The Use of Entropy Maximising Models, in the Theory of Trip Distribution, Mode Split and Route Split

Author(s): A. G. Wilson

THE USE OF ENTROPY MAXIMISING MODELS

in the Theory of Trip Distribution, Mode Split and Route Split

By A. G. Wilson

1. BACKGROUND

In transport planning, a model is often assembled to simulate flows on the transport system. The model usually consists of four submodels concerned respectively with trip generation, distribution, modal split and assignment. This paper is not concerned with generation, and assumes that estimates of trip ends are known; some of the results have implications for assignment, but the paper is mainly concerned with distribution and modal split. A new method of constructing distribution and modal split models, which can be called an entropy maximising, or probability maximising, method, was discussed in a recent paper by the author (Wilson, 1967). Several alternative distribution models were derived in that paper, and some implications were drawn for modal split. The purpose of this paper is to systematically consider extensions of the earlier results, to review a number of models, to discuss the underlying hypotheses for each, and to formulate a programme of empirical testing.

The new method is introduced in section 2, and alternative distribution and modal split models are derived in sections 3 and 4.

In section 5 the underlying behavioural assumptions of these models are examined, and two broad questions emerge. Firstly, is the behavioural parameter which governs the decision on where and how far to travel (trip distribution) the same as that which governs modal choice? The second question is concerned with the costs of travel and how they are perceived: if the costs which are actually measured are the costs of travel by particular modes, is it possible to construct from these a single representative cost of travel which the traveller uses when deciding where to go? In other words, if travel cost is an impedance, how can a composite impedance be constructed out of modal impedances, as required by certain forms of distribution models which have been presented? The answer to the first of these questions suggests a possible alternative form of modal choice model, which is explored in section 6. The answers to the second question are explored in section 7.

It then emerges that a further series of questions can be asked about the nature of a mode and the costs of travel by different modes. It is suggested in section 8 that a "mode" of travel between a pair of points is in fact a collection of routes (alternative road routes, alternative public transport routes, etc.), and that the travel costs which are directly observed are route costs. This means that we must explore the question of route split in relation to mode split, and construct modal costs as composite measures out of route costs. At this stage, really fascinating questions emerge. It appears that the form of model which should be used depends on whether travellers perceive modes in some primary sense and then choose routes, or vice versa. These

108

alternative models give identical answers if, and only if, modal costs are constructed out of route costs on the basis of a particular composite impedance function. It is argued that a programme of empirical testing, which is set out, is needed for these alternative models. The results of these explorations are discussed in section 9 and some conclusions are drawn in section 10.

2. INTRODUCTION TO ENTROPY MAXIMISING MODELS

The method is most easily illustrated by examining a single trip purpose, say the journey to work. In the first instance, let us further assume that there is only one mode of transport and one type of traveller. Suppose the region is divided into zones and that T_{ij} is the number of trips between zones i and j, c_{ij} is the cost (a generalised cost, say, for example, a linear sum of fares or out-of-pocket costs, travel time, and excess travel time) of travelling between i and j, O_i is the total number of trip origins at i and D_j is the total number of trip destinations at j. (It may, of course, be more convenient to define trip ends as productions and attractions, but the methodology is the same). Then, if a gravity model is used to estimate the number of trips between i and j, T_{ij}, in terms of the other variables, then

$$T_{ij} = A_i B_j O_i D_j f(c_{ij}) \tag{1}$$

where f is some decreasing function cost, and

$$A_i = \left[\sum_j B_j D_j f(c_{ij}) \right]^{-1} \tag{2}$$

$$B_j = \left[\sum_i A_i O_i f(c_{ij}) \right]^{-1} \tag{3}$$

in order that the equations

$$\sum_j T_{ij} = O_i \tag{4}$$

and

$$\sum_i T_{ij} = D_j \tag{5}$$

should be satisfied; (2) and (3) are usually solved by some iterative procedure. In the following exposition, it will also be assumed that another constraint equation is satisfied in addition to (4) and (5):

$$\sum_i \sum_j T_{ij} c_{ij} = C \tag{6}$$

This simply says that the total amount spent on these trips in the region, and at the given point in time, is a fixed amount C.

   The basic assumption of the new method is that the probability of the distribution T_{ij} occurring is proportional to the number of states of the system which give rise to this particular distribution, and which satisfy the constraints (4) – (6). Suppose

$$T = \sum_i O_i = \sum_j D_j \tag{7}$$

is the total number of trips. The number of distinct arrangements of individuals which give rise to the distribution T_{ij} is defined as being $w(T_{ij})$ and is the number

109

of ways in which T_{11} can be selected from T, T_{12} from $T - T_{11}$, ... and so on, and so

$$w(T_{ij}) = \frac{T!}{T_{11}! \, (T - T_{11})!} \cdot \frac{(T - T_{11})!}{T_{12}! \, (T - T_{11} - T_{12})!} \cdots = \frac{T!}{\underset{ij}{\pi} \, T_{ij}!} \quad (8)$$

The total number of possible states is then

$$W = \Sigma w(T_{ij}) \qquad\qquad (9)$$

where the summation is over all the distributions T_{ij} which satisfy the constraints (4) – (6). However, the maximum value of $w(T_{ij})$ turns out to dominate the other terms of the sum to such an extent that the distribution T_{ij} which gives rise to this maximum is overwhelmingly the most probable distribution. (And hence the method is a probability maximising method). This maximum will now be obtained.

