The difficulty with the idealized version is the requirement of a demand function for each zone pair for each network. Demand functions are not necessarily the same for the same zone pairs for different networks; if they were the same, the benefit due to a network improvement could be directly computed from only the interzonal volumes and interzonal separations for the initial and improved networks. This computation and the possibilities for its use will be described in the next section. Nevertheless, interzonal travel demand functions are being formulated by other investigators, so it appears it may be possible in the near future to use the idealized version (9, 10).

USE OF THE METHOD WITH CURRENT MEANS OF TRAVEL ESTIMATION

Interzonal Trips as the Basic Commodity Demanded

It is possible to use the method described in this paper with information from current means of travel estimation. Two modes of use are described and both provide approximations of the results that would be obtained from the idealized version. The first of these modes still utilizes the concept of demand on a specific zone-to-zone basis. Its major drawback is that it requires the assumption that the demand curve for travel between any zone pair remains fixed between alternative networks. The reason for this requirement being a serious drawback can be more easily presented after a description of this version of the method is given.

The major advantage of this version is its ease of computation when used with current means of travel estimation. The version involves approximation of enough of a fixed demand curve to give a measure of the difference in the user benefits accruing over two alternative networks from travel between a given pair of zones. The ability to make such an approximation requires the assumption that the demand curve for travel between two zones is the same over each of the networks being compared. Therefore, since the equilibrium point falls on the demand curve, the equilibrium points for two different networks for travel between the same two zones would fall on the same demand curve. A line segment drawn between these two points approximates that portion of the demand curve, as shown in Figure 9.

Figure 10 shows the six possible cases of the relative orientation of the two equilibrium points. It also shows that the change in user benefit accruing from network 2 as compared to network 1 may always be computed by the following formula:

$$B = q_1 \cdot (p_1 - p_2) + \tfrac{1}{2}(q_2 - q_1) \cdot (p_1 - p_2)$$
$$= \tfrac{1}{2}(p_1 - p_2)(q_1 + q_2)$$

where

B = change, or increment, in user benefit;
q_1 = equilibrium number of trips over network 1 between the given zones;
q_2 = same as q_1 but for network 2;
p_1 = equilibrium separation or price of travel between the given zones over network 1; and
p_2 = same as p_1 but for network 2.

As with the idealized version, the sum of the incremental benefits over all zone pairs will give the total user benefit accruing from one network when compared with another for the time period concerned. It is also necessary to sum over those time periods between which demand patterns differ significantly.

The ease of data acquisition and computation makes this version appear highly desirable. The following analysis is intended to illustrate why demand functions for many, if not most, trips are not the same, however, before and after a given improvement.

Demand curves such as those illustrated in this paper show the number of units of a certain commodity that will be consumed as a function of the price of that commodity.

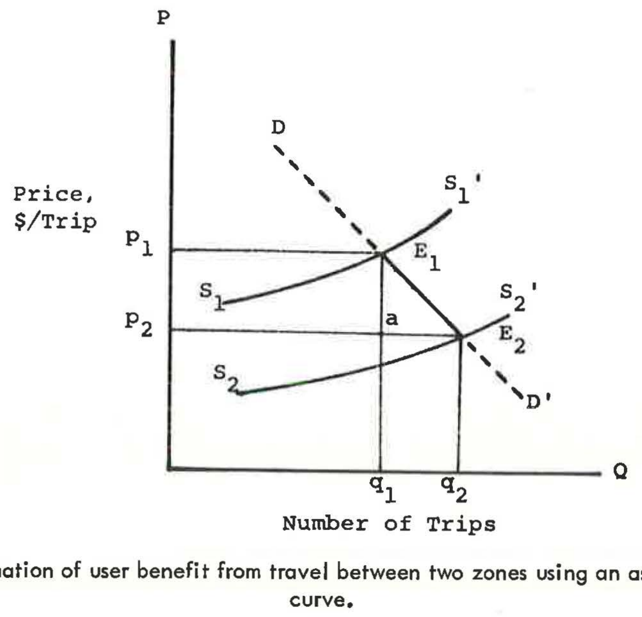
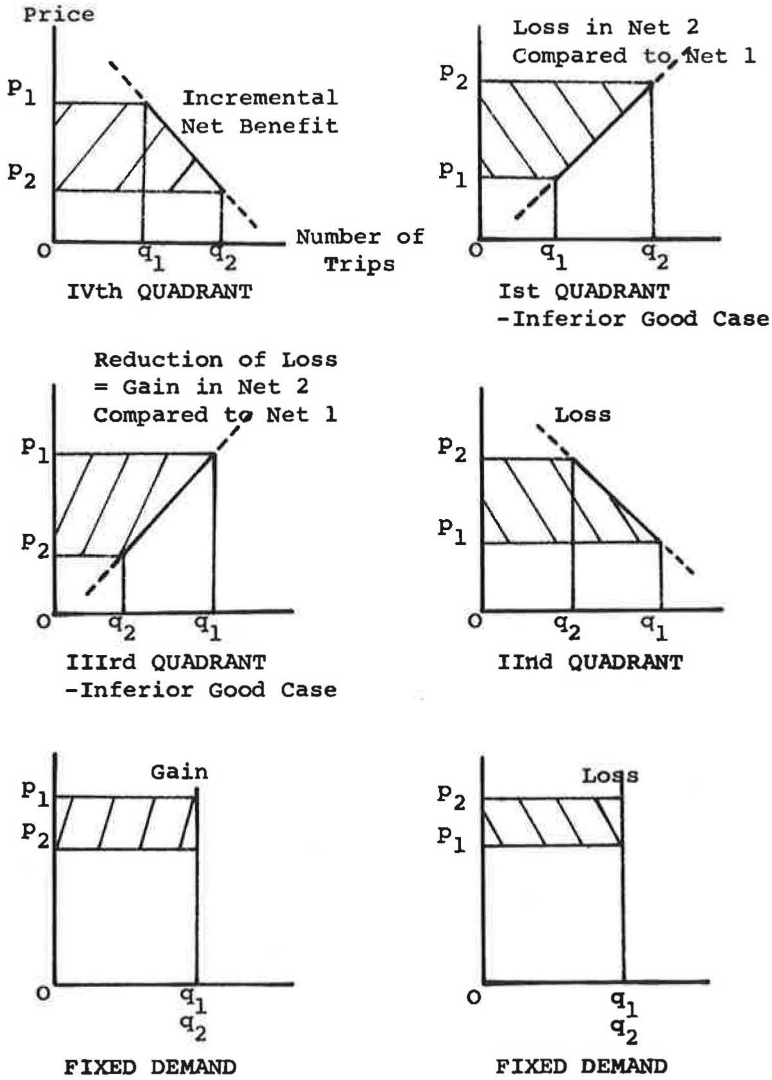


Figure 9. Approximation of user benefit from travel between two zones using an assumed fixed demand curve.

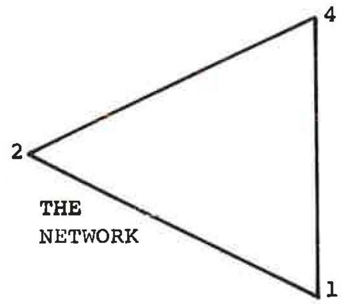


IN EACH CASE, INCREMENTAL NET BENEFIT

$= q_1 \cdot (p_1 - p_2) + \frac{1}{2}(q_2 - q_1) \cdot (p_1 - p_2)$

Figure 10. The six cases of "gain" and "loss" with a fixed demand curve.

However, the number of units that consumers will purchase depends not only on the price of the commodity itself but on the prices of other commodities as well, particularly those that can be substituted for the commodity in question (11). For example, if the price of coffee goes down far enough, tea drinkers may drink a lot less tea because they will substitute coffee for it even though the price of tea remains the same. This problem can be overcome in many economic analyses by assuming that the prices of all other commodities except the one in question remain constant. However, an improvement in a network usually affects not only the price of the trip between a given pair of zones but also the prices of trips between these and other zones. As will be



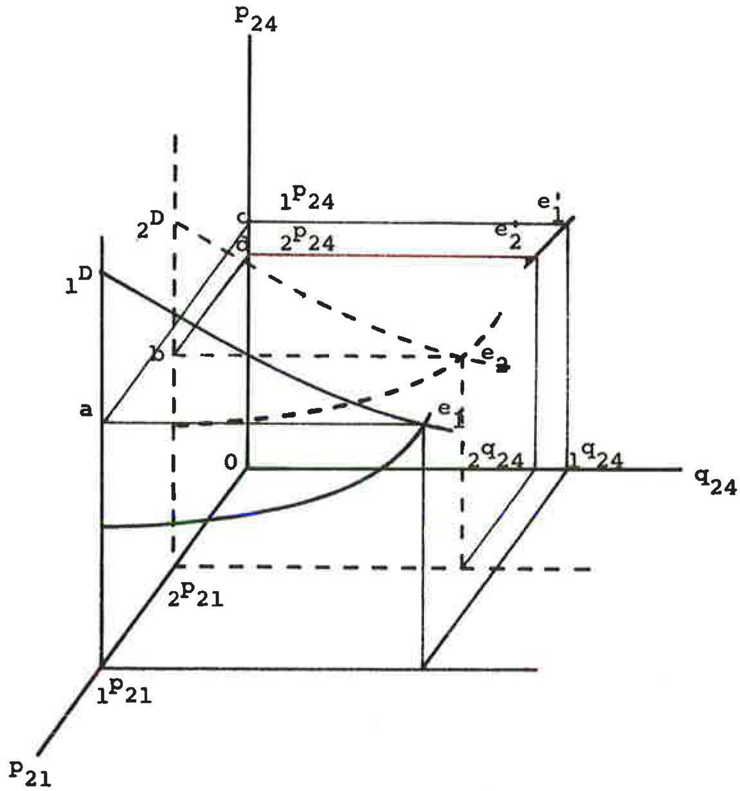
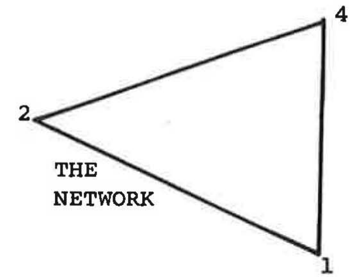


Figure 11. Change in consumer surplus accruing to trips from zone 2 to zone 4 due to change in cost of travel from zone 2 to zone 1.

shown, the changes in prices of trips from the same origin zone as the one in question may well cause the demand curve for the trips in question to shift, thus raising doubts about this version of the method.

The following example refers to the very simple network shown at the top of Figure 11. Travel from zone 2 to zone 4 is under consideration. An improvement is made to link 2-1, shortening the travel time between zones 2 and 1. With regard to individual link supply curves, only that for link 2-1 is changed. This will affect the supply curves of all interzonal trips for which link 2-1 forms a part of the interzonal route. It appears that only the supply curve for trips from zone 2 to zone 1 will be affected. Consequently, a redistribution of trips caused by the improvement and resulting in a change in trips



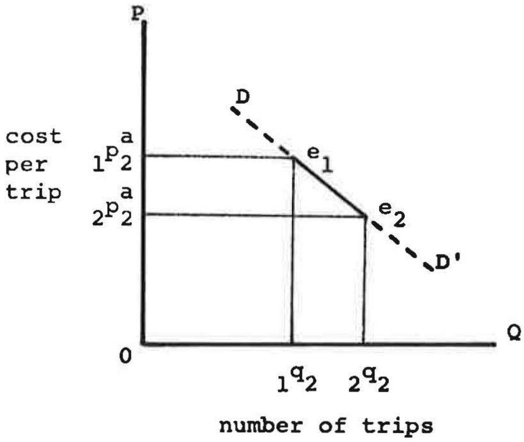


Figure 12. Computation of net benefit accruing from all trips of given purpose originating in zone 2.

from zone 2 to zone 4 must be due to a shift in the demand curve, not in the supply function, since the supply function for link 2-4 remains unchanged. It should be noted that this very simple example merely illustrates a point. In virtually every real-world case, one would expect the supply curves to shift due to differences in the pattern of network loading caused by the differences in networks.

The illustration does seem plausible for trips of certain purposes, particularly shopping and social-recreational. Such activities could be engaged in by people from zone 2 at either zones 1 or 4. If it becomes less expensive to go to zone 1 due to an improvement in link 2-1, it is reasonable to expect some of the trips being made from 2 to 4 for these purposes to be diverted to zone 1, assuming the activities at 4 and 1 are equally desirable on other counts.

The lower portion of Figure 11 illustrates the effect of such substitution on the computation of the change in benefit accruing from trips between zones 2 and 4 due to the improvement. Travel demand between zones 2 and 4 is represented in three dimensions instead of two—the number of trips, q_{24}, the price of travel between zone 2 and 4, p_{24}, and the price of travel between zones 2 and 1, p_{21}. The demand curve for travel between zones 2 and 4 before the improvement is the solid line, ${}_1D$ - ${}_1D'$, shown at ${}_1p_{21}$, the price of travel between zones 2 and 1 before the improvement. After the improvement, the price of travel between zones 2 and 1 falls to ${}_2p_{21}$, causing fewer trips to be demanded between zones 2 and 4 at each level of price, p_{24}. The new, shifted demand curve is represented by the dashed line ${}_2D$ - ${}_2D'$. Note that the supply curve for trips between zones 2 and 4 is the same both before and after the improvement. The actual change in consumer surplus (or benefit) is area ${}_2De_2b$ minus area ${}_1De_1a$, which appears to be a reduction.

The simplified version of the method results in a reduction in benefit defined by area $ce_1'e_2'd$ in plane q_{24} - p_{24}. Although the change in benefit indicated by the simplified version has the correct sign (negative), there is no guarantee of equality between the quantities of benefit change computed by the two versions; area $ce_1'e_2'd$ is not necessarily equal to area ${}_2De_2b$ minus area ${}_1De_1a$.

The magnitude of the error cannot be determined until more is known about the actual shapes and shifts of the interzonal travel demand curves. It is also unclear at this point how well present trip generation and distribution models reproduce changes in interzonal travel patterns induced by network improvements. Any use of the simplified version of the method with present travel estimation techniques should be done with clear recognition of these limitations.

All Trips From a Given Origin Zone as One Commodity

The second version of the method for use with data from current means of travel estimation requires a somewhat different way of looking at travel. In this case all trips of a given purpose from a given origin zone are considered to be one commodity. This point of view differs from that of the previously described versions in which all trips of a given purpose from a given origin zone to a given destination zone were considered to be one commodity.

Figure 12 illustrates this new point of view. The example is very similar to that of Figure 3. The primary difference is that the price is now the weighted average of all the trips (of the purpose concerned) from the given origin zone. For the very simple network shown, the average price is computed as follows:

$$ {}_1p_2^{\,a} = \frac{{}_1p_{24} \cdot {}_1q_{24} + {}_1p_{21} \cdot {}_1q_{21}}{{}_1q_{24} + {}_1q_{21}} $$

where

${}_1p_2^{\,a}$ = average price of trip from zone 2 over network 1;
${}_1p_{24}$ = price of trip from zone 2 to zone 4 over network 1; and
${}_1q_{24}$ = number of trips between zones 2 and 4 over network 1.

The general formula for average price is, therefore,

$$ {}_kp_i^{\,a} = \frac{\sum_j {}_kp_{ij} \cdot {}_kq_{ij}}{\sum_j {}_kq_{ij}} $$

where

${}_kp_i^{\,a}$ = average price of trip from zone i over network k;

${}_kp_{ij}$ = price of trip from zone i to zone j over network k; and
${}_kq_{ij}$ = number of trips from zone i to zone j over network k.

The difference in benefit accruing from two alternative networks may now be computed as shown in Figure 3. Using the terminology of Figure 11, the computation is as follows:

$$ {}_{21}B_2 = \tfrac{1}{2} \left({}_2q_2 + {}_1q_2\right) \left({}_1p_2^{\,a} - {}_2p_2^{\,a}\right) $$

where

${}_{21}B_2$ = net gain to trips from zone 2 on network 2 as compared to network 1;
${}_1q_2$ = total trips originating from zone 2 over network 1; and
${}_1p_2^{\,a}$ = as described above.

The total net gain accruing from travel is the sum of the gains from each origin zone,

$$ {}_{21}B_{Total} = \sum_i {}_{21}B_i $$

where

${}_{21}B_{Total}$ = total gain from net 2 as compared to net 1; and
${}_{21}B_i$ = gain to trips from origin zone i for net 2 as compared to net 1.

The point of view of the travel market represented in this version of the method means that changing the prices of trips that are potential substitutes for the one concerned is no longer a source of error. The shifting of trips between destination zones is reflected in the price of the trip itself since it is the composite or weighted average trip with which we are concerned. Consequently, the benefit computation is less uncertain than was that by the previous version of the method. It can still be accomplished using only the interzonal volumes and separations from each alternative network, but it will still contain whatever uncertainty is introduced by the trip generation, distribution, and assignment models employed.

The major disadvantage of this version of the method is that it is not possible to identify the amount of difference in benefit accruing from travel between each zone pair. The change in benefit is computed on the basis of zone of origin, not zone pair. It can be shown, however, that if only a redistribution of trips results, that is, no trips are generated or deleted, the quantity of benefit difference for trips from a given origin zone is the same when computed by either of the two simplified versions of the method. Thus, for such a situation, the use of an average price for all trips from a given origin eliminates the problem due to trip-substitutability merely by changing the concept of the commodity being demanded. It should be noted that the change in benefit computed by the two methods does differ whenever the total number of trips is not the same from a given origin zone before and after the change in the network.

USE OF THE METHOD IN AN EXAMPLE PROBLEM

Of the three versions of the method presented, one provides the best combination of (a) computability using data output from the assignment process and (b) the ability to determine with reasonable accuracy just who benefits and who loses because of a given network alteration. That version is the one that views demand for travel on the basis

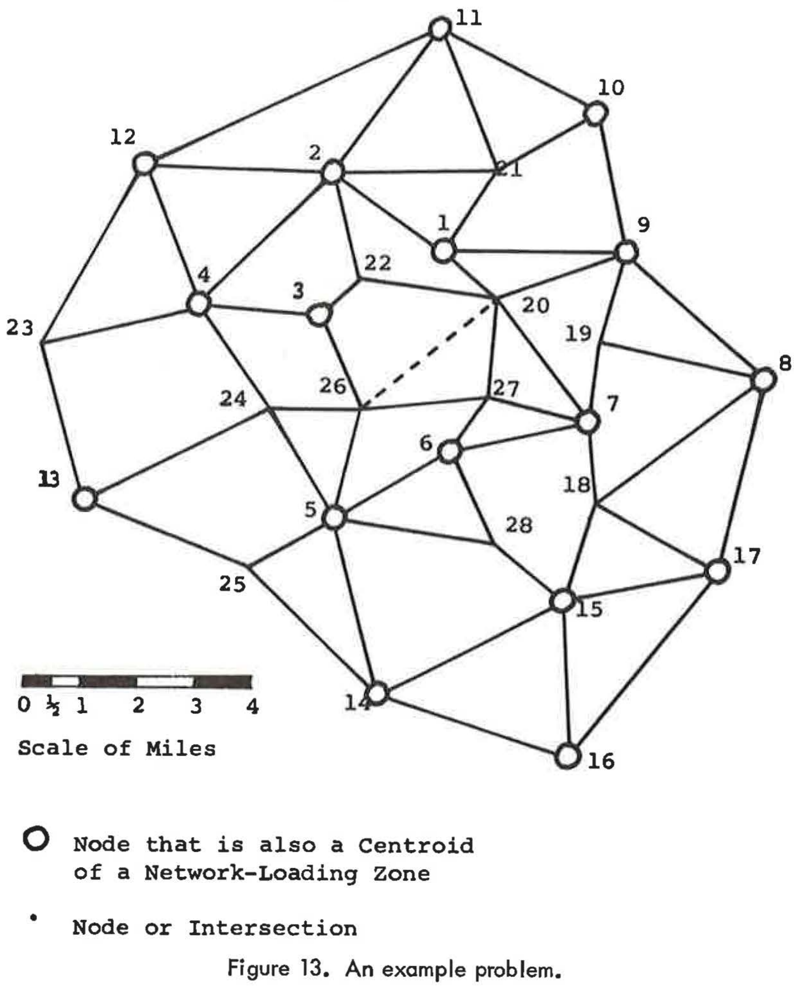


Figure 13. An example problem.

TABLE 1

USER BENEFIT BY ZONE OF ORIGIN

Zone	User Benefit in Vehicle-Minutes	Number of Vehicle Trips	Average Benefit per Trip in Minutes
1	+2654	5307	+0.5
2	-2970	7425	-0.4
3	+1238	4127	+0.3
4	+3508	7015	+0.5
5	+4637	5152	+0.9
6	+6124	7655	+0.8
7	+ 300	3003	+0.1
8	+ 189	1887	+0.1
9	+2344	3348	+0.7
10	+5500	9166	+0.6
11	-3406	8516	-0.4
12	+3913	9782	+0.4
13	+4914	6142	+0.8
14	+2809	7022	+0.4
15	0	7984	0
16	- 395	3954	-0.1
17	+ 652	6516	+0.1
Total	+32,011*	104,001	+0.3

*+32,011 vehicle-minutes per hour of morning peak period = 534 vehicle-hours per hour of morning peak period, or +0.3 minutes per vehicle-trip.

of origin zone rather than zone pair. Its use will be illustrated in an example case. First, however, a brief note is in order on the relationship of the method to present means of travel estimation.

Relationship to Trip Generation, Trip Distribution, and Traffic Assignment

With reference to Figure 10 and the first four cases presented there, the different number of interzonal trips demanded on network 2 compared with network 1 is assumed to be due to (a) the changed distribution of trips caused by differences in interzonal separations on network 2 compared with network 1, and (b) changes in trip generation when trip generation is dependent upon interzonal separation as well as on other parameters more of a socio-economic nature.

The different interzonal prices are the different interzonal separations over the alternative loaded networks. The different prices are due to:

1. For the first four cases of Figure 9, the change in demand, as outlined in the preceding paragraph, resulting in a changed level and distribution of the overall network load;
2. For all six cases of Figure 9, the change in minimum paths between zones due to the network differences.

For absolute consistency, the interzonal separations used to determine trip generation and/or distribution should be the same as those resulting from the loaded networks, the p_1 and p_2 of Figure 10. Since the trip generation and distribution phases normally occur before any traffic assignment (from which the separation measures result), some means of iteration of these several phases may have to be accomplished to assure that the interzonal separation measures finally used in each of the phases are in reasonable agreement.

An Example Problem

The example network is shown in Figure 13. The problem is to estimate the net benefit accruing from the addition of a new 8-lane divided highway to the network, link 20-26. In the problem, the measure of interzonal separation is merely travel time, but it could be out-of-pocket expense plus value of travel time or some similar more comprehensive measure.

The results of the application of the method to the evaluation (for only the morning peak period) of the addition of link 20-26 (and 26-20) to the network are shown in Table 1. The user benefit accruing to trips originating in each of the several zones is shown for the period concerned, the typical morning peak hour. Although the net user benefit is positive (534 vehicle-hours gain per morning peak hour), trips from all origin zones did not benefit. Thus the incidence of the benefits and losses by origin zone is revealed by the method. This attribute of the method is especially helpful when it is desired to improve travel from some particular geographical area or to ascertain the geographical distribution of the user benefits.

The benefits must be summed over all hours of the period for which the assigned traffic is typical. For the case shown in the example problem, the resulting 534 vehicle-hours/hour of benefit must be multiplied by the length of the morning peak period in hours in order to get the total benefit accruing during that period. The same sort of calculations would be required for peak and off-peak or other periods in which the network load varied significantly during the average day. Then, the daily benefit thus obtained could be summed for the year or other length period. Furthermore, the growth in demand over time should be accounted for (12).

Proper discounting and summing of the benefits over the expected life of the proposed facility (link 20-26) will give the total user benefit to be compared with the estimated cost of the facility. Such a comparison would be on a present value basis. Benefit-cost ratio techniques could also be used.

LOGIC FOR A COMPUTER ROUTINE TO PERFORM THE METHOD OF EVALUATION

The proposed subroutine would take the output of a traffic assignment program and would then compute the incremental net user benefit of one alternative network as compared with another.

If data on the cost differences between alternative networks is available, the values of comparison criteria such as benefit-cost ratio or present worth could also be computed with the routine.

Any number of pair-wise comparisons can be made on the same run, provided specifications are correctly made and the required data and computer time are available. The following outline provides a more detailed explanation:

A. Input required

1. Specifications for run
   a. Number of networks to be compared
   b. Names of networks
   c. Designation (O&D) of interzonal transfers to be analyzed for each or all network pairs
   d. Desired criteria of evaluation (if any)
   e. Input data mode (cards, disk, etc.)

2. Input data
   a. For each alternative network
      (i) For each interzonal transfer, the volume and price at completion of assignment
      (ii) Costs (maintenance and construction) associated with each alternative network
      (iii) Interest rate(s) to be used in comparison criteria computation

B. Computation required

1. Pick first pair of alternative networks (alternative networks should be in order of increasing cost)
2. For network pair
   a. For each origin zone
      (i) Compute average price, $p_i^a = \sum_j p_{ij} q_{ij} / \sum_j q_{ij}$
      (ii) Compute net benefit [say, NETBEN (I) = $\frac{1}{2}({}_1q(I) + {}_2q(I))\,({}_1p^a(I) - {}_2p^a(I))$]
      (iii) Increment net benefit for this pair [say, BENSUM = NETBEN (I-1) + NETBEN (I)]

   b. Compute values of comparison criteria

3. Repeat 2 for each successive pair of alternative networks

C. Output: For each pair of network alternatives
1. The names of the network
2. The incremental net user benefit
3. The values of the comparison criteria

A flow chart of the subroutine as outlined appears in Figure 14.

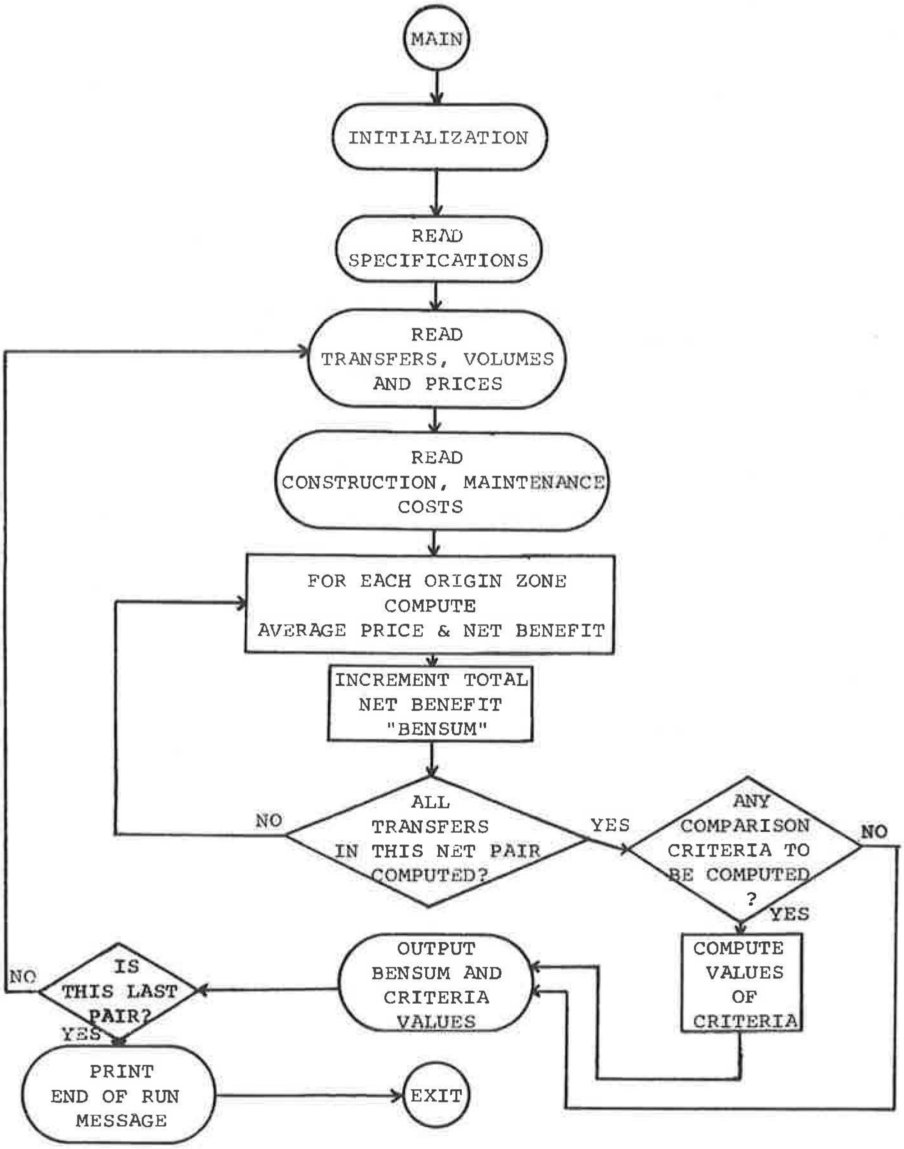


Figure 14. Example flow chart—computation of user benefit by origin zone.

CONCLUSION

The method of evaluation described in this paper makes it possible to compare networks even though the interzonal demand for travel over the two networks may differ. Of course, the differences in demand must be due to the different networks. Thus the effect of alternative networks on travel demand can be accounted for. Constant or fixed travel demand need not be assumed.

The method is based on a consistent economic rationale. It uses consumer surplus as a measure of benefit. The measure of net value produced by the method is only as comprehensive as the data on which the values of interzonal separation are based. There is, therefore, a need for more comprehensive measures of interzonal separation. It is desirable that those items of travel cost perceived by travelers and assessed by them in making travel decisions should be identified and included in some manner.

Although only two networks may be compared at a time, two of the versions of the method require estimation of only a partial segment of the demand curve for travel from each zone. Further investigation is needed to determine if the idealized version may be practically used with the multi-dimensional interzonal travel demand functions being developed by others.

Finally, the version advocated for current use requires only the interzonal volumes and the interzonal separations for each network being evaluated. It is not constrained, therefore, to any particular technique of traffic assignment, or for that matter, even to traffic assignment if reliable volume and separation data are available from some other source.

ACKNOWLEDGMENTS

The research that resulted in this paper was sponsored by the Massachusetts Department of Public Works in cooperation with the U. S. Bureau of Public Roads. The author is indebted to Professors Joseph H. Stafford of MIT and Edward K. Morlok of Northwestern University for valuable suggestions on the content and scope of the paper. A procedure suggested by Dr. Gerald Kraft of Charles River Associates has been incorporated. The author retains responsibility for any errors and omissions which remain in the paper.

REFERENCES

1. Harvey, T. N. Logic of a Subroutine for the Approximation of Incremental User Benefits on Highway Networks Using the Output of a Traffic Assignment Algorithm. Memorandum ITP 4, MIT Department of Civil Engineering, March 1, 1966.
2. Pollard, W. S., Jr. Long Range Urban Traffic Planning. In Traffic Engineering Handbook. Institute of Traffic Engineers, Washington, 1965, p. 703-723.
3. Haikalis, G., and Joseph, H. Economic Evaluation of Traffic Networks. HRB Bull. 306, p. 39-63, 1961.
4. Marshall, Alfred. Principles of Economics, Ninth (Variorum) Edition, Macmillan Co., 1961 (First Edition, 1891).
5. Samuelson, Paul A. Foundations of Economic Analysis. Atheneum, 1965, p. 195-202.
6. Mohring, Herbert D., and Harwitz, Mitchell. Highway Benefits—An Analytical Framework. Northwestern University Press, 1962, p. 23-26, 52-56.
7. Wohl, Martin. Demand, Cost, Price, and Capacity Relationships Applied to Travel Forecasting. Highway Research Record 38, p. 40-54, 1963. [See for discussion of the difference between marginal and average costs in a travel context.]
8. Soltman, Theodore J. Effects of Alternate Loading Sequences on Results From Chicago Trip Distribution and Assignment Model. Highway Research Record 114, p. 122-140, 1966.
9. Domencich, T. A., Kraft, G., and Valette, Jean-Paul. Estimation of Urban Passenger Behavior: An Economic Demand Model. Presented at the 47th Annual Meeting and included in this Record.
10. Quandt, R. E., and Baumol, W. J. The Demand for Abstract Transport Modes—Theory and Measurement. Jour. Regional Science, Vol. 6, No. 2, Winter 1966.
11. Samuelson, Paul A. Economics, Fifth Edition. McGraw-Hill, 1961, p. 436.
12. Wohl, Martin, and Martin, Brian V. Evaluation of Mutually Exclusive Design Projects. HRB Spec. Rept. 92, 1967, p. 50-51.

Estimation of Urban Passenger Travel Behavior: An Economic Demand Model

THOMAS A. DOMENCICH, GERALD KRAFT, and JEAN-PAUL VALETTE,
Charles River Associates, Cambridge, Massachusetts

•THIS paper describes an urban transportation demand model that has a number of attractive features in evaluating the effects of alternative transportation systems. The model is derived directly from the theory of consumer demand in the economic literature. [For a discussion of the theory of consumer behavior consult one of the standard intermediate texts on price theory, e.g., (1).]

To an economist, urban transportation is simply another commodity—in principle no different from any other good or service, although perhaps in practice far more complex and multi-faceted than other commodities. Thus, it is natural for an economist to approach the task of developing an urban transportation demand model in much the same way that he would attempt to model the demand for any commodity. (To be precise, transportation is a derived demand, and therefore the general approach would be the same as that for any other derived demand commodity.) In doing this, he is likely to draw on elements of the theory of consumer behavior. This body of economic theory has been tested with numerous empirical studies, including at least one study of intercity travel demand (2), and provides a well-founded basis for a model of urban travel demand.

This paper describes the urban passenger travel demand model we have developed based on economic theory. A discussion of the reasoning underlying the model leads to a presentation of the general specification of the model, including a description of the relevant variables, the mathematical form taken by the model, and the statistical techniques used in estimating its parameters. Empirical estimates of the model's parameters were obtained using data for Boston. In the final section, selected results are presented and their implications for transportation investment planning are discussed.

We feel that by taking a fresh view of travel demand and by approaching it from an economist's viewpoint we have developed a model that has several important advantages over the existing generation of demand models. Upon further consideration, and after exposure to a wider range of outside review and criticism, it may turn out that the differences between our approach and the extant travel demand models are neither important nor advantageous. We may be presenting, as it were, the same contents in a different package. At present we doubt this; the model seems to be conceptually sounder, a better tool for forecasting, and more useful in evaluating policy alternatives than the previous models that have been developed. To be sure, many of the elements of the model are familiar. This is not surprising since most of the variables used in investigating travel demand are likely to be relevant within the context of any particular model. In any event, approaching this problem from a new point of view can only enrich our knowledge, for if it results in important advantages, the state of the art is that much advanced, while if it only confirms what we already know, our confidence in the existing techniques can be that much stronger.

BACKGROUND

Since the model is to be used in evaluating transportation system alternatives, at the outset it is useful to consider what policy questions we would like to be able to explore with the model. The following are indicative of those that should clearly be included.

Paper sponsored by Committee on Transportation System Evaluation and presented at the 47th Annual Meeting.

What effect will increases in population, personal income, and car ownership have on travel demand and consequently on future traffic congestion? What effect would changes in travel time or cost have on total travel demand and on the demand for the various modes? What effect would changes in the spatial distribution of homes, jobs, and retail establishments have on future traffic flows?

To obtain a deeper level of understanding, we would also like to know whether travel demand is more sensitive to some policy variables than to others. For example, are out-of-pocket costs such as parking charges or toll fees more onerous to auto users than vehicle operating costs? Are transfer or access times more onerous to transit users than in-vehicle line-haul times? Are savings in costs more important to travelers than savings in time?

In approaching the problem of developing a model of urban travel demand, it is useful to begin with consideration of the individual. Although we are concerned with aggregates of people, their behavior can probably best be understood by considering the behavior of individual travelers. We would like to know what decisions the traveler faces in his travel behavior, and what factors influence these decisions.

The individual traveler has a set of choices to make. He must decide whether to make a trip at all, where to go, which route to take, which mode to use, and when to go. Each of these choices has an associated set of values and costs (in money and time) to the individual; the values themselves will vary with the trip purpose and sometimes with the time of day. On a moment's reflection, it is clear that these choices are not independent. The costs of the various modes influence not only the choice of mode but also the selection of destination and the determination of whether the trip should be made at all. For example, an improvement in the freeway system that reduces the travel time to the downtown area may not only divert shoppers from regional shopping centers to downtown and shift travelers from transit to auto, it may also stimulate an increase in the total number of shopping trips. Mounting congestion, on the other hand, may reduce the total number of shopping trips by making each trip more effective and well planned.

Similarly, the attractiveness of a destination may influence both the distribution of trips between destinations and the number of trips that are made. Rejuvenating the downtown area, for example, or building a new stadium or concert hall, may not only redistribute shopping, social, and recreational trips between zones; it may also draw housewives or families out of the home and thereby increase the total number of trips made.

These points may become clearer by considering the following extreme but useful example. A woman living on a relatively isolated coastal island may make very few trips to the mainland because of the time, cost, and general difficulty of making the trip. One would expect that each trip would be well planned and executed. There is little opportunity to "run out to the store" to get the item she forgot. On the other hand, a woman living a few doors from a shopping district may make a large number of trips without nearly as much care. If she forgets an item, the resulting cost and inconvenience are relatively minor.

Although the example is extreme, it spans a wide variety of circumstances that occur in real life. It illustrates the point that the alternatives available to individuals determine not only their selection of modes and destination zones, but also the total number of trips that they make. If conditions are favorable, the individual may make many trips; if all the available alternatives are poor, he may make few trips or even no trips at all. Because the value of the trip depends on its ultimate objective, we might expect shopping, personal business, social, and recreational trips to be more sensitive to these conditions than work trips. But the fact that trips made for the former purposes comprise a large and growing percentage of all trips makes it essential that we analyze the responsiveness of passenger travel demand to time and cost conditions. In most cities nonwork trips probably constitute the majority or close to the majority of trips.

The approach usually taken in the analysis of urban travel demand separates the problem into elements of trip generation, attraction, distribution, assignment to routes, and modal split. As the foregoing discussion indicates, however, these choices are so intertwined that they are best treated as being made simultaneously rather than separately.

The separate treatment of these elements in the currently popular models has several related consequences. First, it results in the implicit assumption that the number of trips generated is independent of the performance of the transportation system. That is, it is assumed that changes in travel time or cost can influence the modal split or distribution of trips between zonal pairs, but cannot change the total number of trips generated. By assumption, the models assert that the policies implemented by transportation planners have no effect on the total number of trips made! While it is possible that changes in the transportation system will not affect trip generation, there is no good reason for making this assumption a priori. It seems better to avoid making this assumption altogether; we have everything to gain and nothing to lose by letting the question be settled through the results of empirical estimation.

Second, the separation of these elements could lead to improper measurement of the effect of the independent variables in the individual trip generation and attraction equations. To clarify this, recall the comparison of the woman on the coastal island with the woman living next to the shopping center. Assume that the woman on the coastal island has a large family, is relatively wealthy, and that the family has several autos, while the second family is moderate in size, relatively poor, and has one automobile. Other things the same, we would expect more shopping trips from the larger family with the higher income and greater number of autos. If the effect of car ownership or family size on trip generation were measured from these two observations, however, the opposite would appear to be true. The reason for this is, of course, that the time, cost, and related inconvenience of travel have been left out of the comparison, and their effects on travel behavior have been attributed to socioeconomic variables. As the example illustrates, improperly specifying the trip generation and attraction equations by omitting relevant variables may cause the effects of the variables actually included in the equations to be confounded with the effects of the omitted variables and consequently cause them to be improperly measured (3).

The next section describes how all these elements of choice can be represented in a single model. To the extent that the model can be made to represent the effects of each element of travel behavior, the model can provide answers to the significant policy questions posed. Furthermore, by incorporating the transportation system characteristics explicitly in the model, the planner can investigate the consequences of alternative designs on tripmaking behavior, facilitating the accomplishment of design objectives.

MODEL SPECIFICATION AND VARIABLES

Economic theory provides us with useful guidelines for specifying a demand model: first, because it identifies in a broad, general way the variables that influence demand; and second, because it specifies the general nature of the relationship between these variables and demand. The variables identified by the theory of consumer behavior as relevant in a study of demand are the price of the good or service being investigated, the prices of competing or complementary goods or services, and income.

For urban auto passenger demand the subject commodity has at least two prices that must be considered—automobile travel time and cost. The prices of competing goods are the times and costs of travel by the available transit modes. For transit passenger demand, of course, the prices of the subject commodity are transit cost and travel time, while the relevant prices of substitutes are the times and costs of travel by auto. For auto, the prices of complementary goods are parking charges, toll fees, and the walking time to and from the car. For transit, they are the times and costs of access to and from the transit station.

Economic theory tells us that demand will be negatively related to the prices of the subject commodity and positively related to the prices of substitutes. Demand will be negatively related to the prices of complements.

The relevant income variables include both the incomes of individuals (or households) in the urban area and various measures of output of the activities that attract trips. Demand for most goods is positively related to the incomes of the individuals in the market for the good or service. However, this need not be so, and there are examples of goods for which the demand decreases as income goes up. People substitute a more desirable commodity for the good in question as their incomes rise. For instance, the demand for cheaper cuts of meat may decrease with a rise in incomes because people shift to more expensive cuts of meat.

The need for measures of output of the activities that attract trips stems from transportation's role as a derived demand commodity. That is, transportation is usually not desired for its own sake but rather because it enables the traveler to satisfy another demand such as shopping, work, or personal business. Thus, some measure of the level of operations in the activity from which the demand for transportation is derived is needed in the transportation demand function. This requires disaggregating the trips by trip purpose and specifying the relevant measure of activity for each trip purpose—sales, employment, etc. These measures of activity are the usual attraction variables. All else equal, we expect the demand for transportation to be positively related to the level of operations of the activities served by transportation.

