USER BENEFIT IN THE EVALUATION OF TRANSPORT AND LAND USE PLANS*

By H. Neuburger

User benefit is of critical importance in the evaluation of transport and land use plans; but until recently the problems associated with its measurement have received little attention. This paper attempts to make a contribution to this field, which will be related to measurement solutions in earlier literature.

Three methods of estimating user benefit in complex transport networks are recommended, their use depending on the level of sophistication desired. The description of the most sophisticated method leads on to a discussion of the economic interpretation of certain forms of transport demand model. This in turn leads to the description of a method of evaluating land use plans. There is also a description of a method of evaluating approximately links or sets of links which form part of a network improvement plan.

The discussion in the main body of the paper omits the important issue of income distribution. This is not because I wish to ignore it, but because in the sequence of evaluation its practical impact comes after the measurement of user benefit. As indicated below, most of the measures may be disaggregated to the level at which data on income and wealth are available, and it is after these disaggregated valuations have been obtained that allowance may be made for distributional objectives or for any redistributional bias of which the measures are suspected. The use of the term "distribution" in the subsequent discussion is reserved for the distribution of trips from origins to destinations rather than for income distribution.

The purpose of economic evaluation is to help to determine whether a particular plan is justified. The following items usually enter into the evaluation of transport projects:

(i) Construction and maintenance costs of facilities
(ii) Vehicle operating cost
(iii) Journey times
(iv) Accidents
(v) Accessibility and comfort and convenience
(vi) Environmental and other non-user effects.

This paper is concerned only with items (ii), (iii) and (v). The meaning of items (ii) and (iii) is obvious. Item (v) is slightly more complicated. When a transport network is improved, some places become relatively easier to reach – they become more accessible – and more travel to them generally ensues. The benefits derived by users

*The author wishes to acknowledge the help of the Ford Foundation in supporting this research.

from this extra travel are defined in this paper as the benefits of improved accessibility. Similarly some modes of travel are regarded as more comfortable and convenient than others. Thus one mode may become relatively cheaper as a result of an improvement. Travellers will then receive a benefit from a switch to this mode. (This could of course be a negative benefit if the mode were less desirable; but that would be offset by benefits under items (i) and (ii).)

Economic evaluation will always involve the comparison of two situations, one of which will often be the result of "doing nothing". The "do nothing" situation must be as fully analysed as the situation which arises from the proposed change, since it provides the benchmark from which the effects of doing something can be measured. This may be contrasted with certain forms of engineering assessment, where it may be regarded as sufficient to look at, for example, the expected volume of traffic on a proposed new road in order to decide whether or not it is necessary. All the items quantified in an economic evaluation represent differences in magnitude rather than absolutes.[1]

The items with which we are particularly concerned in this paper are sometimes called user benefits or costs, since in general they affect the users of the facilities concerned. Economic theory suggests that the benefits that people derive from changes in the quality of transport facilities provided for them may be deduced from the way they behave, and may be expressed in terms of money. A particular case of this is the derivation of the "value of time" by observing travellers' choice of mode of transport or route, and deducing from various different time/cost combinations the amount of money they are prepared to spend to save a certain amount of time. We show here how this principle may be extended to the case of interactive network changes.

Estimates of benefit derived simply from users' behaviour may have to be modified. To show why this is necessary we introduce a distinction between "behavioural or perceived costs" and "resource costs". For example, car drivers tend not to take full account of running costs in estimating the cost of making a trip. Studies have shown that it is possible to estimate from engineering costs the "perceived" and "resource" costs of a journey (see Quarmby [11]). They may therefore be regarded as known quantities. When estimating net user benefit it will normally be appropriate to use perceived costs, since it is these users have in mind in deciding whether, when and where to make a trip. However, in determining the economic value of a project it is appropriate to consider the actual costs as well, *i.e.*, the resource costs.

In the following descriptions I shall assume that there has been established a single value of time. Thus the unit cost to a traveller of making a journey will normally consist of a time cost and either an operating cost or a fare. Resource costs and perceived costs will both include time at this same valuation. The argument is simplified by these assumptions, but remains substantially the same if they are altered.

---

[1]The failure to recognise this point causes some confusion in Martin and Wohl [8]. They draw a distinction between comparing mutually exclusive projects and deciding on the value of any plan in the abstract. If one regards doing nothing as an alternative action (mutually exclusive with doing something), this distinction disappears. They seem to assume implicitly than the costs of doing nothing are either infinite or zero (it is not quite clear which). But the consequences of doing nothing are finite and, in principle, predictable.

METHOD I

A method I evaluation is carried out as follows. Present demand for journeys (or demand expected to obtain immediately before the introduction of the proposed change) is examined. Although any transport improvement must alter the demand for trips, the essence of method I is that these changes may be ignored. Thus it is supposed that the same number of trips will take place between each origin and destination pair by each mode, both with and without the project, but that with the project journeys made will follow different routes to use the new facilities where appropriate. In this case the only available measure of benefit is the overall change in user cost between the "with" and "without" situations. (It will be clear later that in fact all the other methods reduce to method I in these circumstances.) We may use Figure 1 to illustrate this. The horizontal axis represents the number of trips between a particular origin and a particular destination by a particular mode, and the vertical axis the cost of that journey. Since the number of trips is assumed unchanged, the "demand curve" *BF* is vertical. The perceived benefit to the user is thus the reduction in perceived costs *ABFE*. On the other hand, the change in the value of resources used – the change in resource cost – is given by *CDHG*. As we have seen, *ABFE* is perceived by the users as a benefit; thus the additional unperceived benefit (or unperceived cost) is *CDHG – ABFE*. Thus we may combine the perceived and unperceived benefits, and the result is a net benefit given by *CDGH*. It will be observed that it is thus possible to measure the net social benefit without knowing the perceived costs. This arises from the fact that demand is unaffected by cost changes. Thus it should not matter how users perceive costs, since they do nothing about it. This may be contrasted with later methods. Thus net user benefit in the method I valuation is given by the sum of areas such as *CDHG* for all journeys by all modes represented in the system. In general some of these will be positive, very many of them will be zero, and a few will be negative where particular journeys have become more costly (for example, as a result of congestion or street closure).

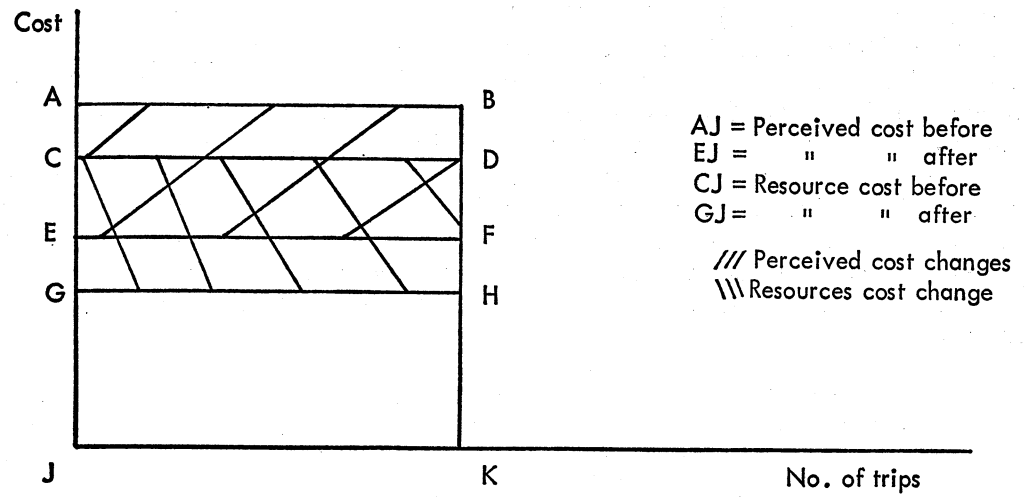

FIGURE 1

Computational Aspects

The total time and operating cost in a network is a normal output of an assignment procedure when computer models are used. If the assignment procedure is manual, the differences in cost may be easily calculated during the procedure. In this method it does not matter for the purposes of the final result whether the sum is arrived at by aggregating over individual journeys or over individual links in the network. One of the major advantages of method I is its computational simplicity.

Some Limitations of Method 1

It will be observed that method I ignores the effects of improved accessibility; that is, the users are assumed not to take advantage, by altering their pattern of travel, of the fact that some destinations have become more accessible. Provided such changes are not large this omission is not likely to be serious. On the other hand it tends to overestimate the value of congestion relief, since it predicts less crowding on roads than is likely to occur. Thus, in this respect, the new networks will appear in an unduly favourable light. These two biases will often work in opposite directions. Both effects are likely to be small with small changes. But method I is permissible only where it may be assumed that the changes have negligible effect on the pattern or extent of trip making. If this is not the case method II or method III must be used.