To obtain the set of T_{ij}'s which maximises $w(T_{ij})$ subject to the constraints in (4) – (6), the function M has to be maximised where

$$M = \log w + \Sigma_i \lambda_i^{(1)} \, (O_i - \Sigma_j T_{ij}) +$$
$$\Sigma_j \lambda_j^{(2)} \, (D_j - \Sigma_i T_{ij}) + \beta(C - \Sigma_i \Sigma_j T_{ij}) \qquad (10)$$

and where $\lambda_i^{(1)}$, $\lambda_j^{(2)}$, and β are Lagrangian multipliers. Note that it is more convenient to maximise $\log w$ rather than w, and then it is possible to use Stirling's approximation

$$\log N! = N \log N - N \qquad\qquad (11)$$

to estimate the factorial terms. The T_{ij}'s which maximise M, and which therefore constitute the most probable distribution of trips, are the solutions of

$$\frac{\partial M}{\partial T_{ij}} = 0 \qquad\qquad (12)$$

and the constraint equations (4) – (6). Using Stirling's approximation, (11), note that

$$\frac{\partial \log N!}{\partial N} = \log N \qquad\qquad (13)$$

and so

$$\frac{\partial M}{\partial T_{ij}} = - \log T_{ij} - \lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij} \qquad (14)$$

And this vanishes when

$$T_{ij} = \exp \left[- \lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij} \right] \qquad (15)$$

Substitute in (4) and (5) to obtain $\lambda_i^{(1)}$ and $\lambda_j^{(2)}$:

$$\exp \left(- \lambda_i^{(1)} \right) = O_i / \left[\, \Sigma_j \exp \left(- \lambda_j^{(1)} - \beta c_{ij} \right) \right] \qquad (16)$$

$$\exp \left(- \lambda_j^{(2)} \right) = D_j / \left[\, \Sigma_i \exp \left(- \lambda_i^{(1)} - \beta c_{ij} \right) \right] \qquad (17)$$

To obtain the result in more familiar form, write

$$A_i = \exp \left(- \lambda_i^{(1)} \right) / O_i \qquad\qquad (18)$$

and

$$B_j = \exp \left(- \lambda_j^{(2)} \right) / D_j \qquad\qquad (19)$$

110

and then

$$T_{ij} = A_i B_j O_i D_j \exp\left(-\beta c_{ij}\right) \qquad (20)$$

where, using equations (16) – (19),

$$A_i = \left[\; \sum_j B_j D_j \exp\left(-\beta c_{ij}\right) \;\right]^{-1} \qquad (21)$$

$$B_j = \left[\; \sum_i A_i O_i \exp\left(-\beta c_{ij}\right) \;\right]^{-1} \qquad (22)$$

Thus the most probable distribution of trips is the same as the gravity model distribution discussed earlier, and defined in equations (1) to (3), and so this statistical derivation constitutes a new theoretical base for the gravity model. Note that C in the cost constraint equation (6) need not actually be known, as this equation is not in practice solved for β. This parameter would be found by the normal calibration methods. However, if C was known, then (6) could be solved numerically for β.

This statistical theory is effectively saying that, given total numbers of trip origins and destinations for each zone for a homogeneous person-trip-purpose category, given the costs of travelling between each pair of zones, and given that there is some fixed total expenditure on transport in the region at the given point in time, then there is a most probable distribution of trips between zones, and this distribution is the same as the one normally recognized as the gravity model distribution, though with the negative exponential function appearing as the preferred form of attenuation function. Students of statistical mechanics will recognise the method as a variation of the micro canonical ensemble method for analysing systems of particles, for example, the molecules of a gas. (*Cf.* Tolman, 1938). The validity of the method is discussed in detail in the previously cited paper by the author (Wilson, 1967).

It is important to note that the principle on which the derivation is based is quite general. The only assumption is that the probability of a distribution occurring is proportional to the number of states of the system which give rise to that distribution, subject to a number of constraints. It can easily be seen that if there were no constraints whatever all the T_{ij}'s would be equal, and would in fact have an equal share of the total number of trips. In other words, it is the constraints which have the effect of giving a distribution of trips other than the trivial one. Thus, to build a model, it is necessary to define a set of variables which defines the system (in this case, the T_{ij}'s), to express hypotheses by writing down constraint equations on the T_{ij}'s (to restrict behaviour in the way it is known not to be random), and then to obtain the maximum probability distribution subject to these constraints.

It remains to explain why the method can be usefully called an entropy maximising method. The log of the quantity w in equation (8), which we maximise subject to constraints, is often defined in statistical mechanics as entropy. What is in some ways more important for our general understanding, however, is the fact that this quantity differs only slightly from the definition of entropy used in information theory (see, for example, Theil, 1967). It can be shown that the model building approach, developed here by analogy with methods in statistical mechanics, is equivalent to the information theory approach. The author was greatly helped in his understanding of this equivalence by a paper (Jaynes, 1957) on the equivalence of statistical mechanics and information theory. If we define

$$p_{ij} = T_{ij}/T \qquad (23)$$

then the information theorist's definition of entropy is (Shannon and Weaver, 1963)

$$H = - \sum_i \sum_j p_{ij} \log p_{ij} \qquad (24)$$

and it can easily be checked that maximising H subject to the constraints (4) to (6) gives the same answer as the approach of this paper. The information theory approach is instructive, and will be discussed in more detail in a further paper.

3.   GENERALISED DISTRIBUTION MODELS

It is now possible to make our model more realistic by introducing several person types and several transport modes. The person types to be identified here are those which have different sets of modes available to them. For example, car owners (this term is shorthand for the set of people who have cars available to them) have access to both car and public transport modes, while non-car-owners have access only to public transport modes. This categorisation, at least, seems essential. It may also be useful to bear in mind the possibility of dealing in this way with people in different income groups or social classes. If the car owner/non-car-owner division is not made, we are likely to be faced with models which allocate some non-car-owners to car trips, or else we are forced to distribute car-owner and non-car-owner trips separately and hence to forecast trip attractions for each group separately; and this is usually neither possible nor desirable.

There is a need, then, to extend our notation: let T_{ij}^{kn} denote the number of trips between i and j by mode k and person type n; let O_i^n be the number of trip origins at i generated by persons of type n; let c_{ij}^k be the cost of travelling from i to j by mode k; and let the other variables be defined as before. Let $M(n)$ be the set of modes available to type n people, and let $k \varepsilon M(n)$ denote one such mode; $\sum_{k \varepsilon M(n)}$ denotes summation over such modes. Note that we should really define a quantity $M_{ij}(n)$ as the modes available between i and j to type n people, but we shall allow the i and j subscripts to be understood on the assumption that usually all modes will be available between all pairs of zones. The subscripts can easily be reintroduced if necessary.

