TOWARDS AN INTEGRATION OF MACRO TRANSPORT MODELLING AND MICRO-ECONOMICS

D.M. GRAY & W.S. BOWEN

ABSTRACT: *The paper is addressed to some problems in the use of macro transport models (large network models) for social cost-benefit analysis. After outlining the components of such models and some desirable properties for them, the welfare-theoretic basis for evaluation is sketched and the consumers' surplus approach extended to include the perceived/resource cost distinction. Issues arising from this distinction are discussed. A model of demand for trips is proposed and some results of an application given. Deficiencies in macro transport models are considered and an approach to land-use interaction suggested. Finally a number of aspects on the interface between operations research and economics are suggested as requiring further work.*

MACRO TRANSPORT MODELS

INTRODUCTION

The term "Macro transport models" (MTM's), to us, encompasses the class of models which deal with a large network and are used to evaluate alternative additions to that network. Typical elements of such models are: trip generation; trip distribution; modal split; and assignment. An example of the application of such a model is described in the Sydney Area Transportation Study (SATS) reports (Volume 2).

It would seem that such models have been developed from an operations research (O/R) branch of transport engineering; they tend to be highly technical, computer-oriented and discussion about them is often replete with jargon and references to competing algorithms.

The object of this paper is to address some aspects of these models from the perspective of the economist, with a view to indicating areas in which economists could make further substantial contributions and to which operations researchers and traffic engineers might direct some attention. Whilst we cannot provide an integrating framework, we hope that the perspective given in this paper will be of interest to transport engineers as well as economists. We will be primarily concerned here with the use of macro transport models in social cost-benefit analysis rather than with their other uses such as direct traffic prediction.

Acknowledgements. The permission of the Chairman of the Bureau of Roads to present this paper is gratefully acknowledged; however, the opinions contained herein are those of the authors, not necessarily those of the Commonwealth Bureau of Roads. Stimulus from Chris Nash and John Stanley has been invaluable.

This paper proceeds by first outlining the components of conventional MTM's, and secondly setting down some desirable properties of models to be used in evaluating large-scale networks. Thirdly, some basic consumer theory is sketched, to provide a basis for a more detailed incursion into various approaches to economic evaluation. Fourthly, a model of demand for trips is proposed and an example of its application given. Then, a number of the problems to be faced in the interface between the O/R and the economic aspects are surveyed. Finally, some emphasis for future work using MTM's are proposed.

Whilst many variants have been developed of the so-called traditional transportation model it can be generally characterised as a process of estimating and forecasting travel through the sequential application of four stages: trip generation; trip distribution; modal split; and assignment. A broad description of each of these stages is given below.

The Trip Generation Stage. The first component is trip production. This is the process of estimating the number of trips of (the n) various types, 0_{in}, arising per period from each zone, zone i, of the area under study. This often takes the form of multiple regression models, the explanatory variables of which may include zonal populations and such socio-economic variables as zonal income measures, zonal indices of car ownership or availability, and employment types and rates. In most and probably all cases the propensity to make a trip is not related in these models directly to characteristics of the transport system.

A second component may be the estimation of trip attraction,[1] that is, the estimation of the number of trips of each type D_{jn}, arriving at each zone (i.e., the jth destination zone). The explanatory variables may normally include: zonal populations and employment and variables oriented to land-use such as indicators of zonal activity in retailing, commerce, manufacturing of service industries.

At the simplest level these relationships are often not estimated in the statistical sense, but just hypothesized. For example, the number of work trips originating each day may be hypothesized to be equal to the number of people in each zone who are in the work-force adjusted appropriately by the proportion of working days in the year and an allowance for absenteeism. The requirement of conservation of trips is applied viz. $\sum_i O_{in} = \sum_j D_{jn}$ for each trip type n.

The Distribution of Trips. The purpose of this process is to distribute or allocate trips to particular zone pairs (ij's) based on estimates of the zonal production of trips, O_{in}, zonal attractions for trips, D_{jn}, and some measure of inter-zonal impedence (cost, distance or time).

A complete literature exists on various approaches to the distribution of trips.[2] However, the usual approach is by way of a doubly-constrained version of the gravity model.

For the n^{th} type of trip this model can be written in the general form $T_{ijn} = A_{in}B_{jn}O_{in}D_{jn}f_n(C_{ij})$ where C_{ij} denotes

---

(1) i.e., when using destination-constrained models.

(2) See, for example, Wilson (1969) and Vol. 14, No. 1, (1970) of Transportation Research for several articles.

the impedence (cost, time, distance) from the i^{th} to the j^{th} zone. A and B are the balancing factors which introduce the origin-constraint $O_i = \sum_j T_{ij}$ and the destination-constraint $D_j = \sum_i T_{ij}$, respectively, in the following way

$$A_{in} = \left[\sum_j B_{jn} D_{jn} f_n \ (C_{ij}) \right]^{-1}$$

and

$$B_{jn} = \left[\sum_i A_{in} O_{in} f_n \ (C_{ij}) \right]^{-1}$$

A range of possible forms of $f(c_{ij})$ have been used. The inverse of the square of distance from the i^{th} to the j^{th} zone is the form which, by analogy with Newtonian physics, gave rise to the term "gravity model".

In practice, here meaning what is available in the readily accessible computer-packages for transport planning, the origin and destination constraints are met by estimating, through an iterative procedure,(1) a matrix of factors K_{ijn} in the following equation which ensure approximate correspondence with the constraints.

$$T_{ijn} = \frac{O_{in} D_{jn} K_{ijn} f_n (C_{ij})}{\sum_j D_j K_{ijn} f_n (C_{ij})}$$

The K_{ijn} are often rationalised as factors which incorporate the effect of socio-economic linkages which are not elsewhere encompassed.

---

(1) The "Furness" method.

The interpretation by transport economists of gravity models as "demand models" (models of the demand for trips) should be noted here; it is the interpretation which is given later.

Modal Split - the Allocation of Trips to Modes. The function of this stage is to allocate the trips, estimated to go from the i^{th} to the j^{th} zone, between the available modes of transport. The general form of a model of modal split is

$$T_{ijkn} = (T_{ijn}, L_{ijk}, SEV_{in}, A_{jn})$$

Where T_{ijkn} denotes trips i to j of type n by mode k

L_{ijk} denotes level of service variables (e.g. time or cost) i to j for mode k

SEV_{in} denotes socio-economic variables of the i^{th} zone (e.g. car ownership, income) for trip type n

A_{jn} denotes zonal attractiveness variables of the j^{th} zone for trip type n.

Numerous variants of this model have been estimated, using regression and other statistical techniques, utilizing inter-zonal times and/or costs in the ratio or difference forms, or various other transformations. Usually the zones of origin are stratified by either income or car ownership and regression equations developed for each subgroup. The time variable may be decomposed into a number of elements such as time spent in the vehicle, waiting, transferring modes, gaining access to a mode, while costs may be decomposed into parking charges, fares, non-cash costs, vehicle operating costs, and so on.

Assignment of Trips to a Network. The function of assignment is to allocate (or "assign") the trips going from the i^{th} origin zone to the j^{th} destination zone to particular links (i.e. roads) in the road network comprising a route from i to j. This allocation may be done on the basis of selecting a route which minimises distance or time. A number of alternative algorithms exist to assign trips and all are characterised by being approximate in their solutions. Modern algorithms have a feature called "capacity restraint" which restrains the loading of links in accordance with their nominatedcapacity. With this feature, speeds on each link become a subsidiary output. Incremental loading of trips to links is a feature of most current computer packages.

In this section we will allude to a number of properties that are desirable when attempting to develop and utilize macro transport models. While the list will be by no means exhaustive, attention will be given to desiderata of both economic and operations research types.

General. As MTM's are used for forecasting, and as they are based on cross-sectional data, either changes in the parameters of the model should be capable of being related to predicted structural changes (such as changes in incomes), or these latter changes should be directly incorporated as variables.

Land use should be a function of the state of the transport network.

Trip Generation. The number of trips produced in each origin zone should be a function of the state of the transport network; if the latter is changed then the former should also. When this is not so, the implicit assumption is one of perfectly inelastic total demand for trips. While this is perhaps not erroneous for journeys to and from work, it is to varying degrees for the other trip categories.

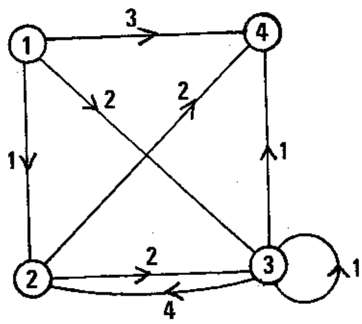
Trip Distribution

In relation to trip distribution, a number of desirable properties of a forecasting model can be listed. The first ones listed will be those with an operations research flavour. Potts and Oliver (1972) include the following:

(i) Conservation Laws. For steady state conditions, the conservation law, or Kirchhoff's law as it is known in other contexts, states that flows are neither created or destroyed. In our previous notation $\sum_j T_{ij} = O_i$ and $\sum_j T_{ij} = D_j$ where D_j is the number of trips attracted to the jth destination and $\sum_i O_i = \Sigma_j D_j$ is the total number of trips. The steady conditions here imply a consideration only of the macroscopic behaviour of traffic rather than the microscopic. Thus models conforming with the conservation law are not modelling, in any sense, phenomena such as platooning or queuing, although these are clearly of interest in congested urban areas.

