WORKING PAPER 362

TRANSPORT AND THE EVOLUTION OF URBAN SPATIAL STRUCTURE

A. G. WILSON

WORKING PAPER
School of Geography
University of Leeds

TRANSPORT AND THE EVOLUTION OF URBAN SPATIAL STRUCTURE *

Alan G. Wilson

Summary

Methods of dynamical analysis developed in location theory are applied to the problem of the evolution of transport systems. It is shown that the influence of transport variables on urban structure can be modelled; here, the reverse is attempted. The analysis proves to be difficult because of the combinatorial problems associated with large networks and a suggestion is explored for making progress using the concept of 'spider' networks.

August 1983

* To be presented at the A.I.R.O. Conference in Naples, September 1983.

1. Introduction : developments in location theory and their implications for transport modelling

The first flush of urban modelling in the 1960s and early 1970s was concerned with population activity and only to a lesser extent with economic activity and the supply of houses, services and so on. This research leaned heavily on earlier developments in transport modelling, mainly again concerned with demand rather than supply. Since the late 1970s, there have been useful developments in modelling the supply subsystems of a comprehensive urban model and it is now appropriate to turn the wheel full circle and to consider the implications of the methodology underpinning these developments for modelling the supply of transport infrastructure.

The new developments in urban modelling to be described here stem from the work of Harris and Wilson (1978) applied to retailing systems (but relevant to a variety of similarly-structured service systems). The developments are reported in more detail in Wilson (1981). It has later been shown that the ideas can be extended to residential location and housing supply and even industrial location and agricultural location - see Wilson (1983) for a review. As we will see, transport-related variables play an important role as exogenous variables in all these models and the problem to be posed is, in effect, how to make them endogenous.

There is, of course, a substantial literature in the transport journals which is relevant to the topic in hand. Key references are the special issues of Transportation Research, B on Transportation network design (Boyce, 1979) and on Transportation supply models (Florian and Gaudry, 1980). The first has a direct relevance to the modelling of network evolution - through seeking optimum additional links, and so on; the second has a broader relevance, particularly in the specification of supply-cost functions, the importance of which we will see later. There is also a useful general paper on transport supply by Manheim (1980) and a later extension of their own framework by Florian and Gaudry (1983).

The objective of this paper is to tackle some of the problems which are common to this literature with a method which is translated from locational analysis. At this stage, it is possible to develop the skeleton of the idea only, and at a later stage it is hoped that some more effective integration will be possible.

In section 2 below, we sketch in barest outline the essence of recent research to date. It turns out that supply-side developments turn on the difference between costs and revenue (or benefits). In section 3, therefore, we set up the appropriate variable for describing transport system supply with particular reference to the nature of costs and benefits. The model developed is summarised formally in section 4. We then consider the fundamental problem of the evolution of network structures (and other supply-side variables) in section 5. This allows us to tackle the problem of transport-land-use interaction and we make some concluding comments in section 6.

2. Urban spatial structure and evolution as a function of transport system variables

A typical urban model which represents both supply-side and demand side behaviour can be put together in a general way as follows (from Wilson, 1983). Consider m-type organisations (or people) in zone i demanding goods or services of type g from organisations of type n in zone j. Then the interaction array - the intersection of supply and demand - can be taken as

$$Y_{ij}^{mng} = Y_{ij}^{mng} (X_i^{mg}, W_{ij}^{mng}, c_{ij}^g) \tag{1}$$

where X_i^{mg} is demand by m for g at i, W_{ij}^{mng} is the attractiveness of (n,j,g) supply for (m,i,g) demand and c_{ij}^g is the cost of travel from i to j for g. This would be a model of consumers' behaviour and it can take a variety of explicit forms.

Let $\underline{h}_j^{ng}$ be a vector of characteristics measuring supply of g by n at j and $\underline{a}_j^{ng}$ a set of corresponding parameters. Then, formally, contracting W_{ij}^{mng} to W_j^{ng},

$$W_j^{ng} = W_j^{ng} (\underline{h}_j^{ng}, \underline{a}_j^{ng}) \tag{2}$$

If X_i^{mg} is measured in money units, then

$$D_j^{ng} = \sum_{im} Y_{ij}^{mng} \tag{3}$$

is the total revenue attracted to the (n,j,g) combination.

Let Z_j^{ng} be the amount of g produced by n at j. Then at least one of the elements of $\underline{h}_j^{ng}$ will be a function of Z_j^{ng}. Let $q_{j\ell}^{ng}$ be the ℓ th input needed to produce Z_j^{ng} and let $p_{j\ell}^{ng}$ be the unit price of that input. Then the cost of production is

$$C_j^{ng} = \sum_\ell p_{j\ell}^{ng} \; q_{j\ell}^{ng} \tag{4}$$

with

$$q_{j\ell}^{ng} = q_{j\ell}^{ng} (Z_j^{ng}) \tag{5}$$

and

$$p_{j\ell}^{ng} = p_{j\ell}^{ng} (q_{j\ell}^{ng}, \ldots) \tag{6}$$

to allow for economies or diseconomies of scale. Then a typical assumption about supply-side dynamics is

$$\dot{Z}_j^{ng} = \varepsilon^{ng} (D_j^{ng} - C_j^{ng}) \; Z_j^{ng} \tag{7}$$

with possible equilibrium states as the solutions of

$$D_j^{ng} = C_j^{ng} \tag{8}$$

There are enormous complexities beneath the apparent simplicity of this formal presentation. When all the appropriate substitutions have been made the systems (7) or (8) - for disequilibrium or equilibrium modelling respectively - are coupled non-linear simultaneous equations in the supply-side variables $\{Z_j^{ng}\}$. Solutions can disappear or change their nature at critical values of parameters and the system can jump between alternative equilibrium states as a result of pertubations. 'Historical accidents' could have a crucial impact on the form of system development.