Some Incorrect Applications of Method 1

Method I is substantially the same as that described in the Road Research Laboratory Technical Report No. 75 or the AASHO Red Book [12], [1]. These methods, however, are often misinterpreted. We describe here three common interpretations.

The first proceeds as follows: The pattern of trip making is estimated in terms of the origin and destination matrix and modal split expected to obtain after the introduction of the project, with allowance for the response of travellers to the new facility. This pattern is then assigned to both the old "do nothing" network and the new network, and the user benefit is estimated as the hypothetical time savings and cost savings between these two situations. This method can vastly overestimate benefits. While in the proper use of method I some benefits are usually underestimated through ignoring changes in accessibility, this is to some extent counteracted by the overestimation of the benefits of congestion relief. In the method described here, on the other hand, these two effects will normally work in the same direction, and both tend to overestimate. In addition, even in the cases where method I is appropriate, the over-valuation of congestion relief is likely to be much larger. Thus where the method is correctly applied the inaccuracy as far as congestion is concerned arises from assigning too little traffic to new roads. In this incorrect application the error is assigning too much traffic to old ones. On congested roads a little extra traffic is likely to have a very marked effect, whereas on uncrowded roads this is less likely. Thus if the "new" traffic is applied to the "old" roads, the congestion is likely to be much overstated, and its relief consequently overvalued.

A second misapplication is found in some assessments, where, after following method I correctly, the investigator then adds an allowance for generated traffic to the "with" case, with the result that a user cost comparison of the "with" and

"without" cases attributes too high a cost to the latter and thus underestimates net user benefit.

The third and most common misapplication is the comparison of user costs in situations where the trip demand is not the same in the two situations being compared and is not assumed to be the same. This will often give highly misleading results. It is quite possible, for example, for the introduction of a new facility to attract trips to a further destination, so that average trip time and cost will increase and more trips will take place. A comparison of the user costs would show that the "net user benefit" was negative, which is clearly absurd. To take an even more extreme example, user cost minimisation, which is the objective implied by this criterion, would select a situation where there was no travel at all as being the best attainable, since that is the situation in which transport cost would be at a minimum.

The reason why this method of evaluation is misleading is that it fails to take account of benefits from changes in accessibility, while allowing these changes to affect demand. It is inconsistent to suppose that something will affect travel behaviour, but that a change in it will have no value. It is not even necessarily true that by using change in cost as one's criterion in this case one will necessarily underestimate benefit. So long as demand is unconstrained benefit will be underestimated, since the effect of transport improvement if demand is free is a change to relatively more costly trips. We may see that this is true by comparing this criterion with the formula proposed in method II, which we bring forward in the simplified form for one mode of travel.

$$UB = \tfrac{1}{2} \Sigma (T^2_{ij} + T^1_{ij}) (C^1_{ij} - C^2_{ij}) \tag{1}$$

where

UB = user benefit

T_{ij} = trips from zone i to zone j

C_{ij} = unit cost of a journey from zone i to zone j.

The superscript 1 and 2 refer to the "old" and "new" networks.

$$CC = \Sigma (T^2_{ij} C^2_{ij} - T^1_{ij} C^1_{ij}) \tag{2}$$

where CC = the change in user cost.

so

$$UB + CC = \tfrac{1}{2} \Sigma (T^2_{ij} - T^1_{ij}) (C^1_{ij} + C^2_{ij}) \tag{3}$$

The sign of this term depends on the correlation between the increase in trips for a particular journey and the mean cost of a trip.

As we have suggested above, this is normally positive in the case where demand is unconstrained. If demand is constrained, as it often is in transport demand models (see below), $UB + CC$ is often likely to be negative. For example, if trips to a particular destination are constrained to a certain number and the destination is made more accessible to more distant origins, this may well mean that makers of trips who used to reach there from nearby origins will be obliged to be content with even more local destinations, which they would have been free to choose before. Therefore they will realise a cost reduction, but only by going to a less desirable destination. Thus the cost reduction will overstate the benefit, if any, which they receive from the transport improvement. (The point is emphasised because Martin and Wohl [8] appear to think that the criterion of change in user cost will underestimate benefits.) This discussion emphasises once more the importance of applying more complicated methods where the transport improvements affect the demand for trips substantially.

METHOD II

A major difference in the approach between methods I and II is that in method I we were concerned with finding the most efficient way of catering for a fixed demand level and pattern, while here we are trying to obtain the optimum level of provision of the things which transport investment is intended to give, *e.g.*, accessibility, comfort and convenience. Hence it is necessary to consider the effects of improvements on the demand for travel. If the prediction assumes that changes take place, the evaluation must allow for this. The evaluation method described here is an extension of the principle referred to above that user benefit may be estimated by observing the way people behave. The value of accessibility, etc., will be assessed in terms of the amount people are prepared to pay in time or operating cost to make different journeys by different modes.[2] The general method for arriving at perceived user benefits is to multiply the cost change for each origin-destination pair, by each mode of travel, by the average of the number of relevant trips made in the "with" and "without" situations. These products are then summed for all zone pairs and modes to give a total perceived user benefit. Algebraically this may be written:

$$UB = \tfrac{1}{2} \Sigma_{ijk} \, ({}_kT^1{}_{ij} + {}_kT^2{}_{ij}) \, ({}_kC^1{}_{ij} - {}_kC^2{}_{ij}) \qquad (4)$$

where

${}_kT_{ij}$ = trips from zone i to zone j by mode k.

${}_kC_{ij}$ = behavioural cost per unit of a trip from zone i to zone j by mode k.

Graphically this may be given by the sum of areas such as *EFGH* in figure 4. In addition it is necessary to calculate the additional change in resource cost which is not perceived by the traveller. Since the travellers are supposed to respond to changes in cost, it is no longer true that the change in resource cost is equal to the change in net user benefit, as in method I. Evaluation now involves the calculation of the perceived user benefit as described above and then the addition to it of the difference between the reduction in resource cost and the reduction in perceived cost. This is shown in figure 2.

Perceived user benefit = *ABCDE*

Fall in perceived user cost = *ABDE*−*DCHG*

Fall in resource cost = *IJLK*−*LMHG*

so

Net user benefit = *ABCDE*+*IJLK*−*LMHG*−*ABDE*+*DCHG*.

We shall now attempt to justify this method. The argument developed here makes the familiar point that unchanged journeys benefit by the full reduction in costs and that "new" journeys benefit by approximately half that amount. The reason why it is presented at such length is that confusion sometimes seems to arise when complicated situations are analysed graphically. In simple situations the demand curve for a particular trip does not shift as a result of the simultaneous change of its own and other prices. Thus the perceived user benefit is equal to the area under the demand curve, and the result is familiar. In transport demand prediction, it is very often

[2]This behaviour is customarily described and predicted with the aid of a computer model. This will involve the production in computible form of trip matrices from zone to zone within a study area and of interzonal time and cost matrices, for the situations both before and after the introduction of a project. These constitute the data for the evaluation process described here.

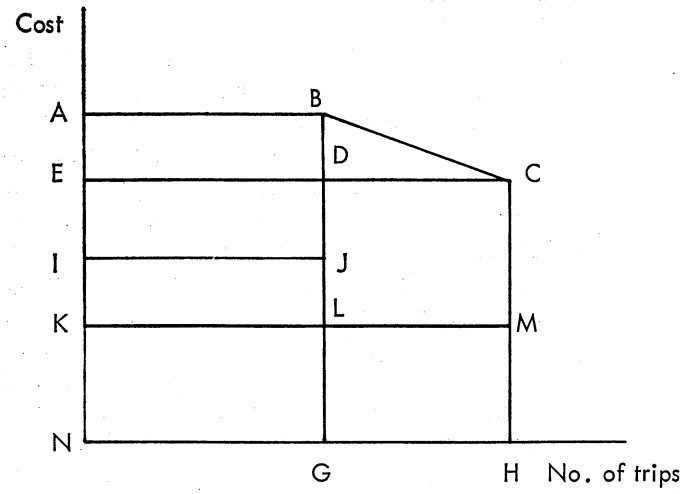

FIGURE 2

found that many costs change simultaneously, causing shifts in the demand curves for individual trips at the same time as movements along them. Figures 4 and 5 represent a simplified case of this. They each represent the demand for trips from the same origin to different destinations. These journeys are substitutes for each other; thus when the costs of both fall the demand curves for both shift downwards. In this case the area under the demand curve is no longer a correct or clear criterion. What the following argument shows is that even under these circumstances the fact that the demand curve has shifted may be ignored for the purposes of evaluation, and the areas under the quasi-demand curves *EFGH* and *JKLM* are still the appropriate measures of benefit, so that the formula of equation (4) is still correct. (The corrections of this process has been denied by Harvey [6], for example.)