The foregoing variables are the appropriate ones to measure individual demand behavior. Aggregate demand will, of course, be positively related to the number of individuals in the market and often will depend as well on various socioeconomic characteristics of these individuals, such as age, occupation, family size, and ethnic background.

These ideas are incorporated in the following equation, which is a general expression for the urban transportation demand model that we have developed:

$$N(i, j, i \mid P_0, M_0) = \phi\,[\underline{S}(i \mid P_0), \underline{A}(j \mid P_0), \underline{T}(i, j, i \mid P_0, M_0), \underline{C}(i, j, i \mid P_0, M_0), \underline{T}(i, j, i \mid P_0, M_\alpha), \underline{C}(i, j, i \mid P_0, M_\alpha)]$$

where

$N(i, j, i \mid P_0, M_0)$ = the number of round trips between origin i and destination j for purpose P_0 by mode M_0;

$\underline{S}(i \mid P_0)$ = vector of socioeconomic characteristics appropriate to purpose P_0 describing the travelers residing in zone i;

$\underline{A}(j \mid P_0)$ = vector of socioeconomic and land-use characteristics describing the level of activity appropriate to purpose P_0 in destination zone j;

$\underline{T}(i, j, i \mid P_0, M_0)$ = vector of travel time components for the round trip from origin i to destination j for purpose P_0 by mode M_0;

$\underline{C}(i, j, i \mid P_0, M_0)$ = vector of travel cost components for the round trip between origin i and destination j for purpose P_0 by mode M_0;

$\underline{T}(i, j, i \mid P_0, M_\alpha)$ = vector of travel time components for the round trip between origin i and destination j for purpose P_0 by each of the alternative modes ($\alpha = 1, \ldots, n$); and

$\underline{C}(i, j, i \mid P_0, M_\alpha)$ = vector of travel cost components for the round trip between origin i and destination j for purpose P_0 by each of the alternative modes ($\alpha = 1, \ldots, n$).

In words, the equation says that the number of directed round trips between any zonal pair for a given purpose and mode is a function simultaneously of the number of individuals (or households) in the origin zone and their socioeconomic characteristics, the appropriate level of activity and other relevant socioeconomic and land-use characteristics in the destination zone, together with the round-trip travel times and costs of the subject mode as well as those of competing modes. Times and costs of complementary services are included in the vectors of times and costs of the subject mode because they are also negatively related to demand and because it is often difficult in practice to distinguish the characteristics of the subject mode from those of its complementary services. There is an equation for each trip purpose and each mode.

Notice that the dependent variable is the interzonal round trip. It is the interzonal trip because, first, this is the quantity of interest rather than the number of trips generated or attracted by a zone; and second, as discussed earlier, the simultaneity of the decisions about whether to make a trip at all, where to go, and which mode to use require that the socioeconomic characteristics of the origin and destination zones be considered together, along with the trip times and costs required to travel between that specific zonal pair. This necessitates examination of zonal-pair combinations.

It is preferable to analyze the round trip because time and cost conditions on both legs of the trip are considered by the traveler in making his trip decisions. Moreover, it is clear that the return trip selection of mode depends on the modal choice made for the outbound trip, and the destination of the return trip depends on the origin of the outbound trip.

The choice of when to travel (i.e., which hour of the day) is not reflected in the foregoing model. This choice was omitted, not because it is unimportant, but rather because it substantially increases the size and complexity of the model. If the day is disaggregated only into its peak and off-peak components, the number of equations is doubled and the number of variables almost doubled. The number of equations is doubled because separate equations are needed for the peak and off-peak times of day, and the number of variables is almost doubled because separate peak and off-peak variables are needed for each travel time and cost variable. Because of the time and budget limitations of the study, it was not possible to consider a model of this size and complexity, so a simple heuristic device was developed to take account of the choice of time of day.

The model allows for consideration of a number of transit modes. In this study, all transit modes were aggregated into a single heterogeneous mode. This was not done by choice but rather because data were available only for the single heterogeneous mode within the time limitation of the study.

The independent variables include the usual socioeconomic and land-use variables used in the current models to measure trip generation and attraction and at the same time include the system performance variables used to measure the times and costs to the traveler of making the trip by each of the alternative modes.

The socioeconomic and land-use variables tested in this study are straightforward and conventional, and need not be described in detail here. They include population and population density (i.e., population per acre), personal income, car ownership, employment and employment density for relevant industry groups, etc.

Since the system variables are the likely policy variables, they require and deserve lengthier comment. First, because in the view of the user all components of the trip probably contribute to its inconvenience, total door-to-door travel time and cost must be examined rather than only line-haul costs or times. Second, because travelers may react differently to different components of travel time and cost, it is desirable to disaggregate the times and costs into their component parts. Answers to the policy questions listed earlier can only be obtained by disaggregating travel costs and times.

The travel time by transit consists of a walk or drive to the station, the wait at the platform, the line-haul time, the time consumed in any transfers that have to be made, the walk from the terminating station to the final destination, and a component we choose to call schedule delay. The schedule delay is any additional time that may be incurred because the arrival time allowed by the transit schedule may differ from the traveler's preferred arrival time. (If he has a 9:30 appointment, for example, and the nearest transit arrival time is 9:00, the traveler has a 30-minute schedule delay.)

Similarly, the travel time by auto consists of the walk to the auto, the line-haul time, the parking time, the walk from the parking place to the destination, and the schedule delay. (If the auto traveler must leave early to get a parking place, for example, he suffers a schedule delay. Schedule delay for automobile also results from high congestion and queuing situations requiring trip-makers to arrive early in order to make their appointments.)

For automobile trips, the travel costs consist of vehicle operating costs, toll charges, and parking fees. The costs of transit trips include both transit fares and any costs incurred in traveling to or from the transit stations.

When the different components of time and cost are taken as separate explanatory variables, it may be possible to bring the effect of policy actions into much sharper focus. The non-line-haul portions of transit travel time, for example, may be far more onerous than the line-haul time. Auto out-of-pocket costs such as tolls and parking charges are more visible and therefore may be more onerous to the driver than vehicle operating costs. The relationships expressed in the model should reflect these evaluations by the traveler because the estimated responses to the onerous portions of the travel time will be greater than those for the less objectionable segments of the trip.

In our empirical research, travel time and cost were disaggregated into the following components: auto in-vehicle time, auto out-of-vehicle time, transit line-haul time, transit excess time, auto line-haul costs, auto out-of-pocket costs, transit line-haul costs, and transit excess costs. These variables are defined in Appendix A.

BEHAVIORAL ASSUMPTIONS AND MATHEMATICAL FORMS

Three basic mathematical forms of the model were tested: logarithmic, linear, and mixed log and linear. Each of these forms can be described in terms of the behavioral assumptions implied. The logarithmic model assumes that equal relative changes in travel times and costs evoke equal responses in travel demand. Thus it assumes, for example, that a housewife will curtail her trips to the supermarket by the same percentage amount whether her travel costs have increased from 10 to 11 cents or from 10 to 11 dollars.

The linear model, on the other hand, focuses on absolute changes. Its shortcoming is that it assumes that reducing a two-hour trip by, say, 10 minutes is as important as reducing a 20-minute trip by 10 minutes.

We generally prefer the mixed form to either the pure log or linear forms because, by including both linear and logarithmic terms for each variable, the effects of both relative and absolute changes in the variable are measured. Because of its greater generality, the mixed log and linear form has been tested for each equation in the model. This procedure provides empirical evidence on whether absolute or relative changes in each variable are important or whether both are important. The difficulty with the mixed form is that the linear and logarithmic values of a variable are closely correlated (i.e., are collinear). This makes estimation of the separate parameters difficult.

In interpreting the results of our empirical research and in comparing the estimated model parameters with our prior notions of traveler behavior, it is useful to introduce the concept of demand elasticity. For our travel demand model, elasticity is the percentage change in the number of trips demanded for a given purpose and mode in response to a one percent change in one of the variables giving rise to travel demand, assuming all other explanatory variables in the equation are held constant.[1] This is a particularly useful concept for comparing the sensitivity of travel demand to changes in a number of explanatory variables because elasticity is dimensionless. Thus, comparisons are not confused by the particular units in which the variables are expressed.

By convention, an elasticity of less than unity (in absolute value) is called inelastic, and one that is greater than unity (in absolute value) is called elastic. In the former case a given change in an explanatory variable results in a less than proportionate change in demand, while in the latter case the change in demand is greater than proportionate. It is also conventional to call the elasticities with respect to the variables for the subject mode direct elasticities, and the elasticities with respect to the variables for competing modes cross-elasticities. We shall employ this terminology in the remainder of the discussion.

---

[1]Elasticity is precisely defined as

$$\eta_x = \frac{\partial N/N}{\partial x/x} = \frac{x}{N}\frac{\partial N}{\partial x}$$

where η_x is the elasticity of travel demand, N, with respect to variable x. It should be noted that, in general, the elasticity is not equivalent to the coefficients in a regression equation. The elasticity expresses a ratio of relative changes while, for example, the coefficient in a linear model expresses a ratio of unit changes. The latter ratio is dependent on the choice of units—minutes vs hours, for example—while the former ratio is independent of the units selected.

TABLE 1

ALTERNATIVE MATHEMATICAL FORMS OF THE MODEL

Model Form	Elasticity	Form Estimated
Logarithmic		
$N = KX^{\alpha}$	α	$\ln N = \ln K + \alpha \ln X$
Linear		
$N = K + \alpha X$	$\alpha \frac{X}{N}$	$N = K + \alpha X$
Mixed log and linear		
$N = KX^{\alpha} e^{\beta X}$	$\alpha + \beta X$	$\ln N = \ln K + \alpha \ln X + \beta X$
$N = K + \alpha \ln X + \beta X$	$\frac{\alpha + \beta X}{N}$	$N = K + \alpha \ln X + \beta X$

N = dependent variable,
X = independent variable,
K, α, β = parameters to be estimated.

Other things equal, we expect the elasticities with respect to the times and costs of travel by the subject mode (the direct elasticities) to be negative. Thus, we expect more trips by a given mode the less the cost and inconvenience of travel by that mode. The elasticities with respect to the travel times and costs of competing modes (the cross-elasticities) should be positive. We expect more trips by a given mode the greater the cost and inconvenience of travel by alternative modes.

The relationship between the model parameters and elasticities for each model form is given in Table 1. The log model implies that the elasticities are constant over the entire range of the variables. The log model is the single case where the elasticities are equal to the coefficients. The linear model implies that the elasticity depends on the level of the variable and accordingly it varies continuously as the level of the variable changes. In the mixed form of the model, the elasticity has both a constant and a variable term.

ESTIMATION TECHNIQUE

The model was estimated by means of constrained multiple regression analysis. This method of estimation consists of estimating parameters by minimizing the sum of squared deviations as with ordinary least squares but performing this minimization while satisfying certain prespecified conditions derived from a priori information. The constrained least squares regression technique used in this study states the problem as an equivalent quadratic programming problem.

One reason for the use of constrained regression analysis is related to the problem of unequal zone sizes. Since zones cannot generally be selected to be of equal size (expressed in terms of either area or population), the model must account for differences, particularly with respect to population. Thus a zone with twice as many people, other things being equal, is likely to produce roughly twice as many trips.

We may consider the problem from another point of view. Suppose we have a model to describe the number of trips from adjacent zones A and B, having similar characteristics, to another zone, C. Let us define a new zone, A′, which is the geographic zone encompassed by zones A and B. The models should predict the same number of trips from A′ to C as the number of trips from A to C plus the number of trips from B to C. This will only be the case if the model is homogeneous in the first degree with respect to the variables describing zone size. Since it is necessary for the model to behave in this way, parameters associated with the zone size-related variables must be made to behave appropriately in the constrained regression formula. This was done by constraining the demand elasticity with respect to the size variables to be unity.

The main problem, however, requiring prespecified conditions on the values of the estimated parameters is collinearity. In this study collinearity can be attributed either to the form of the model or to the nature of the variables. As an example of the first case, a model that contains both the linear and the logarithmic forms of a variable is necessarily subject to some degree of collinearity. The second case of collinearity occurs when trip behavior is independently influenced by two variables which show a close relationship to each other, either structurally or spuriously. Modal choice may, for instance, depend on car ownership as well as income of the trip-makers. The structural collinearity results because car ownership itself is related to income. Because of the statistical problems resulting from collinearity, it is very difficult to assess the individual effect of collinear variables unless some additional information is provided. It is often possible to specify the sign or reasonable ranges of a parameter from a priori knowledge or economic theory. The expected signs of the elasticities with respect to the system performance variables were described earlier. Constrained regression allows the analyst to take advantage of this information. In such a case, this information regarding a variable is explicitly taken into account by constraining the corresponding parameters; it then becomes possible to estimate the individual effect of the collinear variable.

Constrained regression was used to treat collinearity by imposing appropriate sign constraints on the direct elasticities and cross-elasticities of the system variables. This a priori specification of the parameter signs is an application of the economic theory of demand.

DISCUSSION OF EMPIRICAL RESULTS

The parameters for the model were estimated using data for the Boston metropolitan area. Although equations were estimated for several additional trip purposes, we have selected work and shopping trips to illustrate the application of these models. It is important to emphasize the illustrative nature of these results. The empirical work suffered from all of the normal handicaps, such as lack of time and funds for a full exploration, but in addition was dependent on input data that were never intended for this model.

Perhaps the most serious limitation in the available data was the fact that transit trips represent all non-auto trips whether they are commuter rail, subway, or bus. The heterogeneous nature of the transit mode made it extremely difficult to obtain estimates for the parameters associated with the transit variables. Since most research in urban travel is oriented toward highway transportation, it is not surprising that the existing transit data are less carefully compiled than the auto data, but this practice severely inhibits research on transit demand, and because of the interdependencies of auto and transit demand makes research on auto demand more difficult.

Tables 2 and 3 give the elasticities of demand for auto and transit work and shopping trips with respect to each component of travel time and travel cost. The complete auto

TABLE 2

ELASTICITIES OF PASSENGER TRAVEL DEMAND WITH RESPECT TO THE COMPONENTS OF TRAVEL TIME

Trip Purpose	R^2	Auto Trips: Direct Elasticities: Auto In-Vehicle	Direct Elasticities: Auto Out-of-Vehicle	Cross-Elasticities: Transit Line-Haul	Cross-Elasticities: Transit Excess
Work	.41	-.82	-1.437	0	.373
Shopping	.55	-1.02	-1.440	.0950	0

Trip Purpose	R^2	Transit Trips: Direct Elasticities: Transit Line-Haul	Direct Elasticities: Transit Excess	Cross-Elasticities: Auto In-Vehicle	Cross-Elasticities: Auto Out-of-Vehicle
Work	.35	-.39	-.709	0	0
Shopping	.63	-.593[a]		0	0

[a]The available shopping transit trip sample was unsuitable for estimating elasticities for the disaggregated time components.

TABLE 3

ELASTICITIES OF PASSENGER TRAVEL DEMAND WITH RESPECT TO THE COMPONENTS OF TRAVEL COST

Trip Purpose	Auto Trips			
	Direct Elasticities		Cross-Elasticities	
	Auto Line-Haul	Auto Out-of-Pocket	Transit Line-Haul	Transit Excess
Work	-.494	-.071	.138	0
Shopping	-.878	-1.65	0	0

Trip Purpose	Transit Trips			
	Direct Elasticities		Cross-Elasticities	
	Transit Line-Haul	Transit Excess	Auto Line-Haul	Auto Out-of-Pocket
Work	-.09	-.100	0	0
Shopping	-.323[a]		0	0

[a]The available shopping transit trip sample was unsuitable for estimating elasticities for the disaggregated cost components.

and transit work and shopping demand equations from which these elasticities were computed are given in Appendix B. The elasticities in the tables were calculated at the mean value of the variables. In general, the elasticities vary depending on the levels of the variables because the variables are usually expressed in both linear and logarithmic form.

Let us first consider travel time. The results indicate that for auto work trips demand is inelastic with respect to auto in-vehicle time, while auto shopping trips are unitary elastic with respect to auto in-vehicle time. This result is not surprising, since the greater urgency of the work trip would lead one to expect the elasticity of demand for work trips to be less than that for shopping trips.

On the other hand, the elasticities of demand for both work and shopping trips with respect to out-of-vehicle times are nearly identical and substantially greater than the in-vehicle time elasticities. This result lends credence to the generally accepted hypothesis (although generally disregarded in extant models) that out-of-vehicle times are more onerous than in-vehicle times. This phenomenon helps to explain the popularity of the suburban industrial parks and shopping centers, where workers or shoppers can park near their final destinations.

Keeping in mind the problems of the transit data used to estimate the parameter of the system, it would appear that auto work trips are slightly sensitive to transit excess times, i.e., the time required to get to and from the transit system, to wait, or to transfer. All the other cross-elasticities are either zero or nearly so. This indicates that transit travel times do not strongly influence the amount of auto travel, and that the use of the auto mode is more a result of socioeconomic characteristics than of the comparative travel times by transit. (The zero values for these cross-elasticities should not be taken literally, of course. They are zero because the constraints were binding, not because they were estimated to be zero. Thus, they should be interpreted as a lack of empirical evidence in the sample of a positive cross-elasticity rather than as literally zero.)

These cross-elasticity estimates indicate that there is not much promise for reducing auto congestion by improving transit service. The results further indicate that improvement in transit excess travel time will be more consequential in this regard than improvement in transit line-haul times. Of course, the effects of major technological or organizational changes in the transit system cannot be readily inferred from the model as it has been estimated, but the magnitudes discovered may be significant at least for the direction of further research.

If we consider the effects of travel times on transit demand, we find the demand to be relatively inelastic with respect to the transit time components analyzed and, as with the auto results, the effect of excess time is substantially more pronounced than that of line-haul time. Unfortunately, sample considerations made it impossible to disaggregate the transit time components for shopping trips.

In the case of transit travel, all the cross-elasticities with respect to the auto time components turned out to be zero. This result is generally symmetric with the time cross-elasticities in the auto equations and reinforces our observation that the choice of mode is determined more by the socioeconomic characteristics of the traveler than by comparative travel times.

Let us now turn to the effects of travel costs on demand (Table 3). For auto trips, the effect of costs on travel demand appears to be substantially different for the two trip purposes. The demands for both work and shopping auto trips are inelastic with respect to line-haul travel costs (essentially the operating costs of an automobile), but work trips are much more inelastic than shopping trips with respect to this cost component.

When we examine the effect of out-of-pocket expenses (parking and tolls), the difference between the two trip purposes is far more pronounced. The demand for shopping trips is highly elastic with respect to out-of-pocket costs, while the demand for work trips is almost totally inelastic with respect to such costs. These results have some very interesting implications for evaluating an increase in tolls as a means of reducing congestion on a bridge or tunnel. The low elasticity for auto work trips suggests that an increase in tolls would have little effect on morning peak traffic because most of these trips are work trips. If the real problem is the afternoon peak, however, a toll increase may substantially reduce congestion because many of these trips are shopping trips.

It is interesting that shopping trips are consistently more sensitive than work trips to changes in the time and cost of auto travel.

The cross-elasticities of demand for auto trips with respect to transit cost components are, for all practical purposes, zero. We would not place a great deal of significance on the small value of the cross-elasticity with respect to transit line-haul costs for work trips.

Finally, the elasticities of demand with respect to costs for transit trips are highly inelastic and no cross-elasticities appear. This indicates that a decrease in transit fares would not substantially increase ridership and would only add to transit revenue difficulties. On the other hand, it implies that a fare increase would increase revenues because it would cause a less than proportionate drop in ridership.

In the preceding discussion we have drawn a variety of inferences about travel behavior. It is worth noting that it did not require extensive computer simulation to develop these observations; rather they were drawn directly from the model parameters. Many additional inferences about travel behavior could be made, but those already presented should be enough to illustrate the richness of the model in evaluating policy decisions, which is the primary purpose of this presentation. Perhaps the most important finding of the empirical results, however, is the lack of evidence of significant cross-relationships between auto and transit demands. The cross-elasticities for both time and cost are zero for almost all components, implying that socioeconomic factors rather than transportation system characteristics are the principal determinants of modal choice. All of these conclusions are, of course, subject to the qualifications stated earlier regarding the transit data and the sample, as well as to the statistical reliability of the estimates.

STATISTICAL RELIABILITY

In comparing the model presented here with those currently in use, some discussion is in order regarding the statistical reliability of the estimates. In particular, we often look at measures of goodness of fit such as the estimated coefficient of multiple determination (R^2) as an indication of the degree of success in explaining the variations in traffic movements in the base data. We are accustomed to finding very high levels of R^2 for trip generation and attraction equations, suggesting that a high proportion of traffic movements have been explained, but such levels may be extremely deceptive when our interest is in the origin/destination pattern of trips.

In our model the values of R^2 are substantially lower than those generally reported—the values in Table 2 range from 0.35 to 0.63. In comparing these correlation statistics with those generally reported, however, it is necessary to recognize that our results show the percentage of zone-to-zone traffic explained whereas the correlation statistics reported for conventional models relate only to the number of trips leaving or arriving in a zone. It is obviously more difficult to predict interzonal movements than the total number of trips leaving or arriving in a zone. Therefore, lower values of R^2 for our model are not surprising. It should also be pointed out that the values of R^2 obtained with these models are not unusual in economic cross section analysis.

It is not unreasonable to believe that if values of R^2 were obtained for zone-to-zone trips for the existing models, they would be of lower magnitude than those found in our study, particularly if corrections are made for the number of degrees of freedom. The data used in this study, though not very satisfactory, are no worse than those used in other traffic demand studies and there is reason to believe they were used at least as efficiently in our model as they have been used in other demand models. This suggests that the amount of uncertainty in the estimates of interzonal traffic flows in the existing studies may be substantially higher than has generally been recognized.

Some of the high residual variability is likely to be due to inadequacies of the available data and to errors in specifying the model. As was pointed out earlier, readily available data had to be used and these data had not been compiled for use in estimating this type of model, and some variables considered important were not available. The heterogeneous transit trip was the most severe problem of the analysis. We anticipate that the home interview studies do provide a sound data base for the initial exploration of these models, but should be compiled somewhat differently for this application. When the testing opportunities of readily available data have been exhausted, some revision in the data collection process may be necessary to improve the estimates. Such revision should be premised on testing the hypotheses of the model and improving the quantitative estimates of the policy-oriented relationships.

While many mathematical forms of the model were tested in our empirical analysis, time did not permit an exhaustive study of these forms. Some revision of the form may also be useful in improving the results.

The high level of residual error in estimating the total choice mechanism (as opposed to a single aspect) should be regarded as a danger signal by the planner. The result implies high uncertainty in our predictions of the effects of changes in the transportation system. When account is taken of sampling errors and errors in predicting independent variables, in addition to the generally low correlation statistics, it is clear that the uncertainty in predicting origin and destination traffic movements is very great indeed. The planner must therefore be extremely cautious in his decisions and explicitly recognize that his evaluations are subject to this uncertainty.

ACKNOWLEDGMENTS

The model presented in this paper was developed by Charles River Associates as subcontractor to Peat, Marwick, Livingston & Co. as part of a study sponsored by the California Division of Bay Toll Crossings. The data used in estimating the model were provided by Peat, Marwick, Livingston & Co.

The authors wish to thank Martin Wohl of the RAND Corporation, Paul Roberts of Harvard University, and Daniel Brand of Peat, Marwick, Livingston & Co. for their comments and suggestions during the course of the original research, which helped to clarify many of the concepts in this paper. The responsibility for any errors, however, rests solely with the authors.

REFERENCES

1. Stigler, G. J. The Theory of Price. Macmillan, New York, 1952.
2. Demand for Intercity Passenger Travel in the Washington-Boston Corridor, Part V. Systems Analysis and Research Corp., Boston.
3. Griliches, Zvi. The Analysis of Specification Errors. Department of Economics, University of Chicago (mimeograph).

4. A Model of Urban Passenger Travel Demand in the San Francisco Metropolitan Area. Charles River Associates, 1967 (prepared for Peat, Marwick, Livingston & Co. in conjunction with their study for the California Division of Bay Toll Crossings).
5. Kraft, Gerald, and Wohl, Martin. New Directions for Passenger Demand Analysis and Forecasting. Transportation Research, Vol. 1, No. 3, Pergamon Press, 1967.

Appendix A

DEFINITIONS OF TIME AND COST VARIABLES USED IN THE ESTIMATION OF THE MODEL

Time Variables

Transit line-haul time = in-vehicle time spent in the principal transit mode.

Transit excess time = travel spent outside the principal transit mode. It includes time spent in auto, feeder bus, or walking to or from the principal transit mode. It is made up of the following components:

Travel time from origin to principal mode first station;
Waiting time at principal transit mode station;
Transfer time; and
Travel time from last principal mode station to destination.

Auto in-vehicle time = line-haul time from zone centroid to zone centroid plus parking time.

Auto out-of-vehicle time = walk-to-car time at origin of trip and time spent in walk from parking place to destination.

Cost Variables

Transit line-haul cost = fare paid on the principal transit mode.

Transit excess cost = money spent traveling to and from the principal transit mode.

Auto line-haul cost = operating cost of driving an automobile from the zone of origin to the zone of destination.

Auto out-of-pocket costs = tolls plus parking charges.

Appendix B

Tables B-1 through B-4 give the auto and transit work and shopping trip equations from which the elasticities in Tables 2 and 3 were computed. Table B-5 gives the means of the system variables.

TABLE B-1

AUTO WORK TRIPS

$$\text{Dependent variable} = \left\{ \frac{\text{Number of directed work round trips by auto}}{\left[\begin{matrix}\text{Employed labor}\\ \text{force in zone of}\\ \text{residence}\end{matrix}\right] \times \left[\begin{matrix}\text{Employment in zone of}\\ \text{work as a proportion of}\\ \text{total employment in the}\\ \text{region}\end{matrix}\right]} \right\}$$

Independent variables:

Description	Coefficient
Constant	-31.0250
In-vehicle time—auto	*
ln (In-vehicle time—auto)	-1.7973
Out-of-vehicle time—auto	*
ln (Out-of-vehicle time—auto)	-3.1387
Line-haul time—transit	*
ln (line-haul time—transit)	*
Excess time—transit	*
ln (Excess time—transit)	.8153
Line-haul cost—auto	*
ln (Line-haul cost—auto)	-1.0793
Out-of-pocket costs—auto	*
ln (Out-of-pocket costs—auto)	-.1552
Line-haul cost—transit	*
ln (Line-haul costs—transit)	.3034
Excess cost—transit	*
ln (Excess cost—transit)	*
Median income of households and unrelated individuals in zone of residence	.0020
ln (Median income of households and unrelated individuals in zone of residence)	6.1168
Number of cars per capita in zone of residence	13.2677
ln (Number of cars per capita in zone of residence)	.0270
Employment density in zone of work	-.0063

Form of the model:

$$\frac{N}{Y} = \alpha X + \beta \ln X$$

$$\eta_X = \frac{\alpha X + \beta}{N} Y$$

where

η_X = elasticity of demand with respect to variable X

N = number of trips

X = independent variables

$$Y = \left[\begin{matrix}\text{employed labor}\\ \text{force in zone of}\\ \text{residence}\end{matrix}\right] \times \left[\begin{matrix}\text{employment in zone of work}\\ \text{as a proportion of total}\\ \text{employment in the region}\end{matrix}\right]$$

α, β = estimated parameters

*Variables introduced in the model which take a zero coefficient due to the use of the constrained regression technique.

TABLE B-2

AUTO SHOPPING TRIPS

Dependent variable = ln (Number of directed shopping round trips by auto)

Independent variables:

Description	Coefficient
Constant	-2.733324
In-vehicle time—auto	-.024824
ln (In-vehicle time—auto)	-.081710
Out-of-vehicle time—auto	*
ln (Out-of-vehicle time—auto)	-1.439808
Line-haul time—transit	*
ln (Line-haul time—transit)	.095003
Excess time—transit	*
ln (Excess time—transit)	*
Line-haul cost—auto	*
ln (Line-haul cost—auto)	-.878061
Out-of-pocket cost—auto	-.050591
ln (Out-of-pocket cost—auto)	-.853097
Line-haul cost—transit	*
ln (Line-haul cost—transit)	*
Excess cost—transit	*
ln (Excess cost—transit)	*
ln (Number of households in zone of residence)	1.000000
Number of persons per household in zone of residence	.583934
ln (Number of persons per households in zone of residence)	-3.048188
Median income of households and unrelated individuals	-.000029
ln (Median income of households and unrelated individuals)	.304834
Number of cars per capita in zone of residence	15.303761
ln (Number of cars per capita in zone of residence)	-2.341933
Density of employment in retail trade in zone of destination	.086956
ln (Density of employment in retail trade in zone of destination)	-.759571
ln (Employment in retail trade in zone of destination as a proportion of total regional employment in retail trade)	1.000000

Form of the model:

$$N = X^{\alpha} e^{\beta X}$$

$$\eta_X = \alpha + \beta X$$

where

η_X = elasticity of demand with respect to variable X

N = number of trips

X = independent variable

α, β = estimated parameters

*Variables introduced in the model which take a zero coefficient due to the use of the constrained regression technique.

TABLE B-3

TRANSIT WORK TRIPS

Dependent variable = ln (Number of directed work round trips by transit)

Independent variables:

Description	Coefficient
Constant	-12.158232
In-vehicle time—auto	*
ln (In-vehicle time—auto)	*
Out-of-vehicle time—auto	*
ln (Out-of-vehicle time—auto)	*
Line-haul time—transit	-0.005843
ln (Line-haul time—transit)	-0.190862
Excess time—transit	-.025288
ln (Excess time—transit)	.462262
Line-haul cost—auto	*
ln (Line-haul cost—auto)	*
Out-of-pocket costs—auto	*
ln (Out-of-pocket costs—auto)	*
Line-haul cost—transit	-0.002362
ln (Line-haul cost—transit)	.036214
Excess cost—transit	-.005095
ln (Excess cost—transit)	*
Number of cars per capita in zone of residence	1.777146
ln (Number of cars per capita in zone of residence)	-1.163856
ln (Median income of households and unrelated individuals in zone of residence)	1.144006
ln (Employed labor force in zone of residence)	1.000000
ln (Employment in zone of work as a proportion of total employment in region)	1.000000

Form of the model:

$$N = X^{\alpha} e^{\beta X}$$

$$\eta_X = \alpha + \beta X$$

where

η_X = elasticity of demand with respect to variable X

N = number of trips

X = independent variables

α, β = estimated parameters

*Variables introduced in the model which take a zero coefficient due to the use of the constrained regression technique.

TABLE B-4

TRANSIT SHOPPING TRIPS

$$\text{Dependent variable} = \ln \left\{ \frac{\text{Number of directed shopping round trips by transit}}{\left[\begin{array}{l}\text{Number of households}\\ \text{in zone of residence}\end{array}\right] \times \left[\begin{array}{l}\text{Employment in retail trade}\\ \text{in zone of destination as a}\\ \text{proportion of employment}\\ \text{in retail trade in region}\end{array}\right]} \right\}$$

Independent variables:

Description	Coefficient
Constant	-1.976884
ln (Total aggregated time—transit)	-.593240
ln (Total aggregated cost—transit)	-.323692
ln (Number of persons per household)	2.483299
ln (Median income of households in zone of residence)	-.048626
ln (Density of employment in retail trade)	.030759
ln (Employment in personal business activities in zone of destination as a proportion of employment in personal business in region)	-.739325

Form of the model: log/log

$$\frac{N}{Y} = X^{\alpha}$$

$$\eta_X = \alpha$$

where

η_X = elasticity of demand with respect to variable X

N = number of trips

X = independent variable

$$Y = \left[\begin{array}{l}\text{number of households}\\ \text{in zone of residence}\end{array}\right] \times \left[\begin{array}{l}\text{employment in retail trade in zone of}\\ \text{destination as a proportion of total}\\ \text{employment in retail trade in region}\end{array}\right]$$

α = estimated parameter

TABLE B-5

MEANS OF SYSTEM CHARACTERISTIC VARIABLES FOR INTERZONAL TRIPS IN THE BOSTON AREA

Description of Variable	Work Trips*		Shopping Trips
	Transit Sample	Auto Sample	(Single Sample)
Line-haul time—transit (minutes)	34.69	35.24	27.13
Excess time—transit (minutes)	46.84	52.58	47.76
In-vehicle time—auto (minutes)	54.43	49.73	37.15
Out-of-vehicle time—auto (minutes)	5.40	5.15	5.44
Line-haul cost—transit (cents)	56.06	51.69	48.95
Excess cost—transit (cents)	20.01	22.58	15.43
Line-haul cost—auto (cents)	36.88	34.32	20.70
Out-of-pocket cost—auto (cents)	18.31	8.35	16.35

*Separate samples were used for work trips by auto and by transit.

A Theoretical Model for Determination of Expressway Usage in a Uniform Region

STEARNS CASWELL, Planning and Research Bureau, New York State Department of Transportation

•THE purpose of this paper is to present the solution to the problem of computing the theoretical "usage," in terms of mean trip density, of certain high-speed facilities that are placed in a given region containing trip origins and destinations. Certain results based on this solution are also presented. The types of facilities considered are very special, and the assumptions under which the problem is solved are quite restrictive. However, neither the specialization nor the restrictions should completely negate the applicability of the results to certain phases of the transportation planning process.

As the areas in which formal transportation studies are undertaken become larger and more complex, the limits of present planning tools become more clearly defined. These tools rest heavily upon computer simulation techniques—techniques which are unwieldy, time-consuming, and expensive at best, and which are not (at present) wholly applicable in the super-regions for which transportation planning is being attempted.

There are problem areas for which the computer is inadequate for solutions, at least at present. Among them are the basic ones of formulating alternative plans to be tested, of settling upon a desirable network geometry (to say nothing about an optimal geometry), of delineating a reasonable range of facility spacings to be more closely examined, and of solving other problems involving far too many combinations to be dealt with by methods of exhaustion. Now the machine is obviously necessary for any completely practical solution to these hypercomplex problems. The number of cases in which it is not sufficient may possibly be reduced by taking a closer look at the concepts involved in the simulation models we use. One way of doing this is to examine the consequences of the hypotheses of a model in various hypothetical control situations, with the hope that our insight into the relationships implied by the model may be increased. It is toward this gain in understanding that this paper is directed, as have been other papers in recent years (see References).

While the results of these efforts may never be used in actually locating a highway or transit line, they should give some insight into the behavior of trip distribution functions and the relationship between the various parameters inherent in trip-making patterns. It is possible that guidelines for computer model development will be found in them and even that certain broad planning decisions may be based on them.

The basic types of expressway "networks" considered here are (a) an isolated expressway with unlimited access, (b) a sheaf of parallel expressways with unlimited access, (c) an isolated pair of parallel expressways with unlimited access, and (d) an isolated expressway with limited access. The usage of these types of expressways is established rigorously under the following assumptions:

1. The given region is of constant (vehicle) trip density;
2. A constant speed is allowed on the expressway(s);
3. Unfettered movement at a constant speed is allowed throughout the remainder of the region, except that trip distance is measured as right-angle distance and all movement is at right angles;
4. A trip will follow the most economical route at all times: if costs are equal for a non-expressway route and an expressway route, then the latter will be used; if costs

Paper sponsored by Committee on Transportation System Evaluation and presented at the 47th Annual Meeting.

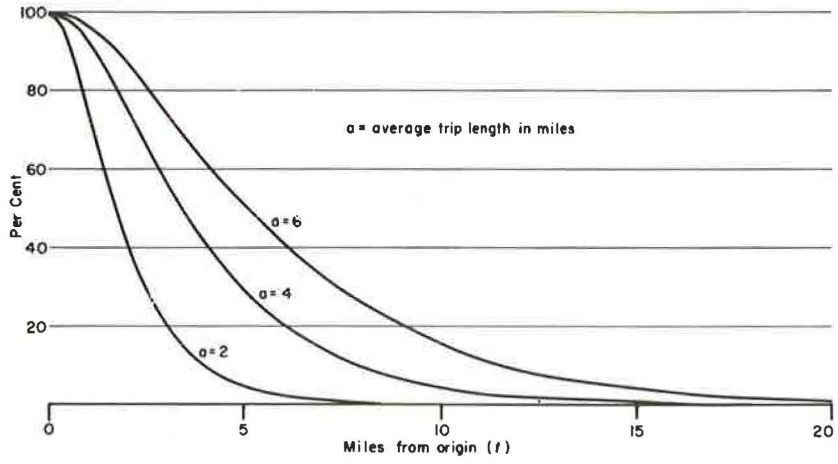


Figure 1. Percent of trips undelivered at t miles from origin.

are equal for two or more expressway routes, then the trips involved are allocated in equal proportions to the expressways concerned; and

5. Trip length is u + v, where u and v are random variables with the joint density function $k^2 \exp[-k(u+v)]$. This assumption implies that the mean trip length is 2/k and that u and v each have the (marginal) density function k exp (-kt) and mean 1/k.

Perhaps the most restrictive of these assumptions are numbers 1 and 5. The assumption of constant trip density is necessary for my peculiar attack, because it assures the requisite uniformity of volume on the various expressway networks. However, there is nothing intrinsically valuable about the method. If another can be devised which will allow alterations in assumption 1, a much more practical set of solutions will result. Of particular interest would be a solution for a region containing a finite number of "point" generators of trips.

There are a priori arguments that can be devised for the use of the trip length frequency function defined in assumption 5 (one can be found in Ref. 1). I choose the argument that this is one of the very few functions amenable to my purpose that even crudely describes available trip data. Examples of this trip length function are given in Figure 1 for various values of k.

A presentation of results and a discussion of some of the applications of these results are given in the next section. There are many implications possible which are only hinted at here (or not mentioned at all). Because a paper of this type is marginally digestible at best, it seemed discreet to hold the list of ramifications to a minimum.

RESULTS AND APPLICATIONS

There are several parameters on which the volumes for all expressway networks considered here depend—namely, expressway speed, arterial speed, average trip length for the region, density for the region, value of time, and operating and accident costs at expressway and arterial speeds. In addition, there are special parameters on which the volumes for some of the special types of networks depend. For example, in the case of an isolated pair of expressways, the distance between them is an explicit parameter, as is the distance between the expressways in a parallel sheaf. In the case of the isolated expressway with limited access, the ramp spacing is a parameter. Each of these parameters is identified in a later section.

Tables have been generated which display the various volumes. Certain data from these tables are presented here in graphical form, accompanied by a discussion. In addition, some obvious applications of these data are discussed and illustrated.

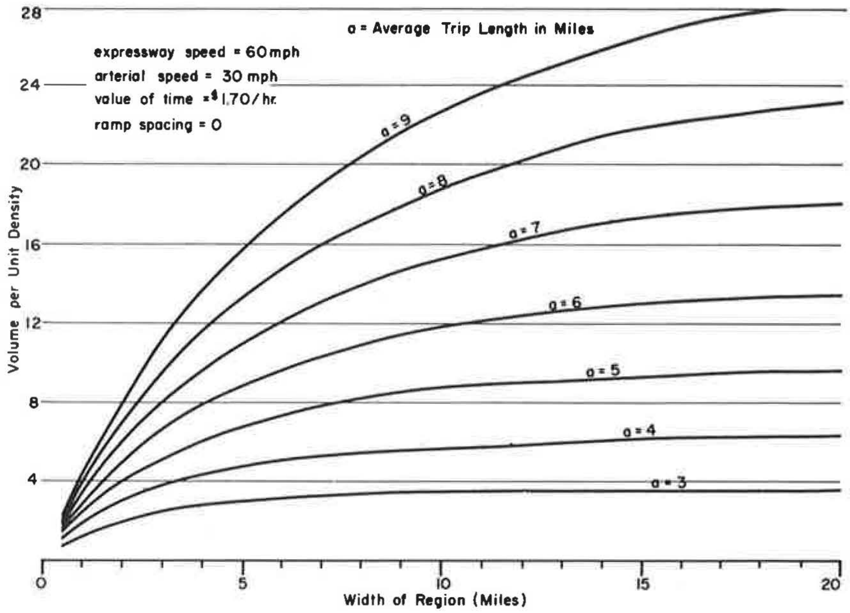


Figure 2. Isolated expressway volume vs width of region from which trips are drawn.

Volume on Isolated Expressway vs Width of Region

Under the assumptions listed previously, the volume carried by an isolated expressway with unlimited access which draws trips from a band of trip origins extending c miles on either side of the expressway is given by

$$
{}_cV = (D/4)a^2 \Big([(R + 2)/R] - \{ [4R^2 - 7R + 2 + 4R(2R - 1)(R - 1)(c/a)] / [R(2R - 1)^2] \} \exp(-4Rc/a) - [4R^2/(2R - 1)^2] \exp(-2c/a) \Big)
$$

where

2c = total width of region,
D = trip density for the region,
a = average trip length for the region,
$R = C_a/\left(C_a - C_e\right)$,
C_a = trip cost/mile on arterials, and
C_e = trip cost/mile on the expressway.

Another interpretation of this formula is that it gives the volumes carried by each member of a sheaf of parallel facilities spaced at 2c miles. In this situation, for each trip that originates in the c band for facility A but uses facility B, there will be a trip which originates in the c band for facility B but uses facility A. Thus, the formula gives the usage of each facility in the sheaf.

The curves in Figure 2 show the usage for both cases. We can identify the horizontal axis with the total width of the region for the first interpretation or with the expressway spacing for the second interpretation. In either case, the vertical axis represents mean trip density on the expressway(s).

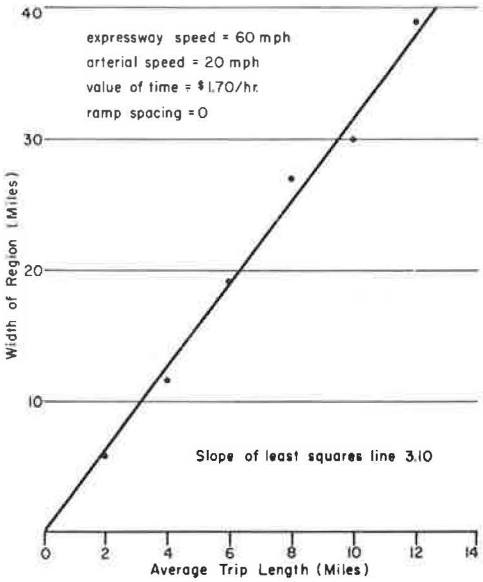


Figure 3. Width of region from which 95 percent of maximum expressway volume is drawn vs average trip length.