There has been much exploration of the kinds of system state which can arise, and the nature of transformations between states, mainly using numerical experiments in relation to idealised systems. Examples of possible states for retail supply as a function of two parameters (α associated with attractiveness, β with ease of travel - the larger α, the more important are consumer scale economies; the larger β, the more difficult in general travel is) are shown in Figure 1 which is taken from Clarke and Wilson (1983).

The transport system has an obvious influence on these models through the arrays $\{c_{ij}\}$. Indeed, in Figure 2, we show modifications to the pivotal case of Figure 1 ($\alpha = 1.3$, $\beta = 3.5$) obtained by factoring city centre costs by 0.95, 0.85 and 0.75 respectively. The scale and nature of the influence is obvious. However, these c_{ij}-variables are exogenous. The next step in the argument is to make them endogenous, first by relating interaction to congestion; and secondly, by making assumptions about the development of transport supply. We define variables and tackle the first (and most traditional) of these issues in the next section. First, however, we make a remark which generalises the formal model presented in equations (1)-(8) above.

Equations (1) and (2) contain a hypothesis about consumers behaviour, (3)-(6) represent the 'production function' and the way it is perceived, while (7) and (8) are alternative hypotheses about supply-side behaviour. The remark is this: any of these components can be modified without changing the essence of the main idea - interdependence and non-linearities will produce bifurcations. In particular, the supply-side equations could be modified in a public sector case to maximise benefits subject to a budget constraint.

3. Making transport-system supply variables explicit

Each element of the $\{c_{ij}\}$ array which appears in the models above should, ideally, though it is not always practicable, be taken as a 'generalised' cost. Suppose we distinguish mode by a superscript, k. Then take, for example

$$c_{ij}^{k} = m_{ij}^{k} + a_1^k t_{ij}^{k} + a_2^k e_{ij}^{k} + p_i^{(1)k} + p_j^{(2)k} \tag{9}$$

where m_{ij}^{k} is the out-of-pocket money cost of the journey, t_{ij}^{k} the travelling time, e_{ij}^{k} forms of 'excess' time - such as waiting time for public transport, and $p_i^{(1)k}$ and $p_i^{(2)k}$ the terminal costs at i and j respectively. For a car driver, m_{ij}^{k} will be marginal costs, such as petrol and $p_j^{(2)k}$ will be a combination of parking charges at j and any time (appropriately weighted) spent walking from car park to final destination. For the public transport traveller, m_{ij}^{k} will represent fares, e_{ij}^{k} the time spent waiting - and so will be inversely related to frequency of service, $p_i^{(1)k}$ will be time spent between home and the public transport facility and $p_i^{(2)k}$ the time to reach the final destination (and each of these may include all the costs of journeys by subsidiary modes). The coefficients a_1^k and a_2^k represent different values of time. Corresponding definitions could be produced for freight transport costs, say c_{ij}^{gk} for type of goods g by mode k, though in the rest of the discussion below we restrict ourselves to person trips for convenience.

This representation is of the <u>average</u> consumer's perception of different components of disutility. This will suffice for present purposes, and clearly begins to make the transport supply variables explicit. We can summarise the position reached in this respect as follows:

(i) characteristics of the links of the road network will determine car (and other road vehicle) travel times - say t_{ij}^{1} if we take k = 1 to be the car mode - through a rather complicated procedure which we describe shortly;

(ii) a combination of network provision and vehicle provision and operating procedures will determine public transport travel times, t_{ij}^{2};

(iii) petrol costs and taxes (say) will determine m_{ij}^{1};

(iv) fares policy will determine m_{ij}^{2};

(v) frequency of service decisions will determine e_{ij}^{2};

(vi) network design in the form of spacing of the routes in relation to housing (and the design of any subsidiary modes) will determine $p_i^{(1)2}$ and $p_j^{(1)2}$;

(vii) the provision of parking spaces and the policy for charging for it will determine $p_j^{(2)1}$.

The next stage in the argument is to show how the travel time element of generalised cost relates to network provision. Consider first car travel times, t_{ij}^{1}. Consider a single (m,n,g) category in equation (1) for convenience and rewrite this as

$$Y_{ij} = Y_{ij}\ (t_{ij}^1,\ \text{other variables}) \tag{10}$$

for car trips. The difficulty is that while Y_{ij} is obviously a function of t_{ij}^1, t_{ij}^1 is equally obviously a function of congestion levels which are determined by the Y_{ij}s. The next step, therefore, is to assign the flows, calculated on a guessed set of t_{ij}^1's, to the network. Let (r,s) be a link of the network and let x_{rs} be a measure of its 'size'; let Q_{rs} be the flow on link (r,s). Let R_{ij}^{min} be the set of links which make up the best route from i to j (measured in terms of generalised cost). Then

$$t_{ij}^1 = \sum_{r,s\ \varepsilon\ R_{ij}^{min}} \gamma_{rs}^1 \tag{11}$$

where γ_{rs}^1 is the travel time for car on link (r,s) and

$$\gamma_{rs}^1 = \gamma_{rs}^1\ (Q_{rs},\ x_{rs},\ \ldots) \tag{12}$$

and, to complete the circle,

$$Q_{rs} = \sum_{i,j\ \varepsilon\ V_{rs}^{min}} Y_{ij} \tag{13}$$

where V_{rs}^{min} is the set of trip interchanges (i,j) for which link (r,s) is on the best route.