The method II measure can be shown to take account of changes in accessibility in the following way. For each mode and zone pair the quantity measured can be divided into two parts. The first part consists of benefits to those who use the same mode to go to the same destination both before and after the change. Clearly these people have not received any advantage from improvements in accessibility. Their behaviour has remained unchanged; therefore their benefit is exactly equal to their change in cost. (We are concerned at this stage only with perceived user benefit.) This may be written as follows:

$$Q = \Sigma_{ijk}\, {}_kY_{ij}\, ({}_kC^1{}_{ij} - {}_kC^2{}_{ij}) \qquad (5)$$

where Q is the benefit to people who do not change their behaviour and

$${}_kY_{ij} = \min ({}_kT^1{}_{ij}, {}_kT^2{}_{kj})$$

The second part of the benefit accrues to those who do change their behaviour. This may be estimated by considering a particular case of those travellers who change from a journey from i to j by mode a to a journey from i to k by mode b. Now

these travellers will figure twice in formula (4). They will be included in ${}_aT^1_{ij}$ and ${}_bT^2_{ik}$. We will assume that the final journey receives a larger cost reduction than the original one. (This is purely for the purposes of exposition.) The benefit to these travellers cannot be greater than the change in the cost of travelling to k by mode b, not can it be less than that of travelling by mode a to zone j. In the former case they would already have changed to their "new" network journey on the old network, while in the latter they would have remained on their old network journey in the new network. The assumption made in method II is that the benefit lies halfway between these two extremes. Formally this is given by

$$R = \Sigma_{ijk}\, {}_k\mathcal{Z}_{ij}\, ({}_kC^1_{ij} - {}_kC^2_{ij}) \text{ where } {}_k\mathcal{Z}_{ij} = \text{mod } ({}_kT^1_{ij} - {}_kT^2_{ij}) \tag{6}$$

Combining these two elements:

$$UB = Q + R \tag{7}$$

which may easily be seen to be the equivalent of the formula in equation (4). It should be observed that no assumption has been made about the shifting of demand curves.[3]

Another problem sometimes arises with method II (and method III) evaluation. This is because many transport demand models are based on time only and do not allow operating costs or fares to affect demand. Experience shows that it is difficult

---

[3]In fact this "demonstration" presupposes that the total number of trips remains fixed. In order to examine the more general case we first formalise this argument and then show how it may be extended to the case where the number of trips varies. (We shall assume one mode to travel. The argument is unaffected and the number of subscripts reduced.) Thus the benefit to those who change from travelling from i to j to going from i to k lies between

$$\text{Max} = ({}_iX_{jk})\ [\max\ (C^1_{ij} - C^2_{ij}),\ (C^1_{ik} - C^2_{ik})] \tag{8}$$

and

$$\text{Min}\ ({}_iX_{jk})\ [\min\ (C^1_{ij} - C^2_{ij}),\ (C^1_{ik} - C^2_{ik})] \tag{9}$$

The mean of these is

$$\mu_{ijk} = \tfrac{1}{2}\ ({}_iX_{jk})\ (C^1_{ij} - C^2_{ij}) + (C^1_{ik} - C^2_{ik}) \tag{10}$$

where ${}_iX_{jk}$ is the number of people formerly travelling from i to j now travelling from i to k. Since the number of trips is constant

$$\Sigma_k\ ({}_iX_{jk}) = T^1_{ij} \quad \text{and} \quad \Sigma_j\ ({}_iX_{jk}) = T^2_{ik} \tag{11}$$

Therefore

$$\Sigma_i\, \Sigma_j\, \Sigma_k\, \mu_{ijk} = \tfrac{1}{2}\, \Sigma_i\, [\, \Sigma_j\, T^1_{ij}\, (C^1_{ij} - C^2_{ij}) + \Sigma_k T^2_{ik} (C^1_{ik} - C^2_{ik})] \tag{12}$$

$$= \tfrac{1}{2}\, \Sigma_i\, \Sigma_j\, (T^1_{ij} + T^2_{ij})\ (C^1_{ij} - C^2_{ij}) \tag{13}$$

as in equation (4).

If we now wish to drop the assumption that the total number of trips is constant, we introduce the possibility of purchasing a non-transport commodity x_{i*}. By assumption the price c_{i*} of this commodity is unchanged by the transport improvements. Thus by the same argument as above the benefit to those who stop travelling is

$$\rho_{ij} = \tfrac{1}{2}\ ({}_iX_{j*})\ [\ (C^1_{ij} - C^2_{ij}) + (C^1_{i*} - C^2_{i*})] \tag{14}$$

$$= \tfrac{1}{2}\ ({}_iX_{j*})\ [\ (C^1_{ij} - C^2_{ij})\ ] \tag{15}$$

Similarly the benefit to generated traffic is

$$\sigma_{ij} = \tfrac{1}{2}\ ({}_iX_{*j})\ [\ (C^1_{ij} - C^2_{ij})\ ] \tag{16}$$

$$T^1_{ij} = \Sigma_k\, {}_iX_{jk} + {}_iX_{j*} \tag{17}$$

$$T^2_{ik} = \Sigma_i\, {}_iX_{jk} + {}_iX_{*k} \tag{18}$$

so

$$\Sigma_i\, \Sigma_j\, \Sigma_k\, \mu_{ijk} + \Sigma_i\, \Sigma_j\, (\rho_{ij} + \sigma_{ij}) = \Sigma_i\, \Sigma_j\, (T^1_{ij} + T^2_{ij})\ (C^1_{ij} - C^2_{ij}) \tag{19}$$

Which is again the same as equation (4).

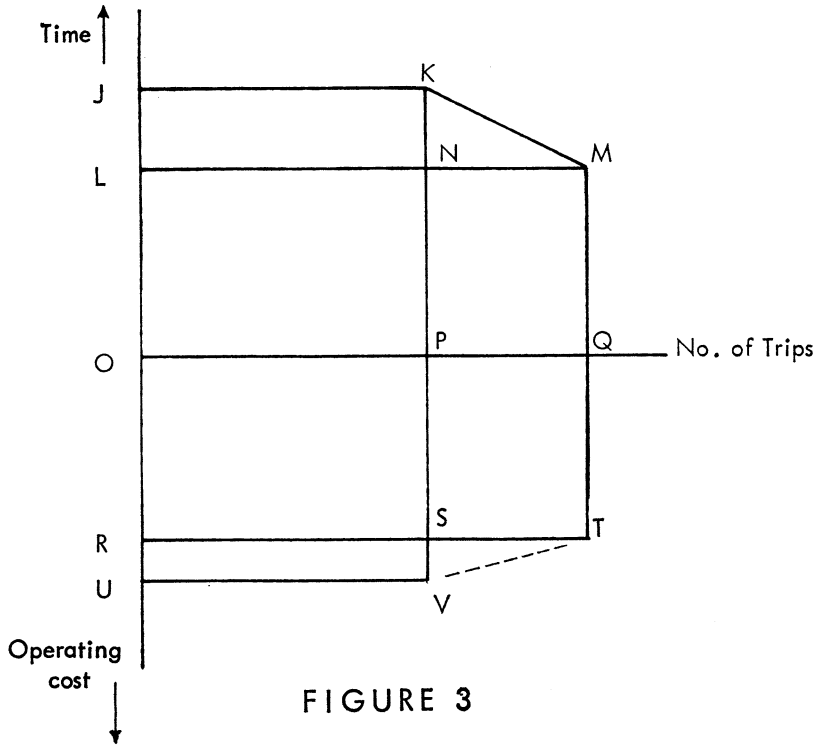

FIGURE 3

to interpret evaluations produced from these models. Thus it is always advisable to use a demand model which is based on generalised cost (*i.e.*, a composite of time and operating costs). The following gives some indication of the reason for this. For the purposes of the argument we assume identity between perceived and resource costs.