As c becomes infinite, ${}_cV$ approaches ${}_\infty V = (D/4)a^2\,[(R + 2)/R]$ (which each curve in Figure 2 approaches as an asymptote for the appropriate value of a).

Practical Meaning of "Isolated"

The term "isolated" has been used throughout this report to describe expressways. This term has been used in the sense that an expressway (expressway network) is isolated if no other expressway (expressway network) is near enough to compete with it for trips. In our theoretical world, this would imply that an expressway is isolated only if it is in an infinite region with no other expressways.

In order to have some practical measure of "near enough," the following calculations were made. The volume (${}_\infty V$) which the expressway would carry if trips were drawn from an unbounded region was considered maximum; 95 percent of this maximum volume was calculated for each value of average trip length, after which the value of 2c (width of the region from which trips are drawn to the expressway) giving this proportion was found. (A linear interpolation was used, even though this is not strictly justifiable here. The error introduced is no more than the same type of interpolation would give in an ordinary table of logarithms.) This procedure was carried out for arterial speeds of 10, 20, 30, 40, and 50 miles per hour. Figure 3 shows the points that resulted for an arterial speed of 20 miles per hour. Because the relationship between 2c and a seemed essentially linear, a least squares line was fitted and included in Figure 3. The same was done for arterial speeds of 10, 30, 40, and 50 miles per hour. The slopes of the least square lines are 3.27, 3.05, 3.14, and 3.22 respectively, with a mean of 3.16.

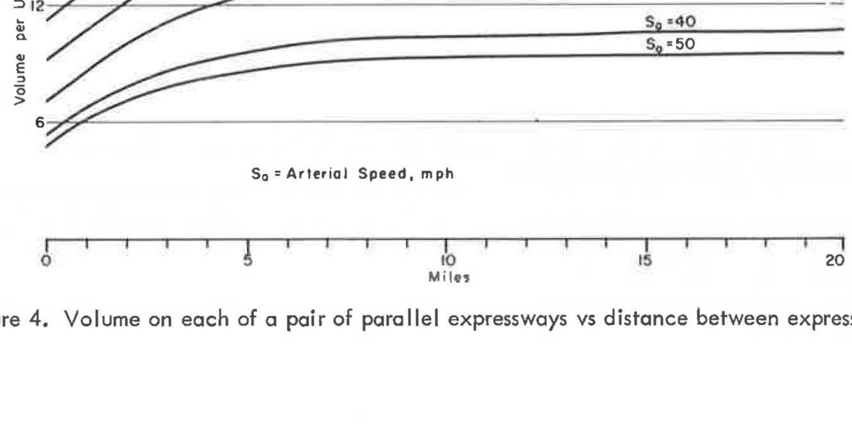


Figure 4. Volume on each of a pair of parallel expressways vs distance between expressways.

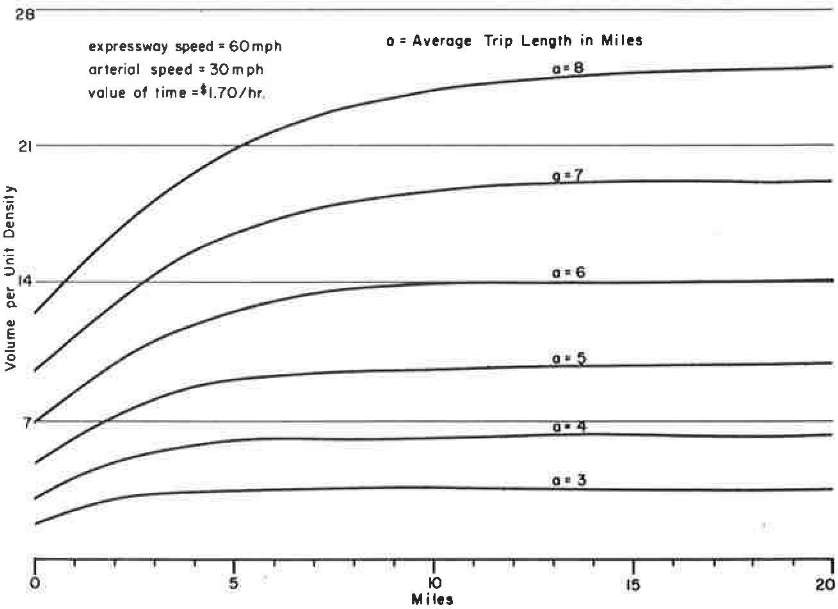


Figure 5. Volume on each of a pair of parallel expressways vs distance between expressways.

Thus, for all practical purposes an expressway could be called "isolated" in the sense needed here if there is no other expressway within approximately 3a, where a is the average trip length for the region.

Expressway Volume vs Expressway Spacing

Figures 4 and 5 show the volume on each of an isolated pair of expressways as a function of the distance between them. In Figure 4, the average trip length has been held constant at 6 miles, while in Figure 5 the arterial speed has been fixed at 30 miles per hour. The formula in this case is

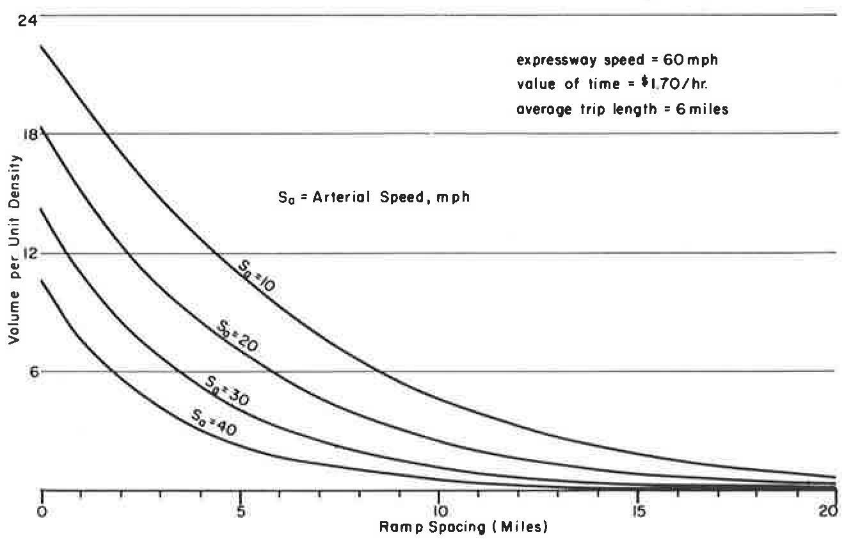


Figure 6. Expressway volume vs ramp spacing.

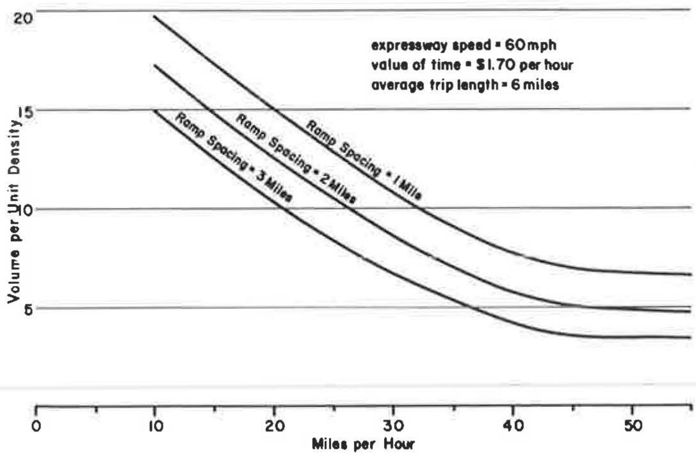


Figure 7. Expressway volume vs arterial speed.

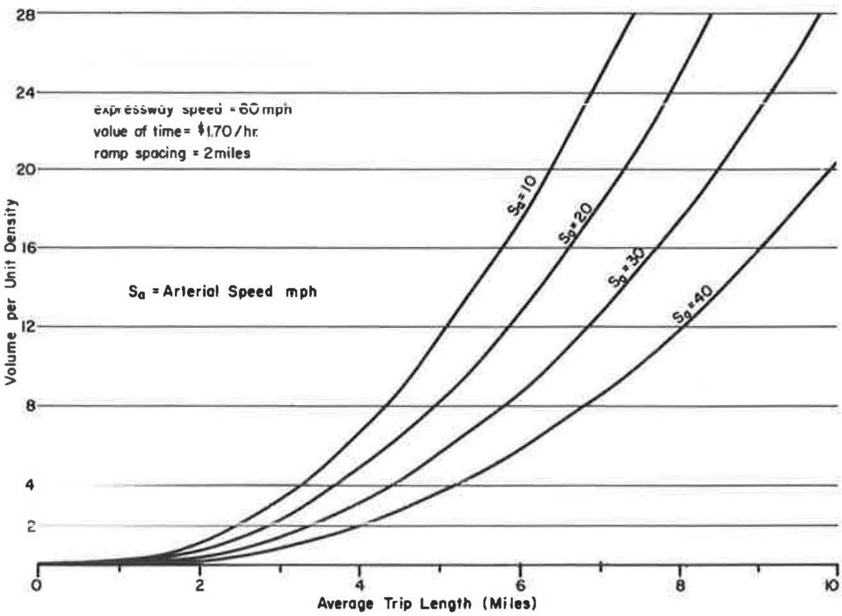


Figure 8. Expressway volume vs average trip length.

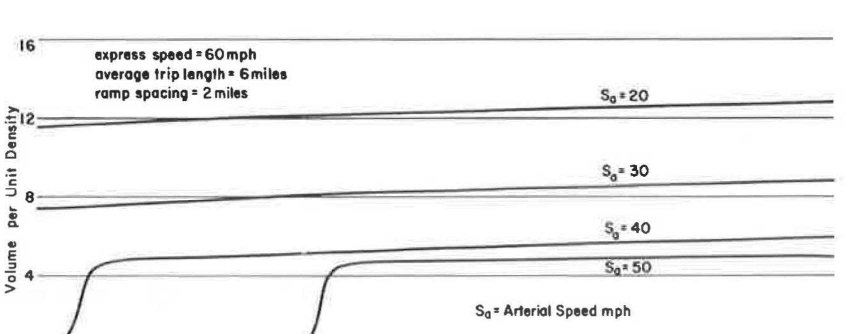


Figure 9. Expressway volume vs value of time.

$$_dV = (D/4)a^2 \Big([(R + 2)/R] - [R^2/2(R - 1)^2] \exp(-2d/a) + \{ [3R - 2 + 2R(R - 1)d/a]/[2R(R - 1)^2] \} \exp(-2Rd/a) \Big)$$

As the distance between the two expressways becomes larger, the volume on each approaches as a maximum the volume ($_\infty V$) which a single isolated expressway would carry for the same choice of parameters. As d becomes smaller, the volume on each of the pair approaches as a minimum one-half of the maximum volume. Fortunately, this behavior is in agreement with intuition.

Effect of Various Parameters on Expressway Volumes

Figures 6 through 9 show the relative dependence of expressway volumes on the various parameters involved: ramp spacing, arterial speed, average trip length, and value of time. The graphs should be self-explanatory. It is worth noting that average trip length is a particularly important parameter, but that volumes are not especially sensitive to the value of time (assuming, of course, that the value is high enough for the expressway to be used at all).

Travel Cost Savings Due to Construction of Expressway

Figure 10 shows the savings in travel cost per trip due to the construction of an expressway in a region. The savings are graphed as a function of ramp spacing. The formula used to compute the savings S is easily derived:

$$S = C_a (Q \div D)/(wR)$$

where

C_a = travel cost per mile at the given arterial speed,
$Q \div D$ = volume of trips past a point on the expressway per unit density of vehicle trips in the region,
w = width of the region (region extends w/2 miles on either side of the expressway), and
R = $C_a/(C_a - C_e)$, where C_e is the travel cost per mile at the given expressway speed.

Because the volumes on the expressway were obtained under the assumption that trips are drawn from an unbounded region, it was necessary to choose w large enough to guarantee that most (approximately 95 percent) of the volume comes from the bounded region under consideration. It was decided to use w = 18 miles because an average trip length of 6 miles was assumed for the region. The assumptions governing travel costs are given later.

Expressway Construction Cost

The construction costs used here are entirely hypothetical. Under no circumstances should the cost analysis given in this report be regarded as anything but a series of examples of the use of the theory in answering certain standard questions. In order to make the examples meaningful, an attempt was made to use cost data that are at least faintly realistic, but a much more nearly precise application of the ideas illustrated here could be made in a particular economic situation.

At any rate, the expressway costs used to obtain the curves in Figure 11 are based on the following rules of thumb:

Right-of-way cost = \$(300,000 + 100D)/mile, where D = number of (vehicular) trips/mile^2;

Main line cost = \$1,400,000/mile;

Interchange costs = \$750,000 per ramp for D = 10,000; \$1,000,000 per ramp for D = 20,000;

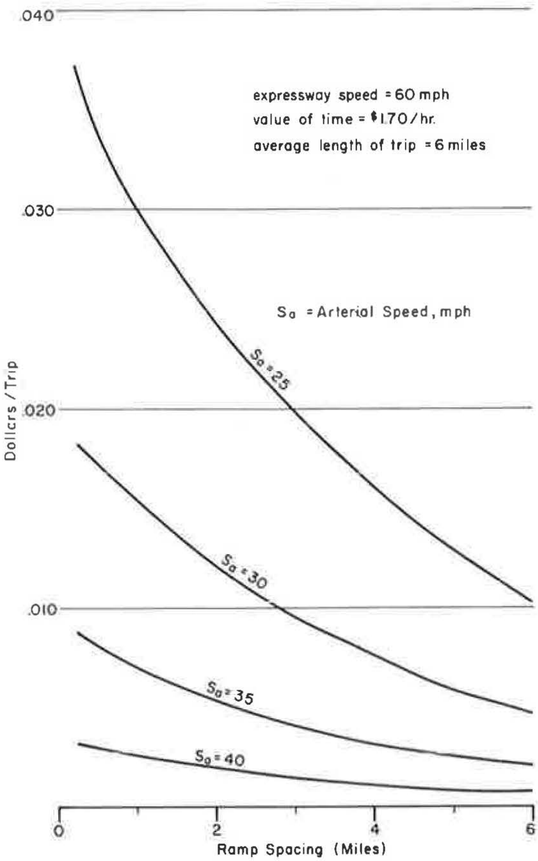


Figure 10. Savings per trip due to presence of expressway.

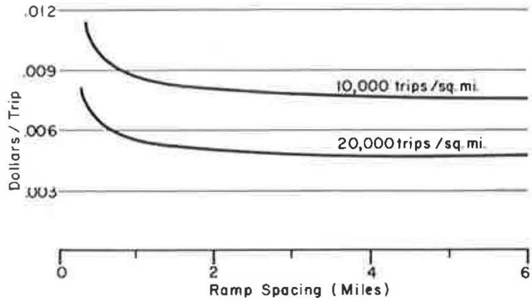


Figure 11. Expressway construction cost per trip vs ramp spacing.

Crossing cost = $1,250,000/mile (assuming 5 crossings per mile at $250,000 per crossing to maintain as closely as possible the "unfettered" movement required by the theory);
Capital recovery factor = .110127; and
Number of days per year = 340.

Benefit/Cost vs Ramp Spacing

Figures 12 and 13 indicate that, from a benefit/cost standpoint, an optimum ramp spacing is determined that is somewhat greater than the minimum spacing dictated by flow theory considerations. Actually, the optimum ramp spacings suggested here are probably too small because no form of capacity restraint has been incorporated in the derivation of the expressway volumes. Because of the crudeness of the cost criteria, the lack of congestion considerations, and the overall restrictiveness of the hypotheses under which the expressway volumes are derived, any estimates of optimality made herein must be of the roughest sort. On the other hand, an analytical crutch is sometimes better to lean on than a less synthetic variant; at least one is propped up enough to identify and examine the assumptions involved.

Despite the foregoing disclaimer, I find it irresistible to point out that the optimum ramp spacing, whatever it may be in absolute terms, is very definite for low arterial speeds. Furthermore, this optimum seems to regress and to become less critical as arterial speeds increase.

Benefit/Cost vs Expressway Spacing

In Figures 14, 15, 16, 17, and 18, average total trip cost was chosen as the criterion for optimum expressway spacing. Trip cost was defined as the sum of the cost per trip of expressway construction and the travel cost per trip. The comments made in reference to Figures 10, 11, 12, and 13 apply to the cost data used here.

Construction costs were defined as follows:

Right-of-way cost = $(300,000 + 100D)/mile;
Main line cost = $1,400,000/mile;
Interchange cost = $3,000,000/mile for D = 10,000,
= $4,000,000/mile for D = 20,000;
Crossing cost = $1,250,000/mile;

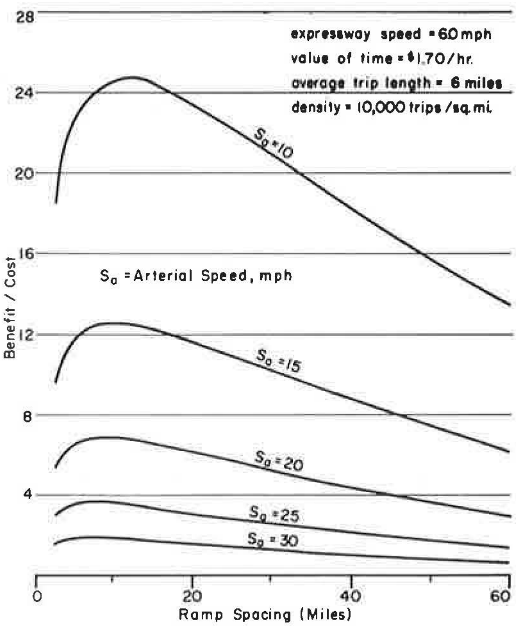


Figure 12. Benefit/cost vs ramp spacing.

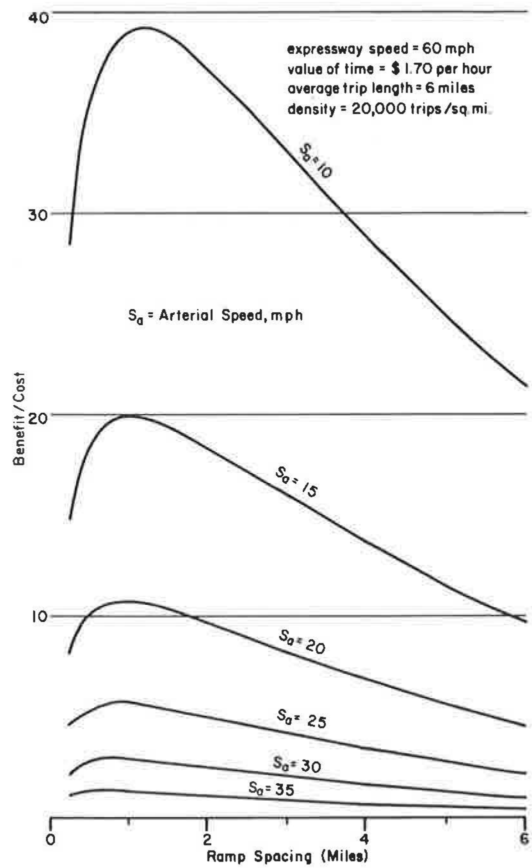


Figure 13. Benefit/cost vs ramp spacing.

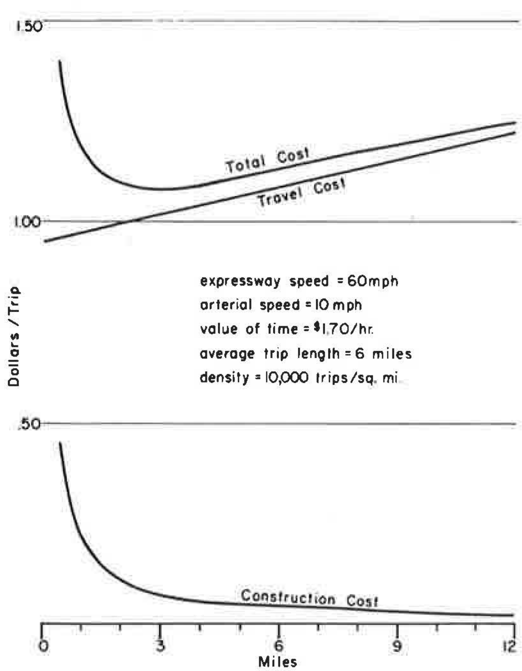


Figure 14. Trip cost vs distance between expressways.

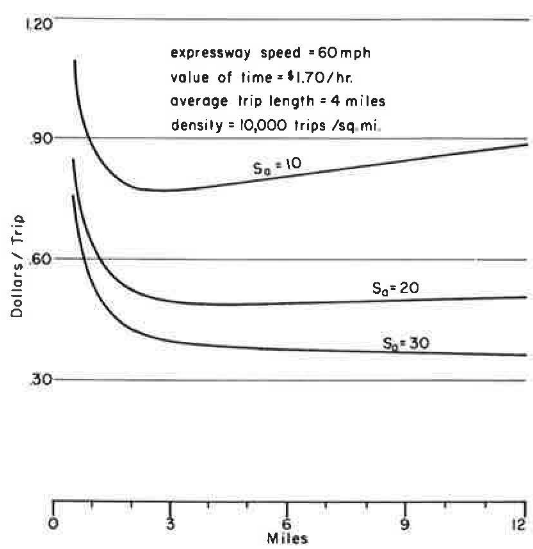


Figure 15. Trip cost vs distance between expressways.

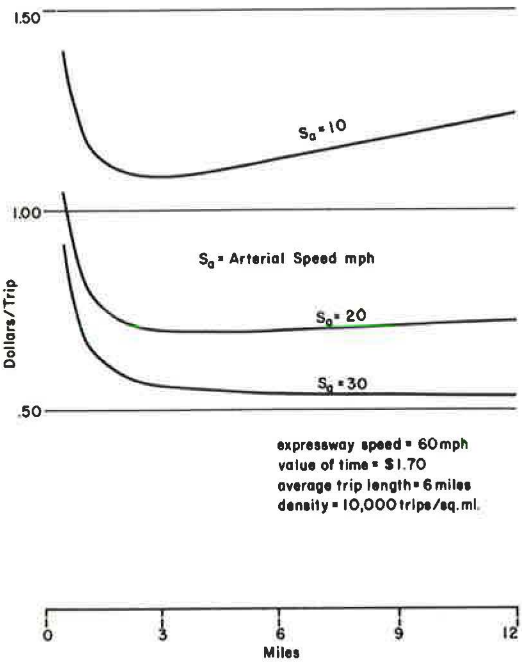


Figure 16. Trip cost vs distance between expressways.

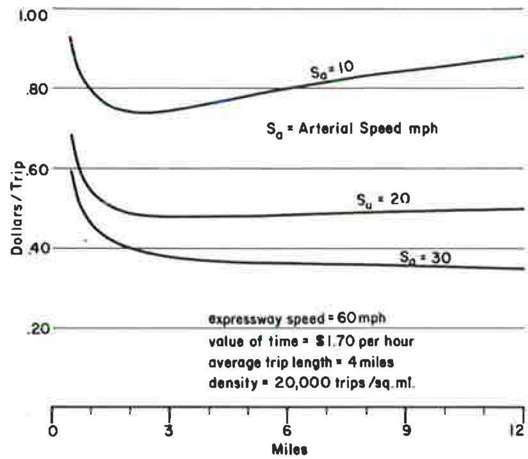


Figure 17. Trip cost vs distance between expressways.

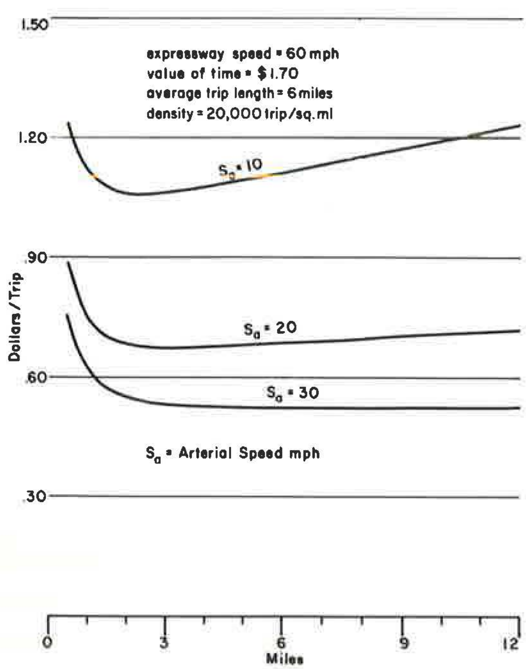


Figure 18. Trip cost vs distance between expressways.

Capital recovery factor = .110127; and
Number of days per year = 340

Travel costs were obtained from the formula

$$\text{Travel cost/trip} = C_a A - S$$

where A is the average trip length, C_a is the travel cost/mile at the given arterial speed, and S is as defined in the discussion for Figure 10, except that Q has been replaced by ${}_cV$.

The components of the total trip cost are shown in Figure 14, but were omitted from the other figures in order to show more clearly the dependence of trip cost on the various parameters.

Again, the crudeness of the assumptions makes the drawing of conclusions about optimum spacing from the graphs a highly questionable practice. However, the hint (obscure as it may be) given in Figures 17 and 18 that optimum spacing in many cases is not at all critical, is one of several in the graphs of this report that may help to frame some sensible questions and to urge further research in the analysis of transportation systems. If this is true, then the real purpose of an investigation of this sort has been realized.

DERIVATION OF FORMULAS

An Isolated Expressway With Unlimited Access

It is convenient to choose coordinate axes so that the expressway coincides with the x axis (see Fig. 19). Consider the element of area ΔA whose centroid is at (0,y), and assume that the trips originating from this element are concentrated at the centroid. Let the trip using the facility originate within c miles on either side of the facility (see Fig. 20). Each trip will be composed of a horizontal segment u and a vertical segment v, each of which can be considered a random variable of one dimension with distribution exp (-kt) and mean 1/k.

We shall classify trips according to the following scheme. Consider the world divided into two regions, 1 and 2 (see Fig. 21a). The trips that originate in region 1 can be typed as in Figure 21b. A trip in 1j will have its origin in region 1 and will be of type j. Note that trips in classes 21, 22, and 23 are mirror images, respectively, of classes 11, 12, and 13 with respect to the expressway. The volume contributed by trips in classes 2_j will be the same as the volume contributed by trips in classes 1_j, so that we may concentrate on the classes 1_j. These may be described analytically as follows:

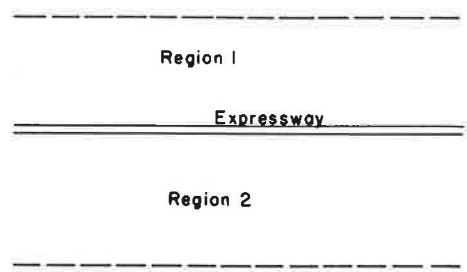


Figure 21a.

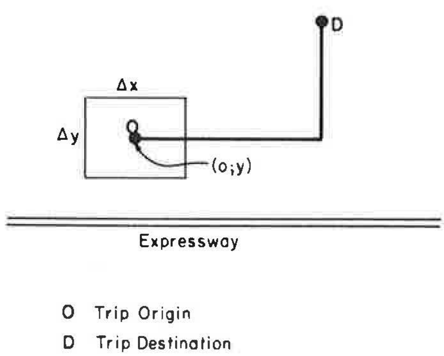


Figure 19.

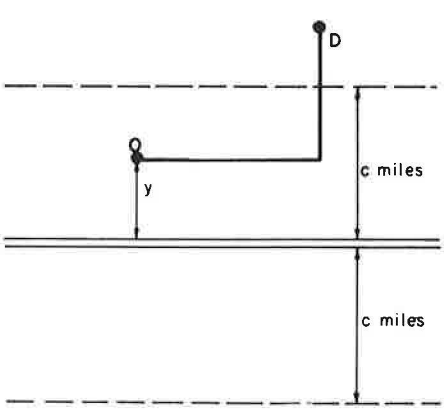


Figure 20.

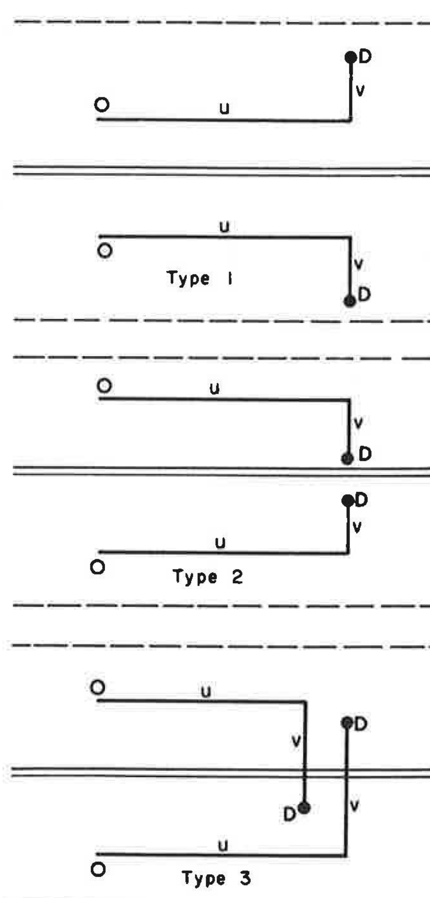


Figure 21b.

Class 11: D above 0

y in ${}_cY_{11} = [0;c]$
v in ${}_cV_{11} = [0;\infty]$

Class 12: d below 0

y in ${}_cY_{12} = [0;c]$
v in ${}_cV_{12} = [0;y]$

Class 13: d below 0

y in ${}_cY_{13} = [0;c]$
v in ${}_cV_{13} = [y;\infty]$

Let ${}_cP_{ij}(y) = \Pr[\text{trip is of class } ij]$ for a given y. Then

$$
\begin{aligned}
{}_cP_{11}(y) &= \Pr[D \text{ is above } 0]\Pr[0 \le v < \infty] = 1/2 \\
{}_cP_{12}(y) &= \Pr[D \text{ is below } 0]\Pr[0 \le v < y] = (1/2)[1 - \exp(-ky)] \\
{}_cP_{13}(y) &= \Pr[D \text{ is below } 0]\Pr[0 \le v < \infty] = (1/2)\exp(-ky)
\end{aligned}
$$

Let ${}_cQ_{ij}(y,v) = \Pr[\text{trip of class } ij \text{ will use expressway}]$ for a given y,v. Let C_a = trip cost per mile at arterial speed and C_e = trip cost per mile at expressway speed. (See note at the end of this paper for a discussion of trip cost as used here.) Then for trips of class 11, the total cost c_a for a non-expressway route is $uC_a + vC_a$ and the total cost c_e for the expressway route is $2yC_a + uC_e + vC_a$, so that $c_e \le c_a$ implies $u \ge 2Ry$, where $R = C_a/(C_a - C_e)$. Thus,

$$ {}_cQ_{11}(y,v) = \Pr[u \ge 2Ry] = \exp(-2Rky) $$

Similarly,

$$
\begin{aligned}
{}_cQ_{12}(y,v) &= \Pr[u \ge 2R(y - v)] = \exp(-2Rky)\exp(2Rkv) \\
{}_cQ_{13}(y,v) &= \Pr[u \ge 0] = 1
\end{aligned}
$$

Let ${}_cm_{ij}(y,v)$ = mean distance traveled on expressway by trip of class ij, given that the expressway is used.

We recall that, for a random variable x with the exponential distribution, the statement $\Pr[x \ge a + b \text{ given } x \ge b] = \Pr[x \ge a]$ obtains. Thus,

$$
\begin{aligned}
{}_cm_{11}(y,v) &= 1/k + 2Ry \\
{}_cm_{12}(y,v) &= (1/k + 2Ry) - 2Rv \\
{}_cm_{13}(y,v) &= 1/k
\end{aligned}
$$

Let ${}_cf_{ij}(y,v) = {}_cP_{ij}(y)\,{}_cQ_{ij}(y,v)\,{}_cm_{ij}(y,v)$, so that for each y, v, ${}_cf_{ij}(y,v)$ = VMT/trip of expressway for trips of class ij. We must find the mean ${}_c\bar{f}_{ij}(y)$ of ${}_cf_{ij}(y,v)$ with respect to v for each ij. Let ${}_cF_{ij}(t)$ be the density function for v for class ij. We have

$$
\begin{aligned}
{}_cF_{11}(t) &= k\exp(-kt),\ t \text{ in } {}_cT_{11} = [0;\infty] \\
{}_cF_{12}(t) &= \{1/[2\,{}_cP_{12}(y)]\}\,k\exp(-kt),\ t \text{ in } {}_cT_{12} = [0;y] \\
{}_cF_{13}(t) &= \{1/[2\,{}_cP_{13}(y)]\}\,k\exp(-kt),\ t \text{ in } {}_cT_{13} = [y;\infty]
\end{aligned}
$$

Then

$$ {}_c\bar{f}_{ij}(y_i) = \int_{{}_cT_{ij}} {}_cf_{ij}(y,t)\,{}_cF_{ij}(t)\,dt $$

so that

$$
\begin{aligned}
{}_c\bar{f}_{11}(y) &= [1/(2k)]\exp(-2Rky) + Ry\exp(-2Rky) \\
{}_c\bar{f}_{12}(y) &= \{(4R - 1)/[2(2R - 1)^2 k]\}[\exp(-ky) - \exp(-2Rky)] \\
&\quad - [R/(2R - 1)^2]\,y\exp(-2Rky)
\end{aligned}
$$

$$_c\bar{f}_{13}(y) = [1/(2k)] \exp(-ky)$$

Let

$$_cS_{ij} = \int_{_cY_{ij}} {_c\bar{f}_{ij}}(y)dy$$

so that $_cS_{ij}$ is the VMT/trip on the expressway for all trips of class ij entering the expressway in Δx. We have

$$_cS_{11} = 1/(2Rk^2) - [(1 + Rkc)/2Rk^2] \exp(-2Rkc)$$
$$_cS_{12} = 1/(2Rk^2) + \{[3R - 1 + R(2R - 1)kc]/[2R(2R - 1)^2k^2]\} \exp(-2Rkc)$$
$$_cS_{13} = 1/(2k^2) - [1/(2k^2)] \exp(-kc)$$

Let

$$\Delta_c V = 2k(_cS_{11} + {_cS_{12}} + {_cS_{13}})\Delta x,$$

so that $\Delta_c V$ = VMT on expressway for trips entering expressway in Δx. Then

$$_cV = \int_0^1 d_cV = \text{mean trip density on the expressway}$$

We have

$$_cV = (D/4)a^2\Big([(R + 2)/R] - \{[4R^2 - 7R + 2 + 4R(2R - 1)(R - 1)(c/a)]/[R(2R - 1)^2]\} \exp(-4Rc/a) - [4R^2/(2R - 1)^2] \exp(-2c/a)\Big)$$

where D is the trip density for the region, a is the average trip length for the region, and $R = C_a/(C_a - C_e)$.

A Pair of Parallel Expressways (With Unlimited Access)

We are concerned here with the case of two parallel expressways that are d miles apart in an unbounded region. We find it convenient in this case to partition the world into four regions (Fig. 22). Trips originating in the various regions are then classified according to their direction, whether or not they cross an expressway. These classes are described schematically in Figure 23 and analytically in the following (notation is that used earlier):

Figure 22.

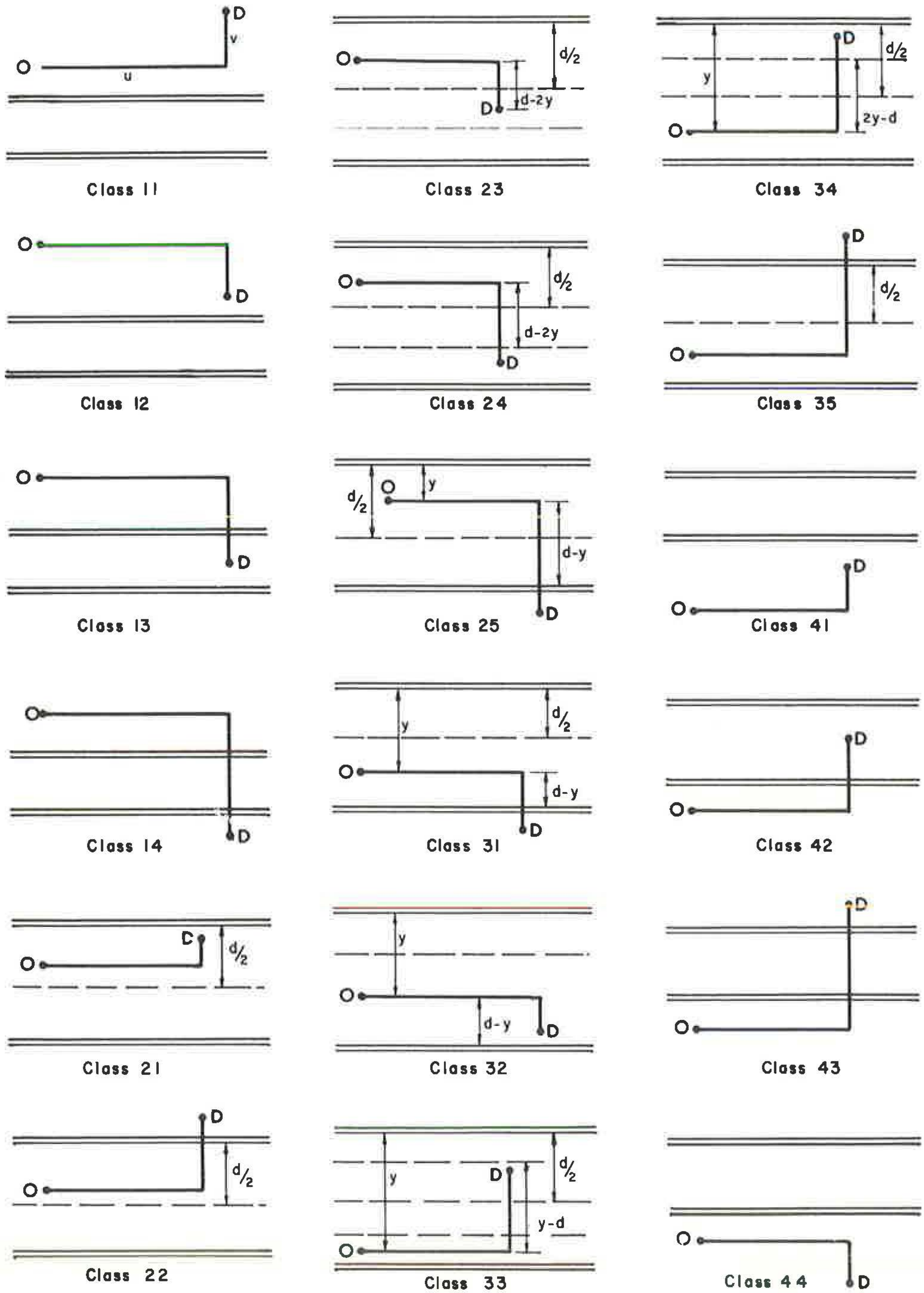

Class 11
Class 23
Class 34
Class 12
Class 24
Class 35
Class 13
Class 25
Class 41
Class 14
Class 31
Class 42
Class 21
Class 32
Class 43
Class 22
Class 33
Class 44

Figure 23.

Class 11: D above 0
y in ${}_dY_{11}$ = $[0;\infty]$
v in ${}_dV_{11}$ = $[0;\infty]$

Class 12: D below 0
y in ${}_dY_{12}$ = $[0;\infty]$
v in ${}_dV_{12}$ = $[0;y]$

Class 13: D below 0
y in ${}_dY_{13}$ = $[0;\infty]$
v in ${}_dV_{13}$ = $[y;y + d]$

Class 14: D below 0
y in ${}_dY_{14}$ = $[0;\infty]$
v in ${}_dV_{14}$ = $[y + d;\infty]$

Class 21: D above 0
y in ${}_dY_{21}$ = $[0;d/2]$
v in ${}_dV_{21}$ = $[0;y]$

Class 22: D above 0
y in ${}_dY_{22}$ = $[0;d/2]$
v in ${}_dV_{22}$ = $[y;\infty]$

Class 23: D below 0
y in ${}_dY_{23}$ = $[0;d/2]$
v in ${}_dV_{23}$ = $[0;d - 2y]$

Class 24: D below 0
y in ${}_dY_{24}$ = $[0;d/2]$
v in ${}_dV_{24}$ = $[d - 2y; d - y]$

Class 25: D below 0
y in ${}_dY_{25}$ = $[0;d/2]$
v in ${}_dV_{25}$ = $[d - y;\infty]$

Class 31: D below 0
y in ${}_dY_{31}$ = $[d/2;d]$
v in ${}_dV_{31}$ = $[d - y;\infty]$

Class 32: D below 0
y in ${}_dY_{32}$ = $[d/2;d]$
v in ${}_dV_{32}$ = $[0;d - y]$

Class 33: D above 0
y in ${}_dY_{33}$ = $[d/2;d]$
v in ${}_dV_{33}$ = $[0;2y - d]$

Class 34: D above 0
y in ${}_dY_{34}$ = $[d/2;d]$
v in ${}_dV_{34}$ = $[2y - d;y]$

Class 35: D above 0
y in ${}_dY_{35}$ = $[d/2;d]$
v in ${}_dV_{35}$ = $[y;\infty]$

Class 41: D above 0
y in ${}_dY_{41}$ = $[d;\infty]$
v in ${}_dV_{41}$ = $[0;y - d]$

Class 42: D above 0
y in ${}_dY_{42}$ = $[d;\infty]$
v in ${}_dV_{42}$ = $[y - d;y]$

Class 43: D above 0
y in ${}_dY_{43} = [d;\infty]$
v in ${}_dV_{43} = [y;\infty]$

Class 44: D below 0
y in ${}_dY_{44} = [d;\infty]$
v in ${}_dV_{44} = [0;\infty]$

Now the volume will certainly be the same on each of the expressways, so that we may concentrate on one of them, say the "upper" one. For this reason, we shall omit classes 24, 25, 31, 32, 33, 41, 42, and 44 from most of our listings below, since these are classes of trips that will never use the "upper" expressway.