(ii) Compressibility. When several centroids are combined to form one centroid (as when zonal boundaries and the origin and destination matrix have been redefined) then a further conservation requirement, compressibility, can be stated. i.e. $T_{ij} = \Sigma T'_{od}$ where the summation extends over the origin/destination pairs (o,d) in the

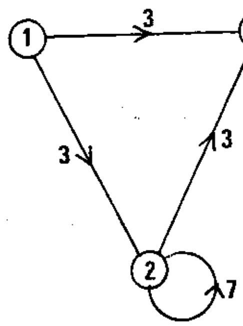
FIGURE 1



node i	1	2	3	4
production a_i	6	4	6	0
attraction b_i	0	5	5	6

destinations \ origins	1	2	3	4
1	0	0	0	0
2	1	0	4	0
3	2	2	1	0
4	3	2	1	0

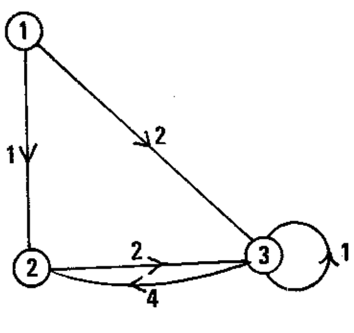
FIGURE 2



node i	1	2	3
production a_i	6	10	0
attraction b_i	0	10	6

destinations \ origins	1	2	3
1	0	0	0
2	3	7	0
3	3	3	0

FIGURE 3



node	1	2	3
production a_i	3	2	5
attraction b_i	0	5	5

destinations \ origins	1	2	3
1	0	0	0
2	1	0	4
3	2	2	1

original 0-D matrix that are combined to form the new 0-D pair (i,j) in the new 0-D matrix. For example, consider the compression of the network (having no intermediate nodes) shown in Figure 1 to the network shown in Figure 2, with nodes 2 and 3 in the former becoming node 2 in the latter, and 4 in the former, 3 in the latter. Compressibility requires that the production and attraction at the new node 2 are the sums of the productions and attractions for the original nodes 2,3.

(iii) Separability. Another "conservation" requirement is that, following the removal of one centroid (or more generally, node) and its connections with other centroids from the network under consideration, the flows i to j for the remaining ij pairs must be unaltered. For example, if node 4 in Figure 1 is removed a new network as shown in Figure 3 is obtained.

(iv) Non negativity (of ij flows).

Kirchhoff's law may seem an obvious mathematical requirement but some distribution models do not conform with it. There is some justification for such non-conformity, as the zonal productions and attractions are not usually known with any great degree of accuracy, and approximate conformity may suffice in the light of other desirable model features. However, both compressibility and separability are properties which are desirable in trip distribution models, because they imply some independence between modelled distributions and the particular way in which the area studied is subdivided into zones.

These and other desiderata may be also quoted in the form given by Bear (1973)[1]:

(a) Reversibility. The final matrix can be transformed into the starting matrix by the same procedure.
(b) Transitivity. The final matrix is the same whether it is derived from the starting matrix by a single transformation, or by way of a number of intermediate transformations.
(c) Exchangeability. If all the initial traffic flows were reversed, the final traffic flows would also be reversed.
(d) Invariance under relabelling. If two rows of the starting matrix (or columns or both) are interchanged, the transformation applied and the same rows (or columns or both) again interchanged, the final matrix is unaltered.
(e) Fractionability.[2] Zones may be combined or split without affecting the traffic predicted to or from other zones.

Potts and Oliver (1972) have stated that the conventional gravity model does not possess the properties of compressibility and separability. However, in a recent article, Beardwood and Kirby (1975) dispute this conclusion and argue that the gravity model does not necessarily have this disadvantage. By suitable averaging of inter-zonal costs, the predictions made after aggregating zones into larger units will be consistent, they assert, with the predictions made with the original zones. They also introduce a further desirable property, that of excludability, and show that the fully-constrained gravity model has this property.

---

(1) quoted in Beardwood and Kirby (1975).
(2) equivalent to compressibility.

In addition to the above properties, in the context of modelling more than one mode we believe it worthwhile to state some conditions or properties more related to the economic context.

Conditions for Perfect Inelasticity of Demand. Taking the distribution and modal split stages together, what the gravity and modal split models imply, inter alia, is that $T_k = f^k (C_1, C_2, \ldots\ldots)$ for the kth mode. The assumption of perfectly inelastic demand for total trips is $\sum_k T_k = T^*$, a constant. In the two-mode case, these imply that $f^1_{C_1} + f^2_{C_2} = 0$ where $f^k_{C_k}$ is the first partial derivative with respect to C_k. Expressed in another way, in terms of the own-price elasticity of demand for mode 1, $E^1_{C_1}$, and the cross-price elasticity of demand for mode 1, $E^1_{C_2}$, this condition reads: $(E^1_{C_1}/E^1_{C_2}) = (T_2/T_1)$.

Other, related conditions may be derived. Compliance with these conditions may be tested, either numerically through model runs, or analytically, where the model form permits.

Other Elasticity Implications. Taking the expression $(C_{ijk}\ \partial\ T_{ijk})\ /\ (T_{ijk}\ \partial\ C_{ijk})$ for given ijk as an appropriate definition for the own-price elasticity of demand for trips by the kth mode, $E^k_{C_{ijk}}$, we suggest that either numerically through model runs, or analytically, where the expression is tractable, estimates of $E^k_{C_{ijk}}$ should be explicitly made. The credibility of these estimates in comparison with the many estimates made through econometric studies gives an indication of what the distribution and modal split phases are implying in behavioural terms.

For instance for the (simplistic) unconstrainted demand model (which combines distribution and modal split) $T_{ijk} = O_i A_j \exp(-BC_{ijk})$ it is not difficult to show analytically that $E^k_{C_{ijk}} = -BC_{ijk}$. We have derived analytical expressions for other model forms; as have Hyman and Wilson (1969).

The Hotelling Conditions on Integrability. These are discussed in the later section on Measurement of Benefits.

CONSUMER THEORY, TRANSPORT AND BENEFIT ESTIMATION

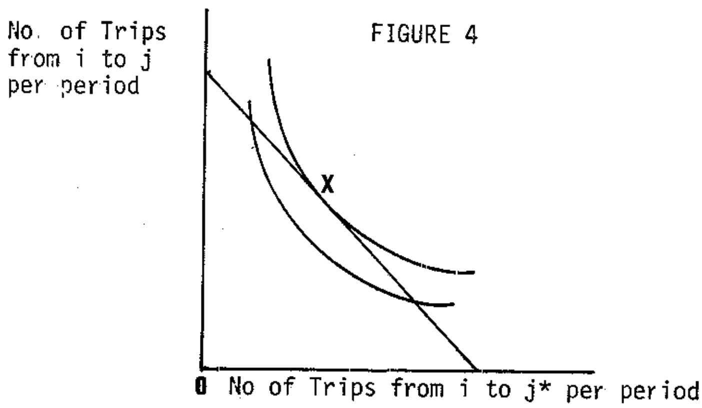
In order to reveal some of the economists' views on evaluation using transport models, we shall first sketch the basis of consumer theory in transport and then discuss various approaches to the measurement of benefits in some detail. The distinction between a perceived and resource costs basis for evaluation is made, and we address a number of issues arising from this.

Basic Consumer Theory

The theory of the consumer was originally explained in terms of diminishing marginal utility (the more one has of a particular good, the less utility will be gained from one more unit of that good). However, the modern explanation of the existence of demand curves is through the use of indifference curves. The latter approach relies only on an ordinal (rather than cardinal) concept of utility where it can be said that A is preferred to B but not by a precise amount (i.e., a preference ordering).

Indifference curves represent combinations of two[1] commodities, which combinations the consumer is indifferent between, that is, regards as being of equal utility. In the context of consumer decisions about transport this could be envisaged as, at each point on the curve, a trip from origin i to destination j being equally preferable to a trip from i to destination j*. In the context of modal choice, trips by one mode would be compared with trips by another. For each level of satisfaction (utility) a corresponding indifference curve can be drawn with the higher levels further from the origin (the curves are normally taken to be convex to the origin)[2]. Assuming the consumer has a budget constraint, this can be represented by a budget line showing the combinations of commodities possible under that constraint. Assuming the consumer's aim is maximization of his utility (satisfaction) subject to his budget constraint, the equilibrium position for the consumer is at the point of tangency of the budget line and the highest possible indifference curve.[3] At this point (point X in Figure 4) the marginal rate of substitution between the two commodities is equal to the ratio of their prices.