We compare the situation where generalised cost was used for prediction (Figure 2) with that where time only was used (Figure 3). In producing an evaluation formula for the latter case we may use one of two alternatives:

(a) Assume that travellers are really responsive only to time; in this case changes in operating costs represent resource costs not perceived, and may be treated accordingly. Benefit is given by B_1

$$B_1 = \mathcal{J}KMNL + RSVU - PQST \tag{20}$$

(b) Assume that travellers are responsive to both time and operating costs but that we have failed to model the response to the latter. In this case it is as if there were another part of the demand curve VT but it was not estimated or used for prediction.

(I should not have considered this case had it not been the assumption made in the London Transportation Study [13].) In this case we take as the measure of benefit the "areas under the demand curves".

$$B_2 = \mathcal{J}KNML + RSTVU \tag{21}$$

The difference between the two measures is clearly

$$B_1 - B_2 = PQTVS \tag{22}$$

Therefore B_1 is usually less than B_2, provided that trips increase relatively where

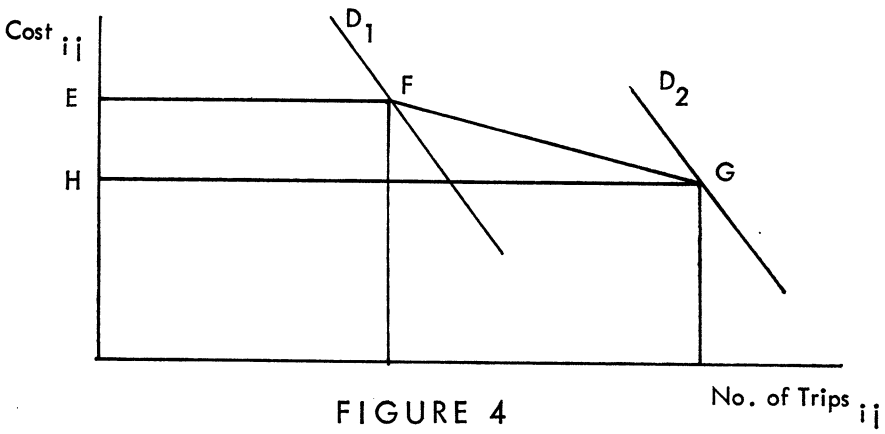

FIGURE 4

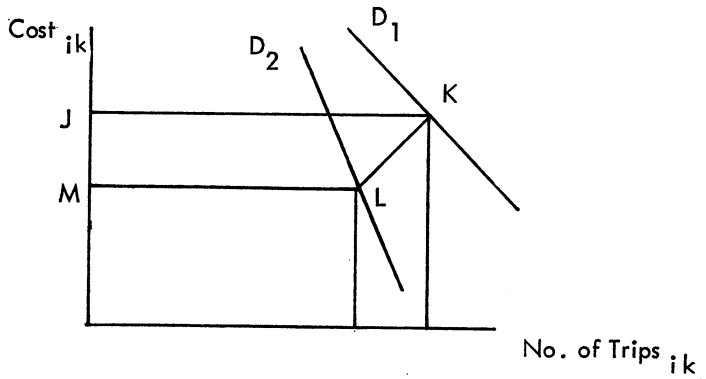

FIGURE 5

operating costs are relatively high, as is usually the case. This may be seen from the fact that $PQTVS$ is the correlation between the change in the number of trips on a particular journey and the operating cost of that journey. More important than this is the relation between either assumption and the evaluation method of Figure 2 (B_3).

$$B_3 = ABEDC \tag{23}$$

Bearing in mind that

$$AC = \mathcal{J}L + RV \tag{24}$$

$$B_3 - B_1 = AC(\mathcal{J}K - AB) - PQST - (GH \times SV) \tag{25}$$

$$B_3 - B_2 = AC(\mathcal{J}K - AB) + BD(PQ - GH) \tag{26}$$

The term $AC(\mathcal{J}K - AB)$ will be positive for fast journeys and negative for slow ones. $PQST$ is likely to be positive. $(GH \times SV)$ is marginally more likely to be negative than positive, since a rise in operating costs indicates a longer or faster journey. This will depend on whether improvements have a time-saving or distance-saving bias. $BD(PQ - GH)$ will probably have the same sign as $(GH \times SV)$. We may see that

it is very difficult to predict the bias of these measures; thus situations which might lead to their use should be avoided.

As stated above, method II is usually employed in conjunction with a computer model so that trip matrices and interzonal times are available in a computible form. It is therefore a simple matter to compute the benefit formula described above. It is possible to represent this in a form disaggregated by origin and destination zone, by type of travellers, and by time of day, etc. With origin and destination zone disaggregation some caution must be exercised in interpreting the results. The extent to which benefits can be attributed to origins and destinations will depend on the form of prediction model used, but this can usually be done satisfactorily. If the model constrains the number of trips at both origin and destination ends, no reliable attribution may be made to either. If the trips are constrained at one end only, attribution is possible at that end. Attribution to mode is unreliable unless the modal split is performed without reference to the network – that is, is based entirely on zonal characteristics. The maxima and minima described above will still apply at this level of aggregations. Reliability of estimates may be judged from these.

Another form of disaggregation which is often useful is the attribution of benefits to individual links or parts of a network for design or phasing purposes. The measures described above will provide an evaluation for a complete network change with a full model run for the two networks under comparison. We may reformulate equation (4) to

$$UB = \tfrac{1}{2}\left[\Sigma_{k \in K}\left(V^1_k C^1_n + V^2_k C^1_k\right) - \Sigma_{m \in M}\left(V^1_m C^2_m + V^2_m C^2_m\right)\right] \qquad (27)$$

where V_{rs} is the volume of trips on link r and c_{rs} is the cost of travelling along it. K is the set of links in the base network and M is the set of links in the new network. Thus we now have the same formula as in equation (4) but expressed in terms of links rather than journeys. For this formula to apply exactly to a part of a network it would be necessary for that part to be completely self-contained; but it is quite often possible to identify sections for which the error in assuming this to be true is not large. A useful approximation may often be obtained as follows. In addition to the normal assignments, the model must contain assignments of trips predicted for the new network to the old network and vice versa. After these assignments have been made the study area may be divided into areas which are reasonably self-contained and which contain parts of the improved network. The sums described in equation (27) may be calculated from these areas from the assignments required.[4]

[4]This method was developed mainly by Angus Niven of Peat, Marwick, Kates Ltd. and Mark Egerton of the Ministry of Transport. The derivation is as follows:

$$UB = \tfrac{1}{2}\Sigma_{ij}\left(T^1_{ij} + T^2_{ij}\right)\left(C^1_{ij} - C^2_{ij}\right) \qquad (28)$$

from equation (4).

$$UB = \tfrac{1}{2}\Sigma_{ij}\left(T^1_{ij}\,C^1_{ij} + T^2_{ij}\,C^1_{ij} - T^1_{ij}\,C^2_{ij} - T^2_{ij}\,C^2_{ij}\right) \qquad (29)$$

since

$$C_{ij} = \Sigma_{K \in n}\, C_k \qquad (30)$$

where n is the set of links on the route from i to j

$$V_k = \Sigma_{ij \in N} T_{ij} \qquad (31)$$

where N is the set of origin destination pairs for which k is a link on the journey between them.

So
$$UB = \tfrac{1}{2}\left(\Sigma_{k \in K}\left(V^1_k\,C^1_k + V^1_k\,C^1_k\right) - \Sigma_{m \in M}\left(V^1_m\,C^2_m + V^2_m\,C^2_m\right)\right] \qquad (32)$$

which is the same as the formula given in equation (27).

It will be observed that this formula is the mean of the Laspeyre and Paasche price indices.

It will be observed that this formula differs from that which would be obtained by replacing trips in the abscissa of a demand curve by the number of trips on a link, and applying the criterion of the area under the demand curve. Thus it is clear that the level of aggregation of transport demand is important in the application of method II. (As we have observed above, it was not important in method I.) This is an area in which transport economists have been very vague [5], [10], [15]. I have assumed that the correct level of aggregation should be the journey from a particular origin to a particular destination by a particular mode. Other possible levels are that the "goods" should be all travel, all travel by a particular mode, or all travel along a particular route or part of a route. It may well be that Winch [15] agrees with my assumption, and indeed when he defines a demand curve it is in terms of demand for journeys, but he also refers on other occasions to the demand for a particular link. Mohring and Harwitz [10] are similarly ambiguous. Indeed, their examples often refer to the demand for a road from *A* to *B* as origin and destination, which appears to sidestep the problem. Friedlander on the other hand, in her evaluation of the Interstate Highway system [5], regards all road travel as one good, irrespective of origin and destination. The result is that she is obliged to define a mean cost per mile on and off the Interstate system with no regard to how this fits in with journey patterns or alignment differences between the two. Also in this case there is too little interchangeability within the "good" defined, so that the level of aggregation is too coarse. On the other hand the demand for travel down a particular link also seems inapposite. So far as the traveller is concerned, travel along a particular link is not a good of final consumption, but an intermediate good which combines with others to provide the final good: a journey. Practical problems arise with the assumption that demand for a link is an economic good. The demand relations are unstable and difficult to determine. In addition the problem of "new goods" is raised unnecessarily where links are present in only one network. (It is clear that the demand for any length of road other than a link is not usable, on account of the possibility of leaving or joining it in the middle.) In addition, the selection of journeys by mode as the appropriate economic good follows the assumption implicit in most engineering practice. This normally routes all journeys on a given mode by the shortest or cheapest route, but does not make all journeys use the quickest mode or make all people travel to the most proximate destination. This presupposes perfect interchangeability between routes but not between modes or destinations.