The key exogenously given transport supply variable is now x_{rs}. Given this, the equations (10)-(13) can be solved iteratively: guess t_{ij}^1, find Y_{ij} from (10), find Q_{rs} from (13), find γ_{rs}^1 from (12) and then t_{ij}^1 from (11). Recalculate Y_{ij} from (10) with new t_{ij}^1, and repeat until convergence is achieved. Equation (12) is the key equation connecting the x_{rs}-variables to this system: it is a time-(or sometimes presented as speed-) flow relationship which obviously depends on the physical nature of the link as represented by x_{rs}. An example of the use of the U.S. speed-flow relationship in the network design problem is provided by Dantzig, Harvey, Landsdowne, Robinson and Maier (1979).

A corresponding analysis can be carried through for public transport. In the case of a rail network, the congestion effects in the form above can be neglected and routes and timetables can be planned directly. This is also true to some extent for bus networks, but in constructing timetables it will be necessary to take account of the impact of car-traffic congestion in its interaction with the buses. There is a complication arising from the fact that the public transport agency plans routes. Suppose the agency operates a set of routes R = 1,2, ... Then we need to define sets analogous to R_{ij}^{min} and V_{rs}^{min}, say $R_{ij}^{(2)min}$ and $V_{rs}^{(2)min}$, anticipating a later development in notation. It might also be useful to define, say, $G_{ij}^{(2)}$ as the set of <u>routes</u> to be used in getting from i to j and $H_R^{(2)}$ as the set of (i,j) pairs which use route R at some stage. A further complication, of course, is that a <u>passengers'</u> route (PR) from i to j may involve more than one agency route (AR). The assignment problem is then analogous to that for car users : travellers should be allocated to the most advantageous route in generalised cost terms. Then, not only can the travel time be calculated in a relationship analogous to (11) (though this should now include any inter-route transfer times), but also the appropriate fare calculation can be made.

In the discussion of assignment above, it has been assumed that travellers take the least (generalised) cost route in each case. In practice, there is more likely to be some dispersion with second-best, third-best, and so on, routes being used to an extent determined by the generalised cost differences between them. For the purposes of this paper, we simply note that such procedures can be incorporated without undue difficulty and would not be expected to change our results here in any significant way.

We can now summarise the discussion so far by noting that the technical supply-side variables to be specified are $\{x_{rs}, e_{ij}^{2}, p_i^{(1)2}, p_j^{(2)2}\}$. The pricing variables to be determined are $\{m_{ij}^{1}$ (the petrol tax part), m_{ij}^{2} (fares), and $p_j^{(2)1}$ - parking charges$\}$. The pricing variables are obviously related to the technical ones and will depend on broader aspects of policy, such as the requirement of the public transport agency to break even or not. Then, bearing this summary in mind, we can proceed to a discussion of the costs of supply.

At least three time scales can be identified over which it is relevant to consider costs : the very long-lasting capital investment in networks; the shorter-term capital investment in, for example, public transport vehicles; and the recurrent running costs of particular systems. Decisions on these different scales are often relatively independent. We need, therefore, the capital costs for building (or extending) a link (r,s) for mode k at 'size' x_{rs}, say $\Gamma_{rs}^{k}(x_{rs})$. Let $\rho_{rs}^{k}(x_{rs})$ be the recurrent costs of maintaining and running such a link. In the public transport case, we have already seen the significance of supply routes and it will be better to relate shorter-term capital costs and non-network running costs to these. Let these be $\Phi_R(y_R)$ and $\phi_R(y_R)$ respectively for providing 'capacity' y_R on the Rth route (though note that these will be dependent on network supply also). We pursue the analysis further in the next subsection with more explicit assumptions about the form of these cost functions.

Finally, we need to define, at least formally, measures of benefit associated with a particular system state. For private car users, the usual measure is consumers surplus (though it could be something simpler, like generalised cost savings). For a recent survey of the problems of benefit measurement, particularly building on the concept of consumers surplus, see Jara-Diaz and Friesz (1982). This can also be applied to public transport users, or alternatively in this case, some market mechanism can be used if this is reflected in the operating policy of the public transport agency. For the present, we can let $B^k(\{x_{rs}\}, \{y_R\}, \{m_{ij}^1\}, \{m_{ij}^2\}, \{p_j^{(2)2}\})$ be the benefit to users of mode k arising from supply-side decisions and policies and we will pursue the consequences of more explicit specifications in section 5 below.