FIGURE 4



(1) or more. For >2 commodities, graphical exposition is is not possible.
(2) Although convex regions, beyond the satiation or "bliss" point, are sometimes assumed to exist.
(3) For convenience, we assume the number of trips (per period) can be a continuous variable.

From this framework a demand curve can be derived. If the price of one commodity is lowered (say the cost of trips from i to j*) then the budget line swings to the right, the individual moving to points A, B and C in Figure 5 corresponding to new points of equilibrium. The price/consumption curve passing through points A, B and C shows how the quantity of trips demanded from i to j* changes as the price is changed. The process can be depicted in Figure 6 which is derived from Figure 5.

It shall be taken as understood that a downward-sloping demand curve such as shown in Figure 6 for the individual consumer can be established. It is relevant however to reiterate one of the assumptions implicit in this approach - that of consumer "rationality". This approach explicitly excludes non-optimizing behaviour. Further the associated assumption of "perfect information" implies that the consumer knows what the prices are, and what his budget constraint is.

Extending these concepts of prices and budgets into the context of transport demand to include time costs as well as cash outlays raises difficulties perhaps more severe in their implications than in the analysis of demand for other commodities.

Suffice it to say that the many studies to estimate demand for trips by various modes have encountered significant problems in explaining travel behaviour strictly in terms of travel costs (time and cash costs). Thus reinforces one's uncertainty about the explanatory power of models based on concepts of rationality and perfect information.(1)

(1) In transport choice not only does the gaining of information about alternatives require expenditure of time and perhaps money but significant uncertainty usually exists about the true price (time and money) of a given journey by a given mode.

FIGURE 5

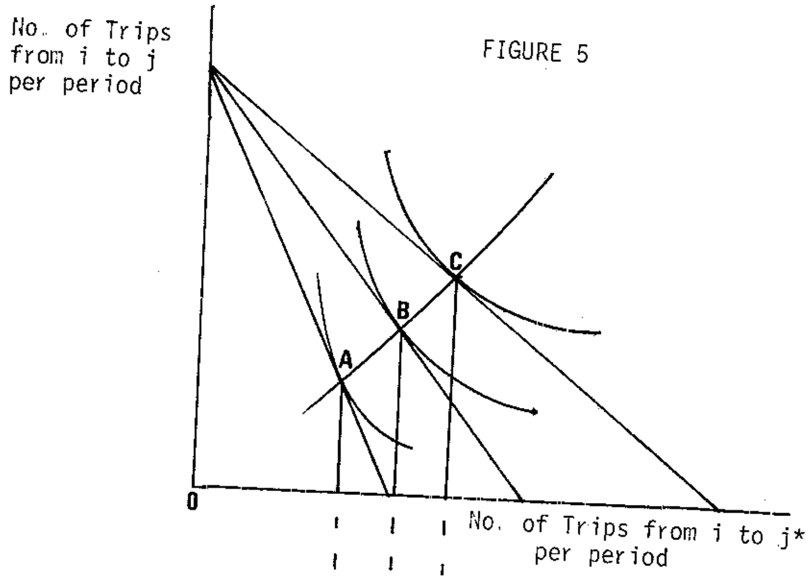
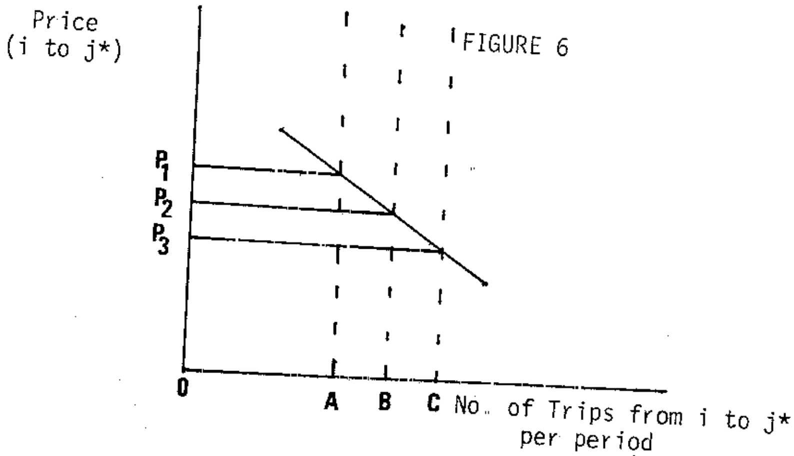


FIGURE 6



A further point to be noted in considering the application of microeconomic theory to transport is that the normal dichotomy between production and consumption may break down; the car-driver consumes and produces simultaneously - one economic agent is involved, not two or more. Optimizing behaviour is hard enough to postulate for actual markets, with transactions between agents. Much of what is said later about misperception and non-rationality may be related to this point, although we have not yet thought out all the implications.

Measurement of Benefits

Although the estimation of user benefits can become complex the basic ideas are simple. They are based on the traditional postulates of welfare economics whereby investments are evaluated in terms of individual preferences embodied in the "willingness-to-pay" criteria. That this basis represents a particular, not unanimously valid, value-position[(1)] is not often realised.

The result of a transport improvement would generally be to lower the generalised cost of travel. "Generalised cost" includes all elements in the cost of making a journey such as in-vehicle time, operating costs, walking and waiting time, and often an allowance for comfort and/or convenience or the lack thereof. Given such an improvement, there are two basic ways whereby the benefits arising may be estimated.

The first method is the "cost-savings" approach where the change in user costs, with and without the improvements, aggregated for all trip-makers, is taken to be the appropriate measure. As sometimes implemented, this is:

$$\text{Benefits} = \sum_{i}\sum_{j} T_{ij} \left(C_{ij}^{1} - C_{ij}^{2}\right) \qquad (1)$$

Where T_{ij} = number of trips from origin i to destination j.

C_{ij} - generalised cost of travel between i and j.

Superscripts 1 and 2 refer to the before and after improvement situations, respectively.

---

(1) See e.g. Nash *et al* (1975).

This approach ignores the existence of any elasticity in the demand curve, and assumes the special case of perfectly inelastic demand for the mode treated in the above fashion.

The primary problem with this approach is that benefits are only comprised of reductions in cost. In the longer run, improved accessibility following a transport improvement may result (and often through gravity models is predicted to result) in trips that are longer and often more expensive than the original set of trips. The total (generalised cost) expenditure with the improvement may exceed that without it. It is considered that where perfectly inelastic demand cannot be reasonable the cost-savings approach must be rejected; consumer theory indicates that individuals would not undertake longer trips if there were not a benefit perceived in so doing.

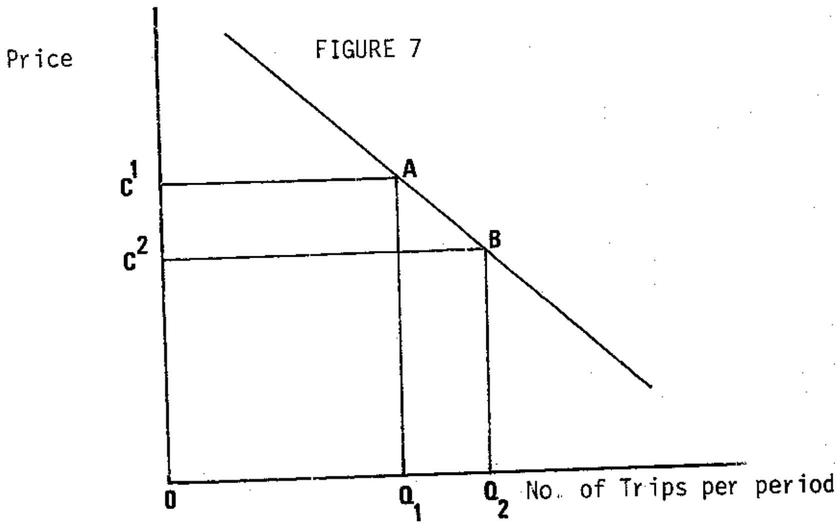
The second approach is the "consumers surplus" approach. This measures the excess over what the consumer actually pays for the goods, as represented by the price facing him in the market, of what he is willing to pay as represented by the demand curve for that good. What this means, in the unlikely case of an improvement which affects generalised costs on only the one interzonal pair, is the evaluation of the integral

$$\int_{C^2}^{C^1} f(C)\, dC \qquad \text{where } f(C) \text{ is the demand curve}$$

for that pair and C^1 and C^2 are the prices without and with the improvement. If the demand curve is linear the measure would then be area C^1 A B C^2 in Figure 7.