METHOD III

The main weakness of method II is the assumption that the benefits arising from improved accessibility lie midway between their possible maximum and minimum. In economic terms this assumption is more or less equivalent to the assumption of a straight line demand curve. In general, however, demand curves tend to be convex towards the origin. Therefore the correct measure of benefit does not lie exactly midway between the two extremes. Method III attempts to determine more accurately where the value should be, by examining the shape of the demand curve implicit

in the models used.[5]

With the kind of travel behaviour generally encountered, if the maximum valuation is less than twice the minimum the error arising from the use of method II rather than method III will be less than 5 per cent. Cases where the required conditions do not hold are likely to occur where very large changes are being evaluated. It is useful if method II is being used to calculate the maximum and minimum valuations as a check on this kind of error.

In practice the value of method III is small where transport plans are concerned. Its use is limited to the comparatively rare cases described above. There are two reasons why I intend to devote some space to it. It raises a variety of interesting theoretical issues. In particular it shows the relation between economic evaluation on the one hand and the notions of accessibility and comfort and convenience on the other: hitherto these have usually been regarded as completely unrelated. It also provides insight into certain aspects of transport demand models. From the theoretical analysis we can show how method III may be used for the evaluation of land use plans. Where there has been any attempt to evaluate land use plans economically in the past, this has been done by using a total cost criterion. This is as erroneous as in the case of transport plans, and for the same reasons. We shall also see that method II is normally likely to give misleading results. We shall examine in detail one particular form of transport demand model and then discuss how this analysis may be generalised. In the remainder of the paper we shall assume that all unit costs of transport may be treated as if they were prices, and that there are no problems of perceived costs, etc. The treatment of all such issues is the same as for method II.

It has been shown elsewhere [4] that there is an analogue for Marshall's measure of consumer surplus when demand is the function of a number of prices which change simultaneously. This is the multi-dimensional equivalent of the area under a demand curve. It was further shown that under certain circumstances it was possible to define exactly a utility function which was in money units, and that this utility function differed from the Marshall measure of surplus by exactly the change in the level of expenditure. In addition these were necessary and sufficient conditions to ensure that the multi-dimensional surplus measure was unambiguous. In general the measure is ambiguous, because its value depends on the path of integration (see for example Hotelling [7]). The demand model we have selected for analysis fulfils these conditions. The multi-dimensional measure of surplus is given by S.

---

[5]The conditions under which method III should be used rather than method II are as follows. The maximum and minimum valuations described above are defined thus:

If
$$ {}_kC1_{ij} > {}_kC2_{ij} $$
$$ {}_kW_{ij} = \max ({}_kT1_{ij}, {}_kT2_{ij}) $$
$$ {}_kV_{ij} = \min ({}_kT1_{ij}, {}_kT2_{ij}) $$

If
$$ {}_kC1_{ij} \leqslant {}_kC2_{ij} $$
$$ {}_kW_{ij} = \min ({}_kT1_{ij}, {}_kT2_{ij}) $$
$$ {}_kV_{ij} = \max ({}_kT1_{ij}, {}_kT2_{ij}) $$
$$ \text{Maximum} = \Sigma_k W_{ij} (C1_{ij} - C2_{ij}) $$
$$ \text{Minimum} = \Sigma_k V_{ij} (C1_{ij} - C2_{ij}) $$

$$S = - \Sigma_i \int_{p^1{}_i}^{p^2{}_i} x_i dp_i \tag{35}$$

where x_i is the quantity of good i and p_i is its price, and

$$U = S + C \tag{36}$$

where C is expenditure on all goods and U is utility measured in money terms. The conditions referred to above are

$$\frac{\partial x_i}{\partial p_j} = \frac{\partial x_j}{\partial p_i} \quad \text{for all } i, j \tag{37}$$

If we assume the Slutsky conditions on the integrability of the utility function, *i.e.*

$$\frac{\partial x_i}{\partial p_j} + x_j \frac{\partial x_i}{\partial C} = \frac{\partial x_j}{\partial p_i} + x_i \frac{\partial x_j}{\partial C} \tag{38}$$

we may see that condition (37) is equivalent to

$$\frac{\partial x_i}{\partial C} \Big/ x_i = \frac{\partial x_j}{\partial C} \Big/ x_j \quad \text{for all } i,j \tag{39}$$

or in words that the income elasticity of demand for all goods of which the prices are changed must be equal.

The demand model we wish to use is as follows. It is a special case of a gravity model which is very popular in transport analysis.

$$T_{ij} = \frac{O_i \, A_j \, e^{-\lambda C ij}}{\Sigma_k \, A_k \, e^{-\lambda C ik}} \tag{40}$$

We suppose for the moment that there exists only one mode of travel, and that demand for a given journey is affected by the cost of that journey and of other possible journeys and of certain characteristics of origin and destination, of which we shall have more to say later. It will be observed that

$$\Sigma_j \, T_{ij} = O_i \tag{41}$$

so that O_i represents the number of trips originating from zone i.

In the analysis that follows we shall be assuming that all persons have the same utility functions apart from additive constants. This assumption enables us to talk about a single utility function for each original zone, since there is no problem in aggregating identical utility functions. Any divergence from this assumption would involve the consideration of distributional issues, which I wish to avoid.[6]

---

[6]We observe that the model described in equation (40) fulfils condition (37)

$$\frac{\partial T_{ij}}{\partial C_{im}} = \frac{O_i A_j \, e^{-\lambda c_{ij}} \, A_m \, e^{-\lambda c_{im}}}{(\Sigma_k A_k \, e^{-\lambda C ik})^2} \tag{42}$$

$$\therefore \quad \frac{\partial T_{ij}}{\partial C_{im}} = \frac{\partial T_{im}}{\partial C_{ij}} \tag{43}$$

which is the same as condition (37). Therefore the Marshall measure of surplus is unique, and equation (36) may be used to derive money measure of utility.

As we have observed, evaluation always involves the comparison of two situations. In the comparison of two transport plans in the context of a single land use plan, these may be defined completely in terms of a matrix of interzonal times$//C_{ij}//$. Thus we compare the transport plans defined by the two matrices $//C^1{}_{ij}//$ and $//C^2{}_{ij}//$ and measure the surplus and utility change between them.

Thus we may define surplus and utility functions[7]

$$U = \frac{1}{\lambda} \Sigma_i \Sigma_j \left(T_{ij} \log \frac{T_{ij}}{G_i A_j} - T_{ij} \right) \tag{44}$$

$$S = \frac{1}{\lambda} \Sigma_i O_i \log ({}_k A_k e^{-\lambda C ik}) \tag{45}$$

$$S = \frac{1}{\lambda} \Sigma_i \Sigma_j \left[\left(T_{ij} \log \frac{T_{ij}}{G_i A_j} - T_{ij} \right) + T_{ij} C_{ij} \right] \tag{46}$$

In order to give some meaning to these functions we shall need first to provide some interpretation for the terms A_i, G_i. These I shall refer to as the inherent attractive and generative factors of the zones with which they are associated. These factors are the same whatever the transport network (as against the trip ends such as O_i which can vary with the quality of the network). If, for example, we are considering the demand model for journeys to work, G_i will refer to the desirability of the zone for residence, ignoring only its accessibility, but considering such factors as density of development, age of housing, etc. Similarly A_j will refer to all aspects of zone j as a place to work, such as the wage level and the type of employment offered, except its nearness to residences. We shall see later that it is not normally necessary to be able to predict these qualities directly, because they can be deduced from predicted origins and destinations and a network with which they are associated. We can now look at the utility function in equation (44). We can see that this is a measure of the degree to which trips are associated with desirable origins and destinations. Similarly the surplus function (45) is a measure of the correlation between number of trips originating from a zone and the accessibility of that zone to desirable destinations. Thus we can see that there is a close connection between economic analysis and the measurement of accessibility, and that the two do not need to be assessed separately.