4. Summary of a formal model

The discussion in section 3 has been informal with the main ideas of each relevant submodel being explained in turn. It is now useful to draw these ideas together more formally, to extend the notation and to make it more consistent where appropriate. As a first step we show the main submodels and their relationships in Figure 3. This is, in effect, a diagram for a comprehensive urban model re-arranged from its usual form to accentuate the relationships which are relevant to the transport subsystem. There are so many obvious feedback loops in the system that there is no clear cycle of causation. One of the interesting issues in the discussion below is how to tackle this question. The main output variables from each stage are shown on the diagram and any amendments to notation will be explained in the summary below.

The interaction variable Y_{ij}^{mng} in equation (1) needs to be broken into (m,n,g) categories for the purposes of modelling urban structure. From a transport viewpoint, however, we are more interested in split by mode, k and we would aggregate over the other indices provided the flows can all be translated into appropriate units - say passenger car units (p.c.u.'s). Here, and in subsequent equations, we use an asterisk to denote summation and also, where necessary, a conversion to appropriate units. We assume, therefore, that equation (1) can be written

$$Y_{ij}^{***k} = Y_{ij}^{***k}(X_i^{mg}, W_{ij}^{mng}, c_{ij}^k) \tag{14}$$

(where we now replace c_{ij}^g by c_{ij}^k).

As a shorthand, we can then make W_{ij}^{mng} a function of provision at j, Z_j^{ng} and travel costs, c_{ij}^k, so that the urban structure submodel can be written

$$W_{ij}^{ng} = W_{ij}^{ng}(Z_j^{ng}, c_{ij}^k) \tag{15}$$

$$D_j^{ng} = \sum_{imk} e_i^{mg} Y_{ij}^{mngk} \tag{16}$$

$$C_j^{ng} = K^{ng} Z_j^{ng} \tag{17}$$

(making another approximation for simplicity, taking costs as proportional to scale of provision, and e_i^{mg} is the expenditure per trip by (i,m) organisations or people for g). If we then concentrate on equilibrium states for illustrative purposes, we can say that Z_j^{ng} is the solution of

$$D_j^{ng} = C_j^{ng} \tag{18}$$

and this is a set of linked non-linear simultaneous equations when all the substitutions are made from (14)-(17) into (18).

The next step is trip assignment to transport networks. Generalised cost can be taken as in (10), which is repeated here for convenience:

$$C_{ij}^{k} = m_{ij}^{k} + a_1^k t_{ij}^{k} + a_2^k e_{ij}^{k} + p_i^{(1)k} + p_j^{(2)k} \tag{19}$$

This makes Y_{ij}^{***k} in (14) dependent on t_{ij}^{k}. We then proceed with a common notation for each of the two main modes, but later adapt this to recognise that it is feasible for public transport agencies to plan routes but that this is not the case for road planners. We also have to confront the problem that buses use the highway network and usually share congestion with cars.

Let $R_{ij}^{min(k)}$ be the set of links (r,s) which form the least generalised cost route from i to j for mode k. In the public transport case, the algorithm will have to be so designed that the successive links are on agency routes, R, and that any transfers are feasible. Let $V_{rs}^{min(k)}$ be the set of interchanges (i,j) for which (r,s) is on the best route. Then equations (11)-(13) can be rewritten :

$$t_{ij}^{k} = \sum_{r,s \,\varepsilon\, R_{ij}^{min(k)}} \gamma_{rs}^{k} \tag{20}$$

$$\gamma_{rs}^{k} = \begin{cases} \gamma_{rs}^{1}\,(Q_{rs}^{1}, Q_{rs}^{2}, x_{rs}^{1}, \ldots),\ k = 1 \\ \gamma_{rs}^{2}\,(Q_{rs}^{2}, Q_{rs}^{1}, x_{rs}^{2}, y_R, \ldots),\ k = 2 \end{cases} \tag{21}$$

$$Q_{rs}^{k} = \sum_{i,j \,\varepsilon\, V_{rs}^{min(k)}} Y_{ij}^{***k} \tag{22}$$

We have subdivided the car and public transport link travel time equation (21) to show a conventional speed-flow relationship for k = 1 (but including total public transport useage as exogenous where this is loaded on to buses sharing the link) and to show public transport times depending on load and route planning (with Q_{rs}^{1} formally present for the congestion link), though we also show γ_{rs}^{2} to be a function of x_{rs}^{2}, the relevant link capacity.

The next appropriate step is to recap on costs and benefits. Link capital costs are

$$\Gamma_{rs}^{k} = \Gamma_{rs}^{k}\,(x_{rs}^{k}) \tag{23}$$

with recurrent costs

$$\rho_{rs}^{k} = \rho_{rs}^{k}\,(x_{rs}^{k}) \tag{24}$$

In theory, it would be possible to combine these into, say, annual costs by the use of an appropriate discounting rate, but network links typically have such a long life that it seems best to keep these separate. It is considered that all other car operating costs are borne by the user; and that public transport route-running costs can, in this case, be combined into a simple recurrent figure

$$\phi_R = \phi_R\ (y_R) \tag{25}$$

where y_R is the level of activity on route R. We assume that waiting time is also a function of the y_R's:

$$e_{ij}^2 = e_{ij}^2\ (y^R, \ldots) \tag{26}$$

Another task which we neglect for the time being, but which can easily be reintroduced, is the fixing of the array of public transport fares $\{m_{ij}^2\}$. For simplicity, we neglect petrol taxes and all origin costs, and assume that parking charges, p_j^1 , are policy variables. The variables to be determined in the transport supply model, yet to be specified, are then

$$\{x_{rs}^k, y_R, p_j^k\} \tag{27}$$

The specification of this model will be the main task of section 5.