However, most improvements (certainly those of the sort addressed by macro transport models) involve changes in

FIGURE 7



price for many interzonal pairs (or routes). This then involves the evaluation of a line integral of the form

$$\sum_{ij} \int_{C^1_{ij}}^{C^2_{ij}} f(C_{ij})dC_{ij}$$ and introduces several complexities.

The basic one is that, in order to achieve a unique integrand[(1)] (that is not dependent on the path of integration), some conditions must be met. These are the Hotelling conditions (Hotelling, (1938)), which are that the cross-price derivatives between all goods whose prices change must be equal. Or,

$$(\partial x_i/\partial p_j) = (\partial x_j/\partial p_i)$$

This will hold (assuming an integrable utility function and expressing in Slutsky Form[(2)]) if:

$$(\partial x_i/\partial p_j)_{dU = 0} - x_j(\partial x_i/\partial M) = (\partial x_j/\partial p_i)_{dU = 0} - x_i(\partial x_j/\partial M)$$

where M is income.

(1) i.e., unique measure of benefits.

(2) See e.g., Green (1971) for elucidation.

This is satisfied if the income elasticities of the goods in question are equal or, alternatively, if

$(\partial x_i/\partial M) = (\partial x_j/\partial M) = 0; \quad (i, j \neq k)$ This latter is the case of zero income effect, where the Marshallian and Hicksian measures coincide, and the k th good is the numeraire.

As Stanley and Nash have pointed out[1] both of these possibilities are very restrictive and will not necessarily hold in practice. Thus it may be inferred that in the multi-price change situation with which the present model is concerned, willingness-to-pay measures of user benefit will generally be ambiguous, depending on the path of integration selected.

Neuberger (1971) has shown that particular cases of the gravity model satisfy the Hotelling conditions. The one which we have used in a recent application is an origin-constrained version

$$T_{ij} = (O_i A_j^a \exp(-BC_{ij}))/(\sum_k A_k^a \exp(-BC_{ik})) \qquad (2)$$

where $\sum_j T_{ij} = O_i$

and T_{ij} = trips from origin zone i to destination zone j

O_i = number of trips originating at zone i

A_j = measures of attraction of destination zone j

C_{ij} = travel costs between zones.

In our recent application we have further generalised this model[2] to determine modal split simultaneously with trip distribution, while preserving the above feature.

(1) J.K. Stanley and C.A. Nash "The Evaluation of Transport Improvements" in D.A. Hensher (ed) Urban Transport (Cambridge U.P., forthcoming).

(2) Based on the work for the Bureau of Roads by Dr. C.A. Nash of Leeds University.

It may be then written as

$$T_{ijk} = 0_i \; A_j^a \; \exp(-BC_{ijk}))/(\sum_{kj}\sum \; A_j^a \; \exp(-BC_{ijk})) \qquad \text{(2a)}$$

where the additional subscript k denotes the k^{th} mode.

Space does not permit full discussion of the advantages or otherwise of this firm, but two features may be stated. Firstly, destination choice and modal choice are, behaviourally, better conceived as simultaneous, rather than sequential processes. Secondly, allowing destinations to vary allows for transport-induced changes in land use.

One aspect mentioned above was elasticities implied by model forms. In this case, the own-price elasticity of demand for trips by mode k, $E^k_{C_{ijk}}$, can be shown to be (to a first approximation) equal to $-BC_{ij},1$.

As an aside to the economists, we perhaps should state that for strictly correct application of the Hicks-Kaldor Compensating Variation Criteria(1), that the relevant individuals have identical, homothetic indifference maps.(2) Aggregation of benefits in the way normally performed may be a logically valid indicator of welfare change only where marginal social utility of income is identical between individuals(3). How severe divergence from these restrictive conditions is for the accuracy of indicators of changes in social welfare is not clear from the literature, but we suppose it would be the least of the inaccuracies.

(1) See, e.g., Nash, Pearce and Stanley (1975) for a discussion of the Hick-Kaldor C.V. Concept.

(2) Smith B. and Stephen F.H. (1975).

(3) Boadway (1974).

Both the form of the demand model utilized and the expression for change in surplus given in equation (8) have further support from a recent article by Cochrane (1975), although such support does not explicitly extend to the mode-split aspect of the model as utilized. Cochrane derives the singly-constrained version of gravity model from the principle that trip-makers choose the trips providing the greatest net benefit to them as individuals and that the pattern of trips reflects the overall probability of particular trips being chosen on this basis. His article thus provides a further, and alternative, basis (which is derived from micro-economic considerations) for the form of the model.

In Figure 7 the number of trips diverted is $Q_2 - Q_1$ and of undiverted trips is Q_1. In the context of this distinction a cost change approach can determine the upper and lower bounds of benefit measurement and the relevant measures for diverted trips. Making the usual assumptions of perfect knowledge and rationality it can be stated that, in a two-mode situation, the traveller must be better off if he changes mode following a price change on one mode. If there is no improvement in welfare, he would not have made the change. At the other extreme, the net benefit cannot exceed the change in cost, for if this were the case, the trip with the improvement would have been made without the improvement. If the assumption is made that the travellers with whom we are concerned are spaced evenly between these upper and lower bounds, then the average benefit gained by these changing their behaviour is one half of those who do not change.

Thus the measure of benefit would be

$$\text{Total benefits} = \tfrac{1}{2}\sum_{ij}\sum (T_{ij}^{1} + T_{ij}^{2})\,(C_{ij}^{1} - C_{ij}^{2})$$

Perceived and Resource Costs

The framework discussed thus far can be extended by making the distinction between perceived and resource costs of travel. This distinction is one which has been established in both the literature[1] and in much of evaluation practice in Australia. Because this distinction is with us and because it requires on the modelling side the development of an additional set of transport costs, it is necessary to discuss the underlying concepts in some depth and to indicate the problems arising in applying it.

We start by setting out what we understand to be the meaning of this distinction. Perceived costs of travel are those costs which the user thinks he bears when undertaking a trip. Resource costs of travel represent the actual consumption of community resources used up in the undertaking of a trip. This would, at the conceptual level, appear to be reasonably clear.

For expository purposes, a further distinction might be made between perceived costs and actual costs, where actual costs are those costs the user actually bears when undertaking a trip and also between actual costs and resource costs. The difference between perceived and actual costs may arise from:

(i) the user having imperfect information about costs actually incurred;

(ii) uncertainties as to the proper allocation of costs (some variable costs may be wrongly regarded as fixed and vice versa);

(iii) the user may not be concerned to assess some of the costs, being indifferent to them because of their small size in relation to his overall budget.[2]

(1) e.g. Quarmby and McIntosh (1970), Harrison (1974).

(2) i.e., especially with respect to cash costs, the user may not have a transport budget, in the consumer theoretic sense.

(iv) straight errors in perception of costs. This is probably important with respect to estimates of travel time.

All of these four factors could be grouped under the heading of misperceptions. It should also be noted that at the point of destinational and/or modal choice the individual, all other misperceptions aside, cannot know ex-ante the true generalised cost of a given journey but faces a random variation in this cost, where the variance is often significantly large. Indeed, the potential size of the variance may influence the choice.

The difference between actual and resource costs may arise from two basic factors:

(i) the existence of transfer elements in the actual costs (where these elements do not represent the consumption of community resources) arising from various taxes and excises.

(ii) the possible divergence between the actual cost to the individual user and the resource cost to the community other than that arising from (i) above. This divergence may be due to the existence of external diseconomies. It may also be due to a rejection of the willingness-to-pay concept in valuing individual expenditures in favour of an alternative view. This latter point will be further discussed below.

While indirect taxes may constitute part of the actual cost of making a journey they do not represent a real cost to the community in terms of resources used, as they are generally transfers from consumers to the government and should be excluded from resource evaluations. This is because we are really concerned with an opportunity cost concept. Furthermore recognising that indirect taxes are incident upon the non-transport, as well as the transport sector, it should be noted that, following a proposed transport improvement, it is the projection of the <u>net</u> transfer effects which is of interest.