It may appear that the term G_i was introduced somewhat arbitrarily into the argument. We shall now try to show how this came about. It was observed that the demand model (40) was such that the total number of trips originating in any zone was not affected by the quality of the network. In order to derive this model again, it is necessary to impose this condition as a special constraint. Once this constraint has been imposed, the same trip pattern will result whatever the value of G_i. Thus it is impossible to deduce any value of G_i from the demand model.

We examine this further in Appendix B by describing a family of models which may all be generated by the same utility function by applying different constraints.[8]

---

[7]For derivation see Appendix A.

[8]Since the utility function is expressed in terms of money, the marginal utility of money term which usually appears in such derivations is equal to unity. Thus it is legitimate to maximise the surplus function directly.

It may be seen that in Appenix B we have derived from the same utility function not only the model originally described (40), but models where there are fewer and more constraints. However, as more of the constraints are introduced fewer of the original features of the utility function are reflected in the final mode, or, to put it another way, the more constraints there are in the model the greater the range of possible utility functions which could have generated it.[9] This means that where constraints are supposed to operate in the model it is not possible to deduce the G_i and A_j parameters of the utility function without some further information on how these constraints affect demand. I shall examine the extreme case where both origin and destination constraints apply, since this illustrates the principles most clearly. We may assume that there exists at least one transport network for which the constraints are not binding, and that the distribution predicted arises entirely from travellers' responses to costs. The same demand would have been predicted by a model of the unconstrained type (82) where G_i and A_j are related to the fully constrained gravity model by the formulae

$$G_i = a_i O_i$$
$$A_j = b_j D_j \qquad (47)$$

The values on the right-hand side of the equations are usually obtained when the demand model is used for prediction. When two transport networks are compared we may make one of three assumptions:

(a) the constraints are not binding in either;
(b) the constraints are binding in one of the networks only;
(c) the constraints are binding in both networks.

We shall suppose that it is always possible to find a network in which the constraints are not binding, and indeed it will very often be one of the networks tested. (This may be supposed from the fact that the main reason why constraints are imposed is a belief that trip ends may be more reliably predicted than origin-destination trips. Thus trip ends are usually constrained to the number of trips which are expected to occur in one of the networks planned.) Thus case (c) may be reduced to two comparisons of the type (b). In these cases we estimate the parameters G_i and A_j, using equation (47) as applied to the situation where the constraints are slack. We may then use surplus function (45) for evaluation.[10] In case (a), if the constraints are slack in both networks, this must arise from a change in the parameters of the utility

[9]It is interesting to observe that in the case of the fully constrained model one of these is Wilson's entropy function [14]:

$$\Sigma = \Sigma_i \Sigma_j T_{ij} \log T_{ij}$$

Wilson generates the same model as is produced in case *C* by maximising entropy subject to the same constraints. It is therefore not very surprising that his entropy function should be one of the admissible utility functions. The resemblance seems to be purely formal, and any attempt to interpret entropy as utility or vice versa is likely to be futile.

[10]A practical point should be noted here. In applying method III it is best to use formulae which are expressed in terms of costs rather than in terms of trips, since trips are often predicted to be zero when they are rounded down. This can often lead to inaccurate and absurd results.

function. This may appear paradoxical, since it is normally supposed that a change in a utility function reflects a change in taste, which is inscrutable to welfare economists. In fact the use of the term "utility function" in this context is not strictly accurate. A utility function cannot strictly be defined over zones as we have done, but must be defined over the attributes of those zones which serve desired ends. The shorthand is acceptable so long as those attributes do not change, but in case (a) we must suppose that the change in the network has brought about exactly that change in zonal attributes which will keep the origins and destinations the same. We express this loosely as a change in the utility function. It is thus legitimate to compare the consumer's surplus arising in the two different situations calculated from the different surplus functions. It should be remembered that, although the utility functions for these models are of the same form, the surplus functions are not, since the changes in expenditure are different with different constraints. Thus for the unconstrained model (82) the surplus function is given by

$$S = \frac{1}{\lambda} G_i \, A_{je}{}^{-\lambda} C_{ij} \tag{48}$$

while in the fully constrained case (91) there is no analytical form of the surplus function. Thus, once the parameters of the utility function have been calculated, one must then return to the form of the prediction model to determine the form of the surplus function. For the fully constrained model the form of surplus function for the origin-constrained model is the most appropriate. Where the results of constrained models are being evaluated it has been found in practice that it makes virtually no difference whether one assumes that any particular comparison is of form (a), (b) or (c).

LAND USE PLAN EVALUATION

Accessibility is also an important element in land use planning. Accessibility may be improved either by improving a transport system or by changing the physical locations of land uses. In terms of the accessibility measure developed here, there is no practical difference between these two methods of changing accessibility: it will measure them both, whether they occur separately or together. We may consider a study area consisting of two towns, A and B. The accessibility of the system may be improved either by improving the road between the two towns or by moving town B towards town A. If the cost of travel between the towns will be the same whichever of these policies is adopted, the accessibility change will be the same, and under method III the same economic benefit will be shown as far as this aspect is concerned. Thus we may use the same benefit formulae appropriately recast to evaluate land use changes. Land use changes may be represented either by a change in the constraints on trip ends or by changes in inherent zonal characteristics, depending on the land use prediction model. We recast equation (45) to show a composite of these

$$S = \frac{1}{\lambda} [O^2{}_i \log (\Sigma_k \, A^2{}_{ke}{}^{-\lambda} C^2 ik) - O^1{}_i \log (\Sigma_k \, A^1{}_{ke}{}^{-\lambda} C^1 ik)] \tag{49}$$

where the A_j are deduced in the way already described using equation (47).[11]

This result is of particular importance because method II cannot be applied, even as an approximation, to the evaluation of land use plans. This may be seen most clearly by returning to our example of the two towns A and B. The normal practice for studying the land use and transport changes would be to divide the study area into three zones for analysis, one for A, one for the original location of B, and one for the final location of B. Thus we can see that if we apply method II (formula (4)) to this situation, we shall obtain a zero valuation of the land use changes, because none of the interzonal costs will have changed. In this simple case it would have been possible to employ a movable zone for town B, so that the change in accessibility would have been reflected in the interzonal times, but in most cases this is not possible. Thus it is usually necessary to use method III.

COMFORT AND CONVENIENCE OF MODES

Hitherto we have assumed the existence of only one mode of transport. If we have a demand model similar to (40) but allowing also for modes:

$$ {}_kT_{ij} = \frac{O_i A_j h_{ke}{}^{-\lambda k C ij}}{\Sigma_{mn} A_m h_{ne}{}^{-\lambda n C im}} \tag{57}$$

Then we obtain utility functions

$$ U = \frac{1}{\lambda} \Sigma_{ijk} \left({}_kT_{ij} \log \frac{{}_kT_{ij}}{G_i A_j h_k} - {}_kT_{ij} \right) \tag{58}$$

and, if we interpret the terms h_m as being factors relating to the comfort and convenience of the mode m, we may see that the utility function is a measure of the

---

[11]We now give a formal demonstration of the legitimacy of this process in the simple case where the origin constraints are unchanged. The argument may easily be generalised along the same lines to give the result shown in equation (49).

$$ T_{ij} = \frac{O_i A_j e^{-\lambda C_{ij}}}{{}_kA_{ke}{}^{-\lambda C_{ik}}} \tag{50}$$

$$ C_{ij} = \frac{-1}{\lambda} \log \left(\frac{T_{ij}}{A_j} \frac{\Sigma_{kj}\ A_k e^{-\lambda C ik}}{O_i - T_{ij}} \right) \tag{51}$$

$$ \frac{\partial C_{ij}}{\partial A_j} = \frac{-1}{\lambda A_j} \tag{52}$$

$$ \Sigma_j \int_{C1_{ij}}^{C2_{ij}} T_{ij}\, dc_{ij} = \int_{A1_j}^{A2_j} \Sigma_j T_{ij} \frac{\partial c_{ij}}{\partial A_i}\, dA_j \tag{53}$$

$$ = \frac{1}{\lambda} \Sigma_j \int \left(\frac{O_i e^{-\lambda C ij}\, dA_j}{\Sigma_k A_k e^{-\lambda C ik}} \right) \tag{54}$$

$$ = \frac{O_i}{\lambda} \Sigma_j \left[\frac{\delta_{je}{}^{-\lambda C ij}\, d\eta}{\Sigma_k A1_{ke}{}^{-\lambda C ik} + \eta \Sigma_k \delta_k e^{-\lambda C ik}} \right]_{\eta = o}^{1} \tag{55}$$

$$ = \frac{O_i}{\lambda} \left[\log \left(\Sigma_k A1_k e^{-\lambda C ik} \right) - \log \left(\Sigma_k A2_k e^{-\lambda C ik} \right) \right] \tag{56}$$

extent to which people travel by comfortable modes. One special difficulty associated with model benefits and their measurement in method III is that the assumptions in equation (39) are much less likely to hold, because the income elasticity of demand for some modes will be positive and for others negative. Thus the requirements for a unique money utility function are not met. In practice this is not a serious problem.

