GENERALIZED COSTS AND THE ESTIMATION OF MOVEMENT COSTS AND BENEFITS IN TRANSPORT PLANNING

P. T. McIntosh, Strategic Planning Directorate, Department of the Environment, London; and
D. A. Quarmby, London Transport Executive

The object of the paper is to provide guidance to transport planners and analysts by describing procedures in two areas: (a) the evaluation of movement costs and benefits consequent to changes in networks and management policies and (b) the estimation of the generalized behavioral and resource cost functions for links and origin-destination pairs that are necessary for this evaluation process and for forecasts of behavior. The procedures are designed for use in situations where the change in network or policy is thought to have strong effects on the trip pattern and individual link loadings. This will generally be the case in the consideration of urban schemes and may be the case for major interurban schemes; in both situations there may be considerable changes in the trip matrices, modal split, and routes used. The emphasis is on operational methods. The precise way in which the benefit expression and generalized costs are calculated will depend on the level of detail and form of particular studies; considerable guidance is given to aid the transfer from concepts to computation.

•TRAFFIC PREDICTION was, for many years, carried out quite independently of the procedures used for assessing the economic value of the possible changes under consideration. In the late 1950s and 1960s, the principal outputs of traffic models were flows of people and vehicles along links of networks in urban areas, and investments were very largely decided on by considerations of physical and technical feasibility (operational evaluation). At the same time, techniques were evolved for estimating the movement costs and benefits arising from the improvement of particular roads, mainly in rural areas; in such situations, the facilities of traffic models—the ability to represent the response of traffic movements over a wide area to changes in the road network—were considered unnecessary in the evaluation procedures.

A growing desire within the responsible authorities not only to obtain value for money in transport investment but also to make comparisons between feasible options has led in recent years to a strong need to integrate the methodology of traffic models with that of economic assessment procedures. The London Transportation Study was probably the first to attempt to do this in network comparisons (1, 2, 3), and a procedure for isolated road schemes was described by the Road Research Laboratory in 1968 (4). The problems were, and largely still are, substantial: To start with, the languages of traffic prediction and of highway investment appraisal were fundamentally different; traffic prediction methods evolved empirically, as a collection of heterogeneously based submodels, with no explicit economic inputs and no economic basis at all; as applied, highway investment methods were aimed at the consideration of individual links or small schemes and, based almost entirely on a "travel time and cost-saving" approach, were unable to handle, other than very simply, consumers' surplus aspects arising from changes in traffic behavior.

Sponsored by Task Force on Application of Economic Analysis to Transportation Problems.

The establishment of a common basis and framework for both traffic prediction and economic evaluation has developed along a broad front within the Department of the Environment (formerly the British Ministry of Transport) and elsewhere as a result of work by both mathematical analysts and economists. The associated methods are now beginning to be used as a matter of routine in planning projects carried out within the Department. The intention of this paper is to disseminate both the thinking and the methodology among local and regional government authorities, public transport authorities, consultants, and others engaged in transport planning so as to establish consistency between studies and to facilitate better predictions and an improved allocation of investment resources.

The layout of the paper reflects the two main themes of this integrated approach to prediction and evaluation. The first part deals with the introduction of an explicity economic content into characterization of space and time in traffic prediction, by means of the concept of "generalized cost"; the nature of the cost function is related first to the factors that influence travel behavior, and then to the consumption of real resources that come about through changes in this behavior. Then there is a section concerned with broadening the cost-saving approach to economic evaluation so as to include the measurement of changes in consumers' surplus that arise through changes in travel behavior as predicted by the traffic models.

THE CONCEPT OF GENERALIZED COST

For some time transportation analysts have been aware that travel time alone is not a satisfactory way of representing the separation between zones as used in transportation studies, particularly for modeling people's travel behavior. For one-mode forecasting (e.g., road) time may be an adequate measure, although the inclusion of high-speed roads in test networks can highlight the problem, inasmuch as the extra mileage and operating costs in a given travel time at a high speed remain hidden. To some extent this effect has been concealed in the general "noise" implied in the present levels of accuracy attained by traffic models.

However, modeling more than one travel mode with an integrated model (5) (as opposed to ad hoc techniques such as distributing trips with road "skim trees" and then using diversion curves) exposes the limitations of time alone rather more obviously, principally because time-cost profiles can differ significantly between travel modes. Furthermore, time-based models are insensitive to changes in the pricing of public transport and of parking facilities, and they can say nothing about road pricing. Operationally, the analyst can express, for example, parking costs as so much extra time on a link, but this is only an ad hoc method of dealing with a problem that it is more satisfactory to tackle in a basic, more general way.

Because the use of integrated distribution-modal split models is increasing, the need is clear to establish a framework for the definition of a more generalized measure of the costs of travel as represented in model-based studies to replace time in the specification of networks for use in distribution, modal split, and assignment procedures.

In addition, the definitions of cost developed to represent how people's travel behavior depends on the characteristics of networks are also relevant in putting an economic interpretation on their behavior. The sections on the estimation procedure show how this "behavioral cost," used to explain travel demand, is also used to attribute "value" received by travelers when networks are improved or policies changed.

At the same time, the economic evaluation procedure is concerned with estimating the real resources consumed in travel and transport. A further concept, parallel to behavioral cost, is needed—a "resource cost." This is a unit cost that describes the value of resources consumed in a unit of travel.

Why should these two costs be different? A behavioral cost is that cost function that best explains people's travel behavior (and therefore enables their behavior to be predicted). A resource cost is that cost function that represents a consumption of resources. Thus the following two areas can give rise to difference:

1. People may base their behavior on imperfect perceptions of cost. For example, there is considerable evidence to suggest that people significantly underestimate the

costs of running cars (6, 7, 8). The mileage cost that best explains and models their travel behavior is less than a strict engineering assessment of the marginal mileage cost.

2. The prices that people face (and therefore that determine their behavior) may not reflect the true resource costs. For instance, taxation on fuel is a component of the price but does not represent resources; also, the fares charged on a public transit system may not reflect the actual costs of operating the different parts of the system, and so on.