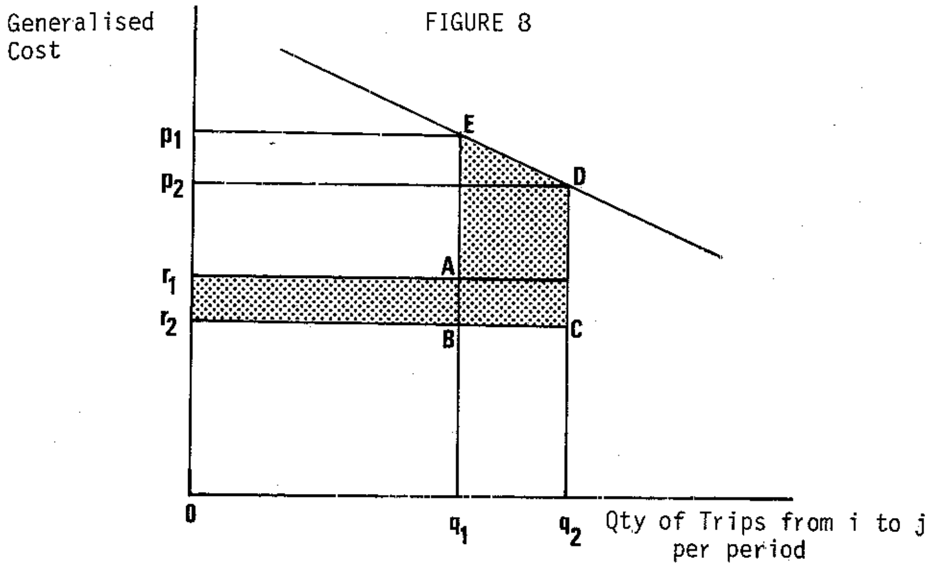
Some qualifications must be expressed about removing the tax component of costs. Firstly, as pointed out by Harrison (1974), the tax should not represent a price (or charge) in some direct or indirect sense, although deciding this is more or less arbitrary. Secondly, the implicit (and usually reasonable) assumption is that the effects of the improvement tax revenue from users would have little or no effect on the government's other taxation or expenditure policies.

The perceived/resource distinction require some modifications to Figure 7. Following Neuberger (1971) user benefits (UB) are given by

UB = (perceived user benefits - fall in perceived costs) +
fall in resource costs. (4)

Depending on the assumptions that can be made about the relationship between the two costs, different diagramatic expositions of various situation result. Assuming that the original and final perceived costs are greater than resource costs before and after the improvement, Figure 8 represents the new situation, the shaded area being user benefits.

FIGURE 8



User benefits (UB) can then be stated as:

$$UB = \tfrac{1}{2}(q_1 + q_2)(p_1 - p_2) + (r_1q_1 - r_2q_2) - (p_1q_1 - p_2q_2) \quad (5$$

where p_1, p_2 = initial and final perceived costs

r_1, r_2 = initial and final resource costs

q_1, q_2 = initial and final number of trips.

The shaded area in the above diagram represents benefits and can be split into two components

(a) resource cost savings on undiverted trips (r_1 AB r_2)

(b) the difference between the money measure of perceived benefit on the generated number of trips and the cost of resources consumed in this travel (BCDE).

Taxation Effects

Adjustments are necessary to make perceived costs adequately reflect resource costs, due to the fact that indirect taxes are considered to be merely a transfer payment. Equation (1) included taxation benefits resulting from generated traffic, as there will be an increase in government revenue due to an increase in expenditure on items that have elements of indirect taxation. However, it is also necessary to consider the goods from which expenditure is diverted, as these will have had some tax elements and hence there will be a tax loss on the goods from which expenditure is diverted. If prices do not equal marginal costs in all other sectors of the economy, gains and losses in other sectors will result, and these effects need to be considered in the evaluation process.

It is assumed that the increased spending on private transport that may result from the investment is attracted proportionately from all other goods, the mean indirect tax rate in the country on other goods can be taken as the tax rate to use in adjusting the measure of user benefits.

Thus, the tax loss in non-transport sectors, if resources are diverted to transport is equal to

$$T.L. = t.\Delta S \qquad (6)$$

where Δs = change in money spending on transport by users (new-old)

and t = mean rate of indirect taxation on goods from which expenditure is diverted to transport.

The user benefit formula would then be

$$UB = \tfrac{1}{2}(q_1 + q_2)(P_1 - P_2) + (r_1q_1 - r_2q_2) - (P_1q_1 - P_2q_2) - t.\Delta s \qquad (7)$$

Non-Linear Demand. There are problems in the assumption of linear demand curve, especially in the case of a large scale project resulting in large cost savings and where (assuming a reasonably elastic demand curve) most of the traffic is generated.

In the context of the non-linear demand relationship expressed in equation (2) above, the appropriate measure of user benefits is as follows. The elements are, for each trip category and mode[1]:

(i) Perceived User Benefits (PUB)

$$(PUB)_i = (O_i/B)\ \ln\left(\left(\sum_j A_j^a \exp(-BC_{ij}^{p2})\right)/\left(\sum_j A_j^a \exp(-BC_{ij}^{p1})\right)\right) \qquad (8$$

where C_{ij}^{p1}, C_{ij}^{p2} = perceived generalised cost in base and improvement cases

(ii) Perceived User Benefit on Undiverted Trips (PUBUT)

$$(PUBUT)_i = \sum_j (\min T_{ij}^1, T_{ij}^2)(C_{ij}^{p1} - C_{ij}^{p2})$$

(iii) Resource Cost Savings on Undiverted Trips (RCSUT)

$$(RCSUT)_i = \sum_j (\min T_{ij}^1, T_{ij}^2)(C_{ij}^{R1} - C_{ij}^{R2})$$

where C_{ij}^{R1}, C_{ij}^{R2} = resource cost plus taxation (i.e., actual expenditure by user)

(iv) Resource Less Perceived Cost on Diverted Trips (RLPCDT)

$$(RLPCDT)_i = \sum_j (T_{ij}^2 - T_{ij}^1)\ (C_{ij}^{R2} - C_{ij}^{R1})$$

(v) Net Taxation Effects (NTE)

$$NTE_i = t\Delta s_i = t \sum_j (C_{ij}^{c2} - C_{ij}^{c1})$$

where the superscripts c1 and c2 refer to actual cash expenditures.

Within each of the above elements, summation should be performed over all modes and trip categories.

(1) Understanding that each trip category and mode has its own C_{ij}'s.

We are now in a position to combine these elements to an overall estimate of user benefits $UB = \sum_i UB_i$

where: $UB_i = (PUB)_i = (PUBUT)_i + (RCSUT)_i - (RLPCDT)_i - NTE_i$

Issues Regarding the Value of Time, Perception, and Resource Costs

One of the key conceptual problems that needs be faced in evaluations with such models is the valuation of travel time savings when perception problems are considered. The problems of deriving values of time from data of behaviour under various circumstances are well known[(1)], but in the present context we are concerned with the problems that arise with the introduction of the distinction between perceived and resource costs. The issue with which we are concerned is whether a misperceived benefit should be counted as a benefit.

The issue derives from the well-known observation that although there may be a transportation improvement that lowers the price that an individual must pay for that service, it is not necessarily true that the individual will be aware of the reduction. This is likely to be the case with small improvements which mean that although travel times are reduced it may not be obvious to the individual that he has saved any time.

If an individual is not aware of the time savings or indeed any other "benefit", why should it have a value placed upon it in an evaluation context? This issue is of considerable significance to the evaluation process in transport planning given the dominance of time savings in user benefit procedures and the fact that many of the schemes with which we are concerned are predicted to have only marginal effects on total travel time. The argument runs along the

(1) See, for example, Hensher (1972).

lines that small changes in time savings (or in other items such as vehicle operating costs) will not be perceived by any individual unless it is a significantly large change. We cannot necessarily assume that an individual will be aware of a time saving of 1 minute out of a total door to door time of 30 minutes. This view has been reinforced by Dr. David Hensher's finding(1) that trip-makers tend to perceive travel time in "lumps" of about 5 minutes.

If a time savings such as this is not perceived is it correct to include that saving in the evaluation procedures? The argument against including this element as a benefit is that the economic basis of welfare theory (upon which conventional evaluation procedures rest) derives from the "willingness-to-pay" criteria, and if people are not aware that a benefit has occurred presumably they will not be willing to pay for it. Why then, should it be valued?

Arguments in favour of going ahead and valuing these non-perceived time savings can then be based on the criteria of benevolent dictation by government (such as the arguments concerning a merit good in public finance theory). However, it would not seem appropriate to say that these arguments are based on traditional welfare economics.

The other way in which a non-perceived (or mis-perceived) benefit could be incorporated into the framework would be to analyse the effects of the misperception in terms of the consumers budget. If a benefit is not perceived to exist by the consumer then presumably his overall budget situation will be affected. There is a slight modification that needs to be made to the argument at this stage concerning

(1) Hensher, 1972.

the way in which costs are measured in the evaluation procedure. It is the generalised cost of travel that is measured, and when time savings are considered we place a value on this in the same way as we do in deriving savings from any other money costs, such as in vehicle operating costs.

If misperception occurs, then we could argue that, in the context of the consumer's overall budget, this will make no difference as the unperceived saving will appear in the budget as additional spending power and hence should be valued; the consumer is better off if he has additional spending power.

This argument may be valid as far as operating costs and other items involving the actual expenditure of money is concerned, but it is less applicable to time saving. That is, small time savings cannot be evaluated in the same way as other monetary concepts because they cannot be accumulated or saved in the same way as operating cost savings can be. This problem goes to the heart of the problems involved in the evaluation of travel time savings.