Let ${}_dP_{ij}(y) =$ Pr [a trip is in class ij] for a given y. We have

$${}_dP_{11}(y) = \Pr[D \text{ is above } 0]\Pr[v \text{ is in } {}_dV_{11}] = 1/2$$

In general ${}_dP_{ij}(y) = (1/2)\Pr[v \text{ is in } {}_dV_{ij}]$, so that

$$
\begin{aligned}
{}_dP_{12}(y) &= (1/2)[1 - \exp(-ky)] \\
{}_dP_{13}(y) &= (1/2)[1 - \exp(-kd)]\exp(-ky) \\
{}_dP_{14}(y) &= (1/2)\exp(-kd)\exp(-ky) \\
{}_dP_{21}(y) &= (1/2)[1 - \exp(-ky)] \\
{}_dP_{22}(y) &= (1/2)\exp(-ky) \\
{}_dP_{23}(y) &= (1/2)[1 - \exp(-kd)\exp(-ky)] \\
{}_dP_{34}(y) &= (1/2)[\exp(kd)\exp(-2Ky) - \exp(-ky)] \\
{}_dP_{35}(y) &= (1/2)\exp(-ky) \\
{}_dP_{43}(y) &= (1/2)\exp(-ky)
\end{aligned}
$$

Given that a trip is in class ij, let ${}_dQ_{ij}(y, v) =$ Pr [the trip will use an expressway route] for given y, v. Now for trips in class 11, the total cost c_a for a non-expressway route is $uC_a + vC_a$ and the total cost c_e for an expressway route is $2yC_a + uC_e + vC_a$, so that $c_e \le c_a$ implies that $u \ge 2Ry$, where $R = C_a/(C_a - C_e)$, as before. (In this case it was clear that the "lower" expressway did not compete for the trip in the sense that c_e for the "upper" was obviously smaller than c_a for the "lower" facility. For some other classes this may not be so obvious, in fact the "lower" facility may be as likely a candidate for the trip as the "upper." Consider, for example, class 14. In these cases the trips were split 1:1 between the two expressways. In all other cases, the reader will observe that the classes themselves have been defined to guarantee that the use of the "upper" expressway will give a smaller c_e than the use of the "lower.")

Thus,

$${}_dQ_{11}(y, v) = \exp(-2Rky)$$

Similarly,

$$
\begin{aligned}
{}_dQ_{12}(y, v) &= \exp(-2Rky)\exp(2Rkyv) \\
{}_dQ_{13}(y, v) &= 1 \\
{}_dQ_{14}(y, v) &= 1 \\
{}_dQ_{21}(y, v) &= \exp(-2Rky)\exp(2Rkv) \\
{}_dQ_{22}(y, v) &= 1 \\
{}_dQ_{23}(y, v) &= \exp(-2Rky) \\
{}_dQ_{34}(y, v) &= \exp(-2Rky)\exp(2Rkv) \\
{}_dQ_{35}(y, v) &= 1 \\
{}_dQ_{43}(y, v) &= 1
\end{aligned}
$$

Let $P'_{ij} =$ Pr [trip in class ij which uses an expressway route will use the "upper" expressway]. We have $P'_{11} = P'_{12} = P'_{13} = P'_{21} = P'_{23} = P'_{34} = P'_{35} = 1$, while $P'_{14} = P'_{43} = 1/2$. (Of course, $P'_{31} = P'_{32} = P'_{33} = P'_{41} = P'_{42} = P'_{44} = 0$.)

Let ${}_dm_{ij}(y, v)$ = mean distance traveled on expressway by trip of class ij, given that the expressway is used. We have

$$
\begin{aligned}
{}_dm_{11}(y, v) &= 1/k + 2Ry \\
{}_dm_{12}(y, v) &= (1/k + 2Ry) - 2Rv \\
{}_dm_{13}(y, v) &= 1/k \\
{}_dm_{14}(y, v) &= 1/k \\
{}_dm_{21}(y, v) &= (1/k + 2Ry) - 2Rv \\
{}_dm_{22}(y, v) &= 1/k \\
{}_dm_{23}(y, v) &= 1/k + 2Ry \\
{}_dm_{34}(y, v) &= (1/k + 2Ry) - 2Rv \\
{}_dm_{35}(y, v) &= 1/k \\
{}_dm_{43}(y, v) &= 1/k
\end{aligned}
$$

Let ${}_df_{ij}(y,v) = (P'_{ij})\, {}_dP_{ij}(y, v)\, {}_dQ_{ij}(y, v)\, {}_dm_{ij}(y, v)$. We need the mean ${}_d\bar{f}_{ij}(y)$ of ${}_df_{ij}(y, v)$ with respect to v for each ij. Let ${}_dF_{ij}(t)$ be the density function for v on ${}_dV_{ij}$, so that ${}_dF_{ij}(t) = [1/2\ {}_dP_{ij}(y, t)]\,k \exp(-kt)$, t in ${}_dV_{ij}$, then

$$
{}_d\bar{f}_{ij}(y) = \int_{{}_dV_{ij}} {}_df_{ij}(y, t)\, {}_dF_{ij}(t)\,dt
$$

We have

$$
\begin{aligned}
{}_d\bar{f}_{11}(y) &= [1/(2k)] \exp(-2Rky) + Ry \exp(-2Rky) \\
{}_d\bar{f}_{12}(y) &= \{(4R - 1)/[2(2R - 1)^2 k]\} [\exp(-ky) - \exp(-2Rky] \\
&\quad - [R/(2R - 1)] y \exp(-2Rky) \\
{}_d\bar{f}_{13}(y) &= [k/(2k)] [1 - \exp(-kd) \exp(-ky)] \\
{}_d\bar{f}_{14}(y) &= [1/(4k)] \exp(-kd) \exp(-ky) \\
{}_d\bar{f}_{21}(y) &= {}_df_{12}(y) \\
{}_d\bar{f}_{22}(y) &= [1/(2k)] \exp(-ky) \\
{}_d\bar{f}_{23}(y) &= [1/(2k)] \exp(-2Rky) + Ry \exp(-2Rky) \\
&\quad - [1/(2k)] \exp(-kd) \exp[-2(R - 1)ky] \\
&\quad - R \exp(-kd)\, y \exp[-2(R - 1)ky] \\
{}_d\bar{f}_{34}(y) &= \{(4R - 1)/[2(2R - 1)^2 k]\} \exp(-ky) \\
&\quad - \{[4R - 1 + 2R(2R - 1)kd]/[2(2R - 1)^2 k]\} \exp[-(2R - 1) \\
&\quad\ kd] \exp[2(R - 1)ky] \\
&\quad + [R/(2R - 1)] \exp[-(2R - 1)kd]\, y \exp[2(R - 1)ky] \\
{}_d\bar{f}_{35}(y) &= [1/(2k) \exp(-ky)] \\
{}_d\bar{f}_{43}(y) &= [1/(4k)] \exp(-ky)]
\end{aligned}
$$

Let

$$
{}_dS_{ij} = \int_{{}_dY_{ij}} {}_d\bar{f}_{ij}(y)\,dy
$$

so that

$$_dS_{11} = 1/(2Rk^2)$$
$$_dS_{12} = 1/(2Rk^2)$$
$$_dS_{13} = [1/(2k^2)]\,[1 - \exp(-kd)]$$
$$_dS_{14} = [1/(4k^2)] \exp(-kd)$$
$$_dS_{21} = 1/(2Rk^2) - \{(4R-1)/[2(2R-1)^2k^2\} \exp(-kd/2) + \{[2(3R-1) + R(2R-1)kd]/[4R(2R-1)^2k^2]\} \exp(-Rkd)$$
$$_dS_{22} = [1/(2k^2)]\,[1 - \exp(-kd/2)]$$
$$_dS_{23} = 1/(2Rk^2) - \{(2R-1)/[4(R-1)^2k^2]\} \exp(-kd) + \{[3R - 2 + R(R-1)kd]/[4R(R-1)^2k^2]\} \exp(-Rkd)$$
$$_dS_{34} = \{(4R-1)/[2(2R-1)^2k^2]\} \exp(-kd/2) - \{(2R-1)/[4(R-1)^2k^2]\} \exp(-kd) + \{[6R^2 - 6R + 1 + R(R-1)(2R-1)kd]/[4(R-1)^2k^2]\} \exp(-Rkd)$$
$$_dS_{35} = [1/(2k^2)]\,[\exp(-kd/2) - \exp(-kd)]$$
$$_dS_{45} = [1/(4k^2)] \exp(-kd)$$

Let

$$\Delta_d V = \left(\sum_{ij} {}_dS_{ij}\right) \Delta x$$

and let

$$_dV = D \int_0^1 d\, _dV = \text{mean trip density on the expressway}$$

We have

$$_dV = (D/4)a^2\Big([R+2)/R] - [R^2/2(R-1)^2] \exp(-2d/a) + \{[3R - 2 + 2R(R-1)d/a]/[2R(R-1)^2]\} \exp(-2Rd/a)\Big)$$

where D is the trip density for the region, $a = 2/k$ is the average trip length for the region, d is the distance between the expressways, and $R = C_a/(C_a - C_e)$.

Isolated Expressway With Limited Access

We are concerned here with the case of an isolated expressway which is placed in an unbounded region and to which there is limited access. The major assumptions are the same as for the preceding situations. Because the techniques for this case differ somewhat from those used for unlimited access expressways, a more nearly complete discussion will be given here than in previous sections.

We assume that access to the expressway is allowed at equal intervals of d miles. One of these access points is selected at random and labeled "exit 0." The origin of a rectangular coordinate system is identified with this point, with the x-axis coincident with the expressway. For ease of discussion, the positive directions of the x-axis and y-axis will be called "east" and "north," respectively (Fig. 24).

The access points are number 0, 1, 2,. . ., q,. . ., looking east. For many of our arguments, it will be convenient to group these exits in pairs, the first pair consisting of exits 0 and 1; the second, 2 and 3, etc. The (k + 1) pair will have its western and eastern exits labeled W_k and E_k, respectively.

The general typing of trips is done as for an isolated expressway with unlimited access, and the same symmetry considerations as in that case allow us here to concentrate

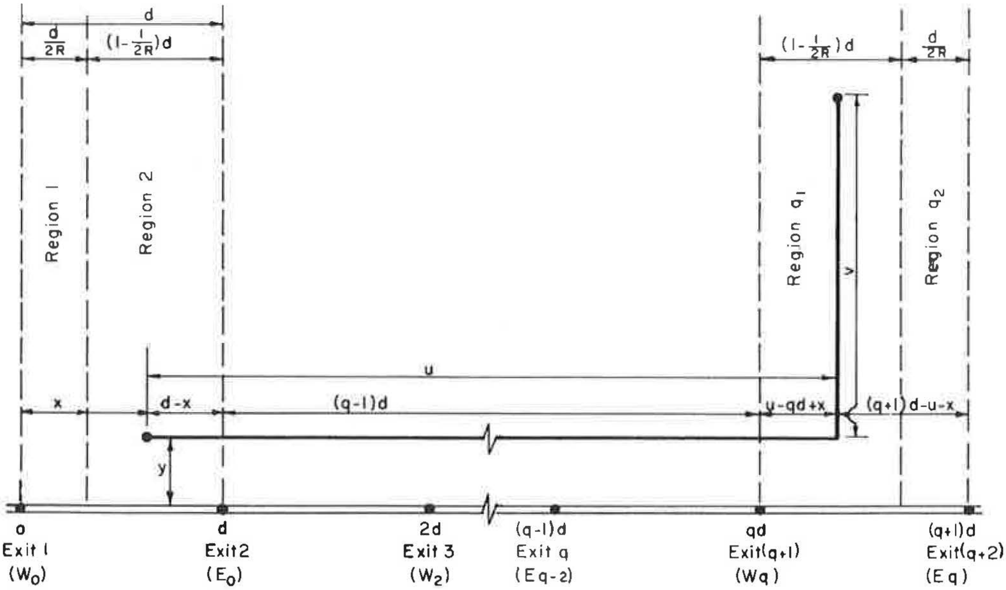


Figure 24.

on trips whose origin is north of the expressway. However, we must stratify the world more finely in this case with regard to the lateral position of trip origins and destinations.

Consider a trip whose origin has coordinates (x, y) with $0 \le x \le d$ and whose destination has coordinates (x′, y′) with $qd \le x' \le (q + 1)d$. Now, if the trip uses the expressway, it may enter at W_0 or E_0 and leave at W_q or E_q. At any rate, the trip (assuming that the expressway is used) will pass the points E_0 and W_q. If, for the leg of the trip from the origin E_0, costs are computed, first directly to E_0 from the origin and second to E_0 from the origin via W_0, it will be observed that trips for which the x-coordinate of the origin lies between 0 and d/2R $[R = C_a/(C_a - C_e)]$ will use W_0 while those for which $[1 - 1/(2R)]d \le x \le d$ will use E_0. A similar computation for the leg of the trip from W_k to the destination (x′, y′) will show that $kd \le x' \le kd + [1 - 1/(2R)]d$ implies that W_k will be used, while $kd + [1 - 1/(2R)]d \le x' \le (k + 1)d$ implies that E_k will be used. Finally, these considerations lead us to a classification of trips from region to region, the regions depicted in Figure 24 and defined analytically by

region 1: $0 \le x \le d/(2R)$

region 2: $d/(2R) \le x \le [1 - 1/(2R)]d$

region k_1: $kd \le x' \le kd + [1 - 1/(2R)]d$

region k_2: $kd + [1 - 1/(2R)]d \le x' \le (k + 1)d$

The classification itself proceeds as follows: Consider those trips whose origin lies within the vertical strip with base from W_0 to E_0 and whose destination lies within the vertical strip with base from W_q to E_q. These trips are classified according to the scheme:

Class 11: D above 0
0 in region 2
D in region q_1
v in $[0;\infty]$

Class 12: D above 0
0 in region 1
D in region q_1
v in [0;∞]

Class 13: D above 0
0 in region 2
D in region q_2
v in [0;∞]

Class 14: D above 0
0 in region 1
d in region q_2
v in [0;∞]

Class 21: D below 0
0 in region 2
d in region q_1
v in [0;y]

Class 22: D below 0
0 in region 1
d in region q_1
v in [0;y]

Class 23: D below 0
0 in region 2
D in region q_2
v in [0;y]

Class 24: D below 0
0 in region 1
D in region q_2
v in [0;y]

Class 31: D below 0
0 in region 2
D in region q_1
v in [y;∞]

Class 32: D below 0
0 in region 1
D in region q_1
v in [y;∞]

Class 33: D below 0
0 in region 2
D in region q_2
v in [y;∞]

Class 34: D below 0
0 in region 1
D in region q_2
v in [y;∞]

We will calculate the VMT on the expressway contributed by each of these classes, add these over q = 1, 2,. . ., multiply by 4 to obtain the total VMT for westbound and eastbound trips as well as those which originate south of the facility, and then divided by d to obtain the mean trip density on the facility.

A detailed discussion of the computation of the VMT on the expressway contributed by class 13 should indicate the general method. Let $P_{ij}(x, y, q)$ = Pr [trip is in class ij for given x, y, given that 0 is in the appropriate region for class ij]. Let $a = \{q + [(2R - 1)/(2R)]\}d - x$ and $c = (q + 1)d - x$. We have

$P_{13}(x, y, q)$ = Pr [D is in region q_2 given that 0 is in region 2] Pr [$0 \le v < \infty$]
= Pr [D is above and to the east of 0] Pr [$a \le u \le c$]
= (1/4) Pr [$a \le u \le c$]

Now for trips of class 13, the trip cost c_a for a non-expressway route is $uC_a + vC_a$, while the trip cost c_e for the expressway route is

$$(d - x)C_a + yC_a + qdC_e + yC_a + [(q + 1)d - u - x] C_a + vC_a$$

(See Fig. 24.) Thus $c_e \le c_a$ if, and only if, $u \ge b$ where

$$b = y - x + \{1 + [2R - 1)/(2R)] q\} d, \; R = C_a/(C_a - C_e)$$

Hence, if we let $Q_{ij}(x, y, q)$ = Pr [trip in class ij for given x, y will use the expressway], then $Q_{13}(x, y, q)$ = Pr [$b \le u \le c$ given that $a \le u \le c$]. Now for $a \le b \le c$, we have Pr [$b \le u \le c$ given that $a \le u \le c$] = Pr [$b \le u \le c$] / Pr [$a \le u \le c$]. Thus

$$Q_{13}(x, y, q) = \begin{cases} \Pr[b \le u \le c]/\Pr[a \le u \le c] \text{ for } a \le b \le c \\ 1 \text{ for } b < a \end{cases}$$

So far, we have that

$$P_{13}(x, y, q)\, Q_{13}(x, y, q) = \begin{cases} (1/4)\Pr[b \le u \le c] \text{ for } a \le b \le c \\ (1/4)\Pr[a \le u \le c] \text{ for } b < a \end{cases}$$

But $b \le c$ implies that $y \le qd/(2R)$, while $a \le b$ implies $y \ge (q - 1)d/(2R)$. Thus, $a \le b \le c$ implies $(q - 1)d/(2R) \le y \le qd/(2R)$. Of course $b < a$ implies $y < (q - 1)d/(2R)$.

Finally, then,

$$P_{13}(x, y, q)\, Q_{13}(x, y, q) = \begin{cases} (1/4)\exp(-kd)\{\exp[kd/(2R)] - 1\}\exp(-qkd) \\ \quad \exp(kx), \text{ for } y \text{ in } [0;(q - 1)d/(2R)) \\ (1/4)\exp(-kd)\exp(kx)\{\exp[-qkd(2R-1)/ \\ \quad (2R)]\exp(-ky) - \exp(-qkd)\}, \\ \quad \text{for } y \text{ in } [(q - 1)d/(2R); qd/(2R)] \\ 0, \text{ for } y \text{ in } [qd/(2R);\infty] \end{cases}$$

Let $m_{ij}(q)$ = distance traveled on facility by trip in class ij, given that the facility is used. We have $m_{13}(q) = qd$. For each q, let $V_{ij}(q)$ = VMT on facility per trip for trips in class ij, so that

$$V_{ij}(q) = \int_Y \int_X P_{ij}(x, y, q)\, Q_{ij}(x, y, q)\, m_{ij}(q)\, dxdy$$

We have

$$V_{13}(q) = \int_0^{(q-1)d/(2R)} \int_{d/(2R)}^{d} P_{13}(x, y, q)\, Q_{13}(x, y, q)\, m_{13}(q)\, dxdy$$

$$+ \int_{(q-1)d/(2R)}^{qd/(2R)} \int_{d/(2R)}^{d} P_{13}(x, y, q)\, Q_{13}(x, y, q)\, m_{13}(q)\, dxdy$$

$$= [d^2/(8Rk)]\{\exp[kd/(2R)] - 1\}(1 - \exp\{[-(2R - 1)/(2R)] kd\})\, q^2 \exp(-qkd)$$

$$- [d^2/(8Rk)] \exp [kd/(2R)] (1 - \exp \{[- (2R - 1)/(2R)] kd\}) \, q \exp (-qkd)$$
$$+ [d/(4k^2)] [\exp [kd/(2R)] - 1](1 - \exp \{[- (2R - 1)/(2R)] kd\}) \, q \exp (-qkd)$$

Let S_{ij} = total VMT on the facility per trip for trips in class ij for all q, so that

$$S_{ij} = \sum_{q=1}^{\infty} V_{ij}(q)$$

In particular,

$$S_{13} = [d^2/(8Rk)] \{\exp [kd/(2R)] - 1\} \left(1 - \exp \{[- (2R - 1)/(2R)] kd\}\right) \exp (-kd) [1 + \exp (-kd)]/[1 - \exp (-kd)]^3$$
$$+ [d/(4K^2)] \{\exp [kd/(2R)] - 1\} \left(1 - \exp \{[- (2R - 1)/(2R)] kd\}\right) [\exp (-kd)]/[1 - \exp (-kd)]^2$$
$$- [d^2/(8Rk)] \left(1 - \exp \{[- (2R - 1)/(2R)] kd\}\right) \exp [kd/(2R)] [\exp (-kd)]/[1 - \exp (-kd)]^2$$

Before listing for the remaining classes the values of the various quantities defined above, we note that symmetry gives $S_{i2} = S_{i3}$, i = 1, 2, 3, and state that it can be shown that $S_{1j} = S_{2j}$, j = 1, 2, 3, 4. With these omissions, then, we have

$$P_{11}(x, y, q) \, Q_{11}(x, y, q) = \begin{cases} (1/4) \left(1 - \exp \{[- (2R - 1)/(2R)] kd\}\right) \exp (kx) \exp (-qkd), & \text{for } y \text{ in } [0; (q - 1)d/(2R)] \\ 0, & \text{for } y \text{ in } [(q - 1)d/(2R); \infty] \end{cases}$$

$$P_{14}(x, y, q) \, Q_{14}(x, y, q) = \begin{cases} (1/4) \exp (-kd) \{\exp [kd/(2R)] - 1\} \exp (kx) \exp (-qkd), & \text{for } y \text{ in } [0; qd/(2R) - x] \\ (1/4) \exp (-kd) [\exp [kd/(2R)] \exp \{[- (2R - 1)/(2R)] qkd\} \exp (-ky) - \exp (-qkd) \exp (kx)], & \text{for } y \text{ in } [qd/(2R) - x; (q + 1)d/(2R) - x] \\ 0, & \text{for } y \text{ in } [(q + 1)d/(2R) - x; \infty] \end{cases}$$

$$P_{31}(x, y, q) \, Q_{31}(x, y, q) = (1/4)(1 - \exp \{[- (2R - 1)/(2R)] kd\}) \exp (kx) \exp (-ky) \exp (-qkd), \text{ for } y \text{ in } [0; \infty]$$

$$P_{32}(x, y, q) \, Q_{32}(x, y, q) = P_{31}(x, y, q) \, Q_{31}(x, y, q)$$

$$P_{34}(x, y, q)\, Q_{34}(x, y, q) = (1/4) \exp(-kd) \{\exp[kd/(2R)] - 1\} \exp(kx) \exp(-ky) \exp(-qkd), \text{ for } y \text{ in } [0;\infty]$$

$$m_{11} = m_{31} = (q - 1)d$$

$$m_{32} = qd$$

$$m_{14} = m_{34} = (q + 1)d$$

$$V_{11}(q) = [d^2/(8Rk)]\,(1 - \exp\{[-(2R - 1)/(2R)]\, kd^2\})^2 (q - 1)^2 \exp[-(q - 1)kd]$$

$$V_{14}(q) = [d^2/(8Rk)]\,\{\exp[kd/(2R)] - 1)^2 (q + 1) \exp[-(q + 1)kd] - [d^2/(8Rk)]\,[\exp(kd/R) - 1]\,(q + 1) \exp[-(q + 1)kd] + [d/(2k^2)]\,\{\exp[kd/(2R)] - 1\}^2 (q + 1) \exp[-(q + 1)kd]$$

$$V_{31}(q) = [d/(4k^2)]\,(1 - \exp\{[-(2R - 1)/(2R)]\, kd\})\,\{\exp(kd) - \exp[kd/(2R)]\}\,(q - 1) \exp(-qkd)$$

$$V_{32}(q) = [d/(4k^2)]\,(1 - \exp\{[-(2R - 1)/(2R)]\, kd\})\,\{\exp[kd/(2R)] - 1\}\, q \exp(-qkd)$$

$$V_{34}(q) = [d/(4k^2)] \exp(-kd)\,\{\exp[kd/(2R)] - 1\}^2 (q + 1) \exp(-qkd)$$

We shall omit a listing of S_{ij}. They are easily gotten, as in the example for class 13, by summing the $V_{ij}(q)$ over q.

Finally, we sum the S_{ij} over all i and j, multiply the result by 4 and divide by d (the reasons for doing this were explained earlier) to obtain the mean trip density Q on the facility. We find that

$$\begin{aligned}
Q = {} & [Dda/(8R)]\,(1 - \exp\{[-(2R - 1)/R]\, d/a\})^2 \exp(-2d/a)\,[1 + \exp(-2d/a)]/[1 - \exp(-2D/a)]^3 \\
& + [Dda/(4R)]\,[\exp(d/Ra) - 1]\,(1 - \exp\{[-(2R - 1)/R]\, d/a\}) \exp(-2d/a)\,[1 + \exp(-2d/a)]/[1 - \exp(-2d/a)]^3 \\
& + (3Da^2/8)\,\{\exp[d/(Ra)] - 1\}\,(1 - \exp\{[-(2R - 1)/R]\, d/a\}) \exp(-2d/a)/[1 - \exp(-2d/a)]^2 \\
& + [Dda/(4R)]\,\{\exp[d/(Ra)] - 1\}^2 \exp(-4d/a)/[1 - \exp(-2d/a)]^3 \\
& + (5Da^2/16)\,\{\exp[d/(Ra)] - 1\}^2 \exp(-4d/a)\,[2 - \exp(-2d/a)]/[1 - \exp(-2d/a)]^2 \\
& + (Da^2/16)\,(1 - \exp\{[-(2R - 1)/R]\, d/a\})^2 \exp(-2d/a)/[1 - \exp(-2d/a)]^2 \\
& - [Dda/(4R)]\,(1 - \exp\{[-(2R - 1)/R]\, d/a\}) \exp[d/(Ra)] \exp(-2d/a)/[1 - \exp(-2d/a)]^2 \\
& - [Dda/(8R)]\,\{\exp(2d/(Ra)] - 1\} \exp(-4d/a)\,[2 - \exp(-2d/a)]/[1 - \exp(-2d/a)]^2
\end{aligned}$$

where

D = trip density for the region,
a = average trip length for the region,

$R = C_a/(C_a - C_e)$, and
d = distance between exits on the facility (ramp spacing).

Note on trip cost: Let

0 (S) = operating cost per mile in dollars for a given speed S in miles per hour,
A (S) = accident cost per mile in dollars for a given speed S in miles per hour,
V = value of time in dollars per hour, and
C (S) = total cost per mile in dollars for a trip segment at speed S in miles per hour.

We have

$$C(S) = 0(S) + A(S) + V/S$$

For the purpose of preparing tables, 0 (S) and A (S) were taken from an earlier paper to which minor adjustments were made (11); V was varied from $1.00 to $2.00 step $.10.

REFERENCES

1. Schneider, M. A Direct Approach to Traffic Assignment. Highway Research Record 6, p. 71-75, 1963.
2. Smeed, R. J., and Jeffcoate, G. O. Traffic Flow During the Journey to Work in the Central Area of a Town Which has a Rectangular Grid for Its Road System. Proc. Second Internat. Symposium on Theory of Road Traffic Flow, London, 1963.
3. Smeed, R. J. Road Development in Urban Areas. Jour. Inst. of Highway Engineers, No. 10, p. 5-26, 1963.
4. Creighton, R. L., Hoch, I., and Schneider, M. The Optimum Spacing of Arterials and of Expressways. Traffic Quarterly, No. 13, p. 477-494, 1959.
5. Fairthorne, D. The Distances Between Pairs of Points in Towns of Simple Geometrical Shapes. Proc. Second Internat. Symposium on Theory of Road Traffic Flow, London, 1963.
6. Friedrich, P. Mathematical Description of Life in Settled Areas and Derivation of Plan Layout Thereout as a Problem of Theoretical Town Planning. Presented at Second Internat. Symposium on Theory of Traffic Flow, London, 1963.
7. Haight, F. Some Probability Distribution Associated With Commuter Travel in a Homogeneous Circular City. Operations Research, Vol. 12, p. 964-975, 1964.
8. Wardrop, J. G. Some Theoretical Aspects of Road Traffic Research. Proc. Inst. Civil Engineers, Part II, p. 325-378, 1952.
9. Holroyd, E. M. Theoretical Average Journey Lengths in Circular Towns With Various Routing Systems. Road Research Laboratory Rept. No. 43, 1966.
10. Tanner, J. C. A Theoretical Model for the Design of a Motorway System. Road Research Laboratory Rept. No. 23, 1966.
11. Haikalis, G., and Joseph, H. Economic Evaluation of Traffic Networks. Chicago Area Transportation Study, 1961.

The Transportation System

An Evaluation of Alternative Land Use and Transportation Systems in the Chicago Area

E. WILSON CAMPBELL, Chicago Area Transportation Study

•THE Chicago Area Transportation Study (CATS) has been involved in transportation planning in the Chicago area for 12 years. In 1962, CATS recommended a highway and mass transit plan. This was the third of a series of reports that described the basic inventories, detailed the development of future forecasts, and explained the techniques developed and used for plan selection and evaluation.

The transportation plan was based on a forecast of future land-use and development patterns with assumptions about economic activity of the region. The forecast was quantitative in nature. It was rigorously controlled with internal consistency checks to insure technical precision. The development pattern represented growth that, in all likelihood, would occur in the area in the absence of a "plan." The estimated new population and additional land uses were distributed throughout the region in the same way that development is occurring today. While the transportation plan was technically sound, CATS recognized the necessity of having a transportation plan that was in harmony with a long-range comprehensive development plan.

The opportunity for testing CATS transportation proposals with a comprehensive plan came in September 1964. At that time, the Northeastern Illinois Planning Commission (NIPC) entered into contract with the Illinois Division of Highways to develop several alternative plans and corresponding land-use and population projections for the Northeast Illinois Metropolitan Area (NIMA) for the year 1990. These alternative projections were to be used as input for evaluation of the CATS transportation plan for the CATS portion of the NIMA. This report describes the CATS efforts in this joint "Land-Tran" project.

PROCEDURE

Evaluation of the adequacy of CATS proposals with respect to the various land-use alternatives required specialized studies and analyses. It was agreed that CATS would test the adequacy of its 1962 Transportation Plan with respect to five separate and distinct land-use alternatives. (This was later revised to four when NIPC eliminated one by policy decision.) It was also agreed that CATS and NIPC would jointly design a transportation plan specifically oriented to the land-use goals of the "best" of the four plans previously tested. This transportation plan would be subjected to the same testing and evaluation procedure followed in testing the original CATS proposals.

The five plans developed by NIPC are as follows (1):

Alternative No. 1—The Dispersed Regional City Plan. The main features of this plan were low-density residential development, dispersed shopping facilities in the form of local neighborhood centers, dispersed jobs in the form of small office clusters and individual manufacturing plant sites, and utilizing areas along major streams and rivers for recreation. The automobile was considered the prime means of transportation with mass transit playing a minor role. (Note: This plan was ruled out by policy decision of the NIPC Planning Committee.)

Paper sponsored by Committee on Transportation System Evaluation and presented at the 47th Annual Meeting.

Alternative No. 2—The Finger Plan. The emphasis of this plan was on radial corridors resulting in a physical pattern resembling the human hand with the area within the Tri-State Tollway forming the palm and urban corridors radiating outward. Residential, shopping, and job centers were developed in the radial corridors. Higher densities were developed along the centers of the corridors adjacent to the high-speed rail and expressway facilities with lower densities at the outer edges of the corridors. This plan emphasized the use of rail transit facilities in the corridors.

Alternative No. 3—The Multi-Towns Plan. This plan emphasized a major urban concentration surrounded by numerous clusters of specialized towns ranging from 10,000 to 100,000 persons. The clusters were bounded, served, and separated by major open spaces, shopping centers, industrial parks, and the expressway network. Main transportation emphasis would be on expressways while recognizing the need for good rail transportation.

Alternative No. 4—Satellite Cities-Greenbelt Plan. This plan channeled growth into four or five large cities of one to two million persons located at a distance of 35 to 40 miles from Chicago's loop. These cities would be separated by large green recreational areas and connected by high-speed transportation facilities.

Alternative No. 5—Trends. This alternative adds growth to the existing land-use pattern in accordance with present zoning laws and existing land development trends.

Detailed land-use data forecast on the basis of each of the foregoing plans by CATS analysis zone was quantified by NIPC and provided to CATS as input for the CATS testing process.

METHODOLOGY

The testing and evaluation performed by CATS involved five distinct phases as follows:

1. Determination of person trip totals—This step involves converting the land use and socioeconomic data to travel demand for each CATS analysis zone for each land-use plan.

2. Determination of mode choice—In this step the person travel is related to particular travel mode. Results of this phase provide transit, auto, and truck trips by analysis zone.

3. Distribution and assignment—In this step, the trip interchange between zones is determined and these volumes are "assigned" to segments of the network that would be used in performing the trip. This step is performed for highway and transit trips separately for each land-use pattern.

4. Network performance evaluation—After the assignments are completed, an analysis is made to determine the quality of service afforded by the transportation network. This involves inspection of total travel time, vehicle and passenger miles traveled, average speed of travel, and relationships of assigned volumes to capacities. Generally, the network that affords the lowest total miles traveled at the highest average speed while minimizing overload situations is the system that provides the highest "quality of service."

5. Economic evaluation—The final step in the evaluation procedure is an economic comparison. This analysis considers the capital investment required to construct the transportation network and the cost to the user of the system. User costs include the operating cost for vehicles, probable accident costs, and personal time costs. CATS' economic evaluation is based on a single criterion—that of least total cost to build and use a system over a period of years. In other words, the objective is to plan a system so that the sum of measurable costs for all travelers and taxpayers in the region will be at a minimum.

To accomplish the five stages of work described, CATS has developed a series of models programmed to utilize high-speed computers which develop the trip estimates, split trips by mode, distribute and assign trips, and summarize outputs of the process. These models will be described in more detail later in this report along with an analysis and summary of the evaluation process.

TRIP ESTIMATION AND MODE CHOICE

Trip estimation is a sophisticated, highly complex process requiring an intimate knowledge of travel patterns as related to land usage and socioeconomic characteristics of an area. The technique used by CATS in this Land-Tran project is best described as an "incremental technique." Trips are estimated in two segments. The first segment contains an estimate of future trips made by the current population. To this estimate are added the trips generated by the increase in population expected in the future year. The estimate of future trips to be made by current population allows for the possibility of increased trip-making due to higher car ownership, more leisure time, and a generally more affluent society. Total trips are determined, then, by adding these future additional trips made by the present population and the trips to be made by the added population to the number of trips actually being made at the present time. Table 1 shows the total trips by mode for each alternative.

This incremental method of trip estimation is completely automated. A large share of the programming was done as part of this joint project. The details of trip estimation in this report will concentrate on the aspects of trip generation that are applicable to the Land-Tran project and not documented in other CATS publications.

EVALUATION OF LAND-USE PLANS FROM TRANSPORTATION STANDPOINT

One of the basic goals of the Land-Tran project was to determine the compatibility of transportation plans proposed by CATS with any or all of the land-use plans being studied. Conversely, it was hoped that traffic simulation studies would provide clues as to which of the land-use plans was best from a transportation point of view. It was reasoned that system measures of average trip length, average speed, total miles of travel, and total travel would serve to indicate the system that satisfied the travel requirements in the most efficient way. A measure of system efficiency was devised to help in the selection of such a system. This index of system efficiency was the ratio of total passenger minutes of travel to total passenger miles of travel. Results of assignments to the total transportation system (including vehicle and mass transit trips) were summed for each land-use plan.

The CATS proposed transportation plans were used for the traffic simulation for each land use to test their ability to serve the land uses and also to provide a common basis for comparing the travel requirements as measured in a network.

There are two cautions. First, since CATS' geographic area of concern did not cover the whole NIMA, a bias could be introduced in favor of land-use plans that had more population (i.e., more travel) outside the CATS area. This is serious in only one case—the satellite cities alternative, since the planned locations of the large cities were mostly outside of the CATS area. While all plans had essentially the same total population for the NIMA, the satellite cities alternative had 600,000 fewer people in the CATS area than were shown in the finger alternative. Fortunately, the satellite cities alternative had other problems which, from a policy point of view, made it less desirable than any of the three remaining alternatives.

Second, there exists the possibility of a bias caused by using a planned trip length as input to the distribution and assignment process. For example, trip lengths were pur-

TABLE 1

TRIP TOTALS BY MODE FOR ALTERNATIVE LAND-USE PLANS

Land Use Alternatives	1990		
	Passenger Trips[a]	Truck Trips	Total Vehicle Trips
1. Dispersed City Regional Plan	10,687,637	2,838,197	13,525,834
2. Finger Plan	10,278,812	2,714,863	12,993,675
3. Multiple Towns Plan	10,192,843	2,674,544	12,867,387
4. Satellite Cities Plan	9,468,248	2,483,843	11,952,091
5. Trends	10,087,721	2,602,299	12,690,020

[a]Auto driver trips plus taxi trips = passenger car trips.

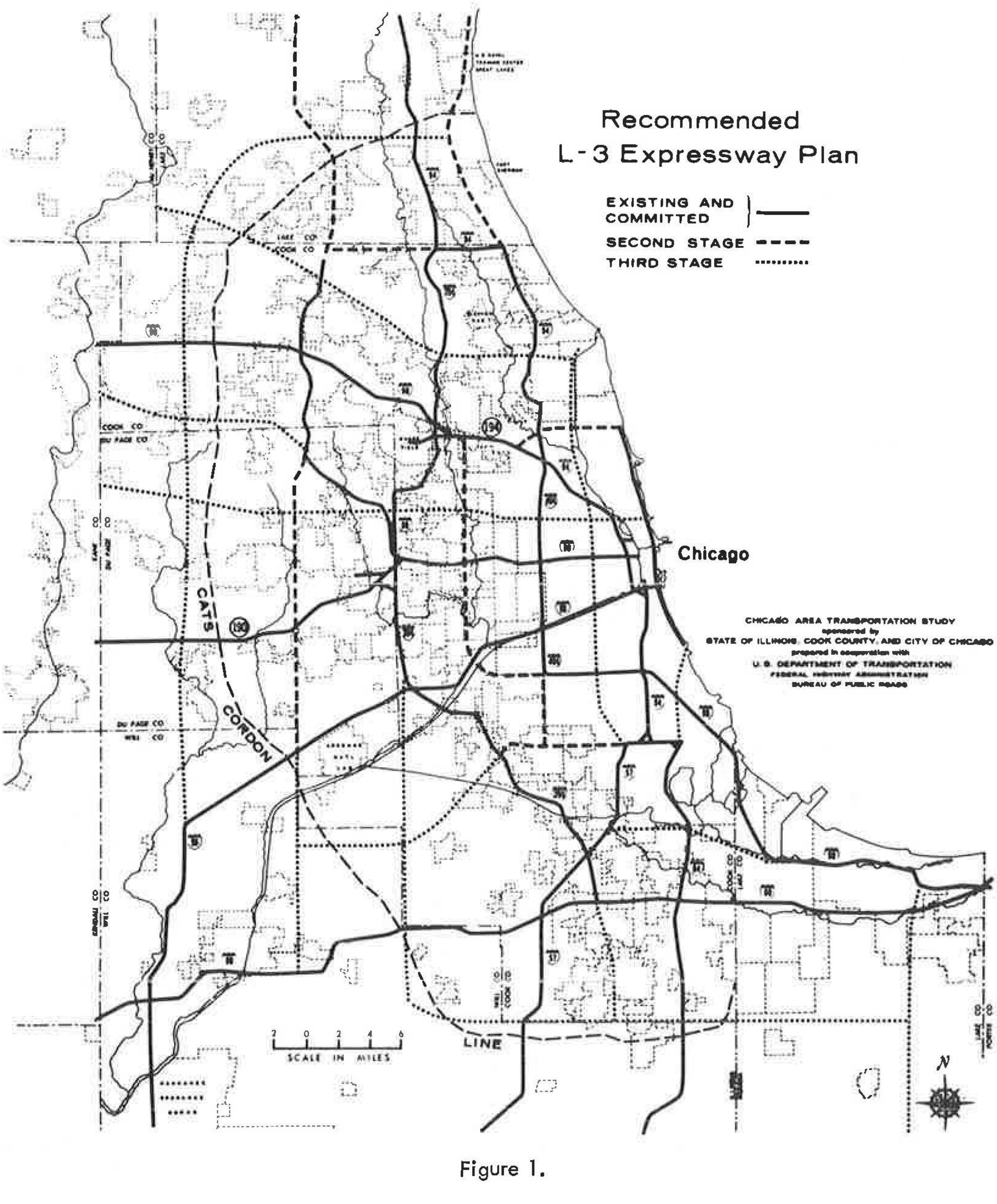


Figure 1.

posely shortened as a planning goal in the finger plan. This may not be a serious bias, however, since the opportunity model can only connect trips where opportunities are available. Trip lengths are somewhat dependent on the distribution of acceptable opportunities for destinations.

Before describing the results of the assignment testing, the transportation networks used in the assignment will be discussed.