Behavioral costs for prediction will normally be measured in equivalent time units. In the evaluation procedure where value received is calculated, a special form of the behavioral costs is needed. This differs in two ways from the prediction value: (a) it is calculated in common monetary units (simply a change of scale), and (b) it may contain assumptions on the values to be attributed by the community to certain items that are different from those placed on those items by the individuals themselves (e.g., a commor value of nonwork time rather than a behaviorally revealed value differing according to income group). In this way the effect of weighting benefits in various ways relatively between groups of recipients can be examined without affecting the forecasting procedures.

Thus for the prediction of travel demand and the appraisal of transport investments and management policies there are the following three kinds of generalized costs:

1. b, behavioral cost for use in prediction models; form of the function is based on the best knowledge about what characteristics of networks influence people's and firms' travel and transport decisions. It takes into account time and costs and is usually in time-equivalent units.

2. u, behavioral costs for use in the benefit estimation procedure; in current practice the form of the function is identical to b except that it may include alternative values of nonwork time as reflections of possible social values and that it is in monetary units.

3. r, resource cost for use in the benefit estimation procedure as society's valuation of the resources consumed by a unit of travel; the form of the function is based on known technical relationships between costs and various transport-related activities. Some items in the function are based on behavioral cost items; they will use the values from u, not b, where they differ.

Theoretical Aspects of the Derivation of Behavioral Cost Functions

Behavioral cost is an expression describing the totality of "cost" or disutility incurred by a traveler in making a zone-to-zone trip by a particular mode of travel; it may well not be the total cost or disutility that the traveler actually incurs, because he may have an imperfect perception of cost. In practice it is simply the cost that best explains his travel behavior within the framework of the model processes in use. And, inasmuch as traffic models are trying to represent people's travel behavior, this is the right sort of cost to use for prediction.

Operationally, all the factors contributing to travel disutility are not known, nor could they all be included in present modeling procedures. In addition, each traveler will implicitly behave according to a unique set of factors (e.g., traveling time, waiting time, fares, interchanges, and comfort) and a unique relative weighting of them. The models used are concerned with the behavior of travelers in aggregate; thus, a decision must be made on the level of generality of the behavioral cost function to be used. Ideally, both the form of the function and the values of its parameters should be relevant to the particular study and possibly determined from data gathered in the area where the transport model is being applied. The main arguments against this are first that most transportation studies cannot sensibly mount the program of survey and analytical work that would be necessary; and second that, with the uncertainties that must invariably attach to estimates of values from individual research studies of this kind, more robust parameter values and function forms can be arrived at by joint consideration of several research studies, providing that the results are expressed in a sufficiently generalized form.

The object therefore is to define a cost function and to suggest parameter values for use in studies. The form of cost function suggested here is

$$b_l = B_1 x_1 + B_2 x_2 + \dots + B_n x_n$$

where b_l is the behavioral cost of travel along a link l of a network by a particular travel mode. $x_1, x_2, \dots, x_n$ are values of factors that are important in determining the overall travel disutility as it affects behavior. $B_1, B_2, \dots, B_n$ are the relative weights of these factors. More complex functions could be suggested, but at this time there is no good reason for not using a simple linear function, especially as it has additive properties that are plausible and simplify calculation.

Normal network manipulation and tree-building programs can then be used to find cheapest routes (using behavioral cost), instead of fastest routes. Trees are skimmed to produce an interzonal behavioral cost matrix, b_{ij}, instead of an interzonal travel time matrix. (If more sophisticated multiple-route finding procedures are used, similar arguments apply.) The behavioral cost matrix is then available for use in distribution, modal split assignment, and evaluation procedures.

In deciding what factors should be taken into account, one can start by including time and add other factors that seem reasonable. Alternatively, one can draw on research in related fields to discover what factors people seem, by an analysis of their behavior, to take into account in their travel. In the last few years, information has come to light from studies of people's choice of travel mode in the journey to work on the factors influencing that choice, and their relative weights (7, 8, 9, 10, 11). In that these factors seem to be taken account of when people compare one mode of travel with another, it is sensible to impute that this is how people see each mode individually.

As a result of this and other empirical work, it is recommended that the factors that should be included in a behavioral cost function are the in-vehicle travel time, the components of excess or outside-vehicle travel time (suggested split is walking time and waiting and transfer time), and the financial cost of travel (including terminal cost, if any). The behavioral cost function for a network link is thus

$$\begin{aligned} b_l = {} & B_1 \times \text{in-vehicle time} + B_2 \times \text{walking time} \\ & + B_3 \times \text{waiting and transfer time} + B_4 \times \text{travel cost} \end{aligned}$$

For any given network link, only some of these variables will be nonzero. For instance, a highway link in a private car network will use only the in-vehicle time and the travel cost (as mileage × cost per mile related to the speed). A terminal link may have a walking time and a parking charge. On a public transit network, an access link may contain a walking time and a waiting time, a route link will contain an in-vehicle time and a fare, and so on. The values of the various coefficients may well vary according to the trip purpose and income group under consideration. This particularly applies to the time values.

It can be seen that B_1/B_4 is the value in travel cost units of in-vehicle traveling time, B_2/B_4 the value of walking time, and B_3/B_4 the value of waiting and transfer time. For any particular study, therefore, the task is to estimate the relative values of the parameters, B, and to decide what units to express them in. It might seem obvious to express b_l in monetary units, but it will be seen in a subsequent section that it may be more appropriate to express the behavioral costs in time-equivalent units when forecasting behavior.

The behavioral costs for evaluation, u_l, will use the same function except, where appropriate, it will incorporate any alternative values under consideration—for example, where the effect of alternative values of nonwork time are being examined. They will be expressed in monetary units.

Resource Costs

A resource cost function is similar in form to a behavioral cost function; it contains personal time, valued at the appropriate rate, and engineering assessments of vehicle or system marginal operating costs, less transfer and other nonresource payments such as taxation. For highway links, the engineering assessments of motor vehicle operating cost can be related fairly easily to the link and thus become part of link unit resource cost.