A further development of notation is that, when an index is replaced by an asterisk, this denotes summation over that index. For example,

$$T_{ij}^{*n} = \sum_{k \varepsilon M(n)} T_{ij}^{kn} \qquad (24)$$

is the total number of trips by n type people between i and j (that is, T_{ij}^{kn} summed over modes k).

We can now begin to write down constraint equations which express our hypotheses about the new situation. Firstly, we must make a hypothesis about trip generation. It is customary (see, for example, Wootton and Pick, 1967) to characterise home-based trip ends (productions) by person type and non-home-based trip

112

ends (attractions) by another zone characteristic such as land use (activity). To put this another way, we might expect O_i to be a function of n, but not D_j, as we make the assumption that different types of people, within a trip purpose, compete for the same attractions. This is why trip origins have been defined above as O_i^n and trip destinations as D_j. This seems useful and sensible, even when our only n categorisation is to car owners and non-car-owners. Then

$$\sum_j \sum_{k \varepsilon M(n)} T_{ij}^{kn} = O_i^n \qquad (25)$$

$$\sum_i \sum_n \sum_{k \varepsilon M(n)} T_{ij}^{kn} = D_j \qquad (26)$$

describe our trip end hypotheses. We might also hypothesise that different person types have different *per capita* expenditures on travel, and hence different travel patterns, and so we write a cost constraint for each person type n as

$$\sum_i \sum_j \sum_{k \varepsilon M(n)} T_{ij}^{kn} c_{ij}^k = c^{(n)} \qquad (27)$$

Now maximise $\log T! / \prod_{\substack{ij \\ k \varepsilon M(n)}} T_{ij}^{kn}!$ subject to (25) – (27) as constraints, and we get

$$T_{ij}^{kn} = A_i^n B_j O_i^n D_j \exp\left(-\beta^n c_{ij}^k\right) \qquad (28)$$

where

$$A_i^n = \left[\sum_j \sum_{k \varepsilon M(n)} B_j D_j \exp\left(-\beta^n c_{ij}^k\right) \right]^{-1} \qquad (29)$$

and

$$B_j = \left[\sum_i \sum_n \sum_{k \varepsilon M(n)} A_i^n O_i^n \exp\left(-\beta^n c_{ij}^k\right) \right]^{-1} \qquad (30)$$

Note that (28) is a linked set of gravity models for each k-n category. The linking is through the B_j's, which each involve all the k's and n's. It is possible to programme this model, and tests are being carried out at the Ministry of Transport in London. There are two ways in which this generalised distribution model can be aggregated: over k and over n.

If it is felt, for example, that one person type is sufficient to categorise a population (for example, when nearly everyone has a car available), we can aggregate over n to get this single category, and†

$$T_{ij}^{k*} = A_i^* B_j O_i^* D_j \exp\left(-\beta^* c_{ij}^k\right) \qquad (31)$$

$$A_i^* = \left[\sum_j \sum_k B_j D_j \exp\left(-\beta^* c_{ij}^k\right) \right]^{-1} \qquad (32)$$

$$B_j = \left[\sum_i \sum_k A_i^* O_i^* \exp\left(-\beta^* c_{ij}^k\right) \right]^{-1} \qquad (33)$$

---

†In these equations, of course the asterisk associated with A_i^* and β^* does not indicate any summation, but merely that new parameters are involved.

H

is the appropriate model. Note that if this model is in turn aggregated over k, we get the original gravity model represented in equations (20) – (22). It remains useful, however, to aggregate the general model (28) – (30) over k to represent a single mode situation, or simply a model for trip interchanges by person type:

$$T_{ij}^{*n} = A_i^n B_j O_i^n D_j \sum_{k \varepsilon M(n)} \exp\left(- \beta^n c_{ij}^k\right) \qquad (34)$$

$$A_i^n = \left[\sum_j B_j D_j \sum_{k \varepsilon M(n)} \exp\left(- \beta^n c_{ij}^k\right) \right]^{-1} \qquad (35)$$

$$B_j = \left[\sum_i \sum_n A_i^n O_i^n \sum_{k \varepsilon M(n)} \exp\left(- \beta^n c_{ij}^k\right) \right]^{-1} \qquad (36)$$

Notice that each k summation is over the modes available to persons of type n.

It is useful to derive one more model in this section. Suppose that, instead of having travel costs presented as an array c_{ij}^k, we have been given C_{ij}^n, the cost of getting from i to j as perceived by an n type person. C_{ij}^n is clearly a composite of the c_{ij}^k's for the k's available to each person type n, and we shall return to this later. Then using the variable T_{ij}^{*n} and maximising log $T! / \pi_{ijn} T_{ij}^{*n}!$ subject to

$$\sum_j T_{ij}^{*n} = O_i^n \qquad (37)$$

$$\sum_i \sum_n T_{ij}^{*n} = D_j \qquad (38)$$

$$\sum_i \sum_j T_{ij}^{*n} C_{ij}^n = C^{(n)} \qquad (39)$$

we get

$$T_{ij}^{*n} = A_i^n B_j O_i^n D_j \exp\left(- \beta^n C_{ij}^n\right) \qquad (40)$$

where

$$A_i^n = \left[\sum_j B_j D_j \exp\left(-\beta^n C_{ij}^n\right) \right]^{-1} \qquad (41)$$

and

$$B_j = \left[\sum_i \sum_n A_i^n O_i^n \exp\left(-\beta^n C_{ij}^n\right) \right]^{-1} \qquad (42)$$

To summarise: we have derived a generalised model represented by equations (28) to (30), two aggregated versions represented respectively by equations (31) to (33) and (34) to (36), and one alternative version of (34) to (36) computed directly and represented by (40) to (42). The implications of these models will be fully discussed in the following sections.

Finally, it should be emphasised that, given any set of assumptions, a maximum entropy model can be constructed by use of the rules outlined in section 2. It has been possible, for example, to find a set of assumptions which give rise to the intervening opportunities model (Wilson, 1967).