OTHER MODELS

We shall now try to relate this analysis of a particular form of demand model to other forms. We consider first other forms of gravity models. In some cases it is possible to develop corresponding measures on exactly the same lines as we have shown. They will then have similar forms to those described. In other cases they do not meet condition (37), and therefore the process is not exact. Nonetheless, as has been indicated in [4], the errors involved in taking a straight line integration path are not likely to be large. Another problem is that often these models do not yield analytic integrals, in which case interpretation along the lines we have indicated is not so straightforward. There is no reason, however, why the principle we have described should not apply in these cases.

Non-gravity model formulations of transport demand models fall into three categories:

1. Those where travel is completely unresponsive to time or cost – *e.g.*, growth factor models and Fratar models.
2. Those where cost or time is minimised subject to certain constraints, *e.g.*, Hitchcock.
3. Those where time or cost is a factor – *e.g.*, Kraft-SARC and intervening opportunities models (and gravity models also).

In the first and second cases reduction in cost of resources is the appropriate criterion. In the first case we have a pure example of method I range of application. In the second case, although the pattern of demand changes in response to cost changes, we know that the "utility function" is negative cost. In other words, the pairing between origin and destination is unimportant as far as the model is concerned, and the total origin or destination in any zone is fixed. Thus the only factor which can affect utility is cost. In both these cases, accessibility in the sense we have used it is of no value. In the third case method II or III will be applicable as appropriate. The majority of models in this class will yield properties like the gravity model. The intervening opportunities model presents certain special problems because its demand functions are unstable and of a strange shape. It is difficult to give it an intelligible economic interpretation.

METHOD III AS APPLIED TO LAND USE AND TRANSPORT EVALUATION

We shall recapitulate by describing how method III might be applied in land use and transport comparisons. We shall assume for the purposes of exposition that a land use plan consists only of the location of homes and workplaces. When land use plans are compared, it should be ensured that they are catering for the same popula-

tion: it should be possible to define a study area accordingly. Normally the information we shall be given will be the number of jobs and the population for each zone. Depending on the form of model used, we may also be supplied with a transport network with which it is associated. (If this is not supplied it will be necessary to define or invent the network for which the constraints are slack.) The only other information required is the form and parameters of the demand function (*e.g.*, in the gravity model). We may then use equations of the form (47) and (91) to find the parameters of the utility function. Taking one land use transport combination as a base, it will then be possible to compare these, using formulae such as (49). The surplus measures thus obtained may then be put into the appropriate context of a full cost-benefit analysis. No additional allowance will be necessary for accessibility or comfort and convenience.

APPENDIX A

The surplus measure may be calculated on the basis of equation (36).

$$\Delta S = \Sigma_i \Sigma_j \int_{C1_{ij}}^{C2_{ij}} \frac{O_i A_j \, e^{-\lambda} C_{ij}}{\Sigma_k A_k \, e^{-\lambda} C_{ik}} \, dC_{ij} \tag{59}$$

We define S_i to be the surplus change for trips originating from zone i. We also define a η such that

$$C_{ij} = C1_{ij} + \eta \delta_{ij} \tag{60}$$

$$\delta_{ij} = C2_{ij} - C1_{ij} \tag{61}$$

In words, we intend to assume a linear path of integration, *i.e.*, all the trips originating from the same zone will make the same proportion of their price changes at the same time. For example, there will be a point at which all costs will be half-way between their original and final values.

$$T_{ij} = \frac{O_i A_j \, e^{-\lambda C1_{ij}} \, e^{-\lambda \eta \delta_{ij}}}{\Sigma_k A_k \, e^{-\lambda C1_{ik}} \, e^{-\delta \eta_{ik}}} \tag{62}$$

$$S_i = - \Sigma_j \int_{C1_{ij}}^{C1_{ij}} T_{ij} dC_{ij} \tag{63}$$

$$= - \Sigma_j \int_0^1 T_{ij} \frac{\partial C_{ij}}{\partial \eta} \delta \eta \tag{64}$$

$$= - \int_0^1 \frac{O_i \, (\Sigma_k A_k \, e^{-\lambda C1_{ik}} \, \delta_{ik} \, e^{-\lambda \eta \delta_{ik}})}{\Sigma_k A_k \, e^{-\lambda C1_{ik}} \, e^{-\lambda \eta \delta_{ik}}} \delta \eta \tag{65}$$

$$\Delta S_i = - \frac{O_i}{\lambda} \left[\log \left(\Sigma_k A_k \, e^{-\lambda C1_{ik}} \, e^{-\lambda \eta \delta_{ik}} \right) \right]_{\eta=0}^{1} \tag{66}$$

$$= - \frac{O_i}{\lambda} \left[\log \left(\Sigma_k A_k \, e^{-\lambda C_{ik}} \right) \right]_{C1_{ik}}^{C2_{ij}} \tag{67}$$

$$= \frac{Oi}{\lambda} \log \left(\frac{\Sigma_k A_k \, e^{-\lambda C2_{ik}}}{\Sigma_k A_k \, e^{-\lambda C1_{ik}}} \right) \tag{68}$$

I shall now derive the same formula in another way. In the derivation (62) to (68) a straight line path of integration was used. Another possible path of integration is seriatim price changes as recommended in Beesley and Walters [2]. In this path one price is changed at a time, while the others are held constant at their original or final values, depending on whether they have been changed already.

In the general case this will not be expected to yield the same results as the straight line case; but here, since this model met condition (37), it will. Where they diverge there are reasons for preferring a straight line path as described in [4]. The seriatim path has the advantage of visual intelligibility. We examine the case of two possible destinations, to both of which the costs of journeys differ between the plans being compared. The case is represented in figures 6 and 7. Beesley and Walters recommend the use of the two areas shaded as a measure of surplus, following the path of integration described.

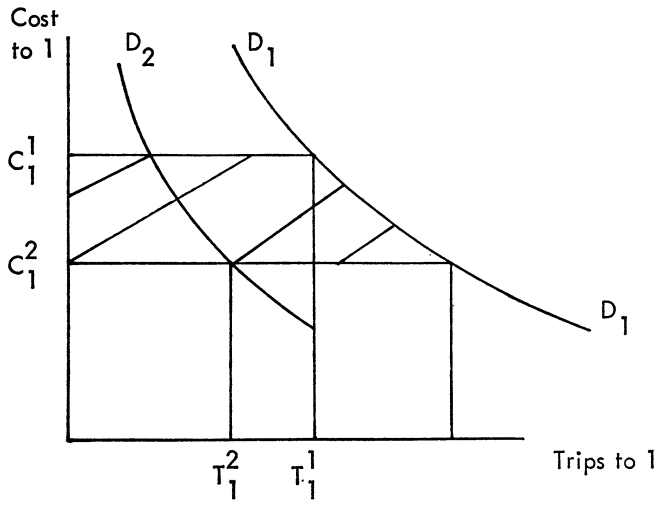

FIGURE 6

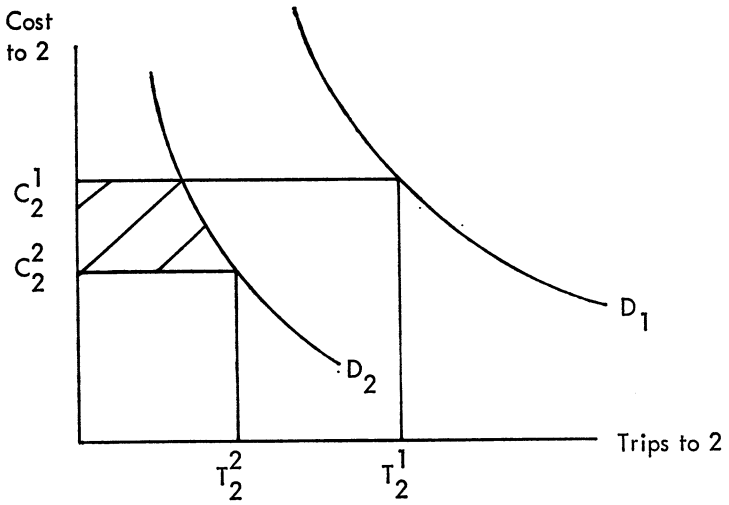

FIGURE 7

The diagrams show how close is the analogy with the two-dimensional area under the demand curve. On the other hand, it is not correct to measure the change in all the areas under all the demand curves.