Transportation Networks

The transportation systems used for testing in the Land-Tran project were CATS' proposed transit and highway plans. These plans were published in 1962 after several years of intensive analysis. As pointed out earlier, the plans were prepared to satisfy

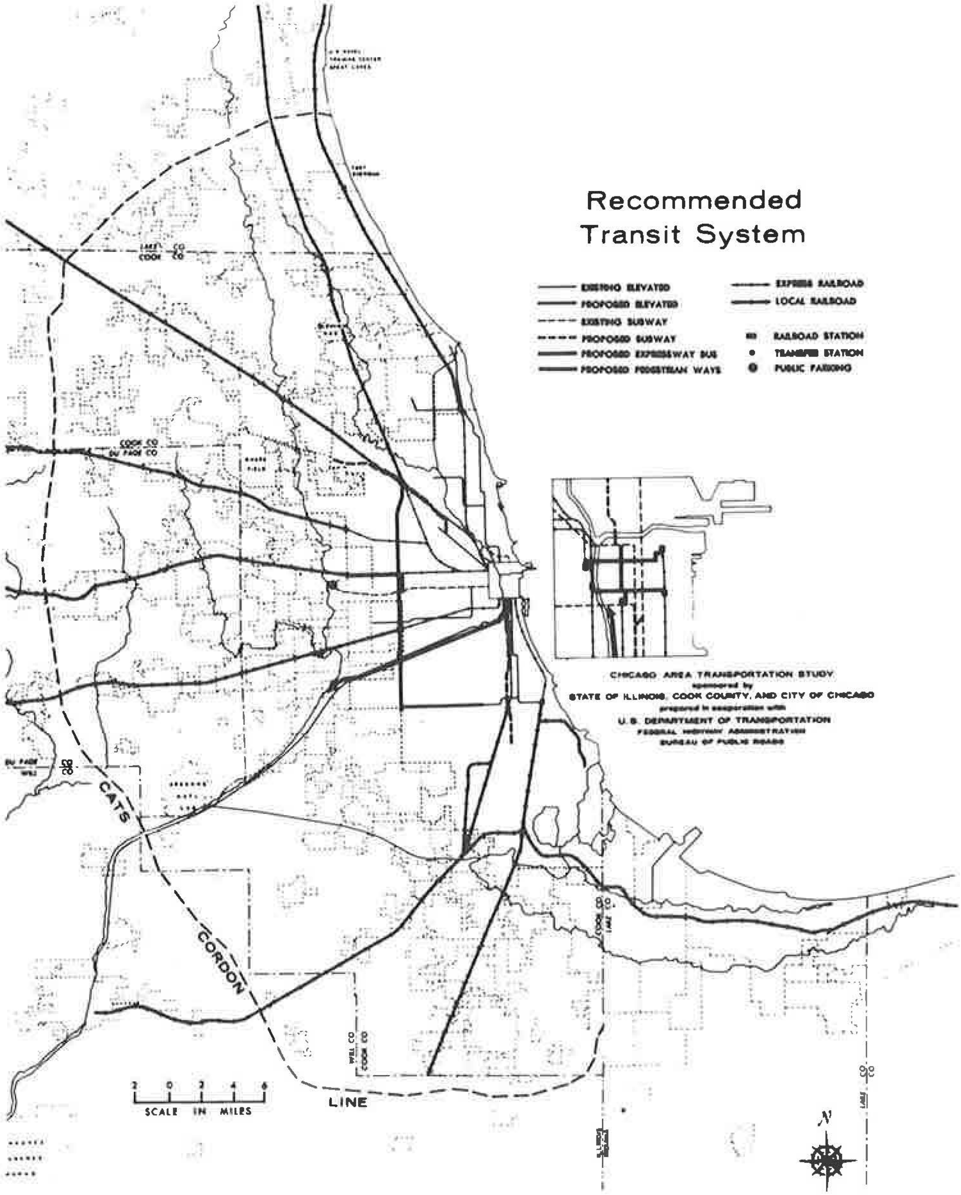


Figure 2.

travel demand based on a land-use forecast and certain assumptions about economic activity. Formulas were developed to assist in selecting an optimal spacing of expressways leading to one that would minimize total travel and construction costs. The recommended plan (designated L-3) was selected after studying a number of alternatives ranging from 300 to 900 miles of proposed expressways. The L-3 plan satisfied the single criterion of least cost. The capital cost was reasonable as gauged by the current annual expenditure for expressway and transit construction in the Chicago area.

The L-3 highway plan (Fig. 1) contains 520 miles of expressway facilities within the CATS cordon line. Included in this total are several miles of toll roads. As a matter of policy, all routes are considered to be freeways in 1990.

TABLE 2

SUMMARY COMPARISON OF HIGHWAY AND TRANSIT DATA WITH ASSIGNMENT RESULTS FOR FOUR ALTERNATIVE PLANS

Factor		Finger Plan Alt. No. 2	Multi-Towns Plan Alt. No. 3	Satellite Cities Plan Alt. No. 4	Trends Plan Alt. No. 5
Population	Internal	7,569,400	7,317,100	6,924,800	7,150,100
	External	1,802,800	2,098,600	2,473,100	2,244,700
	Total	9,372,200	9,415,700	9,397,900	9,394,800
Passenger trips*	Transit	2,366,353	2,297,837	2,184,309	1,948,214
	Vehicle	18,789,110	18,563,521	17,274,977	18,348,203
	Total	21,155,463	20,861,358	19,459,286	20,296,417
Average trip length	Transit	7.76	7.04	7.34	7.55
	Vehicle	4.96	5.56	5.19	5.32
	Passenger	5.30	5.70	5.50	5.55
Passenger miles*	Transit	18,362,094	16,070,904	16,023,659	14,706,521
	Vehicle	93,184,300	103,163,142	90,936,681	97,780,384
	Total	111,546,394	119,234,046	106,960,340	112,486,905
Passenger minutes*	Transit	104,961,206	100,116,716	102,307,849	87,416,203
	Vehicle	193,206,593	207,524,687	188,269,088	239,721,535
	Total	298,167,799	307,741,403	290,576,937	327,137,738
Index of total system efficiency		.374	.387	.368	.343
Passenger travel time (min/mile)		2.67	2.58	2.72	2.91
Equivalence speed (mph)		22.44	23.22	22.08	20.58
Index of transit system efficiency		.175	.159	.157	.168
Equivalence transit speed (mph)		10.5	9.6	9.4	10.1
Index of vehicle system efficiency		.482	.498	.482	.408
Equivalence vehicle speed (mph)		28.9	29.7	28.9	24.6
Minutes/person		39.4	41.9	42.0	45.6

*Includes truck trips.

TABLE 3

SUMMARY COMPARISON OF VEHICLE ASSIGNMENTS FOR FOUR ALTERNATIVE PLANS – INTERNAL AREA

Factor	Finger Plan Asmt. No. 06	Multi-Towns Plan Asmt. No. 07	Satellite Cities Plan Asmt. No. 08	Trends Plan Asmt. No. 09
Sent trips	12,978,504	12,840,560	11,934,720	12,674,280
VEMAV – Arterials	36,750,800	38,884,700	34,717,900	37,726,100
VEMAV – Expwys	30,095,900	35,871,200	29,776,200	32,619,500
VEMAV – Art + Exp	66,846,700	74,755,900	64,494,100	70,345,600
VEMAV/VEMDC – Arterials	0.85	0.90	0.80	0.87
VEMAV/VEMDC – Expwys	0.77	0.91	0.76	0.83
VEMAV/VEMDC – Art + Exp	0.81	0.91	0.78	0.85
Trips per mile – Arterials	13,000	13,800	12,300	13,400
Trips per mile – Expwys	57,900	69,100	57,300	62,800
Trips per mile – Art + Exp	20,000	22,400	19,300	21,000
Op. cost ($) – Arterials	983,000	1,047,000	927,700	1,009,700
Op. cost ($) – Expwys	924,900	1,082,200	910,500	992,300
Op. cost ($) – Art + Exp	1,907,900	2,129,200	1,838,200	2,002,000
Acc. cost ($) – Arterials	541,200	591,200	508,000	557,100
Acc. cost ($) – Expwys	70,100	87,400	70,300	78,000
Acc. cost ($) – Art + Exp	611,300	678,600	578,300	635,100
Time cost ($) – Arterials	2,527,000	2,737,100	2,384,300	2,605,000
Time cost ($) – Expwys	833,100	1,014,200	829,800	914,500
Time cost ($) – Art + Exp	3,360,100	3,751,300	3,214,100	3,519,500
Total cost ($) – Arterials	4,051,100	4,375,300	3,820,100	4,171,900
Total cost ($) – Expwys	1,828,100	2,183,800	1,810,600	1,984,800
Total cost ($) – Art + Exp	5,879,200	6,559,100	5,630,700	6,156,700
Travel cost (cents) per veh. mile – Arterials	11.0	11.2	11.0	11.1
Travel cost (cents) per veh. mile – Expwys	6.1	6.1	6.1	6.1
Travel cost (cents) per veh. mile – Art + Exp	8.8	8.8	8.7	8.8

NOTE: 1. Arterials: Length = 2,822.6 miles, VEMDC = 43,241,500
2. Expressways: Length = 519.5 miles, VEMDC = 39,293,200

The arterial streets (CATS' designation for all roads except freeways) total 2,822.6 miles within the cordon line. For the City of Chicago, the street system selected is the preferential street system as developed by the City of Chicago. Selection of streets is a function of the density of development of the area. In Chicago's Loop, nearly all streets are included. For the remainder of Chicago, arterials are at approximately ½-mile spacing. The spacing then increases to an average of 1 to 2 miles as development becomes less dense. All arterials are considered to be at least four lanes wide in 1990.

CATS' proposed transit system is shown in Figure 2. The assignment network for transit contains 477.4 miles of CTA surface lines, 93.1 miles of CTA rapid transit lines, and 281.2 miles of suburban rail commuter lines. In addition, there are 1,220.2 miles representing express, suburban, school, and service buses, and artificial auto travel links feeding terminal points. The total transit system contains some 2,071.9 miles of facilities.

Results of Traffic Assignments

Highway and transit assignments were performed for each of the alternative land-use plans. Population in the internal area (CATS study area) varies from a low of 6,900,000 in the satellite cities plan to a high of 7,500,000 in the finger plan. Similarly, trips vary in the same manner. The satellite plan has 1.5 million fewer trips in the CATS' study area compared with the finger plan (Table 2).

Trip length varied, the highest being multi-towns with 5.7 miles and the lowest finger plan at 5.3 miles.

The multi-towns plan had the highest total passenger miles and second highest total passenger minutes. The trends plan ran a poor last. With 400,000 fewer people than the finger plan, a million more passenger miles and 29 million more passenger minutes were needed to satisfy the travel requirements generated by location of its people and activities.

The multi-towns plan had the highest index of total system efficiency, with the finger plan second. The finger plan had the highest index of transit system efficiency, however. Its activities were so arranged that the average trip time was lowest of all plans tested. The trends plan, again, was least efficient.

The vehicle equivalent miles of assigned volume (VEMAV) and the ratio of VEMAV to vehicle equivalent miles of design capacity (VEMDC) for arterials, expressways, and all facilities combined were computed for each plan (Table 3). Except for the satellite cities plan (which really cannot be compared), the finger plan has the lowest volume-to-capacity ratio of the three remaining. This means that the finger plan offers a higher quality of traffic service; that is, chances of being delayed by congestion and backups at traffic signals are less so that travel costs are, therefore, probably less. The total daily travel costs are considerably lower for the finger plan than for the multi-town or trends plan.

This analysis showed that the trends plan should no longer be considered as a possible comprehensive plan for the NIMA. It also indicated that the finger plan and multi-towns plan each had desirable transportation features. The multi-towns plan had the highest index of system efficiency and offered the highest average vehicular speed. On the other hand, the finger plan handled travel requirements for a slightly larger population with a minimum average trip time per person and at a considerably lower travel cost to the highway user. The organization of activity on the finger plan made it possible to satisfy trip desires with a shorter trip length. Based on this analysis, it appeared that the finger plan was slightly better than the multi-towns plan from a transportation point of view. Thus, at the December 1966 meeting of the NIPC, CATS reported that the multi-towns and finger plans were both satisfactory but that the finger plan offered the advantages of faster average trip time per person and lower highway travel cost.

EVALUATION OF ALTERNATIVE TRANSPORTATION PLAN FOR FINGER PLAN

The second major goal of this project was to establish whether a new transportation plan would be required for a particular land-use plan—or if the original CATS plan could best satisfy the travel demands of any of the planned situations.

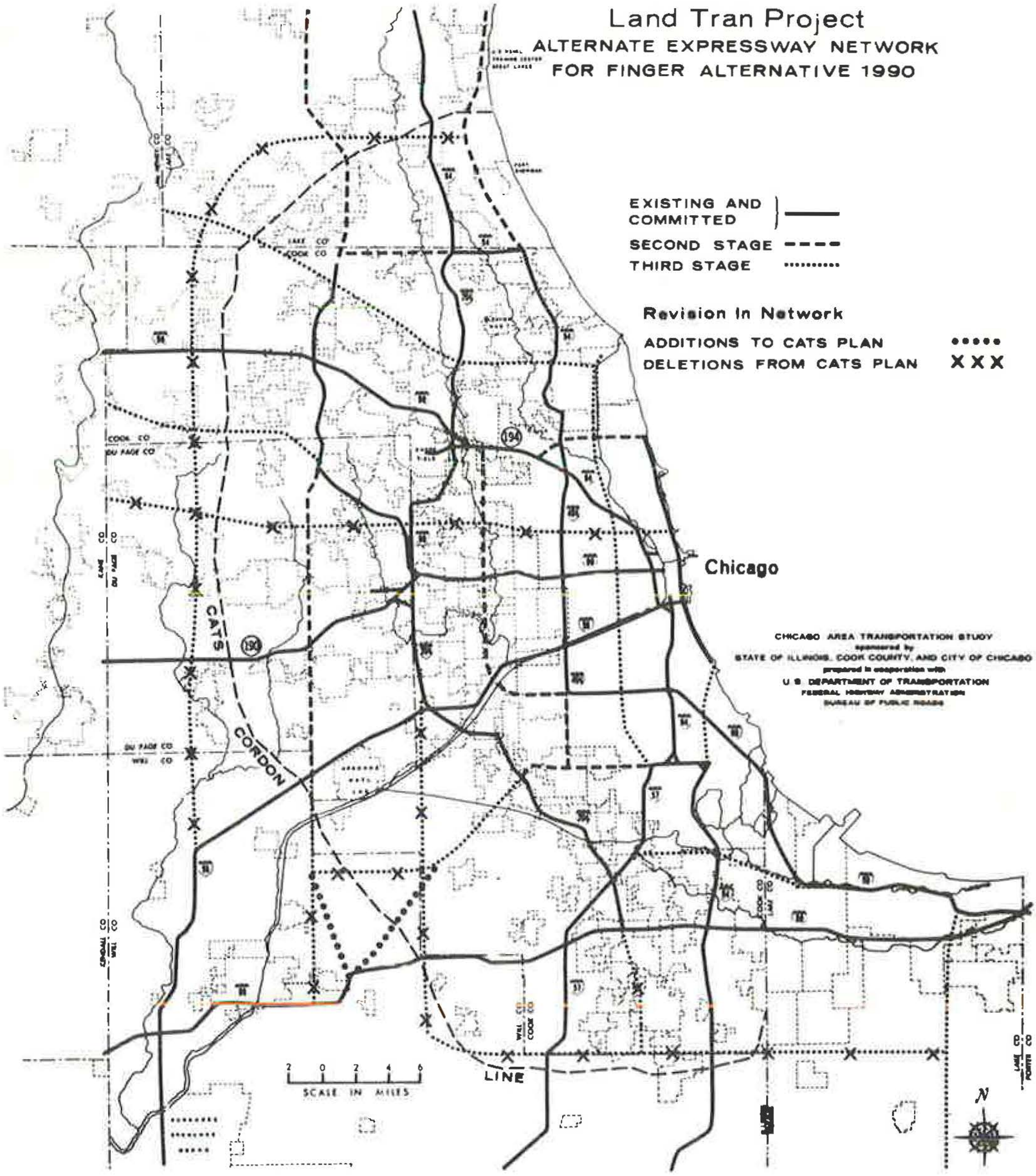


Figure 3.

As a result of the considerable testing of the land-use plans by NIPC and CATS, and the stated preference of those attending a public hearing in late 1966, NIPC chose the finger plan for further testing and development. CATS had agreed earlier to test a transportation network developed especially for the plan favored at the public hearing. The plan was to be developed jointly by the CATS and NIPC staff and certified by the directors of the two agencies.

Alternative Transportation Systems

An alternative highway and transit system was developed specifically to implement the planning goals of the finger plan; that is, the system was planned to encourage growth

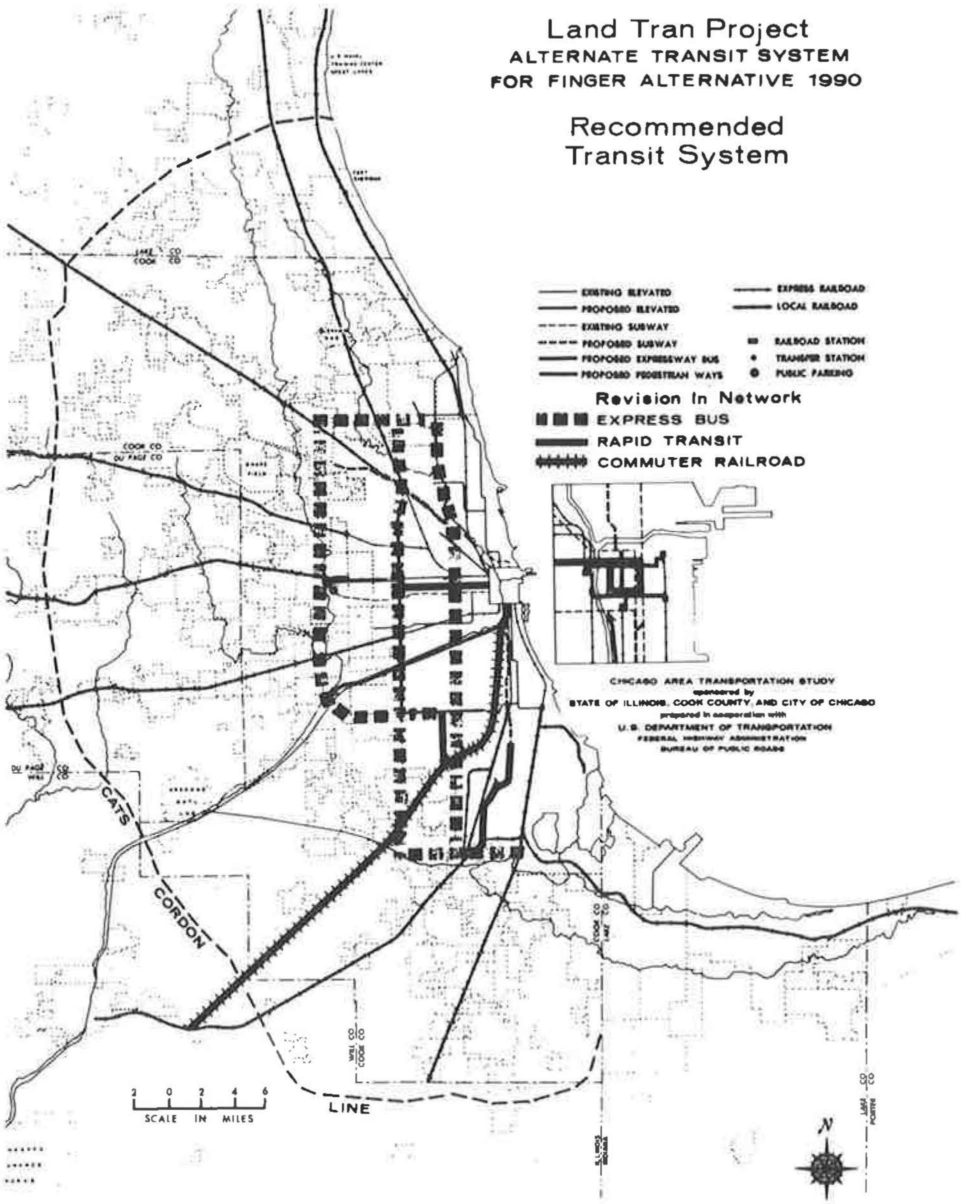


Figure 4.

and accessibility along the fingers and to discourage growth and the construction of transportation facilities between the fingers. Extensive changes were made to both the highway network and transit network. Sixty miles of expressways were deleted from the L-3 plan. To the transit system was added 115 route miles, including 60 miles of proposed express bus routes (Figs. 3 and 4).

Changes in the transit system's level of service required a re-run of the mode split for the finger plan trips. However, the resulting changes were minimal. Only 12,000 daily transit trips were added by the revised plan. This represented an increase of 1 percent in transit travel and a decrease of less than 0.1 percent of vehicle trips.

TABLE 4

SUMMARY COMPARISON OF HIGHWAY AND TRANSIT DATA WITH ASSIGNMENT RESULTS FOR ALTERNATIVE 2 AND REVISED ALTERNATIVE 2

Factor		Alt. No. 2 on CATS (06)	Alt. No. 2 on Revised Network (A8)
Population	Internal	7,569,400	7,569,400
	External	1,802,800	1,802,800
	Total	9,372,200	9,372,200
Passenger trips	Transit	2,366,353	2,378,200
	Vehicle	18,789,110	18,777,245
	Total	21,155,463	21,155,463
Average trip length	Transit	7.76	8.49
	Vehicle	4.96	5.01
	Total	5.30	5.40
Passenger miles	Transit	18,362,094	20,189,756
	Vehicle	93,184,300	94,023,100
	Total	111,546,394	114,212,856
Passenger minutes	Transit	104,961,206	102,903,100
	Vehicle	193,206,593	219,625,864
	Total	298,167,799	322,528,964
Index of total system efficiency		.374	.354
Passenger travel time (min/mile)		2.67	2.82
Equivalence speed (mph)		22.44	21.25
Index of transit system efficiency		.175	.196
Equivalence transit speed (mph)		10.5	11.8
Index of vehicle system efficiency		.482	.428
Equivalence vehicle speed (mph)		28.9	25.7
Minutes-trip total system		14.1	15.2

TABLE 5

SUMMARY COMPARISON OF VEHICLE ASSIGNMENTS 06 AND A8

Factor		06 Alt. No. 2 on CATS L-3 Net.	A8 Alt. No. 2 on Rev. L-3 Net.
Sent trips		12,978,504	12,971,145
VEMAV	Arterials	36,750,800	39,606,700
	Expressways	30,095,900	27,841,700
	Arts + Expwys	66,846,700	67,448,400
VEMAV/VEMDC	Arterials	0.85	0.92
	Expressways	0.77	0.80
	Arts + Expwys	0.81	0.87
Trips per mile	Arterials	13,000	14,100
	Expressways	57,900	60,700
	Arts + Expwys	20,000	20,600
Operating cost ($)	Arterials	983,000	1,064,200
	Expressways	924,900	851,800
	Arts + Expwys	1,907,900	1,916,000
Accident cost ($)	Arterials	541,200	594,400
	Expressways	70,100	65,400
	Arts + Expwys	611,300	659,800
Time cost ($)	Arterials	2,527,000	2,751,600
	Expressways	833,100	774,700
	Arts + Expwys	3,360,100	3,526,300
Total travel cost ($)	Arterials	4,051,100	4,410,200
	Expressways	1,828,100	1,691,900
	Arts + Expwys	5,879,300	6,102,100
Travel cost (cents) per veh. mile	Arterials	11.0	11.1
	Expressways	6.1	6.1
	Arts + Expwys	8.8	9.0

06 VEMDC for: Art. = 43,241,500 Route Length: Art. = 2822.6 miles
Exp. = 39,293,200 Exp. = 519.5 miles

A8 VEMDC for: Art. = 43,145,500 Route Length: Art. = 2817.2 miles
Exp. = 34,626,400 Exp. = 458.75 miles

Results of Assignments

Transit and highway assignments were run for the alternative transportation system. Table 4 compares the results of assigning finger plan travel demand to an alternative travel plan with the results of assigning the same travel demand to the original CATS plan. The population and trips are naturally the same for both plans. However, the trip length is greater for the alternative plan for both highway and particularly transit trips than it is in the CATS plan. In this case, calibration was the same and trip length differences are due only to network differences. Passenger miles and minutes are also greater. Consequently, the index of efficiency is less and average speeds are slower, while travel time and average trip time are greater. The alternative, however, demonstrates greater efficiency and speed for the transit system.

One percent difference in transit passenger volume caused a nearly 8 percent increase in passenger miles. Even with this increase in length, however, there is a decrease in transit travel time because transit speeds are high. Yet, combining transit and highway travel times brings the alternatives' passenger minute time up to a net increase of 8 percent. Clearly, the alternative cannot be favorably compared with the original CATS plan when subjected to the same evaluative criteria.

Results of assignments to the highway network show that the original plan has a lower volume to capacity ratio (VEMAV/VEMDC) and thus provides the best quality of service (Table 5). Again, the chances of congestion, delays, and subsequent travel costs are less on the original system.

The alternative highway plan had the effect of shifting traffic from the expressway back to the arterial system. This is done at a substantially higher per-mile travel cost, as reflected in the cost summaries. Expressway travel is down 9 percent and arterial travel up 8 percent in the alternative plan. On the average, an additional 1100 vehicles have been added to each mile of arterial.

This resulted in relatively the same travel demand being satisfied on a lesser system. The expressways in the revised system show no increase in travel, but rather a 9 percent decrease. Trips were not diverted to the transit system or to other expressways but used surface streets instead. This necessarily increased arterial travel, thus increasing travel costs.

The results of the traffic assignment clearly indicate that the original CATS highway plan (L-3) is superior to the revised plan when considering user costs only. The L-3 plan accommodated the same number of trips as the revised plan, but with a higher average speed, less congestion on the arterial system, and a generally higher quality of service. The overall result of these positive benefits means lower total travel costs to the users.

It has been demonstrated that plan L-3 provides superior travel performance. Since it is axiomatic that the rate of return for each mile of new facility declines as more facilities are built, it must be determined whether the difference in user costs gives a significant return on the additional capital investment required by the L-3 plan when compared to the alternate. This is done by using a standard economic technique known as marginal cost analysis.

Marginal Cost Analysis

Marginal cost analysis is a rather simple, yet effective, analysis tool. This method involves a comparison of the benefits from a marginal (or last item) investment with the additional capital costs required for the investment. As a policy, CATS selected 10 percent as a minimum required rate of return before considering the investment worthwhile. This is a high rate of return compared with the 3 to 6 percent rate return now current for a variety of low risk investments such as insured home mortgages and municipal bonds. The 10 percent rate was chosen purposely as a means of insuring fair investment consideration of a variety of other public works besides transportation facilities. Thus, if an improvement meets this economic test, the public official can be reasonably assured that the project is a worthy investment of public funds.

Capital Costs—As used in this analysis, capital costs are construction costs only. This is the cost to purchase right-of-way and build a facility, but not to maintain and operate it. Also, the so-called intangible social costs are not included. These are the supposed costs of disrupting a neighborhood or relocating people. In point of fact, however, even if these costs could be quantified, they could not be used in the evaluation of a total system. Since the plan considered here is a "corridor plan" with no specific location chosen, it would be meaningless and entirely out of scale to estimate these social costs. This is not to say that the estimation of these social and economic implications of building transportation facilities should be ignored, but rather that these are properly individual route location problems rather than systems indices.

The cost of the highway plan is based on typical per-mile costs of facilities completed since 1960. These per-mile costs are developed by CATS ring and applied to the mileage of proposed expressway in each ring. While these costs are not exact, because no one can say exactly what the costs are until the facilities are built, they are reasonable estimates and are applied in the same manner to each highway system. The application of these ring per-mile costs showed that the L-3 plan would cost $194,870,000 more than the alternative highway plan.

Because of its additional proposed express bus mileage and rail facilities, the revised transit plan had a higher capital cost than the CATS original transit proposal. The capital cost estimate for the revised plan was made conservatively so as not to produce an answer biased toward the CATS plan. For example, it was assumed that no additional construction cost would be required for express buses for the traveled way. However, a cost of $800,000 per station was assessed. It was assumed that where proposed rail lines were near existing rail lines, no additional capital costs would be required, but that the existing line would be used. An example of this is the proposed southwest rail route in the corridor of the Norfolk and Western Railroad line. Adding the proposed 115 miles of new transit facilities would cost an estimated $76,600,000 more than the CATS plan.

User Costs—Daily user costs for the CATS L-3 network are $5,879,300. For the revised plan, daily costs are $6,102,100, or $222,800 higher per day. Assuming 339.5 average weekdays per year, the L-3 plan gives an annual cost savings of $75,640,600 per year.

Approximately 70 percent of these savings is in personal time costs. Whether time savings should be used is a widely debated issue. However, if value of personal time is left completely out of the analysis, the savings in accident and operating expenses total more than $20,000,000 per year. To ignore value of time implies that time savings are unimportant to the system user—an implication which is indisputably incorrect.

TABLE 6

SUMMARIZED COST IN DOLLARS FOR INDIVIDUAL AND COMPOSITE TRANSPORTATION SYSTEMS

Factor	Alt. No. 1 Original L-3	Alt. No. 2 Revised Plan	Savings Alt. 1 Over 2	Add'l. Cap. Alt. 1 Over 2
Capital required				
Transit	116,100,000	192,700,000		-76,600,000
Highway	2,007,000,000	1,812,130,000		194,870,000
Total	2,123,100,000	2,004,830,000		118,270,000
Annual capital required	23,390,193	22,087,212		1,302,981
Annual transit user costs[a]				
Oper.	229,791,900	247,639,800	17,577,900	
Time	570,149,100	448,969,400	-11,179,700	
Total transit user costs	799,941,000	806,609,200	6,668,200	
Annual highway user costs				
Oper.	647,732,050	650,482,000	2,749,950	
Acc.	207,536,350	224,002,100	16,465,750	
Time	1,140,753,950	1,197,178,850	56,424,900	
Total highway user costs	1,996,022,350	2,071,662,950	75,640,600	
System user costs	2,795,963,350	2,878,272,150	82,308,800	

[a]Daily cost times 339.5 average weekdays.

In this analysis, the transit operating costs are equated with user costs since the user must ultimately pay these costs through his fare.

Operating expenses (i.e., fares) would be $52,571 higher per day on the alternate transit plan while, on the other hand, reduced travel time saves the user $32,930 per day. However, this gives an overall daily cost of $19,641 more for users of the alternative transit network. On an annual basis, this is $6,668,120 higher than afforded transit users in the original CATS proposals.

Marginal Rate of Return—The marginal rate of return (MRR) was calculated for the highway and transit plans and, also, for a combined or composite transportation system. It was stated earlier that the CATS highway plan required an additional investment of $194,870,000 as compared with the alternative. However, the L-3 plan yielded annual benefits of $75,640,600 or 38.8 percent—an impressively significant return. Table 6 gives details of user and capital costs of the two systems.

Even ignoring the benefits of time savings, the MRR based on $20,000,000 annual operating and accident cost, showed savings slightly in excess of 10 percent—still meeting the minimum rate of return.

For the transit analysis, the proposed alternative system represented the marginal investment since it was estimated to cost $76,600,000 more than the CATS proposal. In this instance, however, the cost to transit users is greater for the alternative, largely because of increased operating expenses for a larger network of facilities. Thus, a negative MRR resulted from this added transit investment.

Based on this analysis, the transit proposal is not justified. However, since the new plan minimized total travel time, it appears that the CATS transit proposals need restudy. It is possible that addition of new rapid transit proposals could reduce travel times and still meet the test of a 10 percent marginal rate of return. Thus, CATS proposes to review its transit proposals with the possibility of modifying its original recommendations. The improvements in operating efficiency in the proposed transit system were overpowered by the additional capital investment required. If more people could be induced to use transit in preference to the automobile, the investments could be economically sound. Experience has shown that new rapid transit routes cause a greater intramodal shift of transit users than a shift away from the use of the automobile.

An MRR was also calculated for the combined highway and transit networks. This calculation combined investment costs of both systems for each plan and compares the return from the combined user costs. The negative MRR for the alternate transit plan coupled with the high MRR for the CATS L-3 highway plan result in an almost unreal composite MRR of 70 percent.

Results of the traffic and economic analysis indicate that the L-3 plan is the best of the two plans tested in providing transportation service for the land-use distributions in the finger plan. Results are conclusive with respect to the highway system. Conclusions about the transit plan are not as clear. It seems advisable to review the CATS proposal with the possibility of including additional rapid transit recommendations, and suggestions for better coordination and transfer arrangements with the private companies and the Chicago Transit Authority.

SUMMARY AND CONCLUSION

This report has described CATS' role in a land-use transportation study carried out jointly with the Northeast Illinois Planning Commission. It is a companion to the NIPC summary of their work on this project (1).

It was pointed out that this project gave CATS the opportunity to test its transportation plans (based as a forecast) for their adequacy to support a land-use plan or plans. The project involved evaluating the adequacy of CATS plans to serve four (originally five) land-use plans and to determine which plan was best from a transportation viewpoint. It also afforded an opportunity to test CATS plans against an alternative proposal specifically designed to serve one land-use plan (in this case, the finger plan).

The report has described how CATS converted land-use and socioeconomic data to travel demand, distributed and assigned traffic to networks, and prepared traffic and economic evaluation of the results. The methodology and the models were briefly de- described, as were the results of the evaluation.

Results of the testing efforts lead to some definite conclusions:

1. The CATS transportation proposals would adequately serve any of the proposed four land-use plans—within the CATS area.
2. The finger plan was the best land-use plan from the point of view of transportation. It organized land uses so that trip lengths were shorter and had a generally higher quality of service in the networks.
3. The CATS L-3 highway plan was clearly superior to an alternate plan devised for the finger plan. This conclusion is supported by the traffic analysis and by the marginal rate of return economic analysis.
4. The alternative transit plan did not provide marginal benefits to justify the additional capital costs. However, it did have desirable qualities, which suggests that CATS should restudy its transit proposals.

Results of these tests provided a large body of statistics in support of the CATS expressway proposals. It appears that a well-conceived plan based on spacing principles will work under a variety of land-use configurations. While there may be some severe local dislocations due to shifts in density within the CATS area, average trip densities did not vary greatly from one plan to another.

A conclusion not related to the technical aspects of the project is that two separate agencies can, each with particular skills and specialities, successfully work together in preparing and evaluating land-use and transportation plans.

REFERENCE

1. The Plan Study—Summary of Methodology. Northeastern Illinois Planning Commission, Oct. 1967.

Transportation System Development and Evaluation as Practiced in Seattle

STEPHEN GEORGE, JR., Director, Albuquerque Metropolitan Transportation Planning Program[1]

•A FEW months ago, the final published Summary Report (1) of the Puget Sound Regional Transportation Study was distributed. The 115-page document could not have been designed to give full justice to the initial four years' effort of a comprehensive land-use and transportation planning program. The interdisciplinary staff that was assembled in 1961 was able to expand on its previous experiences and sharpen up the tools of system development and evaluation, but the final report does not cover this area. What did we learn from this important exercise that can be applicable to a great number of medium-to-large-size metropolitan areas?

As a preamble to the general conclusions that I wish to emphasize, let me digest the study scope and the area's physical restraints before outlining the system development criteria and system evaluations.

SCOPE OF PSRTS WORK PROGRAM

The overall objective was to "formulate a transportation plan as part of a general development plan for the region." Alternative land-use patterns and alternative transportation systems including mass and rapid transit were developed, tested, evaluated,

[1]The author was associated with the Puget Sound Regional Transportation Study as Deputy Director and Principal Consultant from April 1961 to August 1967.

and an integrated land-use and transportation plan was recommended. The 1961 regional population was 1.5 million, forecast to grow to 2.75 million by 1990. The four-county area of study included more than 40 individual cities. Comprehensive studies of land use, population, economic base, traffic conditions, and travel patterns were the source of the facts and figures, which were stratified by 33,000 grid blocks, 662 analysis zones and 60 analysis districts within the study area of 1500 square miles, of which 200 square miles was in water bodies.

The 65-miles long study area has four separate central cities, varying in population from Seattle, over a half million, to Bremerton, 30,000. The area is divided by the very large body of water known as Puget Sound. Because of the many lakes and the Sound, topography suitable for major transportation corridors is limited and terrain difficult to cross in desired locations.

Over the years, development of the transportation system in Seattle was promoted north and south. With the construction of the north-south Interstate highway, the area found itself with a great number of east-west arterial deficiencies to the freeway interchange points. Because of the extreme depth of the Sound and lakes, the Northwest has pioneered in concrete floating bridges, a more economical but still costly solution to cross-Sound and cross-lake travel needs.

SYSTEM DEVELOPMENT CRITERIA AND STEPS

If the future travel is to be accommodated by planned transportation systems, all components—streets and highways, mass transit, ferry boats, bridges, and parking facilities—must be considered together. Although PSRTS studied each of these components and their interrelationship, only the transportation system made up of streets and highways will be outlined, for simplicity of this discussion. In addition, let us assume that a future trip table is available to us outside of this discussion. Now that we have access to a current and forecast trip table, with which system do we begin the assignment process and what criteria and steps do we follow?

System Development and System Evaluation Are Inseparable

System development is the evolution or transition of today's given transportation system into a planned or recommended system to serve a future time period. This time period could be a short-range plan such as 1975 or a long-range plan fcr 20 to 25 years hence. Systems research concepts are now utilized to analyze, update, modify, or extend a given current transportation system into a more complete network needed for tomorrow.

System evaluation is a process of assessing the adequacy of any given system to satisfy desired criteria and overall goals. Experience of a number of transportation studies has indicated the preference for a continued systems evaluation of a logical series of systems under development.

System development coupled with continued system evaluations can be likened to zeroing in on a bull's-eye from the outer ring to the inner rings directed toward achieving a balanced capacity-demand network. The more we know about system mechanics, the more we realize the need to develop a system that will insure maximum utility of what we have today. Formulating alternative systems for the sake of alternatives, or to appease one pressure group or another, is a time-consuming and costly process that too often develops more data than can be digested and properly utilized in the development process. If one attempted to analyze several alternatives, for example, years could be required to evaluate the actual differences between them.

A continued system evaluations process at PSRTS followed the "minimum assignment-maximum analysis" technique of system development. "Maximum analysis" in this context includes (a) conventional system data plotting techniques, i.e., volume by link and volume-to-capacity ratio calculations by link, and (b) usual tabular summaries to measure overall system performance. By this process, continued iterations from a starting point increases the knowledge of the analyst of his particular system under development. I believe this has been demonstrated to reduce the overall time to reach a recommendation. The process also insures maximizing the full utility of what is given.

Application of Level of Service Concepts and Standards

The traffic assignment procedure usually begins with the assignment of current origin and destination traffic to a base year network. For this step, current travel time or speed data are collected, and usually these same speed data are inserted into the future system under study.

We all know of segments of our own existing networks where obvious improvements have lagged behind the need for one reason or another. In my own city, one of our most congested segments follows the centerline of the City-County boundary line. Delayed improvement has not materialized because of the complexity of developing a joint project to be assessed against the abutting property owners on both sides of the arterial. Using the resulting low speed or level of service will only guarantee perpetuating the condition any analyst would desire to eliminate. A constructive estimate of traffic demand will not materialize if a less-than-desired speed is assumed on a system.

In some cities, the traffic engineer has found himself attempting to increase traffic capacity on facilities improperly spaced due to the lack of a more constructive approach. One of the best means to overcome such deficiencies in today's conditions is to apply a level-of-service speed standard in the traffic assignment package. A level-of-service speed can be developed from the analysis of comprehensive travel time studies. Such standards should be both attainable and desirable speeds by functional classification and can be properly used to help measure existing and planned facility adequacy. The utilization of desirable speed standards in planning for future regional facilities will insure the proper appraisal of all needs without reference to current deficiencies. A level-of-service speed standard, if properly applied, should identify relative desired differences between freeways, major streets and collectors. If one of our goals is to develop "best and optimum" systems that will encourage a desirable environment, design and planning must be separated the same way that "administrative planning" is separated from "scheduling" in the critical path method applications.

Level of service in City A is not the same as in City B. In general, level-of-service speed differentials by functional classification are further apart in the West when compared to the East. This fact is a reflection that urbanization is older in the East and younger in the West and typifies people's choice and desire.

Basic Building Blocks of System Planning

In beginning a future-year system analysis, some studies start by testing the adopted major street and highway plan prepared by a local planning commission and find themselves making drastic surgery to it in the traffic assignment process. Only in recent years have such plans by planning commissions been based on a careful and factual study.

In Seattle, the basic building block for all study systems was the existing base year network of 1961. To this was added the additional facilities that were firmly and definitely committed, judged against uniform criteria across the multi-county study area. The committed system constituted the first future-year system that was also utilized to generate a trip matrix in the trip distribution model.

The Seattle study actually utilized the interim committed system for its accessibility inputs in the land-use allocation model. Since the full land-use impact of a committed transportation network is not felt until after its implementation, intermediate-year system parameters (year 1975 is between 1961 and 1990) were found to best satisfy the land-use allocation and the trip distribution model requirements.

In planning a transportation system, we too often "muddy the waters" by expecting design output from each assignment run. For example, a capacity restraint assignment was not used in the PSRTS development simply because the multiple screen line analysis resulted in significantly deficient capacities across the metropolitan area. The capacity restraint assignment process can only be used if you have some place to divert the traffic and should be postponed to refine the design data needs after a future system has been properly developed. A few of us believe that capacity restraint assignments, in general, have been promoted too early in the planning process and are not a panacea to system development.

Use of Selected Link and Screen Line Analyses

Discussion with other practitioners and computer service bureau staffs indicates that selected link loads and selected zone loads programming options are not used often enough in our computer library. Those of us who have used it extensively feel that we would still be "spinning our wheels" in constructive system development and evaluation without this powerful tool. How can you use it to your advantage?

Selected link or zonal analysis permits the analyst to isolate the network link or zone in question and graphically summarizes the problemized travel pattern. If the link is significantly overloaded, knowledge about the component of the traffic demand creating the overload can be compared with the desires. Such analysis will contribute significantly to evaluating alternative extensions to the network and to selecting the more reasonable link updates to resolve the problem. Too often problems in one corridor are the result of deficiencies in another corridor 90 degrees removed from the problem.

Selected link analysis permits the analyst to define the zone of influence of that link with supplemental knowledge about its probable usage characteristics.

With selected zone loading analyses, the analyst can determine at what level of increase in traffic attractions in a zone interchange design breakdowns will occur. Any desired modification in trip generation can be inserted in selected system design analyses by way of a selected zone loading. Traffic flows can be increased or decreased in special analysis areas, provided the particular distribution pattern is still valid.

Screen line evaluations still appear to be the backbone step of system development. Comparisons of demand vs available and assumed capacity tell the analyst whether the total overload is of freeway or expressway proportions. Evaluations also identify whether an adjacent screen line can absorb the diversion of traffic provided level-of-service standards are modified or linkages are modified.