However, where the prediction units are person trips (rather than vehicle trips) and the vehicle occupancy is not necessarily constant (for instance, with bus operation), it may be unsatisfactory to express the resource costs as so much per person traveling. For fixed track systems, where the marginal and average costs diverge substantially, a unit rate per person trip is nearly always unsatisfactory or impossible to determine. In these cases, therefore, the unit link resource cost will only include personal time, and the change in system operating costs will be estimated separately on the basis of total network assignments or other system characteristics. The resource cost function will then be

$r_l = R_1 \times$ in-vehicle time + $R_2 \times$ walking time + $R_3 \times$ waiting and transfer time + $R_4 \times$ unit resource costs of vehicle operation

Cost Functions in Network Manipulation

The functions are specified as link costs. In this form they can be used to derive zone-to-zone (i-j) values for use in predictive models and in the evaluation procedure. The process is as follows:

1. Estimate link behavioral costs, b_l;
2. Build "cheapest" trees by normal network manipulation and tree-building procedures;
3. Skim these trees to provide zone-to-zone behavioral costs, b_{ij};
4. Using link u-costs, u_l, and link resource costs, r_l, add up each along the behavioral cost paths to obtain zone-to-zone u_{ij} and r_{ij}; and
5. Then the b_{ij} are available for the prediction model, and u_{ij} and r_{ij} are ready for the evaluation procedure.

There is an alternative approach: Some network computer programs only carry link stores for time, t_l, distance, d_l, and possibly speed, with no easy means of subdivision by classes h of time or special treatment for parking, fares, and the like. In this situation it may be useful to convert all the various time elements, t_{lh}, to equivalent "in-mode" time (scale by B_h/B_1) and put these in the time field; similarly cost or distance elements can be treated together and put, to some scale, in the distance record. Then b_l can be formed as needed in the program as a function of t_l and d_l, and b_{ij} can be found as before by building b-trees; t_{ij} and d_{ij} can be found by adding up times and distances along b-cost path, and then, in some circumstances, u_{ij} and r_{ij} may be formed from an appropriate linear function of t_{ij} and d_{ij}.

Difficulties will arise from the treatment of speed in r-costs and with components of the t_l and d_l that have different weights within b, u, and r; e.g., parking charges should appear in full in b but only the resource component should appear in r.

ESTIMATING BEHAVIORAL AND RESOURCE COST FUNCTIONS

Parameters Needed and Classification of Trips

In any particular transportation study, the values of the parameters B, U, and R in the following functions must be estimated:

$b_l = B_1 \times$ in-vehicle time + $B_2 \times$ walking time + $B_3 \times$ waiting and transfer time + $B_4 \times$ travel cost (including terminal and toll charges)

$u_l = U_1 \times$ in-vehicle time + $U_2 \times$ walking time + $U_3 \times$ waiting and transfer time + $U_4 \times$ travel cost (including terminal and toll charges)

$r_l = R_1 \times$ in-vehicle time + $R_2 \times$ walking time + $R_3 \times$ waiting and transfer time + $R_4 \times$ unit operating cost

The values of the parameters in fact depend on a large number of quite specific characteristics of the transportation study and of the models used, such as the units of prediction (e.g., person trips or vehicles?), the base year, the prediction years, where modal split comes into the modeling process, and degree of disaggregation by purpose, person type, and income, and it is therefore not possible to formulate universal sets of standard values. Values of time, particularly of nonwork time, are notoriously difficult to estimate. However, they seem to have some generalized characteristics that help the estimation of local values for particular studies. A report of a British conference on the value of time (16) gives a reasonably up-to-date review of recent empirical and theoretical studies. As a general guide:

> In large urban studies it is usual to build models at one or more future years for person trips and to convert to private car or bus or train movements by applying occupancy factors after the modal split procedure. In some studies peak hour flows only are modelled, and in these the journey to work predominates. Commercial vehicles are ignored or treated separately outside the main modelling process. With this specification, the units would be person-trips, and the purpose-mix (which leads to value of time and occupancy) would reflect the peak period composition. . . . In inter-urban highway studies, it has been conventional, because of lack of data, to model vehicle trips, and to build in assumptions about occupancy and modal choice at the beginning, often implicitly. Often all vehicles (commercial, private cars and buses) are modelled together, so the unit cost would reflect an average purpose mix for the private cars and buses, and an average vehicle mix for the total traffic flows.

Further subclassification of the units of prediction may well be made in the more sophisticated studies, and it is a matter of choice for the analyst whether different cost functions should be developed for each subclassification. For instance, large urban studies may classify journey purposes into home-based work (HBW), home-based other (HBO), and non-home-based (NHB). The population may be stratified into different groups for the purpose of predicting travel behavior. One such classification is into car-owning and non-car-owning households. Another may be by income groups. In both these cases, different values of time will apply because of the different mean incomes. This, in turn, will alter the relative weighting of time and cost in the unit cost functions.

Ideally, therefore, the analyst should use a different cost function for each income group and purpose, which means different costs for any one link in a network and thus different networks. The extent to which this should be done will depend on the particular study and on the analyst's judgment. In many cases such stratification may not be necessary, but there may be circumstances where it is important to represent the fact that networks can really look different to people of different incomes (for instance, long-distance rail commuting to central London where financial costs of travel are relatively high compared with the travel time). In particular, it may often be that different networks should be built at least for car owners and non-car owners, using values of time based on incomes in each category.

Base Year and Rates of Growth

The values of all economic inputs, such as costs and time values, depend on the years in which modeling is attempted—the survey year for calibration and one or more forecasting years—and on the choice of a base year for prices. It is necessary to carry out all economic comparisons at some constant price level. It is conventional to assume that vehicle operating costs (both behavioral and resource) will remain constant at constant prices and that average values of time will rise at some assumed rate of growth of real incomes.

Where some disaggregation into categories of different incomes is adopted, separate rates of income growth should be estimated for each category. One curious effect of this is that the mean incomes of car-owning households and non-car-owning househoulds may both rise at less than the average rate. This comes about because of the acquisition of cars by non-car owners whose incomes tend to be high relative to other non-car owners and low compared with car owners; they "dilute" the car owners as car ownership increases.

Choice of Scales for the Unit Costs

At first sight it would seem obvious to scale the parameters so that the behavioral costs b are in money units. However, projecting values to a future date for forecasting travel patterns exposes a problem of consistency and comparability. As incomes rise, a given cost will carry less weight; it may be better to scale the parameters so that the units of behavioral cost retain some absolute value over time. It can be argued that time has much the same value in terms of personal utility to people of different incomes and to people living now and at some future date. There are, of course, arguments against this proposition, but at least it is probably more tenable than scaling on cost.