5. The evolution of urban spatial structure, including transport systems

We have already seen, in Figure 2 in relation to Figure 1, that transport systems have a major impact on urban structure. It is possible, and interesting, to explore such impacts more systematically. In Figure 2, we explored, in effect, the impact of alternative underlying networks involving different degrees of central orientation. Formally, this could be considered to come about from x_{rs} changes which generate the kinds of c_{ij} change which were investigated. In that kind of single mode situation, we could also consider the impacts on structure of other network configurations combined with alternative policies on parking charges. It would be more interesting, of course, to develop a two-mode model and to explore the effects of alternative policies for the public transport mode, both in terms of networks and pricing. These explorations could be carried out for a structure in a single system - such as a retail system, as used for Figures 1 and 2, or for a more comprehensive base, such as a modified Lowry model. This analysis could be sharpened further by the search for critical values of transport system parameters which demarcate different forms of urban structure. These explorations will all be reported in a later paper. For present purposes, it is more important to turn to new theoretical questions and first look at the question in reverse - how do transport systems evolve? - and then to look at the joint evolution of urban structure and transport systems.

To make progress, the model given in section 4 has to be extended by the detailed specification of cost and benefit functions together with a model-generating hypothesis about how the transport supply agency works. To fix ideas, consider the single (car) mode case and suppose that recurrent costs can be neglected. It also makes sense to assume that links are being added to the network incrementally from a given situation.

We have to imagine the existence of a cost function $\Gamma_{rs}(x_{rs})$ for each possible link (r,s). This will vary most obviously with size, x_{rs}, but also with local topography and land use. It will be lower in flatter rural country, avoiding the need for bridges, avoiding demolition of buildings and so on. The set of Γ_{rs}'s represents a detailed statement of network building possibilities for an area, and of course it is continually changing as the area develops. To make this explicit, let us use the label t to represent a time period, say t to t + 1, and let $\Gamma_{rs}^{t}(x_{rs}^{t})$ be the cost of building a link x_{rs}^{t} in that period. Let $\underline{N}^{t}$ be the set of links already existing up to that period and $\underline{L}^{t}$ a representation of the set of land uses. Then, formally, Γ_{rs}^{t} could be written $\Gamma_{rs}^{t}(x_{rs}^{t}, \underline{N}^{t}, \underline{L}^{t})$ to show its dependence on these quantities.

The aggregate benefit function, B^{t}, will be the benefits deriving from the building of a set of links $\{x_{rs}^{t}\}$ in time period t and this will also depend on the existing configuration: $B^{t}(\{x_{rs}^{t}\}, \underline{N}^{t}, \underline{L}^{t}, \{Y_{ij}\})$. We also show it as dependent on the flows, $\{Y_{ij}\}$. This could be measured, for example, as the additional consumers surplus accruing to travellers from the building of $\{x_{rs}^{t}\}$.

We now need to specify an appropriate policy for the network building agency. Suppose it wishes to maximise benefits, suitably discounted, over a series of periods subject to a total budget constraint operating in each period. This can be represented as

$$\underset{\{x_{rs}^{t}\}}{\text{Max}} \quad B = \sum_{t=0}^{T} B^{t}(\{x_{rs}^{t}\}, \underline{N}^{t}, \underline{L}^{t}\ \{Y_{ij}^{t}\}) / (1+R)^{t} \tag{28}$$

subject to

$$\sum_{r\,s\,t} \Gamma_{rs}^{t}(x_{rs}^{t}) \leq K^{t} \tag{29}$$

where R is a suitable discount rate and K^{t} is the total budget for period t. Since the benefit function would involve a knowledge of all the flows, Y_{ij}^{t}, then the problem represented by (28) and (29) implicitly involves the full networkconstrained transport model as a set of constraints.

It is interesting to compare this formulation with that for the calculation of $\{Z_j^{ng}\}$ described in section 2. In the latter, we focus on Z_j^{ng} for each zone j, and the number of possible configurations, although very great, are much less than involved in the network problem as formulated above. In principle, there is a very large number of x_{rs} combinations. In the problem as formulated in an incremental model, x_{rs} can be taken as varying continuously or discretely. As we noted earlier, both approaches are used in the literature. In the other problem the size at a location can more easily be treated as a continuous variable over time. There are also, of course, fewer variables. It is interesting to attempt to redefine the network problem to make it more like the facility-centre problem, and then alternative modelling formulations could be used to replace (28) and (29) which may be more amenable to analysis. This involves seeking new ways of making approximations.