114

4. MODAL SPLIT IMPLICATIONS

Firstly, consider the model given by (31) to (33), which emphasises modes only. Sum (31) over k, and divide the result into (31) itself to give the modal split as

$$\frac{T_{ij}^{k*}}{T_{ij}^{**}} = \frac{\exp\left(-\beta^* c_{ij}^{k}\right)}{\sum_{k'} \exp\left(-\beta^* c_{ij}^{k'}\right)} \qquad (43)$$

This can be written

$$\frac{T_{ij}^{k*}}{T_{ij}^{**}} = \frac{1}{1 + \sum_{k' \neq k} \exp\left[-\beta^*(c_{ij}^{k'} - c_{ij}^{k})\right]} \qquad (44)$$

This is a modal split model which is implicit in our generalised distribution model. The equation (44), for two modes, represents something rather like a diversion curve, but emphasises cost differences rather than the more usual cost ratios. It can also deal with any number of modes. Further, it is of considerable interest that (43) has the same form as the modal split equation derived from an application of discriminant analysis (cf. Quarmby, 1967), provided that the generalised cost function can be identified with the statistician's discriminant function. This apparent identity suggests that discriminant analysis is a good way of estimating the weights in a generalised cost function of the form

$$c_{ij}^{k} = \sum_{r} a_r X_r \ (i, j, k) \qquad (45)$$

where the X_r's are variables like fares, travel time and excess time, as mentioned earlier.

We can now examine the modal split implications of the general model (28) to (30). By a similar procedure to that outlined above, we can get a modal split for each person type in the form

$$\frac{T_{ij}^{kn}}{T_{ij}^{*n}} = \frac{\exp\left(-\beta^n c_{ij}^{k}\right)}{\sum_{k \varepsilon M(n)} \exp\left(-\beta^n c_{ij}^{k}\right)} \qquad (46)$$

But now the mode summation is over the subset of modes available to persons of type n. The assumptions underlying these models, which are combined distribution and modal split models, will now be made more explicit and discussed in the next section.

5. REVIEW OF UNDERLYING BEHAVIOURAL HYPOTHESES

Firstly, note that $C^{(n)}$, the expenditure on this class of trips by n type people, must exceed some minimum amount in order that a T_{ij}^{kn} can exist which meets the constraints relating the trip matrix to known numbers of trip ends. In fact, as

115

$\beta^n \to \infty$, $\sum_{ijk} T_{ij}{}^{kn} c_{ij}{}^k$ tends to some minimum value, and this is the minimum value of $C^{(n)}$. In so far as $C^{(n)}$ exceeds this minimum, as demonstrated by calibrated values of β^n, the excess represents what people are prepared to pay, and what people actually do pay, partly to travel further than the minimum distances and partly to travel by more expensive modes.

In a logistic modal split formula, of the form (46), β^n measures the sensitivity of type n people to the costs of different modes. If β^n is small, there is little price discrimination between modes; but if it is large, the majority of people travel by the minimum cost mode. A large β^n corresponds to a small $C^{(n)}$, of course.

But β^n also plays another role in the distribution models, say in (28) to (30) for example. It determines the average length of trip, and so also the sensitivity of people to trip length. If β^n is small, average trip lengths are long (and $C^{(n)}$ is large) and vice versa.

The main question to raise is whether a single parameter, β^n, for each person type should do the same job for mode choice and what might be called trip length choice.

A second major point is related to this discussion. The model represented by equations (40) to (42) was derived on the assumption that we were given C_{ij}^n, the cost of travelling from i to j as perceived by an n type person, and it was remarked that this must be some composite of the modal costs c_{ij}^k. This question of determining the form of composite impedance or cost, which is a question of long standing in the construction of transport demand models, turns out to be very important in an analysis of alternative modal split models. At this point, note that the model represented by (40) to (42) can be wholly identified with the aggregate generalised model represented by (34) to (36) if

$$\exp\left(-\beta'^n c_{ij}{}^n\right) = X \sum_{k \varepsilon M(n)} \exp\left(-\beta^n c_{ij}{}^k\right) \qquad (47)$$

where β'^n has been written instead of β^n in the expression on the left hand side of the equation, as the β^n's in the two sets of equations may not be the same. X is an arbitrary multiplicative constant. This then suggests itself as a good functional form for composite impedance, for determining C_{ij}^n as a function of the c_{ij}^k's.

Since C_{ij}^n is constructed out of the c_{ij}^k's, it is an array of quantities whose units are in fact determined by the functional relationship between the two arrays. In fact, it can easily be seen that in equation (47) X determines the position of the zero of the scale, and β^n determines the size of the unit on the scale. Generally, it will be convenient to take X as the reciprocal of the number of modes available (for all travellers), since this ensures that C_{ij}^n will always be positive, and to take β'^n equal to β^n, which will give the C_{ij}^n's the same scale of units as the c_{ij}^k's. We shall see shortly that there are alternative ways of expressing composite impedance relationships.

The earlier discussion above of the roles of β^n implies that alternative modal split hypotheses are possible. The next step is to discuss these alternatives.

116

6. AN ALTERNATIVE MODAL SPLIT MODAL

The obvious alternative is to seek to set up a model structure which has one set of negative exponential coefficients β^n to determine average trip lengths, and another to determine modal split. We can then test empirically which set of hypotheses best fits the facts.

An alternative model can be set up by the entropy maximising method, as follows. Suppose a set of C_{ij}^n's exist, which are the generalised costs of travel between i and j as perceived by type n people. Assume that these are the costs, however they are derived, which are relevant to the distribution of trips, and hence to trip lengths. In the earlier model, an amount $C^{(n)}$ was spent by type n people on these trips. Suppose now that some part of this, $\Gamma_1^{(n)}$, is spent on achieving destinations (that is, trip distribution), so that

$$\sum_i \sum_j T_{ij}^{*n} C_{ij}^n = \Gamma_1^{(n)} \qquad (48)$$

$$\sum_j T_{ij}^{*n} = O_i^n \qquad (49)$$

$$\sum_i \sum_n T_{ij}^{*n} = D_j \qquad (50)$$

in the usual way, which gives as a maximum entropy model

$$T_{ij}^{*n} = A_i^n B_j O_i^n D_j \exp\left(-\beta^n C_{ij}^n\right) \qquad (51)$$