Three principles are involved in this activity, which can be summarized as follows:

1. Initial network evaluations should be viewed as part of normal study procedures to develop staff analyst comprehension of system operating characteristics.
2. Continued system evaluations should be keyed to a demand-type analysis until system balances are resolved as contrasted to performance analysis where individual link operational analysis is emphasized, i.e., capacity restraint.
3. Network performance analysis can then proceed to refine or perform detailed evaluations of the recommended systems and/or their alternatives for implementation purposes.

GENERAL CONCLUSIONS APPLICABLE TO OTHER METROPOLITAN AREAS

The Limits of Future Population and Employment Distributions Are Not the Same

The Seattle study developed a population distribution model based on a statistical evaluation of the factors that determined the growth that occurred in the region between 1950 and 1960 (2). Five factors were isolated as attracting and influencing the locations of population growth. An employment distribution model was also developed (3, 4). Based on this research and application, radically different future population distributions did not result when alternative land-use patterns were tested in 1964. Seattle is now experiencing accelerated growth due to the Boeing expansion program and significant modifications in the distribution of population in the alternative land-use plans are not indicated. Continuing industrial expansion is following the study's land-use plan B concept.

Dispersion of residential growth is a general trend across our land. From Albuquerque to Minneapolis, land ownership characteristics, quality of the homes desired, extent of utilities, and the intangible amenities are becoming more important than access to place of employment.

With regard to employment distributions, the Seattle study demonstrated that the alternative plan responsive to regional goals, if supported by policies to guide regional development, can result in a different distribution of employment location, and thus have a major effect on traffic capacity needs. Based on a more desirable development pattern, less capacity may need to be provided for future travel needs to the CBD's and more capacity may need to be provided on major corridor facilities in today's suburban areas.

The Importance of Alternative Land-Use Patterns

Let me briefly discuss the importance of alternative land-use patterns as they permit flexibility in transportation planning. With a given mileage of committed major facilities, quantification of alternative land-use patterns when translated into system demand can significantly strengthen the basis for transportation facility recommendations by insuring the identification of the major corridors of travel. If committed mileage is significant, as was the case in Seattle (and such committed mileage can be a major shaping tool for land-use activities), alternative land-use or employment patterns will help delimit or confirm the major corridors of travel and provide a workable volume demand range on each facility applicable to the limits of the alternatives studied. A more flexible transportation plan can thus be developed should one part of the study area develop one way while another part follows more closely the alternative pattern.

If, on the other hand, committed mileage is limited, as appeared to be the case in some of the Upstate New York studies, and the impact of transportation on land-use activities is minimal, the analyst will not know if he has delineated all possible major corridors of future travel unless he has studied the limits of the reasonable alternatives that may result from a different distribution of population and employment. Either way, regardless of how extensive our committed transportation systems are, the inclusion of alternative land-use patterns study in our continuing program must be emphasized, and definitely should be included in more land-use and transportation planning programs.

REFERENCES

1. Puget Sound Regional Transportation Study, Summary Report. Puget Sound Governmental Conference and State of Washington, in cooperation with the U. S. Bureau of Public Roads and the U. S. Department of Housing and Urban Development, Seattle, Sept. 1967.
2. Graves, Charles H. Two Multiple Regression Models of Small-Area Population Change. Highway Research Record 102, p. 42-53, 1965.
3. Graves, Charles H. Forecasting the Distribution of 1985 Population and Employment to Analysis Zones for Plan A. Staff Report No. 15, Puget Sound Regional Transportation Study, Seattle, 1964.
4. Joshi, Rajanekant J. Distribution of 1985 Population and Employment to Analysis Zones for Plan B. Staff Report No. 17, Puget Sound Regional Transportation Study, Seattle, 1964.

NOTE

During the formal presentation at the Highway Research Board, the author referred to ten 35-mm slides as graphical examples of system development results extracted from the Puget Sound Regional Transportation Study Summary Report of September 1967. The slides are available on loan from the author. They are not available from the Highway Research Board.

Discussion

THOMAS B. DEEN, Alan M. Voorhees and Associates, Inc.

•ONE must always begin any criticism of another's work on network evaluation with an expression of humility and apology for ignorance in the face of this awesome task, in which everything affects everything else and which covers such a broad spectrum of human activity. Besides, anyone actively practicing as a transportation planner has at one time or another been, and will likely in the future be, forced by lack of time, budget, and knowledge to use techniques similar to or sometimes worse than the ones he criticizes at Highway Research Board meetings. In any case an annual soul-searching and confession is probably useful; otherwise we might complacently fall into the trap of believing that we really do have reliable objective techniques for systems evaluation.

Campbell's paper covers testing of various regional land-use schemes as well as evaluation of networks required to serve a given land-use plan. George's paper covered only the networks and since this session is sponsored by the Transportation System Evaluation Committee, I will restrict my remarks to network evaluation problems. It is interesting to note in passing, however, that Campbell's land-use analysis concerns several radically different regional land-use plans. In contrast, the Puget Sound analysis concludes that substantially different land-use schemes are not possible.

The two papers presented here were a fortunate selection since they are better than average examples of the two evaluation techniques currently in vogue in the United States. George presented the "capacity to meet demand" school of thought while Campbell illustrated the "economic evaluation" school.

The "capacity to meet demand" procedure aims at developing a network that meets the projected travel demand for some future date, eliminating or reducing to a minimum segments with either deficient or excessive capacity, providing directness of movement, and keeping construction costs to a minimum. It involves several implicit assumptions, including

1. That the objective of a transportation study is to devise a network that will accommodate all projected travel demands; and
2. That all travel demands are worth the cost of providing facilities to meet them.

The process employed in Campbell's work implicitly questions whether either of these assumptions is justified, since he goes to great length to determine whether the extra costs of new facilities can be recovered in user savings. Furthermore, the term "travel demand" deserves scrutiny. Demand will be high or low depending on the facilities available and the price charged for their use. Are we really required to meet all travel demands, however trivial, at any cost? Possibly not; "in a world in which resources are limited we make no attempt to meet all demands. An auto manufacturer is not interested in meeting all demands for his cars, but only the demand at the price that will cover his production costs. An investigation aiming to ascertain all projected travel demands neatly avoids considering the level of demand that should be met" (1).

Our greatest problem in transportation systems evaluation stems from our lack of knowledge on "which demands are worth meeting" and therefore "which facilities are worth building." Our problems are different from the makers of autos or the suppliers of electric power, for example, because the price of cars and electricity to consumers is directly related to individual consumption. However, the price of the use of roads or transit facilities in our cities is unrelated to whether one is using a high-cost or a low-cost facility, or to whether that use is during peak or off-peak hours.

When all users pay the same, the result is equivalent to the situation of an electric company that decided to eliminate individual electric meters and bill customers not on the basis of individual consumption but by measuring total power usage and charging each consumer an equal part of the total bill. Not only is this inequitable; more importantly it will eliminate incentive for conserving electricity. Many new homes would be heated with electricity, since an individual's cost would not be increased by a decision to install electric heating. Demand for power would soar, and new investment would be needed for new generating facilities. There would be no real basis for determining the proportion of total resources that should be devoted to power generation.

Such considerations are leading the British into serious thinking of road pricing mechanisms which closely tailor transportation pricing to costs. Whether they can overcome the technological, political, and financial barriers to such a scheme remains to be seen. We probably can agree, however, that the time for such a move in the United States has not yet come. Nevertheless, we must give consideration to the economic elasticity of traffic demands if we are to make meaningful network evaluations.

That George's "capacity to meet demand" procedure, or variations of it, has been employed in many U. S. cities with apparent success can probably be attributed to several factors:

1. We are a wealthy nation and have committed large resources to providing for new transportation facilities;
2. Economically unrestrained travel demands are not so large as to be impossible to accommodate because of the low-density nature of our cities; and
3. Our cities are not so intensely developed but that new rights-of-way can usually be developed without too much trauma.

In other words, in most cities the accommodation of all travel demand is possible, though we must add, not necessarily economic.

But in our largest, densest cities it is becoming increasingly evident that all demand cannot be accommodated at a reasonable cost. Even "capacity restrained" assignments show projected volumes out of proportion to the facilities that can be provided considering political, financial, and sociological realities. In these situations one is forced to re-think goals and objectives. It is clear that in these cases demand is going to be restrained by price, only the price will be in the form of congestion and time losses instead of money. Unfortunately restraint by congestion makes no discrimination between essential and nonessential travel. All are equally restricted.

Campbell's evaluation is a more sophisticated, complex approach to the problem that has better theoretical underpinnings than George's. It attempts to develop the "least cost" system, considering all the user costs including time spent traveling, and then further uses marginal cost analysis to test the least-needed system increments. Its deficiencies seem to be that:

1. It puts heavy emphasis on factors quantifiable in economic terms such as time, operating costs, accident costs, and construction costs. It thus tends to de-emphasize other factors such as neighborhood disruption and displacement costs and environmental aesthetics. Consideration of these costs is relegated to the route location phase; it is suggested they should play no role in system planning. This is a difficult position to maintain, because the status of the highway network in a number of our larger cities is in jeopardy as a result of these factors.
2. It generally assumes no elasticity of demand with supply. Trip lengths are fixed as between alternative networks which must, to some extent, tend to bias the analysis toward larger systems. Time and operating cost savings from larger systems will appear bigger if no travel is induced by the larger system.
3. The assumptions on costs that must be made to carry out such alternative analysis with computational efficiency are sufficiently coarse as to invite questions. Sensitivity analysis on some of these costs to see their effect on conclusions might be justified.
4. All our techniques are coarse when compared with the precision required to evaluate the merits of land-use transportation plans involving carefully structured metrotowns around transit stations. The scale at which we conduct tests of regional plans is in contrast to the "walking-distance scale" critical to the concept of planned towns. In this regard our examinations can be compared to a person searching for a pin in the dark while wearing fur-lined mittens.
5. If an effective trip-pricing mechanism were in use that could input differential costs to users in accordance with their use of individual facilities and for different times of the day, one would assume that the less essential trips might be deferred to less congested time periods, use cheaper facilities, go a lesser distance, or indeed not be made.

If this is so, a different level of travel demand would result, which could produce a different "least-cost" network. While such thinking is only academic until the time that such pricing is feasible, it is useful to reflect that the optimum networks derived through Campbell's procedures might be less than optimum under different pricing and financial policies, and possibly result in even lower total costs.

REFERENCE

1. Roth, Gabriel. Paying for Roads. Penguin Books Ltd., Middlesex, England, p. 16-17.

The Community

Systems Evaluation: An Approach Based on Community Structure and Values

CHARLES C. SCHIMPELER, Technical Director, Louisville Metropolitan Comprehensive Transportation and Development Program; and
WILLIAM L. GRECCO, Professor of Civil Engineering, Purdue University

This paper presents a procedure for the evaluation of alternative transportation system design concepts based on a comprehensive, weighted hierarchy of community development criteria. Existing techniques for alternative plan evaluation are discussed, along with several potentially powerful normative procedures for system design.

The basic decision model relates to the evaluation of alternative design concepts by a single group of professional planners on the basis of a single set of weighted community decision criteria statements. Extensions of the basic model relating to a possible stratification of statements of value by socioeconomic groups and a possible stratification of planners are indicated. Necessary discussion of community decision structure, formulation of community decision criteria, and weighting of those criteria are summarized.

The decision model procedure is applied to three alternative systems design concepts for the transportation plan in the Louisville metropolitan area. Obvious extensions of the research are identified and applications of the procedures in land-use form and plan analysis, transportation corridor analysis, and detailed transportation system evaluation are discussed.

•ALTHOUGH a great deal of sophistication has been reached in the urban transportation planning process, this same level of sophistication has not been reached in plan evaluation. With regard to this general field of research, certain focal points within the problem area have been isolated. They are (a) criteria for evaluating alternatives, (b) techniques for identifying objectives, and (c) use of models in the design process.

This paper presents a technique for utilizing a weighted hierarchy of community decision criteria, or community goals and objectives, in a systematic evaluation of alternative transportation system design concepts. Heretofore, criteria utilized in plan evaluation have generally been easily quantified economic considerations. An alternative to that approach is considered here which utilizes a broader, more comprehensive class of criteria, including social values along with traditional economic considerations. Applied decision theory is used to establish orderly methods of making comparisons between the various alternative design concepts or philosophies.

A group of professional land-use and transportation planners establish effectiveness values for the design concepts relative to each item in a comprehensive statement of community decision criteria. The decision model utilizes these effectiveness values

---

Paper sponsored by Committee on Transportation System Evaluation and presented at the 47th Annual Meeting.

along with utility values associated with each element in the criteria set and proposes for adoption that plan possessing the highest aggregate "plan effectiveness" as defined herein. Various interesting techniques relating to this approach have been published previously and are summarized here. Several proposals for mathematical programming procedures for use in systems design are discussed that may prove useful in eventually structuring a truly optimal approach to system design and evaluation.

The central problem considered in this research is the evaluation of alternative system design concepts by a group of professional planners and engineers on the basis of the probability associated with goal achievement, assuming the adoption of each of the three alternative proposals. Effectiveness values for each alternative with respect to each of the criteria are presented and the rationale associated with the development of these values is given. Detailed presentations relative to community decision structure, statement of community decision criteria, weighting of the elements in a comprehensive hierarchy of community decision criteria, and the statistical analysis of scaling or weighting techniques are beyond the scope of this research; however, results obtained in these areas are summarized insofar as they relate to this paper.

Direct application of these procedures in land-use form analysis, land-use plan evaluation, detailed transportation plan evaluation, and transportation corridor analysis is discussed. Interesting extensions of these procedures in the areas of mathematical programming and more detailed and explicit definition of "yardsticks" for measuring plan effectiveness are presented.

EXISTING TECHNIQUES

Current research and practice have attempted to present means of evaluating alternative proposals based on impact analysis, benefit cost analysis, and other largely economic approaches. Although these purely economic approaches are not proposed for use in this research, their merits are recognized. This research relates to an approach utilizing elementary decision theory (several models are presented later) based on a comprehensive hierarchy of decision criteria rather than a set of decision criteria limited to economic considerations.

Several interesting schemes for system evaluation are presented, followed by a section relating to normative procedures that may prove to be valuable in system design. The central problem in urban planning is the development of planning proposals that satisfy, to the greatest extent possible, the stated goals and objectives of the community within realistic constraints. Mathematical programming provides a framework within which such optimizing or normative planning procedures may be carried out.

Alternative Plan Evaluation

Alexander (2) relates the development of a physical form in a manner consistent with the achievement of stated goals. Although the approach is not recognized as a workable, quantitative tool, it sets the framework for the development of a physical form based on the criteria of achievement of stated planning goals.

The theory of design proposed is plan evaluation based on goals. The form is the design solution which fits the problem, called the context. There are a number of variables contained within the context which the form must satisfy; the better the form meets all these variables (criteria), the better the design solution is for that particular problem (context). Each meeting of criteria is called a fit of that form; each lack of meeting is a misfit. For example, the context of an urban freeway has many variables—beauty, economy, neighborhood continuity, capacity, safety, durability, etc. A freeway (the form) is just as good as it meets (fits) these variables. The difficulty arises in adjusting the form until it best fits all the variables of the context at one time. Adjustments of one aspect of the freeway (e. g., economy) to get a better fit may have ramifications in many other aspects (e. g., beauty, safety) causing a lesser fit overall. Thus these are links between the variables in the form, which may be strong or weak, such that altering one causes the alteration of others. These interconnections may be represented by a mathematical graph. If these links can be recognized, and the system as represented by the graph can be broken down into a series of subsystems, alterations of variables within any subsystem set up ramifications along the links, which are dampened between subsystems because of the weak connections between them. Thus the variables of the form can be intelligently altered without the effects spreading in an unknown pattern to all variables. A mathematical formulation is set out to "optimize" this complete system and satisfy all variables at once as best they can be.

It is an interesting concept and may well have direct bearing on plan evaluation, each (with its objectives) being a variable with potential misfit, and the planned solution (form) must be altered intelligently to best satisfy all criteria.

Klein and Meckling (17) present a study of the application of operations research to development decisions. The fundamental approach relates to the selection of courses of action in initial stages of planning, consistent with a wide range of possible desirable alternative developments in effectuation. By this approach, the choice may be narrowed as decisions are made with progressive development. Due to the foresight of the earlier decisions, this development will be consistent with the overall objectives of the program.

The President's Water Resources Council report (21) considers the problem of policies, standards, and procedures associated with the formulation, evaluation, and review of plans for the use and development of water and related land resources. The publication defines general policies for evaluation at a national level and has no direct applicability in this research. Quantitative policies are not established; instead, broader administrative criteria are considered. The work could have some applicability in regional planning of very large regions.

The Southeastern Wisconsin Regional Planning Commission (23) defines planning as a rational process for formulating development goals and objectives. Development objectives should incorporate the combined knowledge of many people who are informed about the planning region and should be established by elected or appointed officials rather than planning technicians. This is a particularly important point because of the value system implications inherent in any set of development objectives. They have provided for the establishment of an advisory committee to assist the commission and its staff in the conduct of a regional planning program. Only by combining the cumulative knowledge and experience the various advisory committee members possess can a desirable future regional development plan be obtained. To be useful in the regional land-use transportation planning process, objectives must be precisely stated and related in a measurable way to alternative physical development proposals. Two basic types of objectives are (a) those that are difficult to relate directly to development plans, and (b) specific development objectives that can be directly related to physical development plans.

The quantification of specific objectives is facilitated by complementing them with planning standards that are in turn directly relatable to planning principles. A point fundamental to the development of this research is that land-use planning objectives cannot be separated from transportation planning objectives.

The specific objectives adopted for the regional transportation plan are those concerned primarily with a balanced transportation system; those which reduce traffic congestion, travel time, and accident exposure; and those which minimize costs and disrupting influences. An overall evaluation of each transportation plan must be made on the basis of the cost. An analysis may show that one or more of the standards cannot be met practically, and must be reduced or eliminated. No plan will meet all of the standards completely. The extent to which each standard is exceeded or violated serves as a measure of the ability of each alternative proposal to achieve specific objectives. Certain objectives or standards may be in conflict, requiring resolutions or compromise, and meaningful plan evaluations can only take place through a comprehensive assessment of each of the alternative plans against all of the standards.

Hill (13) presents a method for the evaluation of transportation plans. He notes that benefit-cost analysis was developed as a technique for examining plans with respect to their achievement of economic objectives. Although lip service is given to intangibles, they do not really enter into the analysis of many transportation and development plans. Urban objectives may have several dimensions—cultural, political, ethical, aesthetic, and economic. To pursue only one dimension would indeed lead to a suboptimal solution. Hill uses a goal-achievement matrix in his analysis and assumes that community objectives have been identified and relative weights attached to these objectives. The next step, therefore, is the comparison of plans in order to determine which plan best realizes the objectives of a community. An important set of requisites is feasibility, immediacy, and interdependence. The importance of being able to predict the reaction of the existing institutional power structure to various planning proposals is emphasized. The sections of the community to which costs and benefits accrue should be identified. In discussing the determination of weights to be associated with the various goals, Hill suggests the consideration of one or more of the following: community decision-makers, the general public by means of general referendum, a selective sampling of the affected groups, community power structure, public hearings, and the investigation of the pattern of previous allocation of public investment. Hill considers the strong effect of transportation on land-use development by noting possible impacts on neighborhoods and use of transportation facilities to separate incompatible land uses.

Thomas and Schofer (24) state that an inflexible commitment to evaluation strategies relying on the quantification of intangibles such as aesthetics would not constitute an optimal solution in plan evaluation. Major transportation decisions should remain in the hands of political decision-makers. Their review of literature resulted in the following:

1. A particular set of problems is perceived and the need for a solution is noted.
2. A preliminary set of criteria for evaluating alternative solutions is developed. These criteria must be available prior to the formulation of alternative solutions, so that a relevant set of alternatives may be devised.
3. Alternatives are generated.
4. Evaluation of alternative solutions is carried out.

When the characteristics of the alternatives are made available to the public, either at the formulation or evaluation stage, formal or informal interest groups are frequently aroused. If an alternative is found politically and technically acceptable, plans to proceed with design and construction are made. If not acceptable all plans may be rejected and, based on the arguments of the various formal and informal interest groups, a revised perception of the problem, the need for solution, and the evaluation criteria evolve. The process would be expected to cycle in an iterative manner until either a solution to the problem is developed or until the perception of the need indicates that the problem is not so serious as to merit the expenditure of resources required for its solution.

The statement of the plan evaluation problem has been structured as follows:

1. Determine the dimension of the transportation problem as it is viewed by the politicians and interested citizen groups.
2. Determine whether there is a finite set of regularly appearing transportation issues, whether new issues have emerged or old issues have disappeared during the past 20 years, and whether the relative importance of various issues has changed over the last 20 years.
3. Determine the scope and range of the issues relative to comprehensive planning goals. Are interest groups single or multipurpose oriented? Do multipurpose oriented groups emphasize one issue to the virtual exclusion of all others at any particular time? To what extent are non-transportation consequences emphasized? To what extent are the interest groups concerned with the intended consequences of the plan and to what extent are they concerned with the unintended consequences?
4. Determine the relation between published reports and public reaction in the form of isolated response and concerted group efforts.
5. Determine the nature of the political power structure with respect to transportation decision. Is the power structure diffused or centralized? Identify the participants in the transportation planning process.
6. Determine the conceptual model that best represents a process whereby an initial perception of social need is transformed into a political decision followed by implementation of a plan to meet that need.

Efforts are being made to develop criteria sets and to evaluate strategies that will be compatible with a complex political environment. Ackoff (1) studied individual preference for various modes of transportation and applied utility theory in the prediction of modal split. He identified factors affecting choice of transportation mode, such as safety, comfort, convenience, travel time, and economy. It is possible to scale or quantify these factors and consistent relationships can be found between personal preference and modal choice. The work, although directly related to the problem of modal split, indicates potential uses of utility theory in plan evaluation.

Lesourne (18) considers the application of operational research in comparing alternative city plans. Comparing city plans takes the form of comparing sets of hypotheses bearing simultaneously on locations of swellings, locations of industrial areas, and the nature of the transportation structure. From the definition of criterial for comparing urban plans two types of studies may be derived: practical research relating to the plan selection and theoretical research relating to the development of an optimal land-use transportation system.

Jessiman et al (16) present a rational decision-making technique for transportation planning, which is stated as follows:

1. Itemize the objectives which the community hopes to achieve in providing the transportation facility.
2. Select the parameter which best measures each objective.
3. Assign a weight or utility value to each of the objectives which reflects a measure of community value.
4. For each objective, examine the parameter chosen as the yardstick of that objective and determine, by use of a scale such as a utility curve, the value for that alternative.
5. For each alternative, sum the values assigned for all objectives to determine the alternative with the highest total value, that is, the one which best satisfies the complete set of objectives.

The planner must consider all effects of each alternative on the overall community system. Difference in points of views must be reconciled. Trade-offs relative to increased operating balance of public transportation and increased congestion must be objectively evaluated in view of an overall goals structure. In developing yardsticks, economic criteria seem to be the only criteria that are effectively considered. The use of utility values as criteria weights is proposed and the concept of marginal utility is presented. The relationship between incremental amounts of certain facilities and utility weights assigned to these incremental amounts may not be a linear relationship. For example, the first mile extension into a corridor may be more desirable than the second. The marginal utility approaches zero beyond a certain length of extension.

Persons familiar with the value systems of the various interest groups in a community may gain insight into reasons for controversies surrounding a project. An investigation of parametric programming or sensitivity analysis of a given solution relative to slight changes in parameters associated with that solution, such as total budget expenditure, is suggested.

Tendencies to emphasize judgment and subjective probabilities are considered a backlash to rapid expansion in the development of precise computer models. Techniques currently used in personnel evaluation by industry may be of value in alternative plan evaluation.

Similar approaches are considered by Irwin (15) in a discussion of criteria for evaluating alternative transportation systems. Transportation planning standards have far-reaching implications involving philosophy, economics, politics, sociology, engineering, and aesthetics. The purpose of Irwin's research is the definition of criteria on which the plan evaluation process may be based. Selection, definition, and application of criteria for evaluating transportation systems contain much uncertainty. More knowledge is urgently needed about the effect of these uncertainties on transportation planning decisions.

Recent developments in allocation theory may be applied to management decision problems. The work of Dean and Nishry (6) relates to scoring and profitability models for evaluating and selecting engineering projects. They consider problems involving the specification and allocation of manpower, funds, and equipment to projects within a firm. Quantitative measures of organizational performance must be derived that are consistent with corporate goals and that consider relevant resource variables, noncontrollable variables, parameters, and constraints. The authors develop mathematical models that yield solutions for allocating manpower resources to projects. The allocation procedure could be used in the selection of alternative community development plans and in the allocation of public revenues to development proposals.

Pessemier (20) develops a system wherein benefits may be measured in a dollar metric by a prescribed method of making trade-offs between various proposals and the "do nothing" alternative. The procedures require accurate cost data so that intangible or total benefits may be quantified. Although an application of Pessemier's procedures has not been attempted here, such application may greatly strengthen conventional benefit-cost techniques.

Hemmens (12) presents experiments in urban form and structure and states that the evaluation of alternative land development patterns is an important, unsolved task in urban planning. There are many reasons for slow progress in developing methods for evaluating alternative development plans. Among these reasons is disagreement about the proper criteria for evaluation. The solution to a part of this problem may be found in the development of a fairly simple experimental model of an urban community; a model designed specifically for the exploration of the relationships among elements in urban form. The paper is a progress report on a simple model for examining the impact of changes in components of urban form on urban spatial structure.

A distinction is made between urban form and urban structure. Urban form is the physical arrangement of residents, work places, etc. Urban structure is the pattern formed by the connection of these elements in the daily activities of areas of residents.

The author uses a simple linear programming formulation for evaluating urban form on the basis of two criteria: the efficiency of each alternative in terms of minimum travel requirement, and the equity of the alternatives in terms of locational advantage of residential locations. Given alternative distributions of work places, shopping places, residences, and systems of transportation service, and given an allocation rule specifying the manner in which residences will be linked with work places and shopping centers, determine the nature of change in urban form and urban spatial structure. The report examines the relationships among elements of form as a first step toward developing more satisfactory analytical methods of evaluating alternatives.

Dansereau (4) presents an evaluative scheme based on attitudes and economic climate as they affect highway development. The work predicts economic development at selected interchanges, develops alternative land-use plans for interchange protection, and identifies factors conducive to community adoption of reasonable protective regulations. Citizen acceptance of local highway changes is related to acceptance of rational controls and ultimately to implementation of necessary protective practices.

Three types of attitude studies were undertaken: (a) study of attitudes towards local highway developments, (b) study of attitudes toward planning and zoning practices, and (c) study of attitude change toward both development and control practices. Economic considerations that have influenced the findings of the attitude research were studied. The economic analyses consisted of study of land use and land value, study of predictions of interchange development, and study of the economic impact of the interchange development.

Worrall (26) presents an interesting discussion of the use of an urban panel as a longitudinal data source for urban planning. The paper treats data collection as it relates to plan evaluation. Modeling technology is constrained by the characteristics of existing data systems. Data formats developed prior to the current focus on model-building activity are inadequate for present purposes. Present data formats specify an initial level of aggregation considerably in excess of that desired by the analyst. They are predominantly cross-sectional rather than longitudinal in form, and the information content is such that it seldom permits a full-scale evaluation of policy impact.

The paper considers the feasibility of developing a new form of data system for continuous recording of urban information. The mechanism employed is that of a permanent household response panel, an approach frequently used in consumer and market research.

The system has applications as a source of data for future model building and as a general mechanism for urban analysis. The paper emphasizes the application of panel techniques in the study of urban travel. The emphasis is one of convenience, reflecting the particular interest of the author. The discussion might well have been centered on the use of panel techniques for the study of residential location preferences, household activity patterns, or others.

Extremal Methods (Linear Programming): An Optimizing Approach

Hay et al (11) present an interesting use of extremal methods (mathematical programming) in the development of an optimal bimodal transportation system.

A research proposal developed in upstate New York (19) proposes an interesting use of integer programming in the design of a transportation system. The proposal is concerned with the use of operations research (mathematical programming techniques) in the determination of optimal routes and headways for a fixed investment in transit vehicles and/or a fixed level of operating expenditure. The procedure involves the allocation of transit service to existing or proposed route sections in an optimal manner, subject to systems and subsystem constraints. Typical elements of the constraint set are (a) upper limit of available transit system components, (b) lower limit of available transit system components, (c) upper limit on level of transit service on a specific route, and (d) lower limit of transit service on a specific route. Constraints (c) and (d) are considered to be subsystem constraints.

Optimal design techniques could be applied in determining the optimal expansion of an existing system as well as in determining the best overall design of a new system. The optimal design procedure could serve as a highly efficient method for evaluating changes in stated government policy subject to appropriate constraints. Changes in governmental policy could take the form of variations in the parametric values associated with the mathematical programming formulation of the problem. Such variation could be thoroughly evaluated by well-developed and easily manageable sensitivity analysis procedures.

The New York Office of Transportation plans to conduct the research in three phases: (a) an intensive study of transit usage as it is related to transit service and traffic market potential; (b) the development of the mathematical processes necessary to formulate the allocation; and (c) use of the technique in the planning and design of new systems. The transit revenue function, the function to be optimized, will be an expression of the relationship between transit usage and level of service on specified route sections. Level of service is a measure of passenger-carrying capability and is expressed in some unit of capacity per hour. It is assumed that usage-service relationships would differ in areas of different socioeconomic characteristics. The transit usage analysis would develop a temporal usage rate for each service level for each specific route section.

The significance of sensitivity analysis is pointed out. It will be possible to investigate the effect on the "optimal" allocation of service of variations in the usage-service relationships. The formulation must consider the cost of unused equipment and personnel during off-peak periods of demand.

The New York proposal concludes by (a) restating the obvious desirability of "optimal" transportation systems design, and (b) pointing out that only recently have transportation planners acquired, through phenomenal advances in mathematical programming and computer technology, the capability for undertaking such comprehensive transportation systems analyses.

Hitch (14) discusses the problem of sub-optimization. Comments from his article are repeated here because sub-optimization is taking place in many phases of the planning process. Calculating quantitative solutions based on wrong criteria is equivalent to answering the wrong question. The basis for "good" criteria at any level of analysis in operations research is consistency with "good" criteria at a higher level.

Ridley (22) describes an investment policy to reduce the travel time in a transportation network. A transportation network should satisfy traffic demands placed on it and give service to users on the basis of some acceptable criteria, within budgetary, political, and social constraints.

The transportation network is represented by an abstract graph of nodes and arcs on which are defined real-valued variables and functions representing travel times, traffic flows, and money invested. The travel time on an arc is a known function of investment and the assignment of traffic flow on a particular route varies with the travel time on the arcs. Ridley seeks an optimal set of arcs so that investment in these arcs gives minimum travel time. He presents a combinatorial analysis of the transportation planning process. A lemma is proved which puts upper and lower bounds on the minimum travel time in a network for an investment, M. This is then used in a constructive proof of an algorithm which obtains an optimal set of investments for a given budget.

ESTABLISHING A WEIGHTED HIERARCHY OF COMMUNITY DECISION CRITERIA

This section is presented to indicate how a set of community goals and objectives could be formulated and weighted. The weighted community decision criteria are essential to the proposed method of plan evaluation emphasized in this paper. The method proposed assumes involvement of community decision-makers in structuring a list of specific community decision criteria. Professional planners would use the decision criteria, weighted by the decision-makers, in the evaluation of planning proposals. The central problem in this research is the development of analytical methods for plan evaluation, having as input to this evaluation a set of weighted community goals and objectives.

Since the plans to be evaluated were for metropolitan Louisville, a task force from the Louisville Mayor's Citizens Advisory Committee was used as the criteria evaluation or community decision-making group. The task force represented a cross section of highly respected, influential citizens of metropolitan Louisville (the area used as the experimental laboratory). This group is interested in and familiar with the area's community goals and objectives.

Although it was convenient and entirely satisfactory in this research to utilize the committee for criteria weighting, a more general criterion for the selection of such a committee may be stated as follows:

> The committee should consist of direct and indirect influentials including popular public officials and representatives of commerce and industry who are influential in controlling development decisions, and those indirect influentials who, by reason of their personal stature and demonstrated interest, are effective in shaping policy on important community issues.

An alternative presentation of this criterion is the following block diagram:

COMPOSITION OF CRITERIA FORMULATION COMMITTEE

	Influentials	
	Possible Direct	Actual Indirect
Representors	A	C
Implementors	B	D

where the letters are defined as

- A—popularly elected officials;
- B—other heads of public and semi-public bodies, executives of commercial and industrial firms;
- C—unbiased, interested citizens;
- D—other indirect influentials including groups A and B acting outside the area of their direct control.

Procedures used in the establishment of a weighted set of community decision criteria (i.e., specific statements of community goals and objectives) are as follows:

1. Professional planners established a tentative set of community goals and objectives, explicitly and concisely stated.
2. The criteria evaluation group met for general discussion and modification of each item in the statements of community goals and objectives. The end product was a complete statement of community goals and objectives, modified in view of the comments and opinions of the decision-makers or criteria evaluation group. The resulting statements of community goals and objectives are shown in Appendix A.
3. Each member of the criteria evaluation group was asked to individually weight the various sets of criteria by the ranking and rating methods of Appendix B.
4. The decision-makers or criteria evaluation group met and were asked to re-evaluate their initial weighting of the elements of the criteria statements. No committee members changed their initial values.

The aggregated weightings thus obtained, as given in Appendix C, were used in the plan evaluation decision model.

For the two techniques used, the following statistical results were obtained:

1. A high level of agreement among judges using the scaling techniques was observed.
2. Criteria weights obtained by the methods applied were highly correlated in both rank order and interval-level measure. Criteria weights obtained by any given method were highly correlated with criteria weights obtained by averaging all methods.
3. Each judge demonstrated transitivity of preference throughout all methods used.

PLAN EVALUATION: THE DECISION MODEL

Two similar approaches to the development of a decision model used in alternative plan evaluation are the effectiveness matrix approach and a scoring model. This section will develop the mathematics associated with these techniques and will present an actual application of the effectiveness matrix approach. The scoring model extends the effectiveness matrix technique by treating a stratification of judges by background and interest groupings.

The Effectiveness Matrix Technique

At this point, it is assumed that a hierarchy of community planning goals and objectives has been established and that a numerical utility measure or criterion weight has been assigned to each objective statement. Three alternative community plans are under consideration. Outlined in this section is a procedure for evaluating the three alternative proposals.

Definition of Terms—Consider here the set of community planning objectives G_j where $j = 1, 2, \ldots, n$, n being the total number of decision criteria under consideration. Second, three plans are proposed for evaluation. The set of plans under consideration is designated by P_i, where $i = 1, 2, 3$. Associated with each community planning objective G_j is a numerical utility value u_j $(j = 1, 2, \ldots, n)$ which was determined by the procedures of Appendix B. Regardless of the system of decision criteria under consideration, the following equality must hold:

$$\sum_{j=1}^{n} u_j = 1$$

The purpose of this discussion is to describe a procedure for objectively utilizing weighted community decision criteria in the evaluation of physical development plans; therefore, "effectiveness" (e_{ij}) and "plan utility" U_i are defined. Effectiveness (e_{ij}) is a measure of the probability that objective j can be achieved if plan i is adopted. U_i is a measure of the total utility of plan i based on the evaluation of plan i relative to all objectives.

The Effectiveness Matrix—The effectiveness matrix was developed by a committee of planners representing the professional disciplines associated with the comprehensive planning process. The effectiveness value (e_{ij}) is assigned on the basis that an e_{ij} of

1.0 implies that achievement of objective j is assured under plan i, and an e_{ij} of 0.0 implies that achievement of goal j under plan i is practically impossible. If all plans i have no effect on the achievement or prevention of objective j then all e_{ij} associated with that objective are undefined and the unrelated criterion will be dropped from the effectiveness matrix. Values of e_{ij} will be estimated to the nearest tenth by each evaluator, using the previously defined guidelines. Elements of the final effectiveness matrix (e_{ij}) will be documented later in this section with a statement of reason for the numerical value given. In general terms, the effectiveness matrix will have the following form:

EFFECTIVENESS MATRIX

Alternative Plan	Criterion						
	G_1	G_2	G_3	...	G_j	...	G_n
P_1	e_{11}	e_{12}	e_{13}	...	e_{1j}	...	e_{1n}
P_2	e_{21}	e_{22}	e_{23}	...	e_{2j}	...	e_{2n}
P_3	e_{31}	e_{32}	e_{33}	...	e_{3j}	...	e_{3n}
.	.	.	.		.		.
.	.	.	.		.		.
.	.	.	.		.		.
P_i	e_{i1}	e_{i2}	e_{i3}	...	e_{ij}	...	e_{in}
.	.	.	.		.		.
.	.	.	.		.		.
.	.	.	.		.		.
P_m	e_{m1}	e_{m2}	e_{m3}		e_{mj}		e_{mn}

In generalized vector notation, the effectiveness matrix may be represented by E.

The Decision Model—A decision model that determines a total effectiveness for each of the plans P_i (i = 1, 2, 3) with respect to the given decision criteria structure G_j (j = 1, 2, 3, 4, 5) follows. For each plan (i), the total utility is the sum of the products of the individual numerical utility of the plan with respect to objective j (e_{ij}). The model is mathematically stated as follows:

$$U_i = \sum_{j=1}^{n} e_{ij} u_j \quad (i = 1, 2, \ldots, m)$$

where

U_i = total utility associated with plan i,
e_{ij} = probability that objective j can be achieved if plan i is adopted, and
u_j = numerical measurement of utility associated with community planning objective j.

In the generalized vector notation, the decision model may be stated as follows:

$$\begin{bmatrix} e_{11} & e_{12} & e_{13} \dots & e_{1j} \dots & e_{1n} \\ e_{21} & e_{22} & e_{23} \dots & e_{2j} \dots & e_{2n} \\ e_{31} & e_{32} & e_{33} \dots & e_{3j} \dots & e_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ e_{i1} & e_{i2} & e_{i3} \dots & e_{ij} \dots & e_{in} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ e_{m1} & e_{m2} & e_{m3} & e_{mj} & e_{mn} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \cdot \\ \cdot \\ \cdot \\ u_j \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ \cdot \\ \cdot \\ \cdot \\ U_i \\ \cdot \\ \cdot \\ \cdot \\ U_m \end{bmatrix}$$

Or, this can be stated as:

$$U = Eu$$

where $u = (u_1, u_2, u_3, \dots, u_j, \dots, u_n)^T$ is a column vector whose components represent the utility values associated with each of the n community decision criteria, and $U = (U_1, U_2, \dots, U_i, \dots, U_m)^T$ is a column vector whose components represent the plan utility associated with each of the m alternative development plans, and E is the $m \times n$ matrix defined earlier.

The plan possessing the highest total utility would be the alternative plan recommended to the community decision-makers for formal adoption.

A Scoring Model

Work in the area of development of scoring and profitability models for evaluating engineering projects within an industrial firm presents results that may be applicable in alternative plan evaluation (6). A suggested application is presented here.

Previous definitions of G_j, u_j, e_{ij}, and U_i apply here. At this point the scoring concept is exactly the same as the effective matrix technique described earlier. Consider a panel of judges or community decision-makers, individually representative of different and definable socioeconomic sectors of the community. Vogt (25) and others have indicated that community decisions should reflect the makeup of the community relative to socioeconomic group stratification. The model presented previously could be modified as follows:

$$U_i = U_i^1 + U_i^2 + \dots + U_i^k + \dots U_i^p$$

$$U_i^k = a_k \sum_{j=1}^{n} u_j^k \, e_{ij}^k$$

where

U_i = total score for alternative plan i;

U_i^k = score for alternative plan i as determined by the kth socioeconomic group (k = 1, 2, . . ., p);

u_j^k = criterion weight for objective j as determined by the kth socioeconomic group;

e_{ij}^k = value of plan i relative to the criterion j as determined by the kth socioeconomic group; and

a_k = fraction of the area population represented by the kth socioeconomic group.

As a minor but logical modification of this scoring model, one may consider the development of utility values by different socioeconomic groups of citizens while considering only one set of effectiveness values established by one group of professional planners. This problem may be formulated as the following matrix multiplication:

$$\begin{bmatrix} e_{11} & e_{12} & e_{13} \cdots & e_{1j} \cdots & e_{1n} \\ e_{21} & e_{22} & e_{23} \cdots & e_{2j} \cdots & e_{2n} \\ e_{31} & e_{32} & e_{33} \cdots & e_{3j} \cdots & e_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ e_{i1} & e_{i2} & e_{i3} \cdots & e_{ij} \cdots & e_{in} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ e_{m1} & e_{m2} & e_{m3} \cdots & e_{mj} \cdots & e_{mn} \end{bmatrix} \begin{bmatrix} a_1 u_1^1 & a_2 u_1^2 & a_3 u_1^3 \\ a_1 u_2^1 & a_2 u_2^2 & a_3 u_2^3 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ a_1 u_j^1 & a_2 u_j^2 & a_3 u_j^3 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ a_1 u_n^1 & a_2 u_n^2 & a_3 u_n^3 \end{bmatrix} =$$

$$\begin{bmatrix} U_1^1 & U_1^2 & U_1^3 \\ U_2^1 & U_2^2 & U_2^3 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ U_i^1 & U_i^2 & U_i^3 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ U_m^1 & U_m^2 & U_m^3 \end{bmatrix}$$

Or,

$$Eu = U$$

where

$E = (e_{ij})$ is a m × n matrix. The typical element represents the probability that goal j will be achieved if alternative plan i is adopted.

$u = a_k u_j^k$ is a n × 3 matrix. The typical element represents the utility value (criterion weight) for criterion j as determined by socioeconomic group k. In this example 3 socioeconomic groups are considered; a_k is the fraction of the area population represented by socioeconomic group k.