One possibility is to fix the main nodes of the network (perhaps the centroids of origin and destinations), to connect each node to a number of near neighbours and thus to define what used to be called a 'spider network' in the early days of transport planning when it was difficult to handle large networks. An example is shown in Figure 4. The r's and s's, then become i's and j's and R_{ij}^{min} is the set of i''s and j''s which form the best route between i and j, and so on. In Figure 4, $R_{1,14}^{min}$ might be (1,2,7,8,9,14) for example. We could also let S be the set of (i,j) pairs for which notional spider links have been defined. The problem could then become one of determining the capacity of each such link, x_{ij} - bearing in mind that it could, of course, be zero. We could then use a difference equation formulation (choosing this rather than a differential equation because it seems more sensible to work in terms of annual, say, benefits and costs):

$$\Delta x_{ij}^{t} = \varepsilon(\Delta B_{ij}^{t} - R\Gamma_{ij}^{t}(\Delta x_{ij}^{t}))\Delta x_{ij} \tag{30}$$

where Δx_{ij}^t is the increment in capacity over the period t to t+1, ΔB_{ij}^t the gain in benefits from this increment and Γ_{ij}^t its cost. R is a suitable discount rate. If the total budget was exceeded, the problem could be rerun with R set higher. The equilibrium condition would be

$$\Delta B_{ij} = R\Gamma_{ij} \tag{31}$$

to be solved for Δx_{ij} after all the relevant substitutions have been made.

Let us call the models we have developed so far SM1 (supply model 1) and SM2 respectively. In the first case, probably the only reasonable way to proceed is to define a pool of possible projects, $\{x_{rs}\}$. But even then, there will be a large number of possible combinations of these which will complicate the calculation of the benefit function in each case. In the second case, it is important to note that a hierarchical analysis is implied. The increment Δx_{ij}, of capacity to be added to a notional link between i and j has, at a lower and subsequent stage, to be transformed into 'sensible' additions to the actual network. This in turn implies a feedback between the two levels on costs: real costs at the lower, network, scale, have to be translated into (approximate) functions at the higher (i-j) scale.

Numerical experiments would be possible with either SM1 or SM2. However, a better chance of analytical insights seem to be offered by SM2 and so we proceed with that. We spell out the section 4 model for the single-mode assumptions we have been working with and explore the consequences. We simplify further by assuming a given demand sector which is proportional to the population distribution $\{P_i\}$ and a single service sector $\{W_j\}$ which is determined within the model. This can be seen as the first step towards the analysis of the joint evolution of transport systems and urban structure.

The demand for transport is Y_{ij}, say, given by

$$Y_{ij} = A_i e_i P_i W_j^{\alpha} e^{-\beta c_{ij}} \tag{32}$$

$$A_i = 1/\sum_k W_k^{\alpha} e^{-\beta c_{ik}} \tag{33}$$

e_i is the per capita demand for services measured in units of trips per head of population. We also assume that W_j is the size of facilities at j and that this can represent attractiveness. Then revenue at j is given by

$$D_j = f\sum_i Y_{ij} \tag{34}$$

where f is a constant translating trip units into money units and the costs are

$$C_j = k_j W_j \tag{35}$$

for suitable constants, k_j. W_j is the solution of the equilibrium condition

$$D_j = C_j = k_j W_j \tag{36}$$

c_{ij} in these equations is taken as a generalised cost:

$$c_{ij} = m_{ij} + at_{ij} \tag{37}$$

where a is the value of time. It is determined through the assignment model as follows: let R_{ij}^{min} be the set of spider-modes which form the shortest path from i to j; let V_{ij}^{min} be the set of (i',j') trip bundles which use the link (i,j). Let S be the set of links (i,j). Let γ_{ij} be the travel time on a link. Then

$$t_{ij} = \sum_{i'j' \in R_{ij}^{min}} \gamma_{i'j'} \tag{38}$$

$$\gamma_{ij} = \gamma_{ij}\,(Q_{ij},\, x_{ij} + \Delta x_{ij}) \tag{39}$$

$$Q_{ij} = \sum_{i'j' \in V_{ij}^{min}} Y_{i'j'},\ (i,j)\varepsilon S \tag{40}$$

Capital costs are given by

$$\Gamma_{ij} = \Gamma_{ij}\,(\Delta x_{ij}) \tag{41}$$

The benefit function is

$$\Delta B_{ij} = \Delta B_{ij}\,(\Delta X_{ij}, \{Y_{ij}\}) \tag{42}$$

The $\{\Delta x_{ij}\}$ to be chosen are then the solutions of

$$\Delta B_{ij} = R\,\Gamma_{ij} \tag{43}$$

Three functions remain to be specified to complete the model: γ_{ij} in equation (39), Γ_{ij} in (41) and ΔB_{ij} in (42). In the case of the first two, we can show what we would expect in graphical form, and this is done in Figure 5. In the case of ΔB_{ij}, we can use a result first derived in the context of random utility theory that change in consumers surplus is:

$$\Delta B = \frac{1}{\beta} \log \frac{\sum_{ij} e^{-\beta c_{ij}^{(1)}}}{\sum_{ij} e^{-\beta c_{ij}^{(0)}}} \tag{44}$$

where the superscripts (0) and (1) on the cost terms denote 'before' and 'after' respectively. There is a difficulty that this relates to the whole system. We would have to define ΔB_{ij} as the value of ΔB resulting from a change Δx_{ij} with all other Δ_{ij}'s zero. This, as we will see later, exposes a version of the 'backcloth' problem - as in Wilson and Clarke (1979). Thus

$$\Delta B_{ij} = \Delta B\,(\Delta x_{ij} \mid \Delta x_{i'j'} = 0,\ i' \neq i\ ,\ j' \neq j) \tag{45}$$

Let us now cycle through a number of imaginary runs of the model. Suppose the population is growing and spreading and this is reflected in $\{P_i\}$; let $\{e_i\}$ be growing even more rapidly to signify a period of increasing car ownership; suppose the m_{ij} are fixed; let α be increasing and β decreasing; assume fixed values of any other constants and known functional forms where appropriate. Assume an initial value for the travel cost array $\{c_{ij}\}$. The crucial decision to take is then how to 'break into' the model and to cycle through it (cf. Figure 3). We adopt one illustration here but recognise that there is a difficult research question involved. We identify the major steps in turn:

(1) Solve equations (32)-(36) for $\{Y_{ij}\}$ and $\{W_j\}$, probably using an incremental procedure for $\{W_j\}$ - meaning that it cannot decrease (or can only decrease by a proportion) from one time period to the next.