The problem in most modelling exercises is that small time savings of many individuals are amalgamated and, according to some of the arguments advanced, this may lead to over-estimates of benefits. Certainly it is arguable that small time savings are not perceived at all (say below 5 minutes).

A modelling procedure we would propose to overcome this is to weight savings in travel time arising between the comparison of the with and without cases by a "misperception" or "utilization" factor. Ranging between zero and unity, this factor would follow a normal cumulative probability distribution with the point of inflexion at around 5 minutes for private time savings.

When the resource basis of small time savings is considered the issue becomes more complex. In the case of commercial time savings it can be stated that there may be some savings in resources if in-vehicle time is reduced, i.e. existing operators may reduce fleet size, there may be savings in drivers' wages and in maintenance, and there is presumably an element of improved efficiency for the whole economy when goods and services are delivered faster and more economically. Even in this sphere, however, there are problems. It is not certain that these time savings will in fact result in savings in fleet size and drivers' wages because of the nature of commercial goods movement. That is, a time savings of say 5 minutes on the road may be partially dissipated at delivery or pick-up stages where institutional and organisational problems militate against the ability to convert the time savings into actual cost savings (increases in productivity). Drivers may just take longer rest periods.

The reservations that exist in the commercial goods area become more pronounced when the issue concerns private time savings as in the current model. What are the resource savings for the community when there is a saving of several minutes on the journey to work in peak hours? This problem proves to be of considerable difficulty, given the institutional structures of private and public transport and uncertainty as to the way reactions might occur.

Considering private transport first, it is fairly obvious that a time savings for the private journey to work will not necessarily result in a resource saving to the community in terms of a reduction in the numbers of cars on the road. There may be a small reduction due to many factors, but it is difficult to attribute much significance to the travel time savings in this reduction. A similar argument applies to public transport, for if large investments improving travel times were implemented, it is not obvious that they will be converted into a saving in resources (either in terms of the number of trains, or in drivers' wages or other items of expenditure).

Another argument involves the assertion that there will be an increase in efficiency, however measured, from a reduction in private travel time. This could presumably result from increased efficiency of working time(1) because of less time spent on the frustrations involved in travelling to work. This may perhaps be true, but it seems somewhat tenuous.

The issue really revolves around whether the appropriate social criteria for valuation of private travel time is one which involves individuals' willingness-to-pay (as modified for their misperception), one based on a concept of the opportunity cost of time not saved in the do-nothing, or one which requires some evidence that real non-transitory resource savings are achieved which are transmitted into enhanced economic efficiency. In the context of the perceived/resource distinction, one could be forgiven if the semantics of the distinction draws one to prefer the latter.

The significance of these conceptual arguments should not be underestimated, for they bear upon the economic rationale of much of traditional transport planning and have

---

(1) or time in other activities of high opportunity cost.

some bearing upon the ranking of candidate projects. The position that we would have adopted in practice is that resource benefits be calculated by treating the amalgamating small time savings, weighted by the above-mentioned "utilization" factor[(1)], as a correct measure of the resource benefits to society. What the current debate shows is the difficulty of applying theoretical concepts to a practical situation, and the underlying caution needed in interpreting the results.

EXAMPLE OF MODEL APPLICATION

This section reports the results of applying the model described in equation (2a) to a consideration of alternative possible road construction strategies in urban Sydney.

Selection of Strategies. The road investments compared were two sets of construction projects, one generally radial with respect to the Sydney Central Business District, and one generally non-radial (or, as we have termed it, circumferential). These sets of projects were selected largely from projects proposed at some time by the Department of Main Roads N.S.W. and a few other projects were derived from the Sydney Area Transportation Study Report.

The general purpose of the analysis was to compare, at the broad strategic level, the relative advantages, under different pricing regimes, of constructing further extensions to the road system which are either radial or non-radial in orientation. It should be noted that we have not evaluated radially and circumferentially-orientated networks per se but rather roads of those orientations which are additional to the existing system. We have not attempted to specify an optimal or ideal system, but to analyse two meaningful

---

(1) and valued at their opportunity cost.

if extreme systems.

The two sets of projects which have been compared are roughly comparable in length, the radial set comprising 118 km and the circumferential 130 km. The construction and acquisition costs of the radial set have been estimated to be about $916 million (in 1974-5 dollars) and the circumferential system about $788 million.

We have, for cost and other reasons, restricted our detailed analysis to those journeys which are by people from their residence to their work. Given that these are rather more strongly oriented to Sydney's Central Business District than other trip types, the exclusion of the latter from economic analysis is considered to either favour the radial network or be at least neutral with respect to the radial/circumferential comparison. Furthermore, journeys-to-work comprise a significant proportion (33%) of the trips each day in Sydney. We have also constrained our analysis to travel in the peak period for the reason that data about the Sydney road network in the "off-peak" period was not available, nor to our knowledge, in existence. This analytical constraint we regard as being quite severe, as many important types of trips in urban Sydney occur predominantly in the off-peak (viz., goods movement, shopping trips, trips by businessmen and salesmen) and these may be less radial in orientation than journeys-to-work. The needs for road construction derived from an analysis of off-peak trips would almost certainly be different from a peak-based analysis.

Space does not permit either a full description of methods employed in forecasting variables or in model runs. We will just indicate the type of results obtained.

Comparison of Results

Under all of the regimes or sets of assumptions tested, car usage in the peaks by persons going to and from work, is estimated to be greater under the radial configuration of additions to the road network than under the circumferential. Both construction alternatives would result in greater car usage than the do-nothing or 'no-build' alternative.

The level of user benefits for journeys to and from work arising from road construction was found to be dependent to quite an important degree upon the price levels and state of the public transport systems which were assumed for the future year of analysis. Table 1 summarises the results produced under the various assumptions. The highest level of benefits from road construction was estimated after assuming that, in 1990, public transport fares were at a level 50 per cent greater in real terms than in 1975 and car operating costs 15 per cent greater than in 1975. The next highest level of benefits came from assuming that prices in 1990 in real terms would be the same as in 1975. A lower level of benefits from road construction came from assuming that investments (to an unspecified cost) were undertaken in public transport which, in 1990, would reduce average waiting times and in-vehicle times for public transport by 15 per cent and 20 per cent, respectively. The next lowest level of benefits was estimated under the assumption that, in 1990, public transport fares would be in real terms 50 per cent of their 1975 levels while car operating costs would be at their 1975 levels. The lowest level of user benefits estimated arose from the assumption of constancy in the real price of public transport fares, taken together with the assumption that, over the period 1975 to 1990, car operating costs would increase by 50 per cent.

TABLE 1. THE COMPARATIVE ECONOMICS OF CERTAIN RADIAL AND CIRCUMFERENTIAL NETWORKS: SYDNEY, 1975

(All in 1974 - 75 dollars)

	Length of New Roads[1] km	Estimated Road Construction and Acquisition Costs ($m)[2]	Present Value of Construction and Acquisition Costs ($ m.)[3]
Radial Network (R)	118	916	464
Circumferential Network (C)	130	788	400

Pricing Regimes	Present value of Benefits to Work Journeys ($ m)		Present value of Benefits to Work Journeys per Kilometre Constructed ($m km-1)		Present value of Benefits to Work Journeys as a Ratio of present value of Construction and Acquisition Costs	
	(R)	(C)	(R)	(C)	(R)	(C)
1. 1990 prices the same in real terms as 1975	81.1	61.0	0.69	0.47	0.175	0.152
2. Public transport fares reduced by 50% compared with 1.	66.7	52.9	0.57	0.41	0.144	0.132
3. Public transport fares increased by 50% and car operating costs by 15% in 1990 compared with 1.	90.5	69.4	0.77	0.53	0.195	0.174
4. Public transport receives investment which reduced average waiting time and in-vehicle time by 15% and 20% respectively	78.0	56.7	0.66	0.45	0.168	0.147
5. Vehicle operating costs increased in real terms by 100% in 1990 compared with 1975.	54.7	33.2	0.46	0.26	0.118	0.083
6. Prices as in 1 above Model parameter α set at 0.45 and 0.55 for Manufacturing and Non-Manufacturing Work trips, resp.	89.3	63.4	0.76	0.49	0.192	0.158

1. including the length of upgraded roads.

2. excluding acquisitions already made as at 1971

3. assuming costs are spread uniformly between 1975 and 1990 and 10% discount rate (0.507 of previous column)

The average level of user benefits arising from the construction of the radial system and its use for work journeys in the peak hours ranged under the various assumptions, from a high of $770 thousand per kilometre of freeway or expressway constructed to a low of $460 thousand per kilometre. For the circumferential system, the equivalent range was $350 thousand per kilometre to $260 thousand per kilometre. The estimates of the present value of construction and acquisition costs were upwards of $3.93 million per kilometre constructed for the radial system and $3.08 million per kilometre for the circumferential.