$U = (U_i^k)$ is a m × k matrix. The typical element represents the aggregate utility (score) assigned to alternative plan i by socioeconomic group k.

By summing U_{ik} value for each row (i) of the U matrix, a utility value (score) for each alternative plan (i) may be obtained. The values will be weighted in a manner consistent with the socioeconomic group composition of the community.

Model Application and Presentation of Results

The effectiveness matrix technique for plan evaluation has been applied and the results are presented here. The model is described in vector notation U = Eu. The transposed effectiveness matrix E^T is given in Appendix D. The columns represent the 3 alternative plans evaluated and the rows of the matrix represent the 35 criterion statements or community planning objectives. Two professional planners from The Falls of the Ohio Metropolitan Council of Governments (the regional planning authority for the Louisville metropolitan area) and three professional transportation planners from the Louisville Metropolitan Comprehensive Transportation and Development Program participated in the plan evaluation process. The e_{ij} values of Appendix D represent the consensus of this group of professionals.

The components of the column vector u are the utility values associated with the 35 decision criteria or community objectives. This vector is given in Appendix C in the column headed Average Values, u_j. As stated earlier, the Task Force 5 values were used in the plan evaluation model because this group formulated the statements of goals and specific objectives and was, therefore, more familiar with the criteria as well as the community involved. Note that the vector u is a 35-component column vector.

The 3 × 35 matrix E was multiplied by the 35 × 1 column vector u to produce a 3 × 1 vector U. That vector is stated as follows:

$$U = (U_1, U_2, U_3)^T$$

or

$$U = (0.38, 0.52, 0.60)^T$$

where each of the values U_i represent the aggregate planned utility associated with each of the three alternative development plans.

Because these plans are transportation system design concepts only, they have not been developed in sufficient detail to provide cost estimates. This precluded the possibility of doing a complete cost-effectiveness analysis.

The aggregate results indicate that the least preferred alternative is plan 1 (Appendix E). That plan is based on extensive improvements of existing at-grade arterial facilities. The second proposed alternative design concept, plan 2 (Appendix F) is based on extensive construction of freeway facilities with no rail mass transit. Plan 2 possesses an aggregate utility approximately 37 percent higher than that possessed by alternative design concept 1. Finally, the most preferred alternative is plan 3 (Appendix G), based on a balance of new freeway construction and rail mass transit. The rail mass transit-oriented alternative possesses an aggregate plan utility 58 percent higher than that of the first design alternative and 15 percent higher than the freeway-oriented design concept.

In the remainder of this section the reasoning involved in the determination of various e_{ij} values is discussed and "yardsticks" for use in determining the respective effectiveness values are identified. The objective statements are shown as quotations and appropriate comments follow.

"Insure safe public facilities." The transit-oriented system was judged most effective in assuring safety, with the freeway alternative second. A yardstick for the measure of effectiveness here may be a study of accident records on various types of transportation facilities, particularly the study of such accident records on facilities in the metropolitan area.

"Provide for adequate public safety regulations and their enforcement." The high effectiveness for plan 2 indicated that the experts felt enforcement of freeways was by far the easiest type of enforcement. Numerous accident or friction points exist in plan 1, while significant policing problems in transit vehicles and stations exist with plan 3.

"Provide for the removal of contaminants (solid, liquid, and gaseous)." The transit-oriented alternative was most preferred here because of the fact that it removes many vehicle miles of travel from the surface street system, thereby reducing air pollution caused by vehicular exhaust. A yardstick to be used in a measurement of effectiveness here may be aggregate vehicle miles of travel. This statistic is highest on a surface street-oriented alternative (plan 1) and, therefore, that alternative is the least desirable.

"Minimize maintenance costs of public facilities." Wide rights-of-way make the freeway alternative less desirable than the surface street alternative; however, maintenance would be most expensive in a transit-oriented system. A yardstick in determining this effectiveness could be the development of maintenance cost records by type of facility.

"Insure maximum effectiveness of public utilities (including transportation facilities) by design and locational considerations." The freeway-oriented alternative was most desirable in this case, furnishing good access to many major public facilities. The inflexibility of mass transit is reflected in the lower effectiveness value of plan 3. Aggregate hours of travel could be a yardstick relative to this objective as well as the accessibility index produced as a part of the standard trip analysis procedures.

"Develop a balanced, effective, and integrated transportation system which provides for the accessibility requirements of each land use." Balance is implied by transit orientation in transportation system development and this implication is reflected in the high effectiveness value of plan 3. The surface street concept is the least effective of these three plans. Yardsticks may be developed in this area, such as analysis of travel by various modes, measurements of delays and frequency of service, and determination of aggregate travel time and aggregate travel costs.

"Develop public improvement programs within available financial resources." Here, plan 1 and the freeway-oriented plan have the highest effectiveness values. The low effectiveness value associated with the mass transit concept reflects the customary subsidy associated with that type of program. The existence of a financing system, such as the highway trust fund based on road user taxes, reflects a system development within available financing.

"Maintain highest equitable property values." Studies have indicated a skyrocketing of property values in freeway and mass transit corridors; however, accessibility by any means seems to enhance property values. The effectiveness values reflect this greater activity in transit corridors.

"Insure effective utilization of mineral, vegetation, air, and water resources." In the opinion of the professionals developing the effectiveness matrix, this objective is not related to or affected by transportation system design concepts.

"Establish a strong economic base through commerce that will bring money into the community." The effectiveness values indicate that a transit-oriented system is stronger relative to inducing a new industry into a community. A freeway-oriented system providing access to suburban areas for industrial park and new plant development was the second preferred, while the alternative based on improvement of existing facilities received a low value for this objective.

"Establish trade development that provides maximum convenience to consumers." The effectiveness values indicate an edge for a transit-oriented alternative over a freeway-oriented alternative with the improvement of existing street concept receiving a somewhat lower value. Although improvement of existing streets provides for more convenience to neighborhood shopping centers, possibly it impedes access to regional and central business district type facilities.

"Insure the optimal utilization of all land." Again, the transit-oriented alternative received an edge reflecting that this system, a transit-freeway system, provides the best access to land in an urban area. The freeway-only alternative was second and the improvement of existing street facilities was the least preferred or the least effective alternative.

"Achieve increased disposable income for all people." Due to the greater accessibility to work locations for all of the population, the mass transit alternative possessed the highest effectiveness value. Again, for reasons of overall accessibility, the freeway-oriented alternative was second. The planners felt that a street system would not provide access to job centers, particularly for that element of the population that could not afford to maintain an automobile.

"Preserve historic sites and areas of natural beauty." Although plan 1 requires less new right-of-way, it was felt that it was the least desirable alternative because it would result in overloaded conditions or street facilities serving historic sites and sites of natural beauty. Proper alignment of a mass transit line could provide mass access to these facilities, thereby resulting in that alternative's receiving the highest effectiveness value.

"Promote adequate public libraries, museums, and cultural activities." Again, the greater overall accessibility provided by a transit-oriented system resulted in that system's receiving the highest effectiveness value.

"Protect meaningful local tradition and encourage civic pride." The greater accessibility of the freeway-only and transit-freeway alternatives results in the high effectiveness for these two plans. The professional planners felt that civic pride is encouraged by a good transportation system, another reason for the high effectiveness values of plans 2 and 3.

"Establish the mechanism for adequate preventive and remedial health programs and facilities." This objective is not related to or affected by transportation system design concepts.

"Develop educational facilities and opportunities for citizens at every level." Again, the high accessibility provided by a mass transit system resulted in that system's receiving the highest effectiveness value. The second highest value is possessed by the freeway-oriented system, with a very low effectiveness value assigned to plan 1, which would not provide good access to high school and higher education activities and facilities.

"Eliminate injustice based on discrimination." In this case, the more accessible systems, plans 3 and 2 respectively, receive the lowest effectiveness values. The planners reasoned that this type development encouraged the development of ghettos for impoverished minority groups.

"Develop needed public welfare programs." The planners indicated that this objective was unrelated to transportation system development.

"Encourage development of religious opportunities." Again, the high accessibility systems as depicted in plans 3 and 2 respectively received the highest effectiveness values.

"Develop an aesthetically pleasing environment." Although this objective is mostly sensitive to urban design concepts, the panel felt that by placing mass transit systems in subways in congested areas, aesthetics could be realized more readily. Also, heavy travel on surface streets was judged not to be consistent with pleasing aesthetic values.

"Establish open-space programs." Concentration of traffic on rail or on limited-access freeways was judged to be most consistent with the establishment on open-space programs.

"Provide adequate recreational facilities utilizing parks, rivers, and lakes." A surface system was judged to provide the greater accessibility to the type of recreation described in this objective. The inflexibility of the mass transit system resulted in its receiving a low effectiveness value.

"Improve the framework (channels, systematic use) for citizen participation in governmental functions." This objective is unrelated to transportation sytem development.

"Establish equitable taxation policies (bases, mixes, rates)." This objective should be applied in transportation system analysis to assure that equitable cost-sharing is established between users and nonusers and to assure that transportation facility development costs are equitably distributed between participating agencies charged with the responsibility for developing these facilities. The low effectiveness value for the mass transit system indicated that the subsidy normally associated with this type system development is a taxation inequity.

"Achieve efficient governmental administration, representative of all citizens." This objective is not related to transportation system development.

"Develop adequate government staffs and personnel programs (high job standards, reasonable salary ranges, effective delegation of authority)." This objective is not related to transportation system development.

"Establish sound governmental fiscal programs." Again, the subsidy normally associated with mass transportation is regarded as not a sound fiscal program.

"Develop an effective, long-range, metropolitan-wide planning process." This objective implies that transportation and development policies must be coordinated and that studies of both lead to the development of a planning process and implementation devices which accomplish the goals for the least expenditure of direct and indirect costs. The development of an integrated system as reflected in plan 3 seems to be most consistent with this objective.

"Establish effective control mechanisms." This objective is unrelated to transportation system development.

"Encourage rehabilitation and conservation neighborhood programs." The low effectiveness of the transit-oriented alternative implies that many neighborhoods cannot be effectively served by an isolated transportation system such as a mass transit system. The development of a street system coordinated with urban redevelopment projects is an obvious technique implied in the implementation of this objective.

"Provide adequate low-cost housing." The transit-oriented alternative received the highest effectiveness value because the planners establishing these values felt that low-cost, high-density housing could be served best by a transit-oriented transportation system.

"Develop neighborhood units." The surface street system providing good transportation access to neighborhoods was judged to be most effective. A yardstick to be used in a measurement of the compliance of various plans with this objective could be the measurement of vehicle-pedestrian conflicts at the neighborhood level and the measurement of through traffic within neighborhoods.

"Promote a wide variety of housing types as required within the community." The high effectiveness for plan 3 reflects the planner's opinion that rail mass transit could serve high-density residential corridors and promote most effectively the wide variety of housing mentioned in this objective.

As will be stated in the next section, the area of developing yardsticks for measuring the extent to which a plan is compatible with the various objectives presents a most challenging area of further research. This section has provided some examples or guidelines for the development of quantitative and effective yardsticks, along with comments concerning the thinking of the professionals in arriving at the effectiveness values.

SUGGESTED FURTHER APPLICATIONS AND EXTENSIONS

It is anticipated that continuing application and refinement of these techniques will be made a regular part of the Work Program of the Louisville Metropolitan Comprehensive Transportation and Development Program. Obvious applications of the techniques to the work in Louisville are (a) for improvement of the existing recommended plan, (b) for use in the evaluation of alternative land-use forms now under consideration by development planning agencies, and (c) for use in the analysis of selected transportation corridors.

Improvement of Selected Plan

The study consultant will recommend a transportation plan to the Transportation and Development Program. It is proposed that the plan evaluating schemes of this research be applied to that selected plan in a diagnostic manner. The recommended plan will be analyzed in detail relative to each of the community objectives in the weighted hierarchy of community goals and objectives given in Appendix C. On the basis of this evaluation, an analysis of the recommended plan can be made. In the areas where the plan is weak with respect to certain goals and objectives, action to remedy such shortcomings in the plan will be considered.

Currently, a study by The Falls of the Ohio Metropolitan Council of Governments is concerned with the development of a more complete set of community goals and objectives. The goals and objectives study will be carried out over the next two years and will result in a more comprehensive statement of goals and objectives than presented here. At that time, the scheme for evaluation will be repeated subsequent to the weighting of the decision criteria. Again, modifications of the transportation plan will be considered on the basis of the results of the study.

Alternative Land-Use Forms

The current transportation planning program in the Louisville metropolitan area has been based on a single land-use form, defined by the Louisville and Jefferson County Planning Commission as "planned sprawl." Other fundamental land-use forms such as satellite cities, radial corridors, and others are being considered by the development agencies of the area. When a comprehensive plan based on an alternative land-use form is available, a more extensive application of these procedures will be possible. At that time, the procedures may be used to evaluate the alternative land-use forms, the alternative transportation plans associated with these forms, and, finally, alternative comprehensive development plans that encompass both land use and transportation.

Corridor Analysis: Route Planning Studies

In addition to the recommendations relative to new freeway systems and new arterial systems for a metropolitan area, a large effort of the continuing planning process relates to the improvement of existing facilities within that area.

One of the significant tasks associated with this improvement of existing facilities is corridor analysis or route planning studies. It is anticipated that the techniques of this research will be most useful in the development of plans associated with the improvement of existing facilities. Alternative routes may be considered and each of these alternatives may be evaluated in the context of the community goals and objectives structure presented.

An immediate suggestion relative to the application of these techniques in route planning is the development of a pilot study or set of guidelines for the application of these techniques to the planning analysis of an individual corridor instead of a total transportation system.

Defining the Decision Variables: A Work Program Reflecting Specific Objectives

The earlier sections of this research have presented an approach to plan evaluation based on a weighted hierarchy of community decision criteria or goals and objectives. Hopefully, the procedures resulting from this research presented in the earlier sections will provide planners with a straightforward, efficient, and effective methodology for weighting goals and objectives and evaluating alternative plans. It is recognized, however, that the techniques proposed are suboptimal in many respects. Many "givens" are imposed upon the process. Planning is treated as a "second-order" governmental function below the policy-making and financing processes. Possibly, if decisions at the primary level could be guided quantitatively by the weighted hierarchy of community goals and objectives, a truly optimal approach would exist.

Studies of suboptimization (14) indicate that "good" decision criteria at any level are consistent with "good" decision criteria at higher levels. Quantitative solutions based on the wrong criteria (in this case "wrong" givens input to the planning process) are tantamount to answering the wrong question, and this may well apply to the community development process. With most metropolitan governments, well-defined criteria do not exist at the higher level, and this results in suboptimal lower level planning decisions.

This section proposes a procedure for top-level community decision-making using cardinal utility values in an optimal allocation of community resources.

One may consider the mapping of a closed, precisely defined set of community values onto a set of community goals; of goals onto objectives; and finally, a set of community objectives onto a set of items constituting a community work program. Expenditures of public revenues on each of the items of the work program may be considered as the decision variables ($x_1, x_2, \ldots, x_j, \ldots x_n$) of a mathematical programming formulation. For example, decision criterion j may relate to the "establishment of open space programs for metropolitan use"; x_j would then represent the expenditure of public revenues in dollars on open space programs. A set of decision variables would be defined along with items of a work program in such a manner that every community objective would be represented by a work program item (or items) insuring the fulfillment of that objective.

Conceptual Formulation of an Allocation Model

The preceding section defined decision variables. In considering a particular decision variable, x_j, it is possible to associate with that variable a "cost coefficient" indicative of the utility associated with the work program item represented by that decision variable.

It may be considered desirable to maximize the aggregate of the dollar expenditures multiplied by the utility value associated with each individual expenditure represented by the decision variables. The allocation of tax revenues must be performed within certain constraints. Such constraining relationships may be the availability of total money, the availability of other resources such as land, restrictions implied by time factors, desirable minimum or maximum levels of expenditure for various programs, or desirable interrelationships among the various work program items represented by the decision variables. Further, it would be logical to disallow any negative allocation of money.

An Extremal Methods Approach

This section suggests several applications of standard mathematical programming techniques.

Linear Programming Formulation—The definition of decision variables was considered earlier. Consider a class of parameters (u_j) associated with the decision variables (x_j); the parameters represent the utility values associated with the various community work program items defined by the decision variables. That is,

x_j = number of dollars allocated to community work program j; and
u_j = utility associated with community work program j per dollar spent on community work program j.

An optimal allocation of available funds to the various work programs is represented by the following objective function:

$$\text{Max} \sum_{j=1}^{n} u_j \, x_j$$

where there are n possible work program items to which allocation may be made.

Constraints of the following form may be applicable:

$$x_1 + x_2 + \ldots + x_j + \ldots + x_n \leq b_1$$

where b_1 represents maximum available funds;

$$x_k \geq b_2$$

where b_2 may represent an absolute minimum expenditure such as required for education, police protection, or fire protection.

Due to constraints placed by time requirements associated with various projects, maxima may exist such as

$$a_{ij}\ x_j \leq b_1$$

In general, the problem may lend itself to formulation as the general linear programming problem stated as the maximization of a linear objective function subject to appropriate linear equality or unequality constraints. One constraint is that all allocations are non-negative.

Research in progress at Purdue University considers an optimal allocation of land uses. The formulation proposed could incorporate the concept of using criteria weights (utility values) as cost coefficients in the formulation of the objective function of a mathematical programming problem.

Parametric Programming Analysis—An interesting examination of the linear programming model by standard methods of parametric programming appears to be feasible. Changes in the cost coefficients or the utility values, as in this particular application, may be investigated, and the sensitivity of an optimal solution to changes in these criterion weights or utility values may be examined. Further, it may be possible by means of parametric programming analysis to determine the solution with a relaxation of the total money constraint or changes in other parametric values. With slight changes in certain "given" values, a much more desirable solution may be obtained.

Nonlinear Formulation of the Problem—An interesting concept of marginal utility is that additional incremental amounts of a given item are not as valuable as previous increments of the same size. For example, the third or fourth serving of a dessert would not be valued as highly as the first. Bernouilli and others have postulated that the utility function is not linear and may be described by an exponential or quadratic relationship. The methods of nonlinear programming may be applied in the situation of optimal allocation of community resources. The quadratic formulation proposed by Wolfe (5) or the more general convex formulation (10) may be applicable.

Dynamic or Integer Programming Approach—The powerful tool of dynamic programming has been successfully applied in problems where the decision variable is a 0-1 variable, i.e., in a situation where either an allocation is made or it is not made. Integer solutions may be indicated because of the practical situation where it would not be feasible to build a fractional or non-integer portion of a new school.

Plan Evaluation

Many sophisticated techniques developed in the area of economic analyses of plans, particularly transportation plans, must be incorporated in an objective manner in the evaluation process. Much of the work in benefit cost analysis may be applied. Ultimately, an effective means of developing yardsticks to measure compatibility of plans with community values must be researched.

Plan evaluation must be concerned with the manner in which a plan is consistent with community values at a lower level of synthesis than the level of objectives studied here. That is, objectives are often too general and the resulting evaluation may be purely subjective. For use in evaluation of plans of traffic improvement at a more detailed level of analysis, the pertinent objective statements would be further subdivided to establish more meaningful criteria. This may be accomplished within the framework of the procedures presented here.

SUMMARY AND CONCLUSIONS

This paper has presented an approach to the development of a decision model for evaluating alternative transportation system design concepts in the context of a comprehensive hierarchy of community goals and objectives. Various interesting approaches to plan evaluation were discussed as well as several proposals for utilizing potentially powerful normative procedures in system design. Extensive discussion of problems associated with community decision structure, formulation and weighting of goals and objectives, and the statistical analysis of weighting or scaling procedures is beyond the scope of this paper; however, a summary of findings in the areas mentioned is presented. The structuring of several decision models for the evaluation of alternative plans or alternative system design concepts with respect to a weighted hierarchy of community decision criteria is presented. Several immediate applications appear to be feasible and these applications are enumerated. A number of possible extensions of this research are identified. It is concluded that:

1. A decision model for use in systems evaluation may be simply structured to relate utility values associated with each element in a comprehensive statement of community decision criteria with the evaluation of effectiveness of given system alternatives with respect to these criteria. Simple extensions of such a model may provide for the stratification, by socioeconomic categories or other desirable categories, of the group of persons determining the utility values associated with the community decision criteria, or, for the stratification of professional planners, the group determining the plan effectiveness values.
2. In addition to their usefulness in plan evaluation as proposed in this research, weighted community decision criteria or quantified expressions of community values could be useful in system design and capital programming.
3. The procedures structured herein may be useful in the evaluation of alternative land-use forms, detailed alternative land-use plans, detailed transportation system plans, and alternative transportation corridors in addition to the application in evaluation of alternative system design concepts as presented here.
4. Although the community decision criteria considered herein were formulated for general overall community development, 80 percent of these criteria were judged to have a meaningful relationship to a specific problem of transportation system development.
5. The application of the plan evaluation model resulted in the selection of that system design concept based on some improvements of existing at-grade facilities with a balance of new freeway construction and rail mass transit. This plan possesses an aggregate plan utility 58 percent higher than that of the first design alternative (extensive improvement of existing at-grade facilities), and 15 percent higher than the freeway-oriented design concept.
6. Extensions of this work are needed in the areas of capital allocation model formulation and the associated definition of decision variables for such a model, and the development of effective yardsticks for determining plan effectiveness based on a weighted hierarchy of community decision criteria.

REFERENCES

1. Ackoff, Russell L. Individual Preferences for Various Means of Transportation. Report of the National Cooperative Highway Research Program, March 1965.
2. Alexander, Christopher. Notes on the Synthesis of Form. Harvard University Press, 1966.
3. Churchman, C. W., and Ackoff, Russell L. An Approximate Measure of Value. Jour. Operations Research Society of America, Vol. 2, No. 2, 1954.
4. Dansereau, H. Kirk. Highway Development: Attitudes and Economic Climate. Highway Research Record 187, p. 21-32, 1967.
5. Dantzig, George B. Linear Programming and Extensions. Princeton University Press, 1963.
6. Dean, Burton V., and Nishry, Meir J. Scoring and Profitibility Models for Evaluating and Selecting Engineering Projects. Operations Research, Vol. 13, No. 4, 1965.
7. Eckenrode, Robert T. Weighting Multiple Criteria. Management Science, Vol. 12, No. 3, Nov. 1965.
8. Fishburn, Peter C. A Note on Recent Developments in Additive Utility Theories for Multiple-Factor Situations. Operations Research, Vol. 14, No. 6, Nov.-Dec. 1966.
9. Fishburn, Peter C. Methods of Estimating Additive Utilities. Management Science, Vol. 13, No. 7, March 1967.

10. Hadley, G. Nonlinear and Dynamic Programming. Addison-Wesley Publishing Co. Inc., Reading, Mass., 1964.
11. Hay, George A., Morlok, Edward K., and Charnes, Abraham. Toward Optimal Planning of a Two-Mode Urban Transportation System: A Linear Programming Formulation. Highway Research Record 148, p. 20-48, 1966.
12. Hemmens, George C. Experiments in Urban Form and Structure. Highway Research Record 207, p. 32-41, 1967.
13. Hill, Morris. A Method for the Evaluation of Transportation Plans. Highway Research Record 180, p. 21-34, 1967.
14. Hitch, C. J. Sub-Optimization Problems. Jour. Operations Research Society, Vol. 1, No. 2, May 1953.
15. Irwin, Neal A. Criteria for Evaluating Alternative Transportation Systems. Highway Research Record 148, p. 9-19, 1966.
16. Jessiman, William, Brussee, C. Roger, Tumminia, Alfred, and Brand, Daniel. A Rational Decision-Making Technique for Transportation Planning. Highway Research Record 180, p. 71-80, 1967.
17. Klein, Burton, and Meckling, William. Applications of Operations Research to Development Decisions. Operations Research, Vol. 6, No. 3, 1958.
18. Lesourne, Jacques. The Application of Operational Research to Comparing City Plans. Metra, Vol. 2, No. 4, 1963.
19. Optimal Design of a Surface Transit System. New York State Office of Transportation, 1965.
20. Pessemier, Edgar A. Measuring of Project Benefits in a Dollar Metric. Unpublished Manuscript, Purdue University, Aug. 1966.
21. Policies, Standards, and Procedures in the Formulation, Evaluation, and Review for Plans Sources. President's Water Resources Council, U. S. Govt. Printing Office, 1962.
22. Ridley, T. M. An Investment Policy to Reduce the Travel Time in a Transportation Network. Operations Research Center, College of Engineering, University of California, Berkeley, 1966.
23. Forecasts and Alternative Plans, 1990. Southeastern Wisconsin Regional Planning Commission, Planning Report No. 7, June, 1966.
24. Thomas, Edwin N., and Schofer, Joseph L. Criteria for Evaluating Alternative Transportation Plans. NCHRP Quarterly Progress Report, Transportation Center, Northwestern University, 1965.
25. Vogt, Robert S. (Vogt-Ivers and Associates). An Intermediate Report for Project No. 32—Community Goals and Objectives. Louisville Metropolitan Comprehensive Transportation and Development Program, April 1967.
26. Worrall, R. D. The Urban Panel as a Longitudinal Data Source. Highway Research Record 194, p. 62-77, 1967.

Appendix A

STATEMENTS OF COMMUNITY GOALS AND OBJECTIVES

[General community goals (numerals) are subdivided into specific objective statements]

1. Public Safety Program Development
   (a) Insure safe public facilities.
   (b) Provide for adequate public safety regulations and their enforcement.
   (c) Provide for the removal of contaminants (solid, liquid, and gaseous).

2. Public Utility and Transportation Development
   (a) Minimize maintenance costs of public facilities.
   (b) Insure maximum effectiveness of public utilities, by design and locational considerations.
   (c) Develop a balanced, effective, and integrated transportation system which provides for the accessibility requirements of each land use.

3. Economic Development Programs
   (a) Develop public improvement programs within available financial resources.
   (b) Maintain highest equitable property values.
   (c) Insure effective utilization of mineral, vegetation, air, and water resources.
   (d) Establish a strong economic base through commerce that will bring money into the community.
   (e) Establish trade development that provides maximum convenience to consumers.
   (f) Insure the optimal utilization of all land.
   (g) Achieve increased disposable income for all people.

4. Cultural Development
   (a) Preserve historic sites and areas of natural beauty.
   (b) Promote adequate public libraries, museums, and cultural activities.
   (c) Protect meaningful local tradition and encourage civic pride.

5. Health Program Development
   Establish the mechanism for adequate preventive and remedial health programs and facilities.

6. Education Program Development
   Develop educational facilities and opportunities for citizens at every level.

7. Welfare Program Development
   (a) Eliminate injustice based on discrimination.
   (b) Develop needed public welfare programs.
   (c) Encourage development of religious opportunities.
   (d) Develop an aesthetically pleasing environment.

8. Recreation Program Development
   (a) Establish open-space programs.
   (b) Provide adequate recreational facilities utilizing parks, rivers, and lakes.

9. Political Framework
   (a) Improve the framework (channels, systematic use) for citizen participation in governmental functions.
   (b) Establish equitable taxation policies (bases, mixes, rates).
   (c) Achieve efficient governmental administration, representative of all citizens.
   (d) Develop adequate government staffs and personnel programs (high job standards, reasonable salary ranges, effective delegation of authority).
   (e) Establish sound governmental fiscal programs.
   (f) Develop an effective, long-range, metropolitan-wide planning process.
   (g) Establish effective control mechanisms.

10. Housing Development
    (a) Encourage rehabilitation and conservation neighborhood programs.
    (b) Provide adequate low-cost housing.
    (c) Develop neighborhood units.
    (d) Promote a wide variety of housing types as required within the community.

Appendix B

CRITERIA WEIGHTING TECHNIQUES

This Appendix presents a summary of techniques used in obtaining a weighted hierarchy of community goals and objectives. Fishburn (8, 9) lists and classifies 24 methods of estimating utility values. Recent research (7) has evaluated various methods of collecting the judgments of experts relative to the reliability and efficiency of these methods.

Ranking, rating and two variations of the method of successive comparisons are summarized here.

Ranking Technique

Each member of the various judging panels was asked to place a raw rank by each criterion in the given lists of criteria. The most important criterion was assigned a raw rank of 1, the second most important, a raw rank of 2, etc. Criteria weights, or utility values, are developed as follows.

In general, there will be n criteria in a list of community goals or objectives. A converted rank of n-1 will be assigned to the criterion receiving a raw rank of 1, a converted rank of n-2 to the criterion receiving a raw rank of 2, . . . , and a converted rank of 0 to the criterion receiving a raw rank of n. The composite rank (R_j) for a given objective (j) will be determined by summing the converted ranks of all of the m judges; that is,

$$R_j = \sum_{i=1}^{m} R_{ij}, \; j = 1, 2, \ldots, n$$

In this expression,

R_j = composite rank of criterion j,
R_{ij} = converted rank of criterion j established by judge i,
n = number of criteria, and
m = number of decision-makers on the panel of judges.

The composite ranks thus determined are then normalized in the following manner:

$$u_j = \frac{R_j}{\sum_{j=1}^{n} R_j}, \; j = 1, 2, \ldots, n$$

where j = composite weight or utility value associated with community decision criterion j.

Rating Technique

The rating scale technique is the most popular of all procedures used for collecting the judgments of individuals. The numerical type rating scale is used but descriptors are not associated with the integer points on the numerical scale. Appropriate descriptors that would not bias the judges could not be determined.

The lists of criteria to be weighted (i.e., the lists of community goals and objectives) are placed in a column adjacent to a scale marked in units continuously from ten to zero (top to bottom). A rating of zero indicates that there is no value associated with a given criterion and a rating of ten is the highest that may be assigned. Any value along the unbroken continuum may be assigned to any criterion. Even though an approximation will be made of non-integer ratings, the judge was permitted to associate with each criterion an integer or non-integer position on the rating scale. The rating assigned to criterion j by judge i is represented by V_{ij}. Utility values (u_j) or criteria weights for each criterion are determined in the following manner:

$$V_j = \sum_{i=1}^{m} V_{ij} \quad j = 1, 2, \ldots, n$$

$$u_j = \frac{V_j}{\sum_{j=1}^{n} V_j} \quad j = 1, 2, \ldots, n$$

Method of Successive Comparisons

The following procedures (SC-1) are based on the method of successive comparisons (3). The modification of the procedures is as follows.

Step 1 is carried out by placing the criteria in rank order by the utility value determined from the average results of the ranking and rating methods. Step 2 is completed by simply associating with each criterion that average value. The judges then were asked to check the rank order of the criteria as determined by consensus. If the judge agrees, the procedures move to Step 3. If he disagrees, he subjectively reassigns utility values.

Step 1. Rank the criteria according to preference:

$$G_1 \gtrsim G_2 \gtrsim G_3 \gtrsim \ldots \gtrsim G_{n-1} \gtrsim G_n$$

where $G_1 \succ G_n$.

Step 2. Tentatively assign the value $u_1' = 1.00$ to G_1. Then assign preliminary utility measurements u_j' to the remaining criteria in such a manner that u_j' seems to reflect the magnitude of preference for G_j.

Step 3. Compare G_1 vs $G_2 \wedge G_3 \wedge \ldots \wedge G_n$ or G_1 vs $\bigwedge_{j=2}^{n} G_j$

(a) If $G_1 \succ \bigwedge_{j=2}^{n} G_j$ then, if necessary, adjust u_1' so that $u_1' > \sum_{j=2}^{n} u_j'$ and after making this adjustment go to step 4.

(b) If $G_1 \sim \bigwedge_{j=2}^{n} G_j$ then, if necessary, adjust u_1' so that $u_1' = \sum_{j=2}^{n} u_j'$ and after making this adjustment go to step 4.

(c) If $G_1 \prec \bigwedge_{j=2}^{n} G_j$ then, if necessary, adjust u_1' so that $u_1' < \sum_{j=2}^{n} u_j'$. Then repeat step 3 and compare G_1 vs $\bigwedge_{j=2}^{n-1} G_j$; that is, drop the criterion G_n. Continue dropping the least preferred criterion and comparing until situation 3(a) or 3(b) is encountered. This process must terminate since $G_1 \succsim G_2$ from step 1.

Step 4. Drop G_1 from consideration and repeat the entire procedure (steps 1 to 3) for G_2. Continue with G_3 and so on until the comparison of G_{n-2} vs $G_{n-1} \wedge G_n$ is completed. Care should be taken to insure retention of the invariance in u_1', u_2', etc. That is, in adjusting values such as u_2' the relationship $u_2' > u_1'$ must not be accepted in violation of step 1.

Step 5. The values of u_j' obtained in steps 1 through 4 must now be normalized as follows:

$$u_j = \frac{u_j'}{\sum_{j=1}^{n} u_j'}$$

It is to be noted that the numerical values for u_j are relative, hence the deletion or addition of a criterion G_k, where $u_k \neq 0$, would affect the values calculated.

Successive Comparisons Method: An Alternative Approach

An alternative procedure is proposed by Churchman and Ackoff (3) when a large number of criteria (7 or more) are to be considered. This alternative procedure may be useful in the specific application of weighting planning criteria. Churchman and Ackoff suggest the following alternative procedures:

Step 1. Rank the entire set of decision criteria on the basis of the average weights obtained by the ranking and rating techniques.

Step 2. Select the highest ranked criterion from the entire set. Let G_s represent this standard criterion. By random assignment, subdivide the criteria that remain into approximately equal-sized groups of no more than 5 criteria per group. Each criterion, other than the standard G_s, should be included in one and only one group.

Step 3. Insert G_s into each group and assign a criteria weight of 1.00 to G_s (i.e., $u_s' = 1.00$).

Step 4. With modifications made above, follow the procedure of steps 1 through 4 of the preceding section to obtain unstandardized criteria weights (utility values) for the objectives in each of the groups formed in step 3 above. (Note: in adjusting the u_j' values, do not change the value of u_s'.)

Step 5. Compare the ranking obtained for all criteria with the ranking of step 1. If the rank orders differ, reconsider the ranking and, if necessary, repeat step 4 of this alternative procedure.

Step 6. When consistent results are obtained, normalize the criteria weights by dividing the value assigned to each criterion by the sum of the values assigned to all criteria. That is,

$$u_j = \frac{u_j'}{\sum_{j=1}^{n} u_j} \quad j = 1, 2, \ldots, n$$

Appendix C

Criteria (See Appendix A for objective statements)	Weighting Techniques				Average Values		Range	
	Ranking		Rating					
	u_j	Rank Order	u_j	Rank Order	u_j	Rank Order	u_j	Rank Order
1a	0.0142	26	0.0270	15	0.0206	17	0.0128	11
1b	0.0505	6	0.0326	11	0.0415	6	0.0179	5
1c	0.0648	3	0.0335	7	0.0496	5	0.0313	4
2a	0.0000	35	0.0280	14	0.0140	31	0.0280	21
2b	0.0611	4	0.0393	4	0.0502	4	0.0218	0
2c	0.0611	5	0.0449	3	0.0530	3	0.0162	2
3a	0.0217	16	0.0192	20	0.0204	19	0.0025	4
3b	0.0031	34	0.0144	31	0.0087	35	0.0113	3
3c	0.0311	9	0.0168	24	0.0239	12	0.0143	15
3d	0.0248	12	0.0192	21	0.0220	16	0.0056	9
3e	0.0155	21	0.0156	26	0.0155	26	0.0001	5
3f	0.0279	10	0.0180	23	0.0229	14	0.0099	13
3g	0.0279	11	0.0168	25	0.0223	15	0.0111	14
4a	0.0089	31	0.0295	12	0.0192	21	0.0206	19
4b	0.0248	13	0.0348	5	0.0298	9	0.0100	8
4c	0.0069	32	0.0250	16	0.0159	25	0.0181	16
5	0.0925	2	0.1050	2	0.0587	2	0.0125	0
6	0.1555	1	0.1272	1	0.1413	1	0.0283	0
7a	0.0173	20	0.0223	17	0.0198	20	0.0050	3
7b	0.0124	27	0.0186	22	0.0155	27	0.0062	5
7c	0.0035	33	0.0141	32	0.0088	34	0.0106	1
7d	0.0111	29	0.0193	18	0.0152	28	0.0082	11
8a	0.0148	24	0.0340	6	0.0244	10	0.0192	18
8b	0.0148	25	0.0327	10	0.0237	13	0.0179	15
9a	0.0206	17	0.0116	35	0.0161	24	0.0090	18
9b	0.0149	22	0.0145	28	0.0147	29	0.0004	6
9c	0.0218	14	0.0145	29	0.0181	22	0.0073	15
9d	0.0103	30	0.0135	33	0.0119	32	0.0032	3
9e	0.0149	23	0.0145	30	0.0147	30	0.0004	7
9f	0.0206	18	0.0154	27	0.0180	23	0.0052	9
9g	0.0114	28	0.0125	34	0.0119	33	0.0011	6
10a	0.0321	8	0.0330	8	0.0325	8	0.0009	0
10b	0.0195	19	0.0285	13	0.0240	11	0.0090	6
10c	0.0218	15	0.0193	19	0.0205	18	0.0025	4
10d	0.0413	7	0.0330	9	0.0371	7	0.0083	2

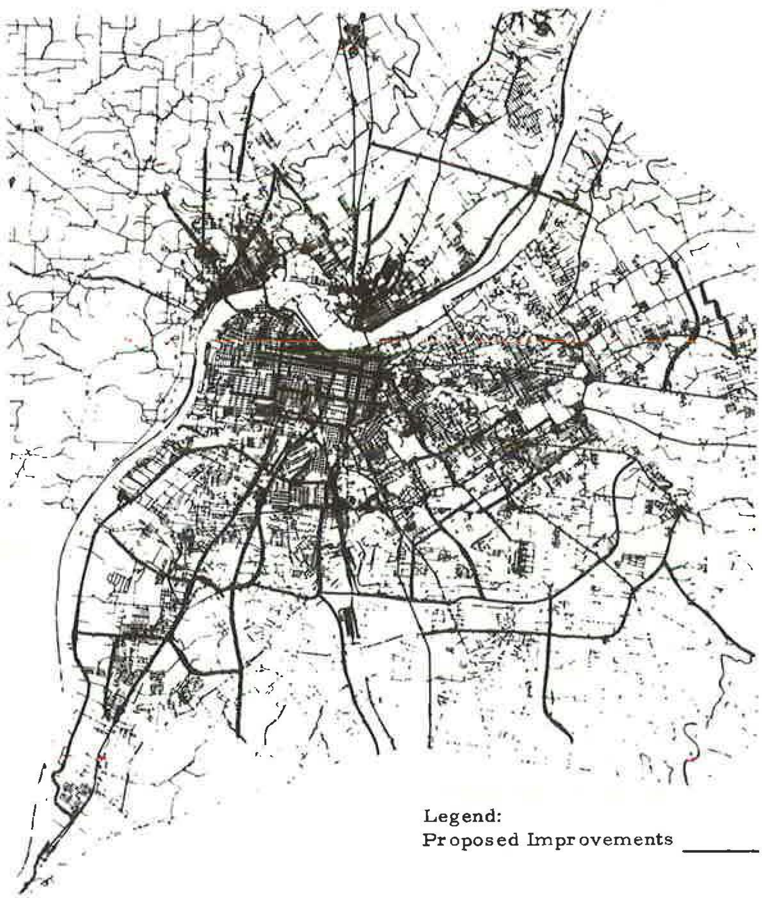
Appendix D

No.	Criterion (Objective) Statement	Effectiveness Value		
		Plan 1	Plan 2	Plan 3
1a	Insure safe public facilities.	0.24	0.56	0.82
1b	Provide for adequate public safety regulations and their enforcement.	0.32	0.76	0.64
1c	Provide for the removal of contaminants (solid, liquid, and gaseous).	0.30	0.44	0.62
2a	Minimize maintenance costs of public facilities.	0.44	0.60	0.62
2b	Insure maximum effectiveness of public utilities, by design and locational considerations.	0.66	0.70	0.62
2c	Develop a balanced, effective, and integrated transportation system which provides for the accessibility requirements of each land use.	0.40	0.62	0.84

3a	Develop public improvement programs within available financial resources.	0.72	0.74	0.54
3b	Maintain highest equitable property values.	0.58	0.60	0.78
3c	Insure effective utilization of mineral, vegetation, air, and water resources.	—	—	—
3d	Establish a strong economic base through commerce that will bring money into the community.	0.44	0.76	0.94
3e	Establish trade development that provides maximum convenience to consumers.	0.62	0.70	0.72
3f	Insure the optimal utilization of all land.	0.62	0.68	0.76
3g	Achieve increased disposable income for all people.	0.45	0.80	0.95
4a	Preserve historic sites and areas of natural beauty.	0.52	0.60	0.72
4b	Promote adequate public libraries, museums, and cultural activities.	0.66	0.66	0.70
4c	Protect meaningful local tradition and encourage civic pride.	0.55	0.70	0.85
5	Establish the mechanism for adequate preventive and remedial health programs and facilities.	—	—	—
6	Develop educational facilities and opportunities for citizens at every level.	0.30	0.70	1.00
7a	Eliminate injustice based on discrimination.	0.67	0.53	0.43
7b	Develop needed public welfare programs.	—	—	—
7c	Encourage development of religious opportunities.	0.30	0.70	1.00
7d	Develop an aesthetically pleasing environment.	0.45	0.52	0.68
8a	Establish open-space programs.	0.45	0.65	0.75
8b	Provide adequate recreational facilities utilizing parks, rivers, and lakes.	0.70	0.66	0.54
9a	Improve the framework (channels, systematic use) for citizen participation in governmental functions.	—	—	—
9b	Establish equitable taxation policies (bases, mixes, rates).	0.68	0.62	0.32
9c	Achieve efficient governmental administration, representative of all citizens.	—	—	—
9d	Develop adequate government staffs and personnel programs (high job standards, reasonable salary ranges, effective delegation of authority).	—	—	—
9e	Establish sound governmental fiscal programs.	0.67	0.67	0.40
9f	Develop an effective, long-range, metropolitan-wide planning process.	0.60	0.68	0.72
9g	Establish effective control mechanisms.	—	—	—
10a	Encourage rehabilitation and conservation neighborhood programs.	0.70	0.62	0.58
10b	Provide adequate low-cost housing.	0.40	0.53	0.80
10c	Develop neighborhood units.	0.64	0.58	0.54
10d	Promote a wide variety of housing types as required within the community.	0.42	0.62	0.82

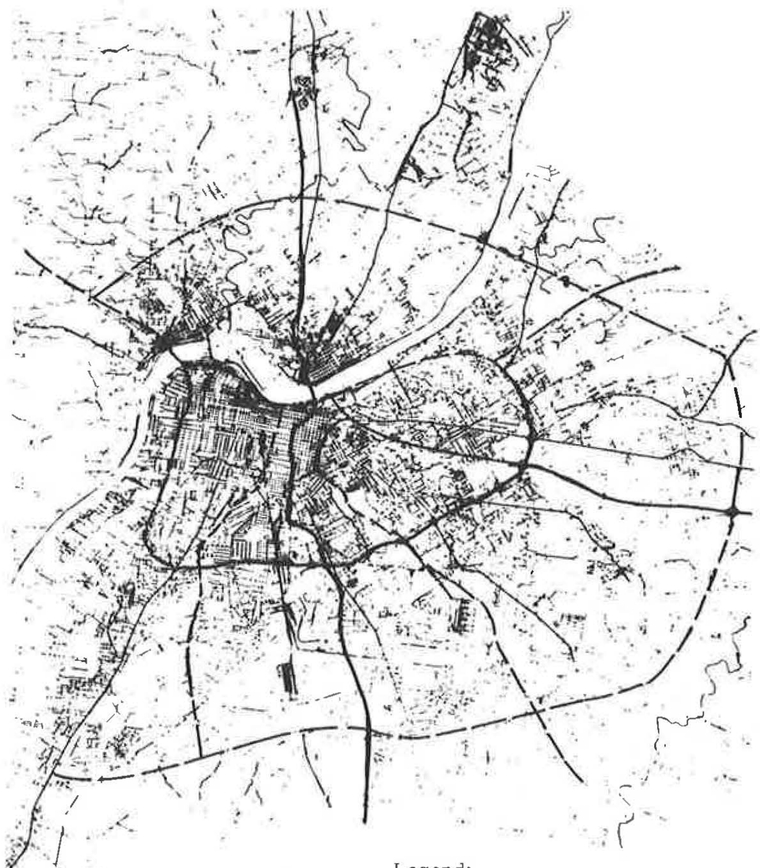
Appendix E

ALTERNATIVE PLAN NUMBER 1: EXTENSIVE IMPROVEMENTS OF EXISTING ARTERIAL FACILITIES



Appendix F

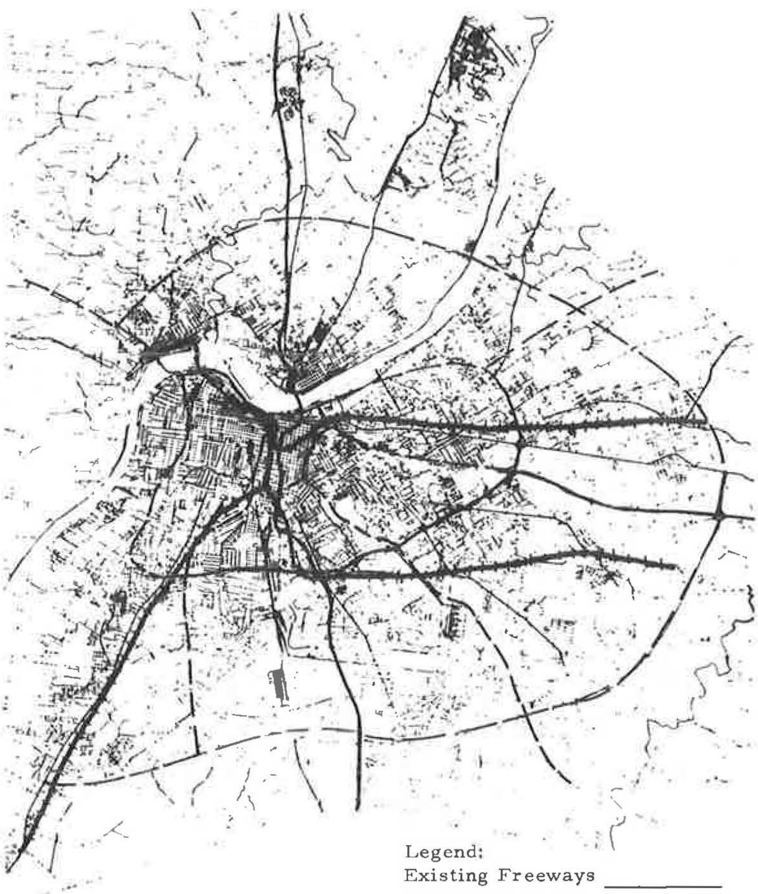
ALTERNATIVE PLAN NUMBER 2: MAJOR FREEWAY CONSTRUCTION, NO RAPID MASS TRANSIT



Legend:
Existing Freeways ___________
Proposed Freeways -----------

Appendix G

ALTERNATIVE PLAN NUMBER 3: MAJOR FREEWAY CONSTRUCTION WITH RAPID MASS TRANSIT



The Rank-Based Expected Value Method of Plan Evaluation

KENNETH SCHLAGER, Chief Systems Engineer, Southeastern Wisconsin Regional Planning Commission

•THIS paper is concerned with the application of a new method of plan evaluation, the rank-based expected value method, to land use transportation plans in urban and regional contexts. The methodology is new in its application to urban plan evaluation, but it has been used for a number of years in corporate long-range planning.