in which A_i^n and B_j are given as in (41) and (42). Suppose now that the balance of the amount $C^{(n)}$, $\Gamma_2^{(n)} = C^{(n)} - \Gamma_1^{(n)}$, is devoted to travel by other than minimum cost modes between each pair of zones. Let n_j represent the minimum cost between i and j, and then we can write

$$\left(c_{ij}^k - n_{ij}\right) T_{ij}^{kn} = \Gamma_2^{(n)}{}_{(ij)} \qquad (52)$$

where $\Gamma_2^{(n)}(ij)$ is the $(i-j)$ part of $\Gamma_2^{(n)}$, and

$$\sum_{k \varepsilon M(n)} T_{ij}^{kn} = T_{ij}^{*n} \qquad (53)$$

Then, maximise $\log T_{ij}^{*n}! / \pi T_{ij}^{kn}!$, subject to (52) and (53), in the usual way to
$k \varepsilon M(n)$

give

$$\frac{T_{ij}^{kn}}{T_{ij}^{*n}} = \frac{\exp\left[-\lambda_{ij}^n(c_{ij}^k - n_{ij})\right]}{\sum\limits_{k \varepsilon M(n)} \exp\left[-\lambda_{ij}^n(c_{ij}^k - n_{ij})\right]} \qquad (54)$$

which reduces to

117

$$\frac{T_{ij}^{kn}}{T_{ij}^{*n}} = \frac{\exp\left(-\lambda^n c_{ij}^{k}\right)}{\sum\limits_{k\varepsilon M(n)} \exp\left(-\lambda^n c_{ij}^{k}\right)} \tag{55}$$

We assume that the multiplier λ^n is independent of i and j. The n_{ij}'s cancel out. Thus, we now have a distribution modal split model which distributes trips by person type in the same way as the aggregated version (40) to (42) of the general model (28) to (30), and which has a modal split (55) of the same form as (46), but with a different set of negative exponential coefficients. These alternative models could, in principle, be tested against survey data to see which gives the best fit.

7. ALTERNATIVE ESTIMATES OF COMPOSITE IMPEDANCE

There is a further degree of freedom, hinted at in section 4, which we must not over-look: the construction of the array C_{ij}^n as a composite imedpance out of the modal costs c_{ij}^k. In section 4 it was shown that

$$\exp\left(-\beta'^n C_{ij}^n\right) = X \sum\limits_{k\varepsilon M(n)} \exp\left(-\beta^n c_{ij}^k\right) \tag{47}$$

is one way of obtaining the C's as a function of the c's. We now see that the generalised model (28) to (30), and the corresponding modal split equation (18), imply that a set of C_{ij}^n's can be defined so that an aggregate model of the form (34) to (36) holds, and that these C_{ij}^n's are related to the c_{ij}^k's by (47) with $\beta'^n = \beta^n$. But we can now consider (55) to be our most general modal split formula, and so we now also see that, if (47) holds as a composite impedance rule, then

$$\beta^n = \lambda^n \tag{56}$$

Here we usually take

$$X = 1/N \tag{57}$$

where N is the total number of modes.

Since we have now admitted that the modal split mechanism can be separate from the distribution model, we also implicitly admit that composite impedance functions may be different. Thus, instead of constructing C_{ij}^n's from (47), people may behave for trip distribution as though they perceived the minimum cost only of all the modes available to them. Thus, for distribution, we might have

$$C_{ij}^n = \underset{k\varepsilon M(n)}{Min}\ (c_{ij}^k) \tag{58}$$

with equations (51) and (55). This may give good results with β^n and λ^n estimated separately, and perhaps not being equal. We could also think of more general forms of composite impedance: for example

$$C_{ij}^n = \sum\limits_{k\varepsilon M(n)} w_{ij}^k c_{ij}^k / \sum\limits_{k\varepsilon M(n)} w_{ij}^k \tag{59}$$

118

or

$$\exp\left(-\beta^n C_{ij}{}^n\right) = \sum_{k\varepsilon M(n)} w_{ij}{}^k \exp\left(-\beta^n c_{ij}{}^k\right) / \sum_{k\varepsilon M(n)} w_{ij}{}^k \qquad (60)$$

where the w_{ij}^k's are weights. As it happens, because of our separate modal split mechanism (55) (which can be calibrated separately from, and in fact before, the distribution process), we already know the proportions of travellers of each type on each trip interchange travelling by mode k, and this enables us to work out the proportion of all trips between i and j made by mode k. Then, we could take w_{ij}^k to be this proportion, so

$$w_{ij}{}^k = \sum_n T_{ij}{}^{kn} / \sum_n T_{ij}{}^{*n} \qquad (61)$$

which can be written as

$$w_{ij}{}^k = \sum_n \left[\frac{T_{ij}{}^{kn}}{T_{ij}{}^{*n}} \cdot T_{ij}{}^{*n}\right] / \sum_n T_{ij}{}^{*n} \qquad (62)$$

Then, substituting from (51) and (55) for T_{ij}^{*n} and T_{ij}^{kn}/T_{ij}^{*n} respectively, we have

$$w_{ij}{}^k = \frac{\displaystyle\sum_{n\ k\varepsilon M(n)} \frac{\exp\left(-\lambda^n c_{ij}{}^k\right)}{\sum \exp\left(-\lambda^n c_{ij}{}^k\right)} \cdot A_i{}^n B_j O_i{}^n D_j \exp\left(-\beta^n C_{ij}{}^n\right)}{\displaystyle\sum_n A_i{}^n B_j O_i{}^n D_j \exp\left(-\beta^n C_{ij}{}^n\right)} \qquad (63)$$

This can then be substituted in (59) and (60), and either equation could be solved iteratively for C_{ij}^n. Note that the solution takes a particularly simple form if $n = 1$ (and we are simply estimating a single composite cost for distribution from modal costs), as C_{ij}^n cancels from the right hand side. We thus have at least four suggestions for functional forms of composite measures of impedance, given by equations (47), (58), (59) and (60). For convenience, call these the $C1$, $C2$, $C3$ and $C4$ forms respectively.