At the present time, therefore, it is recommended that, in forecasting procedures, the behavioral cost functions b be scaled on time, so the units are "equivalent minutes." This means that, for groups of higher income, the value of the cost component in the function will fall; i. e., a particular financial cost means less to those with higher incomes. For instance, if the real income (relative to prices in general) of some category is expected to rise by 50 percent (i. e., 1.5 times), then the coefficients of cost in the b functions fall by $\frac{1}{3}$ (i. e., $\frac{1}{1.5}$ times).

Effectively, as people become relatively better off, time assumes a greater proportion of the behavioral cost b of a trip. (For a typical trip to work by car in U. K. urban areas in 1968, time accounted for just about half the behavioral cost.)

In the evaluation procedure, however, resource costs, r, and modified behavioral costs used, u, should be in common monetary units.

Summary of Behavioral and Resource Costs

As a general rule, therefore, for each group considered, the following holds true:

1. In b-costs: B_1 will be unity, B_2 will be some factor such as 2, B_3 will be some factor such as 2, and B_4 will be one divided by the (averaged) value of traveling time, at the appropriate unit base (persons or vehicles).
2. In u-costs: U_1 will be the appropriate value of traveling time in monetary units, U_2 will be some factor such as $2U_1$, U_3 will be some factor such as $2U_1$, and U_4 will be unity.
3. In r-costs: R_1 will be the same as U_1, plus time-dependent elements of vehicle operating cost, R_2 will be the same as U_2, R_3 will be the same as U_3, and R_4 will be unity.

The factors for B_2, B_3, U_2, and U_3 are current estimates based on behavioral studies. If other locally derived or more reliable factors are available, they could be substituted.

ESTIMATION PROCEDURE FOR MOVEMENT COSTS AND BENEFITS

In transportation planning, economic evaluation involves the estimation of the costs and benefits that accrue as a result of investment in networks or changes in their management; its purpose is to make objective statements relating to the relative and possibly the absolute worth of alternatives. The procedures described here are designed for use in the situations where the change in network or policy is thought to have strong effects on the trip pattern and individual link loadings. This will generally be the case in the consideration of urban schemes and may be the case for major interurban schemes as well; in both situations, there may be considerable changes in the trip matrices, modal split, and routes used.

Individual networks or management policies cannot be sensibly considered in isolation. Comparisons between alternatives are essential for the valid consideration of the individual proposals. One particularly important comparison is between the various possible future systems under consideration and, as a base, the "do-nothing" situation; in this context "do nothing" means including only those changes to the existing situation that are, for all practical purposes, now unavoidable between the present day and the period under consideration.

The principal items giving rise to costs and benefits are

1. Capital and initial costs: (a) construction costs, including interchange facilities, parking provision, etc.; (b) land and property costs; and (c) delays and inconvenience during construction; and
2. Recurring costs and benefits: (a) transport user benefits and costs; (b) operator revenues and costs; (c) user transfer payments (e.g., taxation); (d) change in accidents; and (e) external economies and diseconomies (development consequences, environment, etc.).

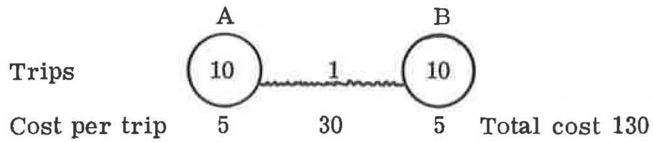
This section concentrates on the measurement of those costs and benefits that arise from changes in the volumes and pattern of movement as a result of changes in networks or management policies. In particular, it deals with the estimation of transport user costs and benefits (item 2a) and their joint treatment with operators's costs and benefits (item 2b) and taxation (item 2c). Neither the estimation and treatment of the other important costs and benefits nor the problem of intergrating all these into a decision-making framework is considered here.

Why Not Simply Compare User Costs?

Many studies have calculated the relative benefits of alternative plans by simply calculating the change in user costs, on the basis of either a fixed or changed trip-making pattern. The following trivial example shows some dangers of this procedure.

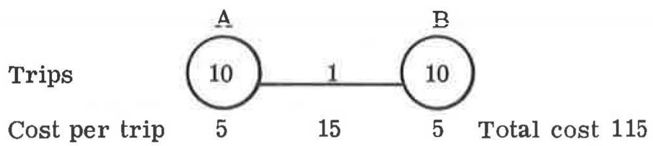
Two towns, A and B, each have 10 internal or intrazonal trips taking place at 5 cost units each. The towns are connected by a poor road. It would cost a traveler 30 cost units to journey between them and, as a consequence, only one does in the time period considered.

Situation 1



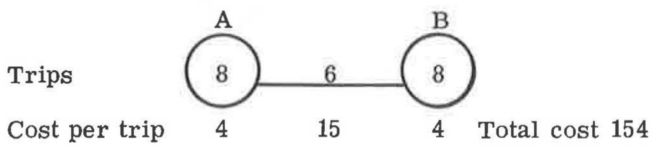
The connecting road is improved so that the cost is 15 units. Assume that the trip pattern is unchanged.

Situation 2



Let the trip pattern and intrazonal cost now change in a sensible way.

Situation 3



The total user cost (situation 2) with a fixed trip pattern is reduced below cost (situation 1) by 15 units. The cost (situation 2) with a sensibly changed trip pattern (as would happen in real life and in a transportation model sensitive to costs) is 24 units more expensive. A minimum cost comparison would then suggest that the improvement to the link was worth 15 units (situation 2) or else was worse than useless (situation 3).

Both solutions cannot be correct. The first solution shows a benefit but takes no account of the resulting change in trip-making; the second allows for the change in trip-making but indicates that the travelers as a whole are worse off.

The procedures described in the following overcome this dilemma by taking into account the benefits stemming from improved choice and the change in trip-making.

Consumer Surplus Approach to the Measurement of User Benefit

In general economic theory, this problem is well known and understood (12, 13). Only in recent years has the methodology been developed to evaluate the benefits in comprehensive comparisons of alternative networks (1, 2, 14, 15).

The measure of user benefits is essentially a measure of change in consumers' surplus. To estimate the benefit of a single plan and hence, by difference, the change in benefit between two plans, we should consider a demand curve for travel. This curve relates the demand q between a particular origin and destination to the behavioral cost of travel between them, b, defined here in monetary units; it is shown in Figure 1.