(2) Solve equations (37)-(40) with $\Delta x_{ij} = 0$ initially. Calculate Q_{ij} from (40), γ_{ij} from (39) and t_{ij} from (38). Then return to step (1) with new c_{ij} from (37) and recycle between (1) and (2) until equilibrium is achieved.

(3) Take a trial set $\{\Delta x_{ij}\}$, rerun steps (1) and (2) to obtain a new $\{c_{ij}\}$, say $\{c_{ij}^{(1)}\}$. Calculate ΔB from (44), and carry out separate runs to estimate ΔB_{ij}, the contribution to ΔB from Δx_{ij}. (This could be done, approximately by running the model system for the Δx_{ij} change only with all other $\Delta x_{i'j'} = 0$). Use (43) to obtain an R_{ij} from

$$R_{ij} = \Delta B_{ij}/\Gamma_{ij} \tag{46}$$

and then scale Δx_{ij} so that

$$\Delta x_{ij}^{new} = \Delta x_{ij}^{old} \cdot \frac{R_{ij}}{R} \tag{47}$$

Recycle with these new Δx_{ij} from step (2). This is analogous to the balancing procedure for calculating $\{W_i\}$.

We can now see whether the model would do what we expect it to do by assuming the scenario sketched at the beginning of this discussion. Let us add that in one particular suburban zone there is a substantial additional increase in population. Steps (1) and (2) of the model would produce new centres and measures of congestion. A possible before and after situation is sketched in Figure 6. In step (3), the procedure should then generate high rates of return on links which relieve this congestion. So the model should deal correctly with basic development.

This sketch does lead to some new notions, however. First, there should be a minimum available transport capacity to anywhere in the region (representing country lanes or whatever) to allow for the possibility of new $\{W_i\}$ development. Secondly, it is clear that a highway agency may be other than simply responsive in the manner assumed here. Particular Δx_{ij}'s may be implemented to facilitate development, for example; and this would have an appropriate impact on W_j's through step (1) above. Thirdly, it is clear that in any real particular case, much work would have to be done on the detailed specification of parameters and functional forms. Finally, it is clear that the model, which is already complicated through step (3), will be immensely more complicated when realistic detail is added back into it.

6. Concluding comments: an ongoing research programme

It is appropriate to conclude with comments under three main headings about the ways in which these ideas can be taken forward. First, we consider the further theoretical advances which are necessary; secondly, we look at how to make the models more realistic; and thirdly, we examine the potential utility of these models.

There are two main aspects to the first heading. First, it would be valuable if the models could be made sufficiently explicit that it was possible to carry out the kind of analysis presented in Harris and Wilson (1978) for retail systems. That is, to explore analytically the nature of ΔB_{ij} and Γ_{ij} as functions of Δx_{ij} - and the equivalent of these in more complex formulations - and to use this as the basis of gaining insight into the nature of the transitions from one kind of equilibrium state to another. We can be certain that the degree of interdependence and non-linearities involved will lead to jumps, for example. Secondly, it will be necessary to be more explicit about the linking of the two levels in the hierarchy - the capacities on the spider network as a more realistic network at a finer level of resolution. As usual, it may well be that significant progress would be possible through the carrying out of numerical experiments on idealised systems.

The second main heading referred to the task of building more realistic models. The centre of this is the specification of the three main functions which will determine much of the outcome: on costs, on benefits and on speed-flow relations (or, in the last case, travel time - investment relations). A further step involves developing the model in realistic complexity - for example, to handle multi-modal situations. In this case, it will be possible to formulate the model in such a way that the focus can be on the effect of particular variables - say public transport fares or parking provision and charging - and this would then connect more directly to some of the previously cited literature on transport supply. It may also be possible to develop the cost functions in such a way that it is possible to represent links of road in different levels of a hierarchy.

Thirdly, we need briefly to review the utility of the approach. There are two aspects to this. First, it would be interesting to look at the long run history of the evolution of network structures in particular places and to attempt to interpret this using the models and concepts developed. Secondly, as with the more traditional literature on transport supply, it should be possible to find ways of using the models in a policy context. The first of these tasks may be in a theoretical sense easier than the second, because forecasting, for example, is difficult because of the influence of perturbations and historical accidents. But it should be possible to seek 'best additions' and also to investigate the stability of structures.

References

Boyce, D.E. (ed.)(1979) Special issue on transportation network design, Transportation Research, 13B, pp. 1-90.

Clarke, M. and Wilson, A.G. (1983) The dynamics of urban spatial structure: progress and problems, Journal of Regional Science, 23, pp. 1-18.

Dantzig, G.B., Harvey, R.P., Landsdowne, Z.F., Robinson, D.W. and Maier, S.F. (1979) Formulating and solving the network design problem by decomposition, Transportation Research, 13B, pp. 5-17.