It is now worth summarizng the alternative models which have now been assembled for testing:

(1) There is the generalized model (28) to (30), with its implied modal split, (46). This has the same set of negative exponential coefficients for determining both average trip lengths and modal split. It implies a $C1$ form of composite impedance for C_{ij}^n if this is defined. Thus, formally, this combined model can be seen as a combination of a person type distribution model, as represented by (40) to (42), with coefficients β^n, and a separate modal split mechanism (55), but with $\lambda^n = \beta^n$, and C_{ij}^n defined and constructed out of the c_{ij}^k's by a $C1$ type composite impedance formula.

(2) We now recognise that the modal split model can be separate, that λ^n is not necessarily equal to β^n (it is a matter of empirical test), and that C_{ij}^n is not necessarily constructed out of the c_{ij}^k's by a $C1$ impedance formula: there are the alternatives, such as $C2$, $C3$ and $C4$.

For convenience, call the modal split mechanism represented by (55) $M1$, and the distribution model represented by (40) to (42) $D1$. Then the four possible models defined so far are shown in Figure 1.

Figure 1
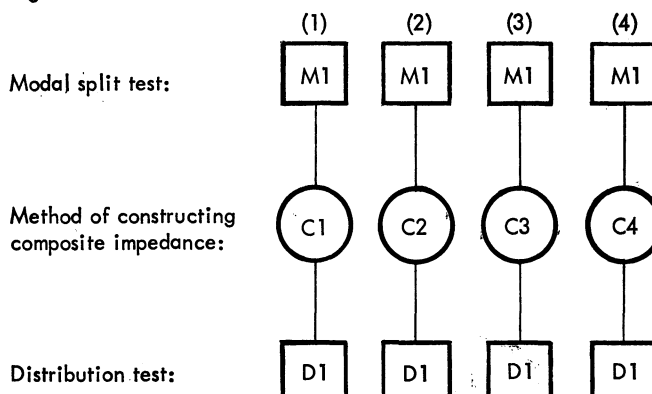
In the first column, the set of β^n's is necessarily equal to the set of λ^n's.

8. DEFINITION OF A MODE; THE CONCEPT OF ROUTE SPLIT

The problem has still greater depth and, unfortunately, complication. c_{ij}^k has been defined as the cost of travelling from i to j by mode k, and we constructed our other costs, the C_{ij}^n's, using composite impedance formulae, out of these. But what is a mode, and what are modal costs? Costs are observed on actual *routes*, and mode k, between i and j, may consist of several routes. For example, if we have two modes, car and rail, there may be several road routes and several rail routes between each zone pair. Thus we must define γ_{ij}^r to be the cost of travelling on the rth route between i and j. This is what is observed. A mode can now be defined as a set of routes. Let $R_{ij}(k)$ be the set of routes between i and j which we define to be mode k, let $r \, \varepsilon \, R_{ij}(k)$ be one such route, and let $\displaystyle\sum_{r \varepsilon R_{ij}(k)}$ denote summation over such routes. It is also useful to consider $M_{ij}(n)$ (with the ij restored) not only as the set of modes available to type n people, but also as the corresponding set of routes, so $\displaystyle\sum_{r \, \varepsilon \, M_{ij}(n)}$ denotes summation over that set of routes.

We can now define the concept of route split, to complement that of mode split, and we can investigate how to construct the c_{ij}^k's out of the γ_{ij}^r's, using again composite impedance relationships, and also the interrelationship of route split and mode split. The concept of route split is, of course, particularly relevant to the assignment part of a transport model. For example, various kinds of capacity restraint procedure, those which generate several routes between each pair of points, need allocation, or route split formulae, to allocate travel between alternative routes. So the work of this section has direct implications for transport models, as well as being important for its relation to modal split concepts.

120

Let S_{ij}^{rn} be the number of trips between i and j by persons of type n on the rth route between i and j. Then note

$$S_{ij}{}^{*n} = T_{ij}{}^{*n} \tag{64}$$

There are at least two possible mechanisms which might determine route split within the maximum entropy model methodology:

(1) That people perceive route costs directly, and that a route split formula can be developed by analogy with (55), but using a parameter μ^n to allow for the possibility of its being different from λ^n. Then,

$$\frac{S_{ij}{}^{rn}}{S_{ij}{}^{*n}} = \frac{\exp\left(-\mu^*\gamma_{ij}{}^r\right)}{\displaystyle\sum_{r\,\varepsilon\,M_{ij}{}^{(n)}} \exp\left(-\mu^n\gamma_{rj}{}^r\right)} \tag{65}$$

is an appropriate route split equation.

(2) That people perceive mode costs directly, and that mode split is determined by (55):

$$\frac{T_{ij}{}^{kn}}{T_{ij}{}^{*n}} = \frac{\exp\left(-\lambda^n c_{ij}{}^k\right)}{\displaystyle\sum_{k\varepsilon M(n)} \exp\left(-\lambda^n c_{ij}{}^k\right)} \tag{55}$$

Route split is then determined *within* modes thus:

$$\frac{S_{ij}{}^{rn}}{T_{ij}{}^{kn}} = \frac{\exp\left(-\mu^n\gamma_{ij}{}^r\right)}{\displaystyle\sum_{r\varepsilon R_{ij}(k)} \exp\left(-\mu^r\gamma_{ij}{}^r\right)} \tag{66}$$

Call these the $R1$ and $R2$ views of route split, for convenience. Once again, as mentioned earlier, we have a composite impedance problem. Consider the functional forms considered previously, and apply them to the problem of constructing c_{ij}^k's out of γ_{ij}^r's:

$$C1: \quad \exp\left(-\lambda^n c_{ij}{}^k\right) = X \sum_{r\varepsilon R_i(k)} \exp\left(-\mu^n\gamma_{ij}{}^r\right) \tag{67}$$

$$C2: \quad c_{ij}{}^k = \operatorname*{Min}_{r\varepsilon R_{ij}(k)} \left(\gamma_{ij}{}^r\right) \tag{68}$$

$$C3: \quad c_{ij}{}^k = \frac{\displaystyle\sum_{r\varepsilon R_{ij}(k)} \gamma_{ij}{}^r w_{ij}{}^r}{\displaystyle\sum_{r\varepsilon R_{ij}(k)} w_{ij}{}^r} \tag{69}$$

$$C4: \quad \exp\left(-\lambda^n c_{ij}{}^k\right) = \frac{\displaystyle\sum_{r\varepsilon R_{ij}(k)} w_{ij}{}^r \exp\left(-\mu^n\gamma_{ij}{}^r\right)}{\displaystyle\sum_{r\varepsilon R_{ij}(k)} w_{ij}{}^r} \tag{70}$$

121

Note firstly that if C_1 is the correct form we can take $\mu^n = \lambda^n$, and then the R_1 and R_2 mechanisms can hold simultaneously. In fact, C_1 is a necessary and sufficient condition for this. To put this the other way round, if C_2, C_3, C_4 or some functional form other than C_1 emerges as the correct functional form for estimating modal impedances from route impedances, and if a mode split model fits the facts, then the R_1 view is untenable. If this turned out to be true, it would be a strong statement about perception. It would mean that people perceived modes as entities in some sense prior to routes.