The definition of the curve indicates that for any particular trip at , say, q' the traveler would have been willing to pay b', thus making a profit or surplus of $b' - b$. When all these surpluses are added, they form the rather ill-defined upper shaded area in Figure 1, which is the consumers' surplus.

Some part of the cost to the users as indicated by their behavior b does not reflect consumption of resources (e.g., taxation or parking charges that may not equal the costs of provision and operation) but is additional surplus transferred either to the community as a whole (through taxation) or to operators (e.g., a parking authority). In Figure 1, then, the nonresource n element of the surplus is the lower shaded area, and the total surplus is the complete shaded area.

Figure 2 shows the benefit arising when two plans are compared. Suppose initially that the cost of travel is b_1, at which cost the demand for travel is q_1, but that as the result of a transport improvement or a different plan the cost of travel falls to b_2 and the demand for travel increases to q_2. The transport users obtain an increased consumers' surplus, illustrated by the upper shaded area in Figure 2, which is approximately $\frac{1}{2}(q_1 + q_2)(b_1 - b_2)$. In addition, there is the nonresource correction ($n_2 q_2$ -

Figure 1. The estimation of movement benefits: demand curve and surplus.

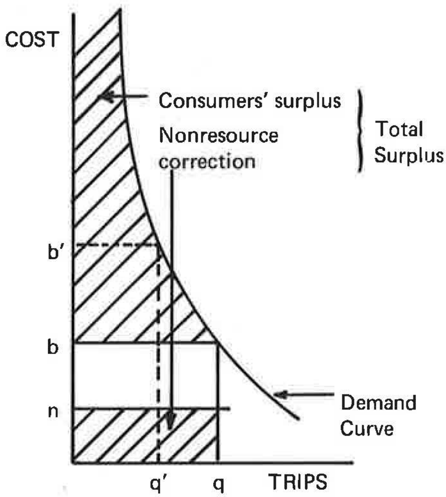
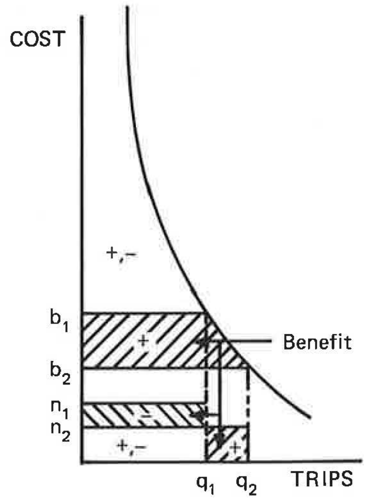


Figure 2. The estimation of movement benefits: comparison of alternatives.



$n_1 q_1$), giving a total change in surplus or benefit of

$$\Delta = \tfrac{1}{2}(q_1 + q_2)(b_1 - b_2) + (n_2 q_2 - n_1 q_1)$$

Various algebraic manipulations that improve understanding and ease computation are possible. In particular we have $n = b - r$, where r is the unit resource cost, and hence

$$\Delta = \tfrac{1}{2}(q_1 + q_2)(b_1 - b_2) + q_2 (b_2 - r_2) - q_1 (b_1 - r_1)$$

or

$$\Delta = \underset{(A)}{\tfrac{1}{2}(q_1 + q_2)(b_1 - b_2)} + \underset{(B)}{[(q_2 b_2) - (q_1 b_1)]} - \underset{(C)}{[(q_2 r_2) - (q_1 r_1)]}$$

In other words,

Δ = increase in user surplus (A) } i.e., increase in gross value received by
+ increase in costs to users (B) } travelers as measured by behavioral costs
- increase in resource costs (C)

In this form it will be seen that the change in resource costs (C) can either be calculated from unit costs per trip or be estimated separately from a consideration of overall costs. For public transit systems, where there are many shared but variable costs, the latter course may be necessary.

This important result is the basis of the current network evaluation procedures, and in practice it is estimated for all the trip classes pertinent to the study; i.e., the expressions are summed over origin-destination pairs, modes, times of day and year, trip purposes, and person type classifications (e.g., income groups or car owners and non-car owners) to give the total direct movement costs and benefits. It is possible to examine partial summations (e.g., separate out origins, destinations, or person types) with a view to learning about the distribution of benefit. The validity and uses of this procedure are still under examination.

Strictly speaking, it is not possible to describe the demand curve for any one of these separate trip classes unless the costs for all the others are kept constant; to this extent the explanation and the figures are simplifications because the traffic models simulate the fact that the costs of travel between many origin-destination pairs will vary simultaneously when networks are altered. However, it can be demonstrated (15) that this treatment is a close approximation to the much more complicated situation where all the demand curves vary together. This simple expression also requires that the land use and socioeconomic assumptions be constant between alternatives.

Application of the Approach to the Example

Returning to the example, assume for simplification that the user costs quoted do in fact equal resource costs. In this situation there is only the increase in user surplus analogous to the attempted calculations in the example. Comparing the before situation with the after situation, which takes account of the changed trip-making and pattern, gives the user benefit of

$$\begin{aligned}\Delta &= \underset{AA}{\tfrac{1}{2}(10 + 8)(5 - 4)} + \underset{AB}{\tfrac{1}{2}(1 + 6)(30 - 15)} + \underset{BB}{\tfrac{1}{2}(10 + 8)(5 - 4)} \\ &= \tfrac{1}{2}(18 + 105 + 18) \\ &= 70.5 \text{ units}\end{aligned}$$

Note that this differs from both of the previous estimates. Previously there was a benefit of 15 units with the fixed trip pattern and a disbenefit of 24 units with the changed pattern. As explained, the difference is due to placing a value on the benefits that stem from the changes.

Behavioral Costs for Evaluation

The previous sections, for ease of presentation, derived the basic evaluation expression in terms of the behavioral costs, b. However, as explained earlier, in the evaluation procedure the behavioral costs need to be in monetary units, and it may be appropriate to study the effect of values of certain elements differing from individuals' values as revealed through their behavior. This allows, for example, the effect of alternative relative weightings of benefit among time, cost, and income groups to be examined and different views on their absolute worth to be considered.