Given that we have only examined peak work journeys in the model, it becomes apparent that benefits to trips other than this category would have to amount to at least the difference between the benefits as calculated and the figures for construction and acquisition costs before either system could be supported on grounds of road user benefits.

It must also be remembered that these calculations do not include an allowance for uncompensated environmental and social costs. Nor do the acquisition costs include opportunity costs of land previously acquired.

LIMITATIONS OF THE TRADITIONAL MACRO TRANSPORT MODELS

Thus far, we have discussed from an economist's view the measurement of user benefits using macromodels. Some further things need be said by way of integration with the O/R aspects.

Desired Accuracy. Firstly, if the sole aim were to produce forecasts of traffic volumes on links in the system especially new links a certain level of accuracy (certainty of the forecast) would be sufficient. However, with the use of social cost-benefit analysis for the justification of projects or groups of projects, the accuracy, not only of the projected costs (being interzonal costs derived from the link speeds developed from a capacity-restrained assignment) but of the projections of exogenous variables becomes more important. The tasks and significance of economic appraisal using macro-transport models vary widely. This has important implications for the degree of accuracy required and finesse applied. For the initial screening of project viability simple methods may be appropriate, whereas for the assessment of a large set of projects more sophistication of method, albeit at a higher level of abstraction as to network details, may be more appropriate. In the latter case, significant effort should be directed not only to more sophisticated modelling but to sensible background assumptions and projections about land-use and city form, population growth, vehicle operating costs and so on.

In order to discuss the limitations of Macro Transport Models in terms of their utility in economic evaluation, it is necessary to first indicate some of the processes whereby outputs relevant to evaluation are obtained. Economic evaluation of alternative states of the road network requires, at a minimum, both cost and trip matrices (C_{ij} and T_{ij}, respectively) in all of those states, including the "do-nothing" state. These matrices must purport to represent cost and times, at the point of steady-state equilibration of supply and demand.

A major problem exists in attempts to interpret the link speeds which are the output of the assignment of trips to the network as equilibrium speeds.

The first reason for this related to the uncertainty as to the most appropriate parameters of the speed-flow relationships which are inputs to capacity-restrained versions of the process of assignment. Parameters must be fairly arbitrarily chosen on at least five criteria: (i) the likelihood of assisting convergence in the iterative assignment algorithm; (ii) the likelihood of producing reasonable traffic flows on links; (iii) the likelihood of achieving reasonable speeds on links; (iv) behaviour at extremal values of flow, particularly under high volume/capacity ratios or "overload"; (v) the degree of sensitivity of speeds (and hence costs) to changes in volumes. This last criterion is particularly important if iterations are to be performed of the distribution, modal split and assignment stages based on inter-zonal costs derived from the previous assignment.

In practice, the speed-flow parameters can neither simultaneously satisfy all these criteria nor be determined empirically except in an approximate way. Calibration of the parameters by non-rigorous comparison of the flows and speeds of the base network (often themselves subject to uncertainty) with those of the reproduced network is the normal approach.

The second reason is that, as normally performed, the elements of the cost matrix which relate to new or improved links have to be arbitrarily guessed prior to running the distribution, mode-split and assignment stages. If these new costs are not guessed accurately, the resulting distribution and modal split will not represent the new equilibrium. Sometimes one iteration back through the last three stages is performed in an attempt to provide a better approximation to equilibrium, but this is normally costly of computer time. It is also subject to tautological results arising from the fact that if parameters of the speed-flow relationship have been chosen such that speeds at high link volumes are insensitive to changes in link volumes (see (v) above) then iterations which reallocate trips and modes are not likely to change the link speeds from those derived from the first assignment.

A further problem exists in the derivation of the average interzonal cost by the skimming of the speeds (times) and lengths, according to a minimum time (or cost) criteria. One route from i to j then purports to represent the "typical" or average route from the point of view of generalised costs. One might expect that costs derived from trip-weighted times and lengths would, conceptually, be better candidates. The spectre of computer costs and feasibility are here too, however.

Forecasting Problems

In using MTM's to forecast traffic conditions in the longer-run, a number of issues arise. Some have to do with the exogenous forecasting of the variables of the model, others have to do with the likelihood or otherwise of structural changes that would invalidate the parameters of the model and with the inherent stability of the model's parameters over time.

The reservations one must have in using cross-sectional data to estimate models used in making predictions are well-known[(1)].

Since time-series data are almost never available,

(1) e.g. Kuh (1959).

estimation of parameters on cross-sectional data introduces inaccuracies of unknown magnitude.

Further, there are reasons to suspect that the cost (or impedance) parameter of gravity models is unstable over time. Fisk and Brown (1975) show that B_i, the origin specific value of B, is approximately equal to the inverse of the mean travel cost out of the ith origin zone. This probably means: firstly, that if any transport improvement changes the mean travel cost then B_i will change; and, secondly, that B_i will change over time due to the effect of income on average expenditure on travel (and hence on average travel cost). In their words: "the assumption adopted in present model applications that the parameter remains constant between base and prediction years is unjustified". A similar conclusion is reached by Hyman and Wilson (1969).

There may be some advantage, from the point of view of prediction, in separating the cash and non-cash (or cost and time) portions of generalized cost in the demand equation i.e., instead of $\exp(-BC_{ijk})$, the cost function would be $\exp(-(B_1C_{1ijk} + B_2C_{2ijk}))$, for example. This is because it is likely that rising income levels would result in greater (willingness to pay) cash travel costs but reduced willingness to incur given time costs. Given that city population growth over time will lead to increased congestion and hence an increase in average time expenditure on travel with incomes constant, the effect of simultaneous growth in income and population would be to mitigate this (a hightened preference for faster modes). The aggregated generalized cost approach used in longer-run prediction (probably erroneously) implies that no important change over a period of years in the trade-off between time and cost occurs. However, separate cash and non-cash B's could be estimated to change over time in proportion to the inverse of the change in cash and non-cash average expenditure on travel, respectively, as estimated from the relationship between these latter and income, and other factors.

Since major transport improvements usually take some years to implement, the vital comparison of the do-nothing (or without) case and the do-something (or with) case is often taken with reference to a future year (the year 2000 has been a favorite). Such an approach focuses attention on the realism of the "without" situation. If used in a simplistic, mechanistic fashion, MTM's often predict more traffic than there is road capacity with the resulting absurd prediction that average speeds on many links would be below 5 mph and hence travel costs very high. With this approach the comparison with the do something cases over estimates the benefits arising. This is because there is nothing in the structure of most MTM's that necessitates consideration of:-

(i) the tendency of trip-makers, faced with congested roads, to shift their trip-making to another time of day (institutions eventually facilitate this) or to choose a form of public transport not affected by road congestion (rail, ferry);

(ii) longer-run changes in land-use (the spatial distribution of economic activities) and hence trip patterns in an ameliorating response to increasing road congestion. Perhaps much of the relative decline of the CBD (in Melbourne and Sydney) as an employment centre for many industries reflects such a response, particularly in the context of a decline in rail services and increasing competition for road space between private and commuter traffic and business and commercial traffic.

Other problems exist with MTM's which have implications for the accuracy of estimated benefits. These include:

(i) the fact that congestion is not a steady-state phenomena in large urban centres, contrary to the assumptions in all MTM's. For instance, at the start of the "peak", flow rates incident on many freeway off-ramps and arterial road intersections exceed their capacities, and queue lengths rapidly increase(1). It is not until flow rates decline below off-ramp or intersection capacities that queues diminish.

(ii) the feedback or second-order effects of a relative decline in public transport on modal split.

As Fairhurst (1975) points out: "Transportation studies have .... generally ignored the possibility that the level of car ownership may be affected by the transport system itself and, in particular, by the quality of public transport. It has been assumed that the quality of public transport may influence modal split, and hence car use, but not car ownership. Indeed, there has been avoidance of any contact between the user cost factors that determine modal split and car ownership levels". His findings from an econometric study based on London data "imply that policies favourable to one mode will have not only an initial effect on modal split" but will also have "significant second-order effects, as marginal households consider whether to own a car. The consequences of

(1) "infinite" queues in the parlance of the traffic theorist.

their assessments will be to produce second-order changes in the usage and financial viability of public transport".

It is believed that MTM's can be a useful tool for social cost benefit analysis of road improvements, but only if the above points are not only recognised but incorporated in the analysis in some approximate way.

A discussion of such appropriate ways is beyond the scope of this paper. However we will dwell briefly on the question of predicted land-use, given its importance.