To summarise: if an R_1 view was justified, perception would be route dominant; if an R_2 view was justified, perception would be mode dominant unless the appropriate form of composite impedance was C_1, in which case the perception of modes and routes would have equal strength. There is one other way in which this could be true: it is just theoretically possible that a composite impedance functional form exists which allows R_1 to be satisfied together with the mode split part of R_2 (that is, equations (65) and (55) respectively). In such a case we could not really say that R_2 is satisfied simultaneously with R_1, since (66) is not satisfied, but we could say that R_1 (65) is satisfied simultaneously with M_1.

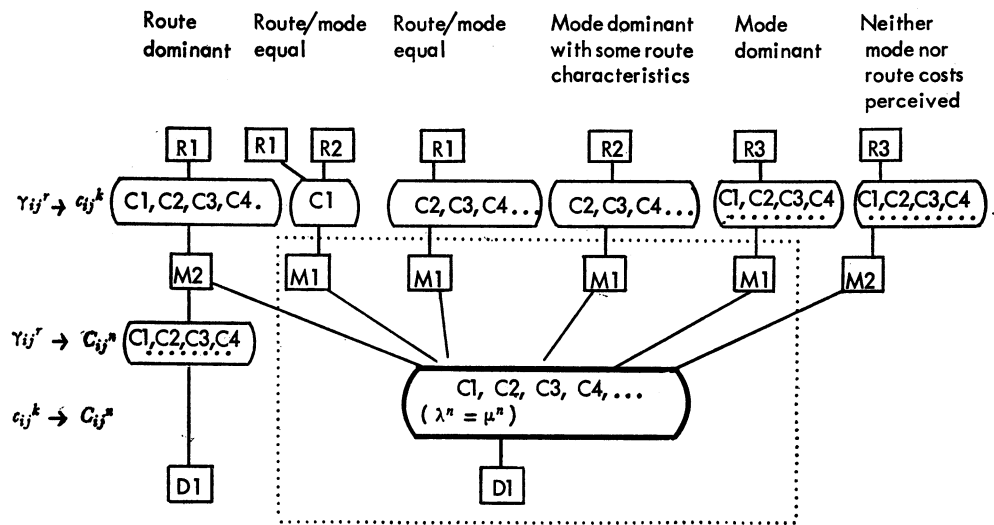Figure 2 summarises the definitions of the various submodel components.

Figure 2

R_1 :   route dominant route split model; equation (65)
R_2 :   mode dominant route split model; equation (66)
R_3 :   neither R_1 nor R_2 fits data successfully; model unknown
M_1 :   mode split model; equation (55)
M_2 : M_1 does not fit data successfully
D_1 :   person type linked gravity model distribution model; equations (40) – (42)
C_1 :   equation (47) or (67)
C_2 :   equation (58) or (68)⎫   alternative forms of composite impedance; can be used to construct
C_3 :   equation (59) or (69)⎬   (i) C_{ij}^n out of c_{ij}^k, (ii) C_{ij}^n out of γ_{ij}^r's; (iii) c_{ij}^k out of γ_{ij}^r's.
C_4 :   equation (60) or (70)⎭

The combinations of these submodels to form complete distribution, mode split and route split transport models, which are to be tested empirically, are set out in Figure 3.

The part of figure 3 enclosed by dotted lines is in fact figure 1, and so this illustrates the additional number of tests which have to be carried out as a result of considering the concept of route split, and the problem of either constructing mode costs or constructing person type costs out of observed route costs. Figure 3 should be seen as a summary and as an outline of the number of empirical tests which have to be carried out on various combinations of submodels. The tree of tests could, of course, be extended in various ways: for example, (i) to include D_1 failures, and the implications of alternative distribution models, (ii) to specify M_2 in more detail by suggesting alternative models, and (iii) by specifying R_3 in more detail by suggesting alternative models; but the extent of Figure 3 is sufficient for at least illustrative purposes.

122

Figure 3

Route dominant	Route/mode equal	Route/mode equal	Mode dominant with some route characteristics	Mode dominant	Neither mode nor route costs perceived

$\gamma_{ij}^r \rightarrow c_{ij}^k$ | R1, C1,C2,C3,C4. | R1, C1 ; R2 | R1, C2,C3,C4... | R2, C2,C3,C4... | R3, C1,C2,C3,C4 | R3, C1,C2,C3,C4

M2 ; M1 ; M1 ; M1 ; M1 ; M2

$\gamma_{ij}^r \rightarrow c_{ij}^n$ | C1,C2,C3,C4

$$C1, \ C2, \ C3, \ C4, \dots \quad (\lambda^n = \mu^n)$$

$c_{ij}^k \rightarrow C_{ij}^n$

D1 ; D1

9. DISCUSSION

(1) A simple example

Suppose there are only two points A and B in our system, and there are two routes between them, say one road (car) and one rail. Suppose the cost of travel on each route is identical. The split of traffic between A and B, route or mode, will be 1:1. Suppose now an additional road route is introduced, again all costs being identical and the same as before. The situations are illustrated in Figure 4.