It is necessary then to calculate and use the variable u, which is identical in form to b except that it contains the alternative values and is scaled in cost units. The resource cost r is only defined in the context of the evaluation and will take account of any alternative values used in u. The expressions for the benefit become

$$\Delta = \tfrac{1}{2}(q_1 + q_2)(u_1 - u_2) + q_2(u_2 - r_2) - q_1(u_1 - r_1)$$

and

$$\Delta = \tfrac{1}{2}(q_1 + q_2)(u_1 - u_2) + [(q_2 u_2) - (q_1 u_1)] - [(q_2 r_2) - (q_1 r_1)]$$

Treatment of Taxation

As indicated earlier, some part of the difference between resource costs and behavioral costs is in terms of indirect taxation, which will account for part of the benefit. This will be transferred through the taxation system to society at large and arises because such taxation payments do not constitute use of resources; this nonresource element in the transport sector has already been taken into account. However, any increase in expenditure on transport would be accompanied by a reduction in expenditure, and hence taxation, on other goods and services. There will be a compensating nonresource element in the remainder of the economy equal to the change in taxation in the nontransport sector. The adjustment consists of estimating the change in expenditure between the sectors ($\Delta T/\phi_t$) and factoring this by an assumed nontransport sector tax rate ϕ_n that, in the absence of any specific knowledge of the alternative consumption foregone, can be assumed to be the average indirect tax rate in the remainder of the economy. Hence

$$\text{Correction} = -\Delta T(\phi_n/\phi_t)$$

where

ΔT = increase in tax paid within the transport sector modeled, and
ϕ_t, ϕ_n = taxation rates in the transport sector modeled and in the remainder of the economy respectively.

Summary of the Estimation Expression

The foregoing sections have indicated that the movement costs and benefits may be estimated as follows:

Total movement benefit = increase in gross value to travelers (as measured by their behavioral costs) (i) and (ii)
- increase in resources consumed in transport (as measured by resource costs) (iii)
- taxation adjustment for rate of tax in remainder of economy (iv)

Algebraically this is

$$\text{Benefit} = \underbrace{\tfrac{1}{2}\Sigma(u_1 - u_2)(q_1 + q_2)}_{\text{(i)}} + \underbrace{\Sigma(q_2 u_2 - q_1 u_1)}_{\text{(ii)}} - \underbrace{\Sigma(q_2 r_2 - q_1 r_1)}_{\text{(iii)}} - \underbrace{\Delta T(\phi_n/\phi_t)}_{\text{(iv)}}$$

where

u_1, u_2 = unit behavioral origin to destination costs for the mode, purpose, etc., under consideration in two alternative situations, incorporating any special or trial valuation of the components;
r_1, r_2 = unit resource costs (also incorporating any alternative valuations), i.e., time costs plus money costs less indirect taxation;
q_1, q_2 = trips;
ΔT = increase in tax paid on transport;

Σ = sum over all modes, purposes, origin-denstination pairs, income groups, etc., under consideration; and

ϕ_t, ϕ_n = taxation rates in the transport sectors considered and in the remainder of the economy.

Note that Σ $(q_2 r_2 - q_1 r_1)$ is the net change in resource, and it may be necessary or more convenient to estimate it in part or whole in an aggregate fashion.

SUMMARY

The objective of this paper is to provide practical guidance to transport planners and analysts by describing procedures in two areas: (a) the evaluation of movement costs and benefits consequent to changes in networks and management policies and (b) the estimation of the generalized behavioral and resource cost functions for links of networks and for origin-destination pairs that are necessary for this evaluation process and for forecasting behavior.

The procedures are designed for use in situations where the change in network or policy is thought to have strong effects on the trip pattern and individual link loadings. This will generally be the case in the consideration of urban schemes and may be the case for major interurban schemes as well; in both situations, there may be considerable changes in the trip matrices, modal split, and routes used.

The emphasis of the paper is on operational methods. There are many theoretical points of potential dispute or of current ignorance; insofar as possible, judgments have been made and procedures determined to cover these.

Individual networks or management policies cannot be sensibly considered in isolation. Comparisons between alternatives are essential for the valid consideration of the individual proposals. One particularly important comparison is between the various possible future systems under consideration and, as a base, the "do-nothing" situation. In this contex, "do nothing" means to include only those changes to the existing situation that are, for all practical purposes, unavoidable between now and the period under consideration. It is essential, in the context of the evaluation analysis described in this paper, that identical land use and socioeconomic assumptions be made for each of the alternatives considered.

One part of the comparative evaluation of alternative transportation networks or management policies requires the estimation of those costs or benefits that arise directly from the changes in costs of movement and the associated changes in the volume and pattern of movement. This paper concentrates on the evaluation of these "movement costs and benefits" and does not consider the estimation of all the other important costs and benefits (e. g., capital and other initial costs, accidents, development consequences, and environment) nor the integration of all these into the decision-making framework.

Individual travelers are often not aware of the true costs of travel by alternative modes and to alternative destinations and, in fact, may not even have an objective assessment of these costs. In this situation behavioral costs, b, are here defined as those costs that when used in appropriate models give the best empirical fits to observed behavior; with this basis there is reasonable satisfaction that such models and costs can be used to forecast patterns of movement in alternative situations. These costs in practice are described as linear functions of the costs (fares, perceived mileage costs, and so forth) and component times for the various stages of the possible journeys.

Because the behavioral costs represent the best available estimate of the individual traveler's "disutility of travel," it is sensible to estimate the benefits, or increase in value, to travelers in terms of these costs. The user benefit is estimated as the extra that travelers would have been prepared to pay over the behavioral costs they experience; this is the concept of "consumers' surplus" and a measure of the net user benefit may be obtained by summing it over all journeys by all users. However, these behavioral costs will often not represent use of resources; for example, they may differ from resource costs, r, because of misperception of outgoings, profits or losses of operators, and taxation (e. g., on fuel). It is necessary then to add to the consumers' surplus received by travelers the amount by which behavioral costs exceed resource costs.

For purposes of evaluating public sector investments or policies, a public authority could examine the effect of alternative values of some items of individuals' behavioral costs—leisure and commuting time, for instance—differing from those of the individuals themselves. This means that the net benefits would have to be adjusted by the use of special behavioral costs for evaluation, u, that contain these chosen values rather than those in the normal behavioral costs, b.