Land Use

We have already said that, in already-congested cities, urban form responds over time to the state of the transport network. Thus, in a city such as Melbourne, the location of employment for both industrial and service activities gradually changes in response to transport factors. In the projection of the land-use associated with the future base or do-nothing network, the analyst must be careful that inconsistencies do not arise between the predicted traffic volumes and speeds and what could be expected in practice. For instance, if significant growth in central employment opportunities were assumed by the analyst, the predicted increase in traffic volumes on the do-nothing network would probably involve traffic crawling along at unrealistically low speeds. If so, the predicted growth in central employment may be strictly contingent on the transport improvements which are being evaluated! Unfortunately, there are no models capable of simultaneously estimating benefits arising from the comparison of a city, having a given pattern of land-use and a given transport system, with the same city with a changed pattern of land-use and a changed transport system.

The approximate approach we would then propose for city-wide analyses in this case is as follows. State I "decentralised" land-use in the table below is a realistic pattern for the "do-nothing" case; State II has more activities "concentrated" in the central city. State II can be achieved only with transport improvements, i.e. by "doing something".

FUTURE YEAR LAND USE			
Do-Nothing		Do-Something	
State	Item	State	Item
I decentralised	1	I decentralised	2
II concentrated	3	II concentrated	4

Comparison of Items 2 and 1 would give benefits arising from the transport system alone, whereas comparison between Item's 4 and 3 gives an erroneous (over-) estimate[1]. In the event that synergism is expected between transport improvements and concentration of land-use (as may well be the case), the transport benefits (possibly costs) arising from concentration may be obtained by comparing Items 4 and 2. The total benefits arising from achieving Item 4 may be given as

= (Non-Transport benefits (less costs) of achieving 4) + (4 - 2) + (2 - 1).

It will be seen that explicit estimates of the non-transport benefits and costs from changing land-use are required.

(1) The comparison is between the trip matrices and generalised cost matrices for the relevant items, using one of the methods discussed above for calculating user benefits.

In passing, we could note that Item 3 compared with 1 may give a rough estimate of the transport benefits (probably costs) of going "concentrated" in the absence of transport improvements.

Some Statistical Implications of Demand/Gravity Model Calibration

The study of "gravity" models is reported to originate from Ravenstein's (1885) analysis of migration flows. Despite the effort expended over the years, some further progress is required in developing statistically appropriate methods of testing hypotheses, according to Cliff and Ord (1975), given the frequent existence of spatial dependence amongst the observations.

In most such models it is implicitly assumed that the effects of trip-making of: (i) the characteristics of origins, and (ii) the characteristics of destinations, are dependent, in the statistical sense. In relation to destinations it is hence assumed that the relative attractiveness for any given origin i of the jth destination with respect to the kth for $C_{ij} = C_{ik}$ is given by $(A_{1j}/A_{1k})=(A_{2j}/A_{2k})= \ldots = (A_j/A_k)$. In other words, it is assumed that the relative attractiveness of any two destinations is independent of the origins. The implication is that T_{ij} is hypothesized to be proportional to A_j. Cesario (1973) has termed this property the Ω property. A similar property, Θ, may be stated with respect to the origins: $(0_{i1}/0_{11})=(0_{i2}/0_{12})= \ldots = 0_i/0_1$ given $C_{ij} = C_{1j}$.

The hypothesis Θ implies that every destination receives trip-makers in the same ratio from every pair of origins, provided that $C_{ij} = C_j$ for all i and j. Alternatively, θ assumes that the relative propensity of any two origins to produce trips is independent of the destinations and, hence, that T_{ij} is proportional to O_i. Tests for these hypotheses are rarely applied.

The effect upon tests of inference of spatially auto-correlated observations should be mentioned. Firstly, spatial independence is required for the standard application of the *t* and *F* statistics for the comparison of means or the construction of confidence intervals. Secondly, the assumption of spatial independence is necessary for the error terms in regression analysis, if the ordinary least square estimators are to be "best linear unbiased". In the case of gravity models based on the assumption of observations possessing spatial independence, little is known about their robustness to departures from the assumption of independence[(1)]. However in a very recent article, Cliff and Ord (1975b) discuss tests for spatial dependence.

Suffice it to say that it is possible to harbour doubts about the statistical validity of the process employed in this and other studies to estimate the "demand" (gravity) model. The substantial precedent for the somewhat rough methodology is comforting, however.

FUTURE EMPHASES

Some of what we suggest in the following as being appropriate directions for future work arise out of the prior discussion. Those that do not (for reasons of space) we hope are reasonably self-evident.

---

(1) Cliff and Ord (1975 a).

Attention should be given to ensuring, after the manner of Beardwood and Kirby (1975), that the compressibility and separability properties are approximately complied with.

The challenging task of assessing the implications for evaluation of modelling the dynamic phenomenon of road congestion as a steady-state (i.e. static) one, should be proceeded with.

Greater emphasis should be directed to the modelling of those important categories of trips (business and commercial, shopping, social and recreational) where the assumption of inelastic demand for trips just does not hold in the longer-run.

Attempts should be made to uncover, through empirical studies, the ways in which and the degree to which travel time savings (of freight vehicles, especially) flow through into increases in economic efficiency.

If assignments onto fairly detailed networks are to be the basis of an interactive procedure though the distribution, modal split and assignment stages in the search for a "better" set of equilibrium speeds, then some attention should be given to several aspects of the assignment process. These include:-

(i) the testing (through the development of appropriate software) of the adoption, as the route-selection criteria, of minimization of cost, rather than time. While this would involve extra computation (to achieve a link cost), it should avoid the phenomenon of trips being loaded onto quicker, but more costly (i.e. longer) routes. It may also assist successful iteration.

(ii) the development, probably through trial-and-error procedures, of speed/flow parameters which are oriented to successful iteration while producing satisfactory flows and speeds.

More attention should be given to the testing of the hypothesis of spatial independence, if statistical inference about means, confidence intervals and so on is to be performed.

CONCLUDING COMMENTS

Given the difficulties and approximations involved in using MTM's, if they are going to be used for social cost-benefit analysis which provides a basis for or support for major transport policies, then there is probably no alternative but to develop adequate skills, experience and common-sense in both the modelling and the economics sides. What is needed are multi-disciplinary persons, not just teams.

The importance of subjecting: (i) predictions about population growth and land use, (ii) speeds in the future do-nothing case, (iii) speeds on the new links in the future do-something case, as well as all components of the analysis, to considerable judgment and commonsense cannot be over-stressed, as these combine to dominate the estimates of benefits. Yet grander and more comprehensive models will not invalidate this. The transport economists should make a contribution in these matters.

In this paper we have tried to lay the foundations for a bridge between the O/R and economic perspectives about Macro Transport Modelling. Others who have attempted the crossing might appreciate the difficulties.

REFERENCES

Beardwood J.E. and Kirby H.R. (1975) Zone Definition and the Gravity Model. *Transportation Research* 9(6) : 363-369.

Boadway R.W. (1974), The Welfare Foundations of Cost-Benefit Analysis, *The Economic Journal* 84(336) : 926-939 December 1974.

Cesario F.J. (1973), A Generalised Trip Distribution Model, *Journal of Regional Science* 13 (2): 233-247, August, 1973.

Cliff A.D. and Ord J.K. (1975a), Model Building and the Analysis of Spatial Pattern in Human Geography, *Journal of the Royal Statistical Society* (Series B), 37(3) : 297-328.

Cliff A.D. and Ord J.K. (1975b), The Comparison of Means when samples consist of Spatially Autocorrelated Observations, *Environment and Planning A* 7(6): 725-73A, September, 1975.

Cochrane R.A., A Possible Economic Basis for the Gravity Model, *Journal of Transport Economics and Policy* 9 (1): 34-49, Jan. 1975.

Fairhurst M.H. (1975), The Influence of Public Transport on Car Ownership in London, *Journal of Transport Economics and Policy*, Vol. IX No. 3, September 1975.

Fisk C. and G.R. Brown, The Role of Model Parameters in Trip Distribution Models *Transportation Research* 9 (2/3) 143-148, July 1975.

Green H.A. (1971) *Consumer Theory*, Penguin, U.K.

Harrison A.J., (1974) *The Economics of Transport Appraisal*, Croom Helm, London.

Hensher D.A. (1972), The Consumer's Choice Function : A Study of Traveller Behaviour and Values, *unpublished Ph.D. thesis*, School of Economics, University of New South Wales.

Hotelling H. (1938), The General Welfare in Relation to Problems of Railway and Utility Rates, *Econometrica* 6:549-532, July, 1938.

Hyman G.M. and A.G. Wilson, The Effects of Changes in Travel Costs on Trip Distribution and Modal Split *The High Speed Ground Transportation Journal* 3 (1): 79 85, January 1969.

Kuh E. (1959), The Validity of Cross-Sectionally Estimated Behaviour in Time Series Applications, *Econometrica* 27(2): 197-224, April 1959.

Nash C.A., Pearce D. and Stanley J.K. (1975), An Evaluation of Cost-benefit Analysis Criteria, *Scottish Journal of Political Economy* 22(2): 121-134, June 1975.