Figure 4

Before   After

A ⟨====⟩ B   A ⟨====⟩ B

All costs equal, before and after.
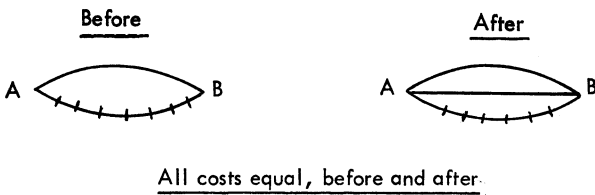
What is the split in the new situation? The $R1$ model gives the route split as 1:1:1 (road, road, rail), and hence the mode split (by summing over appropriate routes, not by applying a formula since there is no formula in this case) as 2:1 (road, rail); so there has been a shift from rail to road (assuming fixed total number of passengers, and neglecting any generated trips). But if $C2$, $C3$ or $C4$ impedance is used, the $R2$ mechanism gives the route split in the ratios 1:1:2 and the mode split as 1:1. The results given by applying $C2$, $C3$ and $C4$ impedance formulae are identical in this simple illustration, but would not be so in a more general case. Thus quite different results can be obtained according to which model is chosen. It should be possible

123

to seek out real life situations which approximate to these examples, and this should allow us to discriminate between the models.

The example suggests that, *a priori*, R_2 is preferable to R_1, since the R_1 model shifts some rail passengers to road with the introduction of a new route, even though no costs have changed. One might consider that the new road was an irrelevant alternative for marginal rail passengers, as they previously had the option of travelling by road at the newly available cost, and did not take it. If this feeling was borne out by empirical work, this would give a lot of importance to what we define as a mode. For example, should we just have car and public transport; or should express and local train services, for example, be counted as different modes; should toll motorways be counted as different from ordinary roads in this sense? A psychological motivation study may throw light on this. This problem was discussed (Quandt and Baumol, 1966) in a paper on modal split, but not really answered.

The choice between R_1 and R_2 models would also have implications for capacity restraint assignment algorithms which allocate traffic to alternative routes. What this work has shown is that (1) the algorithms should allocate among all routes in *joint* public-private networks if the R_1 theory holds, (2) they should allocate between routes on *modal* networks only, with previously determined modal splits, if R_2 holds, and (3) either can be used if and only if the C_1 form of composite impedance relates modal costs and route costs.

(2) Alternative Functions

It has been argued in this paper that the negative exponential function plays a major role in distribution and split models, and the success of the methods advocated in this paper may seem to rest heavily on this particular function being successful in practice. But there are several ways of changing the function while retaining the framework of the paper:†

> (i) The costs with which we are dealing in the models are perceived costs, and there may be a transformation relating these to directly measured costs. This transformation could change the shape of the function.
>
> (ii) Trips could be classified into groups, such as long or short, each considered to be more homogeneous than the complete set, and different functions applied to each.
>
> (iii) Britton Harris (1964) has shown that, if β^n (in the notation of this paper) is a gamma distribution for the population, the negative exponential function can be transformed into a power function.

(3) Alternative Allocation Formulae

There are also alternative allocation formulae that could be considered, in addition to the possibility of changing the functional form of the generalised cost function. If the generalised cost is given, as suggested, by (45), Quandt and Baumol (1966) use an allocation formula which is related directly to the X_r's:

$$\frac{T_{ij}{}^{k**}}{T_{ij}{}^{**}} = \frac{X_1\,(ijk)^{\alpha_1} X_2\,(ijk)^{\alpha_2}\ldots}{\sum\limits_{k} X_1\,(ijk)^{\alpha_1} X_2\,(ijk)^{\alpha_2}\ldots} \qquad (71)$$

---

†I am indebted to Britton Harris for suggesting some of these.

Other functions of the X_r's can also be written down. They all have different elasticity properties, and the representation of these properties, and comparison with observed behaviour, may be a good route into emprical testing.

(4) Procedure for Estimating Models

The methods to be used for estimating the models which have been outlined have not been discussed in detail, but one simple observation should be made for mode split or route split models. Rather that using the form (55), it is better to rearrange the basic single mode formulae and represent the split as

$$\frac{T_{ij}^{k_1 n}}{T_{ij}^{k_2 n}} = \exp\left[-\lambda^n \left(C_{ij}^{k_1} - C_{ij}^{k_2}\right)\right] \qquad (72)$$

and then take logs.

10. CONCLUSIONS

A number of questions have been explored in depth concerning the relation of trip distribution to modal split, modal split to route split, and the form of composite impedance functions which relate directly observed costs to "perceived" (or "theoretically constructed") costs. A range of empirical tests has been outlined, and it would now be of considerable interest to see some of these tests carried out. The results, as well as possibly laying the foundation for better forecasting models, would tell us something about the way in which travellers perceive modes, routes and costs. One difficulty in carrying out a wide range of tests will be that data on route choice is rather difficult to obtain. There remain, of course, a number of tests to carry out on the mechanism of distribution models themselves, especially the generalised model represented by equations (28) to (30) and its various aggregates.

It is hoped that the methods outlined in this paper provide, at the very least, a powerful method of constructing distribution, mode split and route split models. It is the methodology which should be emphasised, perhaps, rather than the particular models outlined.

REFERENCES

Harris, B.: A note on the probability of interaction at a distance. *Journal of Regional Science*, V Number 2. 1964.

Jaynes, E. T.: Information theory and statistical mechanics. *Physical Review*, 106, Number 4, 1957.

Quandt, R. E., and Baumol, W. J.: The demand for abstract transport modes: theory and measurement. *Journal of Regional Science*, 6, Number 2, 1966.

Quarmby, D. A.: Choice of travel mode for the journey to work. *Journal of Transport Economics and Policy*, I, Number 3, September 1967.

Shannon, D. A., and Weaver, W.: *The mathematical theory of communications*. University of Illinois Press, 1963.

Theil, H.: *Economics and information theory*. North Holland, 1967.
Tolman, R. C.: *Principles of statistical mechanics*. Clarendon Press, Oxford, 1938.
Wilson, A. G.: A statistical theory of spatial distribution models. *Transportation Research*, 1, Number 3, 1967.
Wootton, H. J., and Pick, G. W.: A model for trips generated by households. *Journal of Transport Economics and Policy*, I, Number 2, May 1967.

Centre for Environmental Studies, London

126