Part of the value received by travelers is transferred through taxation payments to society at large; such payments do not constitute use of resources and therefore are an addition to total social benefit. However, any increase in expenditure on transport would be accompanied by a reduction in expenditure on other goods and services. It is necessary to make an adjustment to allow for the different rates of indirect taxation on transport and other expenditures; in the absence of specific knowledge of the alternative consumption foregone it can be assumed that it attracts the average rate of taxation on final expenditure in the remainder of the economy.

An expression is developed for the estimation of the movement costs and benefits in terms of the b, u, and r costs and the trip matrices that result from the modeling process. An algebraic summary of the benefit expression is given.

The precise way in which the benefit expression and generalized costs are calculated will depend on the level of detail and form of particular studies. Considerable guidance is contained in the paper to aid the transfer from concepts to computation.

REFERENCES

1. Movement in London. Greater London Council, 1969.
2. McIntosh, P. T., and Powell, T. J. London Transportation Study: Some Aspects of the Economic Evaluation Methods. Planning and Transportation Research and Computation Co. Ltd. Symposium, Sept. 1969, pp. 15-19.
3. Tressider, J. O., et al. London Transportation Study: Methods and Techniques. Proc., Institute of Civil Engineers, March 1968, p. 433.
4. The Economic Assessment of Road Improvement Schemes. Road Research Laboratory, Tech. Paper 75, Ministry of Transport, 1968.
5. Wilson, A. G., Hawkins, A. F., Hill, G. J., and Wagon, D. J. Calibration and Testing of the SELNEC Transport Model. Regional Studies, Vol. 3, 1969, pp. 337-350.
6. Harrison, A. J. Some Results From Using Interview Methods in Time Research. Dept. of the Environment, Highway Economics Unit, Time Research Note 6, 1969.
7. Modal Split: Factors Determining the Choice of Transport for the Journey to Work. Local Government Operational Research Unit, Rept. C32, 1968.
8. Quarmby, D. A. Choice of Travel Mode for the Journey to Work: Some Findings. Jour. of Transport Economics and Policy, Vol. 1, No. 3, 1967.
9. Lisco, T. The Value of Commuters' Travel Time: A Study in Urban Transportation. Univ. of Chicago, Dept. of Economics, PhD dissertation, June 1967. (An abridgment of this paper appears in Highway Research Record 245, 1969, p. 16.)
10. Stopher, P. R. Predicting Travel Mode Choice for the Work Journey. Traffic Engineering and Control, Jan. 1968.
11. Planning for the Work Journey: A Generalised Explanation of Modal Choice. Local Government Operational Research Unit, Rept. C67, 1970.
12. Dupuit, J. On the Measurement of the Utility of Public Works. Annales des Ponts et Chaussées, Second Series, Vol. 8, 1844.
13. Hicks, J. R. Value and Capital: An Inquiry Into Some Fundamental Principles of Economic Theory, Second Ed. Clarendon Press, Oxford, 1965.
14. Quarmby, D. A. Transport Planning in a Multi-Resource Context: The Morecambe Bay Barrage. Operational Research Quarterly, Vol. 21, Special Conference Issue, July 1970.
15. Neuburger, H. User Benefit in Transport and Land Use Plans. Jour. of Transport Economics and Policy, Jan. 1971.
16. Research Into the Value of Time. Dept. of the Environment, Time Research Note 16, July 1970. HMSO, London.

DISCUSSION

Herbert Mohring, Departments of Economics, University of Minnesota and York University

I am in broad sympathy with the proposals by McIntosh and Quarmby for improving the arts of predicting and evaluating the effects of alternative transportation improvements. If factors other than travel time could adequately be taken into account, more accurate predictions would likely result. Using consumers' surplus rather than cost change benefit measures would also be highly desirable. I do not say this because I expect that consumers' surplus measures would yield radically different estimates from other sensibly chosen cost change measures. [I very much doubt that the danger with cost-based measures that the authors discuss has ever led to rejection of a highway improvement proposal. The change in user cost with a fixed trip pattern seems almost invariably to be used in such estimation work.] Rather, I espouse the proposal because it would bring transportation benefit estimation closer to the economic analysis used in dealing with formally similar problems, thereby reducing the still regrettably high frequency with which vast nonuser benefits are associated with such phenomena as land value changes and generated traffic.

Unfortunately, though, this paper is not without its shortcomings. The authors have opened several cans of worms without either sorting the contents out very well or, indeed, recognizing how difficult the sorting process would be. It is to three of these unsorted cans of worms that I would like to point.

First, in regard to their discussion of "generalized costs," it certainly is true that trip attributes other than travel time influence travel behavior. It is also true, however, that the relative importance of these attributes varies substantially from individual to individual. Although the authors do mention this point, they fail to come adequately to grips with it.

Consider a consumer who desires to maximize the utility he derives from spending his fixed income on two commodities—a general purpose good, "purchasing power," and trips from here to there. The consumer enjoys trips not for themselves but rather for what happens once he gets there. Indeed, the time he spends in transit is a source of dissatisfaction, not utility. It can be shown that this sort of consumer will allocate his income between trips and purchasing power as if the price of a trip equals whatever fare he pays plus the value he attaches to travel time times the time required per trip. As a number of studies have shown, the value of travel time varies substantially among individuals and is particularly closely related to their income levels.

Suppose, now, that there are two ways of traveling between here and there. One is fast but involves a high fare; the second is slow but is low priced. Except for time and money, the consumer is indifferent between them. If so, he can be expected to choose the mode (say) that involves the lowest price to him. The difference between the money costs of the two modes divided by the differences between their travel times is, in effect, the price he must pay to save a minute's travel time by using the faster mode. If the value he attaches to travel time is more than this price, he will use the fast, high-priced mode; if not, he will use the slow mode. In such a system, reduction in the fare or travel time for one mode will lead some travelers to shift to it from the other mode—note, some travelers, not all.

In brief, accurate prediction of route and mode choice on a transportation system in which routes and modes have different mixes of travel times, money costs, and other attributes requires taking into account the fact that any given route has as many prices (as many generalized costs) attached to it as there are potential travelers. The authors' procedure does not take this fact into account.