We may now use this path of integration to measure the surplus change:

$$\Delta S_i = \Sigma_j \int_{C1_{ij}}^{C_{ij}} T_{ij} dC_{ij} \tag{69}$$

$$= \Sigma_j O_i \int \frac{A_{je}{}^{-\lambda C_{ij}}\, dC_{ij}}{(\Sigma_{k<j} A_{ke}{}^{-\lambda C2_{ik}} + \Sigma_{k>j} A_{ke}{}^{-\lambda C1_{ik}} + A_{je}{}^{-\lambda C_{ij}})} \tag{70}$$

$$= \frac{O_i}{\lambda} \Sigma_j \left[\log \left(\Sigma_{k<j} A_{ke}\, C2_{ik} + \Sigma_{k>j} A_{ke}{}^{-\lambda C1_{ik}} + A_{je}{}^{-\lambda C_{ij}} \right) \right]_{C1_{ij}}^{C2_{ij}} \tag{71}$$

We may see that all terms except the first and last cancel out, so that

$$= \frac{O_i}{\lambda} \log \left(\frac{\Sigma_k A_{ke}{}^{-\lambda C1_{ik}}}{\Sigma_k A_{ke}{}^{-\lambda C2_{ik}}} \right) \tag{72}$$

which is the same as the formula (68).

Thus the surplus change for the whole study area is given by

$$\Delta S = \frac{1}{\lambda} \Sigma_i O_i \log \left(\frac{\Sigma_k A_{ke} - \lambda C1_{ik}}{\Sigma_k A_{ke} - \lambda C2_{ik}} \right) \tag{73}$$

We may further derive from equations (36), (40) and (73) a utility function thus:

$$\Delta U = \Delta S + \Delta C \tag{74}$$

$$= \Sigma_i \frac{O_i}{\lambda} \log \left(\frac{\Sigma_k A_{ke} - \lambda C1_{ik}}{\Sigma_k A_{ke} - \lambda C2_{ik}} \right) + \Sigma_i \Sigma_j \left(T2_{ij}\, C2_{ij} - T1_{ij}\, C1_{ij} \right) \tag{75}$$

$$C_{ij} = \frac{1}{\lambda} \left[\log \frac{T_{ij}}{A_i O_i} + \log \left(\Sigma_k A_{ke} - \lambda C_{ik} \right) \right] \tag{76}$$

$$\therefore \Sigma_i \Sigma_j T_{ij} C_{ij} = \frac{1}{\lambda} \Sigma_i \left[\Sigma_j T_j \log \frac{T_{ij}}{A_j O_i} - \frac{O_i}{\lambda} \log \left(\Sigma_k A_{ke} - \lambda C_{ik} \right) \right] \tag{77}$$

$$\Delta U = \frac{1}{\lambda} \Sigma_i \Sigma_j \left[T1_{ij} \log \frac{T1_{ij}}{A_j O_i} - \left(T2_{ij} \log \frac{T2_{ij}}{A_j O_i} \right) \right] \tag{78}$$

Bearing in mind the fact that

$$\Sigma_i T_{ij} = O_i \tag{79}$$

we may rewrite (78) introducing an extra term G_i, which will be described below:

$$\Delta U = \frac{1}{\lambda} \Sigma_i \Sigma_j \left(T_{ij} \log \frac{T_{ij}}{G_i A_j} - T_{ij} \right)_{T1_{ij}}^{T2_{ij}} \tag{80}$$

APPENDIX B

We have three cases: (A) no constraints, (B) constraints on the number of trips originating from each zone, (C) constraints on both the number of trips originating and the number having destinations in a zone.

Case A: No Constraints

Maximising the surplus function (46), we have

$$\frac{1}{\lambda}\log\frac{T_{ij}}{G_iA_j} = C_{ij} \tag{81}$$

or

$$T_{ij} = G_iA_je^{-\lambda Cij} \tag{82}$$

This we call the unconstrained model.

Case B: Origin Constraints

Maximising the surplus function (46) subject to the condition

$$\Sigma_j T_{ij} = O_i \text{ for all } i \tag{83}$$

we have the Lagrangean

$$L = \frac{1}{\lambda}\Sigma_j\left[T_{ij}\log\frac{T_{ij}}{G_i A_j} - \Sigma_i T_{ij}C_{ij} + \alpha_i(O_i - \Sigma_j T_{ij})\right] \tag{84}$$

$$T_{ij} = G_i A_j e^{-\lambda(C_{ij}+\alpha_i)} \tag{85}$$

$$\Sigma_j T_{ij} = e^{-\lambda\alpha_i} G_i(\Sigma_j A_j e^{-\lambda Cij}) \tag{86}$$

$$\therefore T_{ij} = \frac{O_i A_j e^{-\lambda C_{ij}}}{\Sigma_k A_k e^{-\lambda C_{ik}}} \tag{87}$$

Destination constraints have similar effects. It may be observed that we have recovered the original model (6).

Case C: Origin and Destination Constraints

Maximise the surplus function (46) subject to the constraints

$$\begin{aligned}\Sigma_j T_{ij} &= O_i\\ \Sigma_j T_{ij} &= D_j\end{aligned} \tag{88}$$

We have the Lagrangean

$$L = \frac{1}{\lambda}\Sigma\left(T_{ij}\log\frac{T_{ij}}{G_i A_j} - T_{ij}\right) - \Sigma T_{ij}C_{ij} - \alpha_i(O_i - \Sigma_j T_{ij}) - \beta_j(D_j - \Sigma_i T_{ij}) \tag{89}$$

$$T_{ij} = e^{-\lambda(Cij+\alpha j+\beta j)} \tag{90}$$

Solving for the constraints we have the fully constrained distribution model

$$\begin{aligned}T_{ij} &= O_i D_j a_i b_j e^{-\lambda C_{ij}}\\ a_i &= (\Sigma_j D_j)/\Sigma_j b_j D_j e^{-\lambda C_{ij}}\\ b_j &= (\Sigma_i O_i)/\Sigma_i a_i O_i e^{-\lambda C_{ij}}\end{aligned} \tag{91}$$

REFERENCES

[1] American Association of State Highway Officials: *Economic Evaluation of Highway Projects.* Washington, 1961.

[2] Beesley, M. E., and A. A. Walters: "Some Problems in the Evaluation of Urban Transport Investment", *Applied Economics*, Vol. I, No. 4, 1970.

[3] Chicago Area Transportation Study: *Final Report Vol.* 3. Chicago, 1962.

[4] Foster and Neuburger: *The Foundations of Cost Benefit Analysis.* Not yet published.

[5] Friedlander. *The Interstate Highway System.* Amsterdam, 1965.

[6] Harvey: *A Method of Network Evaluation using the Output of the Traffic Assignment Procedure.* Highway Research Record No. 238.

[7] Hotelling: "The General Welfare in Relation to Taxation and of Railway and Utility Rates", *Econometrica*, July 1938.

[8] Martin and Wohl: *Evaluation of Mutually Exclusive Design Projects*. Highway Research Board Special Report No. 92, Washington, 1968.

[9] Merseyside Area Land Use and Transportation Study. *Technical Report no. 17*. Liverpool, 1970.

[10] Mohring and Harwitz: *Highway Benefits; An Analytical Fromaework*. Evanston, 1962.

[11] Quarmby: "Travel Mode for Journey to Work." *Journal of Transport Economics and Policy*, September 1967.

[12] Road Research Laboratory: *Technical Report no. 74*. Crowthorne, 1968.

[13] Tressider, Meyer, Burrell, Powell: *Proceedings of the Institution of Civil Engineers*. March 1968.

[14] Wilson: "Statistical Theory of Spatial Distribution Models", *Transportation Research*, November 1968.

[15] Winch: *The Economics of Highway Planning*. Toronto, 1963.

Massachusetts Institute of Technology