The plan evaluation problem may be simply stated. Given a set of alternate plans, plan evaluation is concerned with the selection of one of the alternate plans that best fulfills the objectives of the planning project. Plan evaluation, by its very definition, assumes that a number of alternative plans have already been synthesized, and that these plans have been screened to eliminate plans that are infeasible because of certain defined constraints in the planning objectives.

Although the method to be discussed is applicable to the entire field of urban and regional planning, it is probably most useful in the evaluation of a land-use plan. The facility plans that are based on the land-use plan may also be evaluated with the rank-based expected value method, but the benefits are not as pronounced as they are in land-use planning. Generally, land-use plans are more difficult to evaluate than facility plans since the objectives of such plans are more qualitative in nature. In the usual case, then, it is expected that the method would be applied either to the land-use plan or to a comprehensive set of land-use and facility plans as part of an urban master plan.

PROBLEMS IN PLAN EVALUATION

The traditional approach to plan evaluation is the benefit-cost method. Benefit-cost evaluation involves a tabulation of all the benefits and costs of the project followed by a comparison of the sum total of all the benefits and costs in order to arrive at an estimate of the value of the project. Although in a theoretical sense the benefit-cost method would seem to provide a satisfactory solution to most plan evaluation problems, in practice it suffers from the following difficulties:

1. It is quite difficult to evaluate intangible benefits. The estimation of the intangible benefits can often distort the whole value estimation process.
2. The expense of obtaining all of the benefit and cost data needed to apply the benefit and cost method is often prohibitive.
3. Benefit-cost analysis is not easily related to indirect costs of related projects or programs.
4. It is difficult to allow for various uncertainties of implementation in benefit-cost analysis.

If another method is to represent an improvement over benefit-cost analysis, it must overcome, to some degree at least, the difficulties listed. What is really needed is a method that would handle intangible benefits and indirect costs with less data collection and analysis effort.

RANK-BASED EXPECTED VALUE METHOD

The rank-based expected value method is quite simple in both concept and application. Application of the method involves the following steps:

1. The rank ordering of plan objectives;
2. The rank ordering of plans under each objective; and
3. The estimation and assignment of a probability of implementation for each of the plan alternatives.

The philosophy of rank ordering relates to the concept that intangible benefits and costs are easier to rank in preference order than they are to assign a scalar benefit or cost value. The concept of rank ordering is proposed to overcome the difficulties of scalar estimation inherent in the intangible benefits and the indirect costs of planning

Plan	Specified Development Objective	Balanced Allocation of Land	Natural Resource Conservation	Facility Costs	Plan Value, V $V = p \Sigma (n_1m_1 + n_2m_2 + n_nm_n$
		Rank Order Value of Objective n = 2	Rank Order Value of Objective n = 3	Rank Order Value of Objective n = 1	
		Rank Order Value of Plan, m	Rank Order Value of Plan, m	Rank Order Value of Plan, m	
1	Probability of Implementation p = 0.6	3	1	3	0.6 [(2 × 3) + (3 × 1) + (1 × 3)] = 7.2
2	Probability of Implementation p = 0.5	2	2	1	0.5 [(2 × 2) + (3 × 2) + (1 × 1)] = 5.5
3	Probability of Implementation p = 0.9	1	3	2	0.9 [(2 × 1) + (3 × 3) + (1 × 2)] = 11.7

Figure 1. Theoretical application of the method.

projects. The probability of implementation concept is designed to introduce the aspect of uncertainty into plan evaluation.

The detailed methodology of the rank-based expected value method may be understood from a more detailed description of the sequence of activities:

1. All objectives, n in number, are ranked in order of importance and assigned values of n, n minus 1, n minus 2, ..., to n minus (n - 1) in descending rank order.
2. The alternative plans, m in number, are ranked under each of the specific land-use development objectives and assigned a value of m, m minus 1, m minus 2, ..., to m minus (m - 1) in descending rank order.
3. A probability, p, of implementation is assigned to each of the plans being ranked.
4. The value, V, of each alternative plan is then determined by summing the products of n times m times p for each of the specific development objectives:

$$V = p \Sigma (n_1m_1 + n_2m_2 + \dots + n_nm_n)$$

The matrix table shown in Figure 1 illustrates a simple theoretical application of the method for three specific development objectives. In the hypothetical plan evaluation shown in the table, Plan 3 would be selected as that plan which best meets the development objectives.

HISTORICAL DEVELOPMENT

The development of the rank-based expected value method is best understood in terms of its application in the business arena, since it was first developed for application to corporate planning in a large multi-product manufacturing firm.

It was applied to that top management level of problems usually designated as corporate strategy. Corporate strategy involves the selection of new products in new markets and the modification of old products and markets to establish a continuing dynamic product-market portfolio for a business. In its corporate application, the development of such a strategy involves the screening of a large number of alternatives for feasibility prior to the direct application of the rank-based expected value method. After this initial screening, the corporate objectives are ranked, the plan alternatives are ranked, probability of implementation is estimated, and each of the plan alternatives are then evaluated as described in the foregoing urban planning application.

The similarities of corporate and urban planning are many. Like urban planning, corporate planning must deal with intangibles. Long-range profitability is not the only goal of most business firms, and even long-range profitability, since it is not possible to obtain precise estimates of future revenues and costs, is really best evaluated as an intangible itself. There are, in addition, many social, political, and legal constraints on a firm that must be treated as intangible objectives. Corporate planning is also fraught with uncertainty. Most product and market ventures are doomed to failure, so it is vital to introduce the concept of probability of the implementation into any methodology of corporate plan evaluation. Although the concept of probability of implementation is somewhat different in the corporate and urban planning functions, it must be dealt with in a direct and straightforward manner in both.

APPLICATION OF THE RANK-BASED EXPECTED VALUE METHOD IN THE SOUTHEASTERN WISCONSIN REGIONAL PLANNING COMMISSION

The potential and problems of the rank-based expected value method are better understood from an application in a specific plan evaluation project. This method was applied to the evaluation of the land use-transportation plans developed by the Southeastern Wisconsin Regional Planning Commission.

In this application, the first set of problems related to a clarification of the meaning of objectives and standards. Urban planning objectives often tend to be vague, overlapping, and even irrelevant. To relate the method to actual objectives and their resulting design standards entails a need for discipline in the precise definition of each plan objective and standard. To avoid semantic difficulties, the following definitions used in southeastern Wisconsin will be assumed:

Objective: A goal or end toward the attainment of which plans are directed.
Standard: A criterion used as a basis of comparison to determine the adequacy of plan proposals to attain objectives.

Further confusion in the definition of planning objectives results from the indiscriminate mixing of different levels of objectives as they affect the preparation of regional plans. Some planning objectives are important only at the level of regional plans, others affect planning at the community level, while still others are important only at the neighborhood level. To eliminate this source of confusion, a hierarchy of objectives and standards was classified in southeastern Wisconsin. This made it possible to utilize only those objectives and standards that applied for a particular level of plan evaluation.

The procedure used in applying the method was basically the method previously described with some modification. The plan evaluation activities involved were the following:

1. Each of the three plans was evaluated for each of the design standards and, since each design standard applied to only one objective, the ranking of the plans under each objective was determined from the composite ranking of the plans for the design standards of the objective.
2. Each of the three objectives was ranked as described previously.
3. The probability of implementation of each plan was estimated.
4. The value of each plan was calculated as previously described.
5. The plan with the highest plan value was selected for recommendation by the planning staff.

The response of both the planning staff and the Commission members to the rank-based expected value method was quite favorable. The staff found it a useful tool in clarifying the objectives and design standards and for introducing a discipline of thought into the plan evaluation process without the need for extensive benefit and cost data analysis. Criticisms of the method revolved primarily about the concept of probability of implementation.

It is apparent from the evaluation plan values shown in Figure 2 that the probability of implementation is an important factor in plan selection. The Satellite City Plan would have ranked higher if it were not for the low likelihood of implementing this plan. Although some philosophical problems of the purpose of planning are introduced by this probability concept, most members of the staff and Commission agreed that this concept was quite crucial in plan selection in southeastern Wisconsin.

Plan	Major Group Objective	Provide for a Balanced Allocation of Land[a]	Provide for an Appropriate Spatial Distribution of Land Uses[a]	Meet the Design Requirements of the Major Land Uses[a]	Plan Value
		Rank Order of Group Objective = 3	Rank Order of Group Objective = 2	Rank Order of Group Objective = 1	
		Rank Order Value of Plan[b]	Rank Order Value of Plan[b]	Rank Order Value of Plan[b]	
Controlled Existing Trend	Probability of Implementation = 0.6	2	3	1	7.8
Corridor	Probability of Implementation = 0.3	1	2	2	2.7
Satellite City	Probability of Implementation = 0.1	3	1	3	1.4

[a]Includes the objectives listed under this group in Appendix Table A-31, SEWRPC Planning Report No. 7, Vol. 2.
[b]Based on the rank order value as shown in Appendix Table A-29, SEWRPC Planning Report No. 7, Vol. 2.

Source: SEWRPC

Figure 2. Land-use plan selection criteria.

There is little question that further development of the rank-based expected value method is necessary before it could be applied on a wide scale. The most important need is for the development of computer programs and techniques for sensitivity analysis. Confidence in the method would be increased if it were possible to obtain easily some of the sensitive values at which the plan selected would change. By determining these sensitive points, the planner could better evaluate whether his information is sufficient to recommend the plan suggested by the method. Personnel at the consulting engineering firm of Consoer, Townsend in Chicago have been applying the method in this manner. In an application of the method to the evaluation of alternative route locations, a number of sensitivity analyses were performed to develop confidence in the route location that was recommended. Documentation of this application is not yet available but is expected in the near future.

Although the rank-based expected value method is not offered as a panacea to all problems of urban plan evaluation, it is believed that the method provides a contribution to a difficult aspect of urban and regional planning.

Discussion

BYRON D. STURM, Assistant Director, Akron Metropolitan Area Transportation Study

•PLAN evaluation involves the examination and analysis of alternative plans to select the plan that best satisfies the goals and objectives of the study area. The technique employed in the plan evaluation phase of a transportation planning program should be the object of much concern, early in the planning process. Each study area should subscribe to a comprehensive plan evaluation technique that will assist plan selection and, most important of all, insure implementation.

Two papers have been presented that describe techniques for plan evaluation. Both papers point out a real problem that exists in the transportation planning process today, and they are attempts to provide solutions to the problem. The problem pointed out is that, while very sophisticated techniques are available to inventory, analyze, and forecast travel patterns, land use, and socioeconomic activity in developing a transportation-land use plan, the same sophistication is not available to evaluate plan alternatives to insure the selection of the best plan and its implementation.

THE RANK-BASED EXPECTED VALUE METHOD OF PLAN EVALUATION

Mr. Schlager's paper deals with a form of plan evaluation that is new to urban planning, but that has been used before in industry. Basically he has sought to provide a practical means of communicating the advantages and disadvantages of a set of alternative land-use and transportation plans to seek the selection of the best plan.

The paper points out the fact that the common practice of cost-benefit analyses cannot take all factors into account because of the difficulty of quantifying intangible criteria. The author proposes a simple ranking of alternative plans with regard to the manner in which they meet a ranked set of regional planning objectives. In other words, through necessary contact with technical and policy decision-makers, the area's planning objectives are ranked according to a consensus of preference, and then each plan is ranked according to its ability to meet each goal or objective. With this approach, much is accomplished to insure the implementation of the selected plan because the people of the area have been involved in the decision-making process.

Further, the author has added a factor to the decision-making process called the "probability of implementation" which tends to temper "optimistic or unrealistic" plans with an appropriate air of certainty. Many times, worthwhile projects have not been constructed for lack of funds, inadequate promotion, inadequate technical justification, or because their benefits were diffused by unrealistic projects.

Therefore, the technical and policy sessions that were held to identify goals and standards and alternate transportation and land-use plans, to rank goals and objectives in order of preference, and to rank plans according to their ability to satisfy specific goals or objectives were invaluable. This procedure should do much to provide a dynamic and successful transportation planning process in the southeastern Wisconsin area.

SYSTEMS EVALUATION: AN APPROACH BASED ON COMMUNITY STRUCTURE AND VALUES

The paper by Messrs Schimpeler and Grecco is an approach directed toward the same problem. However, the techniques for ranking the regional goals and objectives and determining the effectiveness of the various plan alternatives were developed through the application of decision-making theory and operations research that has been the subject of considerable attention over the past 15 years, mostly in industry.

This paper, like Mr. Schlager's, proposes extensive contact with technical and policy decision-makers in establishing statements of goals and objectives and ranking these statements in order of a consensus of preference. Through ranking and/or rating techniques, a utility value is determined relative to the importance of each goal to the region. The professional staff then determines the effectiveness of each alternative plan in satisfying each goal and objective. Total plan effectiveness is then measured for each plan through a decision model by summing all products of each plan effectiveness value times the utility value for each regional goal or objective.

Schimpeler and Grecco also point to the use of this technique in capital programming and route location studies. An approach similar to this has been used in the Akron Metropolitan Area Transportation Study program to solve a route location problem, and the City of Akron is considering a similar technique in capital investment programming.

The authors suggests that additional research will be required in the development of a succinct, comprehensive, well-defined statement of regional goals and objectives. Statements of general goals do not work because it is impossible to draw definite distinctions between how various plans affect them, much less attempt to rank them in order of preference, because all seem equal.

Most important, however, once a goals statement is defined, is the establishment of measuring devices for each goal or objective so that the effect of each plan can be compared objectively. The various interpretations of goals must be understood and agreed upon by all concerned in the planning process. And, finally, the professional task force that evaluates the effectiveness of each plan must be composed of a good cross section of the professional disciplines required to develop the plan.

SUMMARY

The success of any transportation planning process in the country can be measured in very simple terms by answering the question: "Is the plan being implemented?" In other words, are the transportation facilities recommended by the plan being constructed on schedule and have the development policies set forth in the plan been used to guide development and growth in the urban area?

Many factors will determine whether or not a plan will be implemented, but probably the most significant factor will be how the plan was presented and accepted. Were the proper decision-makers in the region made part of the plan development and evaluation process? If they were not, they may feel that the plan is impractical or not well-founded and, therefore, will not work toward its implementation.

An objective for any study staff, in order to insure plan implementation, is to establish a good channel of communication with the policy-makers in the area. In this business, this can only be achieved by talking at the level of the policy-maker who is not necessarily concerned with gravity models, modal splits, O-D surveys, all-or-nothing traffic assignment, capacity restraints, and the rest of the technical jargon. He is more concerned with his transportation needs and the implications of various solutions to those needs. If a staff can show the policy-maker his needs and related improvement costs and involve him in the methodology of measuring how alternate solutions to his needs satisfy the various criteria in selecting the best alternate, then that study will be successful.

In effect, the policy-maker must understand the plan evaluation procedure. Further, the policy-maker must be given an active part in the development of and application of the plan evaluation procedure to determine the best transportation-land use plan for the region.

The Decision-Making Forum

Improving the Decision-Making Process

ROBERT C. EINSWEILER, Metropolitan Council, Minneapolis-Saint Paul, Minnesota

•THERE IS a trend in the evaluation of transportation facilities that is quite apparent from the papers presented in this RECORD. That trend is ever-increasing scope and comprehensiveness.

Analyses have advanced from routes treated individually to systems. Forecasting demand has advanced from an extrapolation of past use of a route to the simulation of traffic on systems based on the land development patterns and travel behavior of people in urban areas. But the concern here is not with either of these aspects. It is with the evaluation of whether, when, and how to build future facilities—individual routes or systems—regardless of how the demand is determined.

NEXT STEPS IN TRANSPORTATION SYSTEM EVALUATION

It is helpful to divide the trend in evaluation into four phases. In the first phase the critical factor—and sometimes the only one—used in reaching a decision was construction cost. Few are using this method today. In the second phase, the predominant evaluative technique in use today, construction costs are still employed but decisions are based primarily on user benefit-cost analyses. The papers of George and Campbell are examples. The third phase, just emerging, encompasses the other two and adds evaluation of the impact of the highway on the community, or non-highway user benefit-cost analyses. The papers of Schlager and Schimpeler are examples. Although the third phase is only beginning, and although it has been hailed as a great step forward, the fourth phase is already overdue and we must get to it. This is the one we are attempting to pursue in the Twin Cities. This phase views a transportation facility as an investment of scarce resources in competition with other needs. This gets to the question not only of where the route should be built, but whether it should be built at all. It asks whether the money would return a greater benefit to society if spent on something else.

Fully perfected techniques for this type of analysis do not exist today, although the basic concept of economic utility is employed. We are pushing ahead as rapidly as time and funds permit to develop workable measures. These evaluative techniques are needed. They are needed far more than techniques to improve our ability to forecast traffic or development. We already have better forecast data than our current metropolitan decision-making apparatus can use.

We need not wait until the techniques are developed, however. We can ask questions of those who make development decisions today so they may assess how valid their decisions are. For example, in the Twin Cities Metropolitan Area one of the counties levied $6.5 million in real and personal property taxes in 1967 for its road and bridge fund. Many of these dollars were spent on a highway that goes through a metropolitan park site. This park could not be acquired by the park reserve district in that county because it ran out of funds. The park district is also supported by a real and personal property tax levy. The total purchase price of the park was $2 million. To add further to the discussion, the highway is being built as a freeway when, according to metropolitan

Paper sponsored by Committee on Transportation System Evaluation and presented at the 47th Annual Meeting.

spacing criteria, it should be developed as a lower level facility. The park is being threatened by development and is one of the highest priority sites for acquisition. The question is whether some of the funds should have been spent acquiring the park, a one-time acquisition, or whether they were really best spent on the highway. We do not have the answer, but we certainly consider it worthwhile to put the question to those who make the decisions.

As this trend in evaluative techniques indicates, great advances are being made. But do the techniques meet our needs? Most are geared toward evaluating complete plans—and the plans leave much to be desired—rather than day-to-day planning. They are geared toward evaluating the assumed results of a comprehensive set of unstated future decisions, rather than assessing the worth of current, individual decisions in a comprehensive context. Most are not sufficiently decision-oriented.

WHY SUCH PALLID PLANS?

Many metropolitan plans today look like present trends extrapolated or mildly organized sprawl. Why is this so? One eminent transportation planner has pegged it as "the decline of the world-changers."

One reason for this is increased understanding of how urban areas function, a situation made possible by the computer, large amounts of research money, and the contributions of the social scientists among others. The result has been a restraining influence on the designer. The result is good.

Another reason is that it is common practice in planning and transportation studies to accept the present method of making decisions as given. We cannot accept this. We cannot even accept the trend to an increasingly more organized means for making decisions on the metropolitan level. We must, as concerned professionals, actively work in our research and planning programs to propose and to implement proposals to create an improved capability for making metropolitan-level development decisions.

We must not give up the idea of shaping our destiny rather than extrapolating the future from the past. The current rate of change in large metropolitan areas cannot be accommodated by the historic process where each decision is made in response to the immediately preceding one. The old process is incapable of producing a pattern of development that will allow a style of life anywhere near what the people in the Twin Cities, at least, have said they would like to have. Therefore, those who are making plans for metropolitan areas, whether functional plans or total comprehensive plans, have as much responsibility to study and to propose changes in the capabilities of those who must make development decisions as they have for the content of the decisions themselves.

An example of this from the Twin Cities experience is the spacing of interchanges on freeways. Present interchanges are spaced much more closely than proposed, professionally responsible standards recommend. The reason is simple. A local municipality must give its consent to any highway construction within its boundaries. Access to the freeway and the possibility of commercial or industrial development in the municipality are paid for by property taxes. Add these up and the community wants more interchanges.

To prepare a plan using the desired interchange spacing standards requires a concomitant action to diminish the "veto" power of the municipality and to diminish the inter-municipal rivalry for tax base by new tax legislation or metropolitan development controls. Without these accompanying actions, the plan will be just so much paper. If, instead, the plan is drawn to fit the present decision-making mechanism, desirable development will not occur.

THE JOINT PROGRAM RESPONSE

The Joint Program is a continuing planning program for the Twin Cities Metropolitan Area undertaken collaboratively by 13 existing public agencies. Work was begun in the spring of 1962. The objective of the Program is to encourage development decisions that will enhance both the livability and efficiency of the metropolitan environment.

As do most new ventures, the program to this point has fallen short of aspirations. But the material produced under this approach has proved sufficiently effective in handling referrals and other current development decisions that it is being sharpened and expanded.

Purpose of Planning to Guide Development Decisions

The statute that created the Twin Cities Metropolitan Planning Commission in 1957 stated that "The Commission shall make plans for the physical, social, and economic development of its metropolitan area with the general purpose of guiding and accomplishing a coordinated and harmonious development of the area." Too frequently the purpose of planning is viewed as making the plans rather than guiding development. Plans should be viewed as one of a number of tools to guide development decisions and to make rational decisions about how to use scarce resources—dollars, man-hours of skilled people, land, and others.

The Metropolitan Plan

If a metropolitan plan is to be a tool in guiding development decisions, it must contain agreed-upon rules for day-to-day decisions. The new "policies plans" do this. But if the plan is to be accepted, it must project some image of where the community will be in the future if it follows the rules. The map-oriented master plans or "blueprint plans" do this. There is a third approach, "incrementalism," where components are added to urban systems to meet daily needs with no long-range view in mind.

Map-oriented plans gather dust and die. They do not show how to reach the desired future state, so public officials ignore them. And they show so precisely how pieces of land will be affected that citizen opposition occurs. The incrementalists do not step on individual toes because no long-range proposals are made. But they do solve current pressing problems so they are relied upon by public officials.

We need a better approach, a blending of the policies, blueprint, and incremental approaches. The Joint Program plan, to avoid confusion with standard master plans and to emphasize its purpose, is titled the Metropolitan Development Guide. Its focus is on major metropolitan development—large centers of commerce, industry, and government; large open spaces; and the systems of transportation and utilities that shape and serve those developments. The guide envisions making the major decisions at the metropolitan or state level while leaving the remaining decisions to the local level. The guide contains maps but does not show how each parcel of land should be used.

Goals—Policies—Programs

The Metropolitan Development Guide contains three main elements: goals, policies, and programs. Goals are seen as the ends toward which we strive. Policies are the settled courses of action toward the goals or the decision rules that will be applied in moving toward the goals. Programs are the allocation of resources by type, time, amount, and location in line with established policies to achieve the goals.

The goals-policies-programs approach arrives at decisions by going from the general to the specific and getting agreement at each stage. When we agree to goals or ends, we are dealing with statements in which the values of the individuals of the community generally will be consistent and agreement will be fairly easy to achieve.

When we go to the next step, policies, we find differing values. Differences arise in political philosophy and the extent to which decisions should be made in the interest of total society rather than the individual. Differences between the values of producer and consumer are revealed. These must be reckoned with, argued out, and resolved at this point.

When we get to programs, we are for the first time talking of specific pieces of land and specific dollars of investment. We have, by this time, achieved substantial agreement on the objectives of investments and the rules for making investments. We now have a firm enough base of agreement to take this last difficult step, the step at which the blueprint approach to planning has failed in the past. In the blueprint approach, there was no opportunity to discuss overall goals or objectives or the rules by which those who make decisions should be bound. Each individual could look at a map and see exactly how his individual interests were going to be affected. He reacted to those individual interests first and to overall considerations second.

The Joint Program approach may sound like the planning-programming-budgeting system (PPBS) advocated by the federal government. The purposes are identical—to insure the most effective use of scarce resources in meeting stated goals or objectives. The methods have similarities except for one important element. The PPBS technique starts from stated goals or objectives. We had to go back one step and formulate goals or objectives—a difficult task.

Goals and policies will be revised at least once every five years. Programs will be revised annually. To have some measure of progress toward goals, one- and five-year objectives will be set each year and annual programs designed to achieve these objectives.

Our use of goals is different from some other studies around the country. Some say that goals conflict. We do not hold this view. We believe that goals are sufficiently general by nature that they should not conflict, but that conflict arises when one begins to allocate resources to achieve the goals. In other words, people agree that they want ease of movement throughout the metropolitan area. But the disagreement occurs when they allocate dollars for highways as opposed to transit or for transportation as opposed to parks or schools. This is not a conflict in goals, it is a conflict in how much weight a given goal should receive or how each goal should be pursued. It is this conflict in weighing the goals that must be settled by a community, not a professional, decision.

Some believe goals should be used as tools for a community debate. Others believe they should be prepared by the professionals to guide their own later actions. We believe that goals can only be adequately understood and integrated when they are extended in terms of policies and programs. When the pursuit of a goal is expressed in terms of dollars from the pocket or property rights or some other item close to the individual, he can adequately assess how strongly he feels about the goal. Therefore, the final goals of the Joint Program and of metropolitan planning in the Twin Cities area will not be established until we have gone all the way through the process to adopted programs.

The Joint Program's use of alternatives was quite different from the standard use of alternatives by professional engineers to arrive at a "best" solution. The engineer lays out several precise alternative schemes and then bases his evaluation of them on specific accepted criteria, such as cost-benefit analyses. The work is all done within his offices by the engineer.

Our purpose was to discover what individuals in the community value. We did this by obtaining responses to (a) the total pattern of development and (b) specific development policies. Four schemes were prepared that seemed to bracket the range of choice. In each of these we asked the community to say which "direction" it preferred rather than which specific pattern. For example, would they prefer to move strongly or slowly toward dispersion, as shown in the Spread City scheme, depending on the acceptability of the related development controls, tax policies, and transportation policies? The Radial Corridors scheme, building large concentrations in the downtowns, could have been even stronger through tighter controls of the use and development of land. Therefore, in our meetings with the public, we asked not only which scheme they preferred but whether the scheme went far enough, too far, or not far enough in the general direction it represented.

Specific development policies were constructed to bracket the possibilities. In the four schemes, if we were to look at the size of centers of retailing and office employment, we find the smallest ones in Spread City, next larger in Present Trends, next larger in Radial Corridors, and the largest suburban or outlying centers in the Multiple Centers scheme. We noted that as one moved toward a large number of small centers, convenience increased and the choice decreased at any given center. Conversely, as one moved toward a limited number of large centers, choice increased and convenience decreased. This was described as a "value couplet" in which each individual had to balance choice and convenience.

Obviously, no one was willing to choose either extreme convenience with no choice or a single center that would be very inconvenient but in which total choice would be available to the region. In fact, we found that neither the Spread City nor the Multiple Centers commercial pattern was acceptable. But we did find that choice was more important than convenience, suggesting that the centers should be relatively large. The upper limit on size was set by two factors: development controls and taxes. If the development controls had to be stringent to obtain large-size centers, they would be unacceptable. Also, if taxes are not redistributed to provide some benefits to the communities that otherwise might have had commercial development, the larger centers would not be acceptable. If these two points are taken into account, however, the public prefers choice to convenience and large centers to small ones.

The important point is that the planners and engineers must discover what the community values and not substitute their own personal preferences. To make a choice, individuals must be informed of the consequences of their choices. This was our prime purpose in development of alternatives. We also received a liberal education in understanding some of the forces at work in the region.

Guiding or Managing Change

Which comes first, the land-use "chicken" or the transportation "egg?" The cliche is apt because, like chickens and eggs, the ways we use parcels of land and the transportation systems we build to connect them are each products of the other (1). And we can't plan adequately for one without considering the other.

In its simplest form, this relationship is shown in the cycle diagram of Figure 1 (we could use the same basic diagram for other capital expenditures such as sewers). Being a continuous cycle, we can enter it at any point, but let us start at 1, land use. Whether the land is used for shopping, manufacturing, residences, or parks, the activities on the site generate trips (2). These trips are depicted on planning maps by straight lines called "desire lines" that connect point of origin and point of destination. Desire lines are the basis for identifying highway needs (3). Construction of a highway or other transportation facility (4) to meet these needs creates accessibility (5). No site in any area is going to develop if people can't get to it, so through the provision of access you help create land value (6).

Land value, in turn, completes the cycle by helping to determine land use (1). For example, it is an exceptional person who can afford to build his home on the highest value land in the city. Nor is this the likely spot for a marginal operation like a junkyard that cannot afford a big capital investment. It is usually the site of the city's largest department store or a wealthy prestige office building. The name that planners and land economists use for it is the "100 percent corner," the theoretical point of greatest activity. Thus, more trips are generated, the desire lines are drawn heavier, more highways are built to provide more accessibility, and so on, perhaps until the cycle spirals out of balance.

If we accept this as an abstraction of how urban areas work, we can also see from a diagram how we can apply controls to make certain that the cycle does not spiral out of balance and to make certain that it does meet the needs of the region. To keep the cycle in balance, we have traditionally used controls to manage the changes. Returning to the cycle diagram, we see that government regulates land use

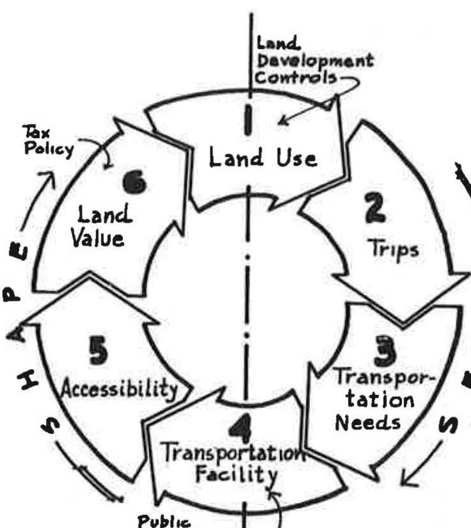


Figure 1. Land use-transportion cycle (Source: Ref. 2).

(1) by various means such as zoning and subdivision regulations. Highways are built with public funds. Thus, they are part of the public capital investment (4), which can be altered as needed. Land value (6), although affected by accessibility, is also affected by tax policies, which can be altered as needed.

So policies affecting the land, expenditure of public funds, and the handling of taxes offer three ways of managing development or change. This gets us back to the planning process and shows us why it is important to agree on goals or objectives, that is, where we want to go, and then to agree on policies concerning how we wish to get there.

Shape or Serve?

There is one other aspect of the diagram that we must cover, and that is the extent to which we will guide development, i.e., the extent to which we will use the controls noted in the foregoing. Traditionally, highways have been built to respond to forecast needs derived from land use of the future. That is, the highway engineer has traditionally used only the right-hand side of the diagram. But if the construction of transportation facilities does affect land use, which is what the left side of the diagram says, can we afford to ignore this fact?

The Joint Program adopted the position that, because of the rate of growth and change in the Twin Cities metropolitan region, we had to use all possible methods of control compatible with the communities' values. The Coordinating Committee of the Joint Program recommended that public capital investments be used to shape land use as well as to serve it. Community leaders accepted this after considerable discussion.

It should be pointed out that there are many planners and engineers in the country today who contend that transportation facilities should not be used for such purposes, while others indicate flaws in the old approach. It is our view that this is not a professional decision to make, but rather a decision for the community, because it involves the balancing of individual values against the values of the group and the benefits to be gained by the one at the cost or expense of the other. It is the role of the professional to show what can be done and to leave it to the community to decide which approach it favors.

We feel fortunate that our community has decided to use public capital investments to shape as well as to serve development. If we are to shape development, it is then mandatory that we agree on objectives for the development of the region before we make any of the public capital investments. It is no longer possible to merely build highways, transit, or parks on an incremental basis. They must be designed to meet goals.

IMPROVING METROPOLITAN DECISION-MAKING

During the last two years, we have been working on the Metropolitan Development Guide and its implementation at the same time. Contrary to the information published in behalf of the growing array of metropolitan councils of governments, the critical metropolitan decisions cannot be made by a voluntary association of local governments. Most require action of the state legislature. Earlier it was noted that our concern was with the major metropolitan systems. What is metropolitan and what is local?

Metropolitan vs Local Interests

The two are generally thought of as being in conflict. And they are if we define "metropolitan" as the total community and "local" as the individual community as in the example of freeways. From the metropolitan viewpoint, freeways must be designed to serve metropolitan high-speed, long-trip movement and major concentrations of activity. But the individual community sees access to the freeway as an enticement to tax-producing development. If extensive local access is provided, the metropolitan purpose is thwarted.

From the vantage point of the individual resident who benefits from both ease of travel on high-speed metropolitan facilities and from easy access to the freeway, it is not metropolitan vs local. It is not a question of either-or. It is a question of how much of each. As this fact is better understood by the average citizen, better development decisions are being made. This leads to the Joint Program view that metropolitan facilities are those that local governments cannot provide but which metropolitan citizens desire and need.

Who Makes the Decisions?

All too frequently, the planners and engineers working for agencies that design highways, transit systems, parks, and other public improvements think of their task as a technical one. But it is not. When a specific speed is assigned to a highway or transit link, public policy is involved. A decision to provide for higher speeds means more opportunities to the citizen within a given amount of time, but at a higher cost in tax dollars. Thus, the highway planners, the citizen, and the elected official all have an interest in such a decision.

It is easy to say that there is a relationship among the planner, the elected official, and the citizen. It is difficult to formally organize the relationship. The planners and engineers find it difficult to program breaks in their work when elected officials are reviewing and reacting to proposals. There are no metropolitan elected officials, and only recently have we had metropolitan organizations of local officials. There are few metropolitan organizations of citizens, and those that do exist are recent additions to the urban scene.

The Twin Cities area's Joint Program made a start by putting together professionals from a variety of agencies on its Coordinating Committee and Technical Advisory Committee. The chief elected officials of each unit of government in the metropolitan area (about 300) were brought together in an Elected Officials Review Committee. Business, labor, and other community interests were brought together in a Citizens Advisory Committee. Metropolitan area legislators were briefed as a group. While much was done, much more must be done in the future. The area needs a stronger metropolitan council to make major public development decisions. It needs effective citizen participation in metropolitan affairs. And last, but certainly not least, it needs more professionals who see their role as advising on and carrying out development decisions—not making them.

The 1967 Legislature: Activities and Results

By the summer of 1966, the meetings with the local groups and legislators to obtain a sense of direction were over. One point seemed clear. The concept of guiding metropolitan development by means of coordinated control of selected major elements was accepted. We began simultaneously drafting the plan policies and position papers on legislation.

These position papers included three on government—local consent, the Minnesota Municipal Commission, and metropolitan government. We obviously cannot claim to be solely responsible for what passed the legislature, but we do feel our assistance had some effect.

The local consent provision in the state constitution had to be removed (which the legislature could effectively do) in order that any metropolitan bill could be passed as a special law without needing the consent of up to 200 local governmental units. It passed.

The Minnesota Municipal Commission rules on all political boundary changes in the state. A major feature of the bill would have allowed the MMC to initiate action rather than respond to petitions of local governments. The bill failed. The staff of the Metropolitan Planning Commission, now the Metropolitan Council, serves in an advisory capacity to the MMC.

The position on metropolitan government sought the creation of an organization with regulatory or operating powers in highways; transit; sanitary sewage; open space; airport land; land, water, and air development control; and comprehensive planning and programming. This would have effectively controlled the major element specified in the Metropolitan Development Guide.

The staff also prepared a "back-off" position in which control was achieved not through operation but through veto of plans and programs. The Metropolitan Planning Commission, as a matter of political strategy, decided to publish the strong position only together with a resolution that the MPC be abolished if a metropolitan council were created. The strong bill failed by four votes in the House and one vote in the Senate. The bill that passed was similar to the "back-off" position.

In addition to the position papers, statements were prepared on sewage collection and treatment, metropolitan parks, mass transit, and the highway local consent problem mentioned earlier. No material was prepared on the control and development of major centers because adequate research had not been completed. Of these, the only one to pass was transit. The transit bill was drafted to precisely carry out the policies identified in the draft of the Metropolitan Development Guide. Although the other bills did not pass, the Metropolitan Council achieved a degree of control over each one except highways and over major centers through its own legislation.

Because legislators in Minnesota do not have paid staffs, the MPC offered its staff services on metropolitan legislation. The work involved drafting anonymous single-copy detailed critiques of various bills on short notice and in some instances drafting selected portions of key bills. We feel this is a necessary and proper part of planning and plan implementation.

SUMMARY

If we are to succeed in meeting the ever-increasing problems of metropolitan growth and rebuilding, we, as professionals, must adapt our techniques for and must participate in the decision-making process. Where adequate capability does not exist to make the decisions needed, we must concentrate as heavily on creating that capacity as on the content of the needed development actions. In all this we must remember that in the realm of public policy, professionals advise and elected representatives decide.

REFERENCES

1. Metropolitan Balancing Act. The Joint Program, Program Notes No. 1, March 1964.
2. Transportation and the City. Architectural Forum, Oct. 1963, p. 89.