Second, a minor point in the paper, but one worth mentioning because of its importance to the literature on modal choice, is that the authors suggest stratifying the population by, inter alia, automobile ownership in predicting travel behavior. Implicit in this suggestion is what seems to me to be an erroneous assumption of causality except in the very short run. A simple but perhaps not unusual example follows: A husband and wife both work. He must use a car on the job. The choice by her to drive rather than to use an available bus is therefore effectively a choice by them to buy a second car. In deciding on her travel mode, they will presumably weigh the time she will save by driving against the additional money outlay doing so will require. In these calculations, it is worth noting, the relevant money cost is that of owning and operating a second vehicle, not just out-of-pocket operating costs. Here, clearly, automobile ownership is determined by choice of mode, not the other way around. More generally, the choices of mode, automobile ownership, and, indeed, residence and work place location are interdependent. The process is not one in which the last three variables are exogenous in a single equation that serves to determine modal choice.

To do justice to the last group of problems in this paper that I want to discuss would require a complicated and lengthy discussion, far more than I can provide in this limited space. All I can do here is to point out the existence of the problems and assert that solutions to them do exist. [That is, benefit measures closely related to McIntosh and Quarmby's consumers' surplus techniques can be developed that require no more information than do their measures and that avoid the problems to be discussed (17).]

Generally, the amount of a commodity any individual consumes depends not just on its price but also on his income and the prices of complementary and substitute products. If the price of one commodity changes, consumer demands for other commodities and, quite likely, their equilibrium prices also change. Thus, improving one highway and thereby lowering the price of trips on it will likely serve to divert traffic from other highways. The result will be lower congestion and lower trip prices and, hence, benefits to the users of these facilities that are clearly not directly reflected in the demand schedule for the originally improved highway.

An obvious extension to the authors' analysis suggests itself to handle this problem: Add up changes in consumers' surpluses not just for the originally improved highway but also for those that are benefited through traffic diversion. The problem is that the position of the demand schedule for one highway, the area under it, and the measured benefit depend on the price of trips on other highways.

Formally, this extension of McIntosh and Quarmby's proposal involves evaluation of a line integral along some particular path. It can be shown that the value of a line integral will depend on the path chosen to evaluate it unless certain integrability conditions are satisfied. It can also be shown that these conditions are not normally satisfied for the sort of demand schedule dealt with by economists generally and by the authors in particular. This being the case, it is quite possible that, using their techniques, the rank order of two alternative improvements would depend on the specific path used in evaluating their benefit line integrals—clearly an unhappy state of affairs. At least one consumers' surplus type of benefit measurement technique that does not suffer from this disability exists, but it is not that implied by McIntosh and Quarmby's paper.

Reference

17. Mohring, H. Alternative Welfare Gain and Loss Measures. Western Economic Jour., Dec. 1971.

AUTHORS' CLOSURE

We would like to thank Mohring for his remarks. They bear on several theoretical points that, although generally covered in the references supporting the paper, are not fully discussed in the body of what was intended to be a paper for practical guidance.

His points may be split into two main areas: those dealing with the use of generalized cost in modeling and forecasting individual and group behavior and those connected with the calculation of benefit. We accept completely that different people will individually have different generalized costs for the same journey. This is fully recognized in our paper and in the forms of models that are generally used in the United Kingdom for modal split and distribution and for which this paper is intended to help provide inputs. These models do not predict behavior on the basis of minimum cost but rather by the use of logistic-exponential probability splitting functions, dividing any group at risk among the origins, destinations, and modes available.

Such models give deterministic estimates of aggregate behavior, but they are based conceptually on probabilistic hypotheses about individuals' behavior so that variations in behavior among individuals are implicitly allowed for. In practice, to split the population into subgroups according to, for instance, income groups, levels of car ownership, or journey purposes will move toward homogeneity within groups and will reduce the effects of the overall interpersonal variations. But some variations between individuals within groups will remain and are allowed for.

There are, of course, procedures being developed for joint modeling of mode, automobile ownership, and residence and work place locations that use various concepts of utility, generalized cost, and dynamic behavior, but we did not try to include them in a paper deliberately restricted to generally available and fully practicable procedures. To make firm recommendations on how the automobile ownership-choice of mode relationship should be treated (other than by ignoring it) would be dangerous in view of the relatively limited research that has been carried out in this area.

In suggesting that our procedure should be extended to cover traffic benefited by diversion, we think that Mohring may have missed an important concept of our paper, inasmuch as benefits arising from all changes in travel behavior are counted, as well as benefits from simple cost reductions. We deal with the totality of all trips affected; the integration, or summation, of the benefit expression is over all origin-destination pairs and modes and thus takes account of changes between them on account of cost-induced behavior changes. The generalized cost is the aggregation of total origin-destination costs, not the element of cost on a particular element of highway or transit system; similarly, the origin-destination trip is the "good," not the trip from one end of a highway link to the other. Traffic between other origin-destination pairs benefiting from congestion reduction is thus allowed for.

The issue relating to the path of integration and the form of the benefit expression is very complex, and we accept that it is not fully resolved. The point is briefly discussed in our paper and in some detail elsewhere (15). A summary of the argument is given here. Traditionally the argument in favor of using the trapezium measure as an approximation in the case where only one price falls is as follows:

For those people who continue to consume the same amount of a commodity before and after a price change, the benefit must be exactly equal to their change in expenditure (or, in our case, generalized cost). For those who change their consumption pattern, their benefit per unit cannot be greater than those who do not change their behavior; otherwise, they were in a nonrational position before the price change. Also it cannot be less than zero because, otherwise, they would have moved to a nonrational position after the price falls. If we assume that their benefit lies midway between the possible extremes, we obtain the trapezium measure.

We now come to the case of a multiple price change. We divide expenditure into that which continues to be devoted to the same good and that which is switched from one good to another. The first group, as in the one price case, receives a benefit equal to the change in expenditure. By the same argument as in the sample case we may show that the benefit per item of switched expenditure lies approximately half way between the price changes of the goods between which the switch is made. This in turn can be shown to add up to the sum of the trapzium measures taking the demand diagram for each good separately. Clearly this is an approximation in the same way that the single price change case is an approximation.

Theoretically it has been shown that the problem of the path of integration is closely related to variability in the marginal utility of money and is of the same order of significance. In the context of transportation models normally used, it has been shown that very extreme changes of transport provision and costs are required before the issue becomes of a more than negligible importance. Other measures, which avoid these approximations, require enormous amounts of normally unobtainable data for their computation; we feel that these cannot be considered as available for general and routine use.