Florian, M. and Gaudry, M. (1980) Special issue on transportation supply models, Transportation Research, 14B, pp. 1-220.

Florian, M. and Gaudry, M. (1983) Transportation systems analysis: illustration and extensions of conceptual framework, Transportation Research, 17B, pp. 147-153.

Harris, B. and Wilson, A.G. (1978) Equilibrium values and dynamics of attractiveness terms in production-constrained spatial-interaction models, Environment and Planning, A, 10, pp. 371-388.

Jara-Diaz, S.R. and Friesz, T.L. (1982) Measuring the benefits derived from a transportation investment, Transportation Research, 16B, pp. 57-77.

Manheim, M. (1980) Understanding 'supply' in transportation systems, Transportation Research, 14A, pp. 119-135.

Wilson, A.G. (1981) Catastrophe theory and bifurcation: applications to urban and regional systems, Croom Helm, London; University of California Press, Berkeley.

Wilson, A.G. (1983) Location theory: a unified approach, Working Paper 355, School of Geography, University of Leeds.

Wilson, A.G. and Clarke, M. (1979) Some illustration of catastrophe theory applied to urban retailing structures, in M. Breheny (ed.) London Papers in Regional Science, Pion, London, pp. 5-27.

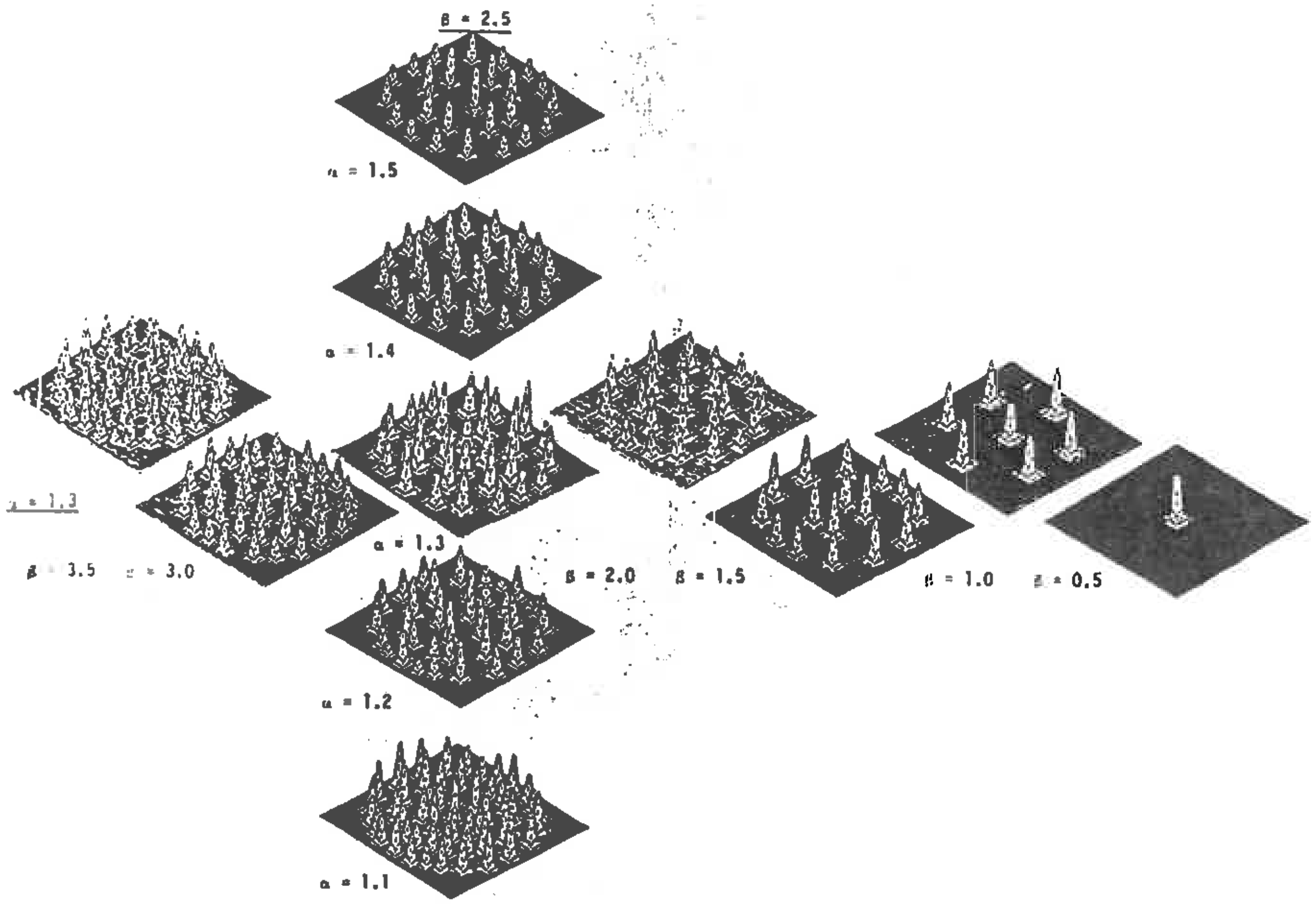

FIGURE 1 : Retail Patterns for Various Alpha and Beta Values

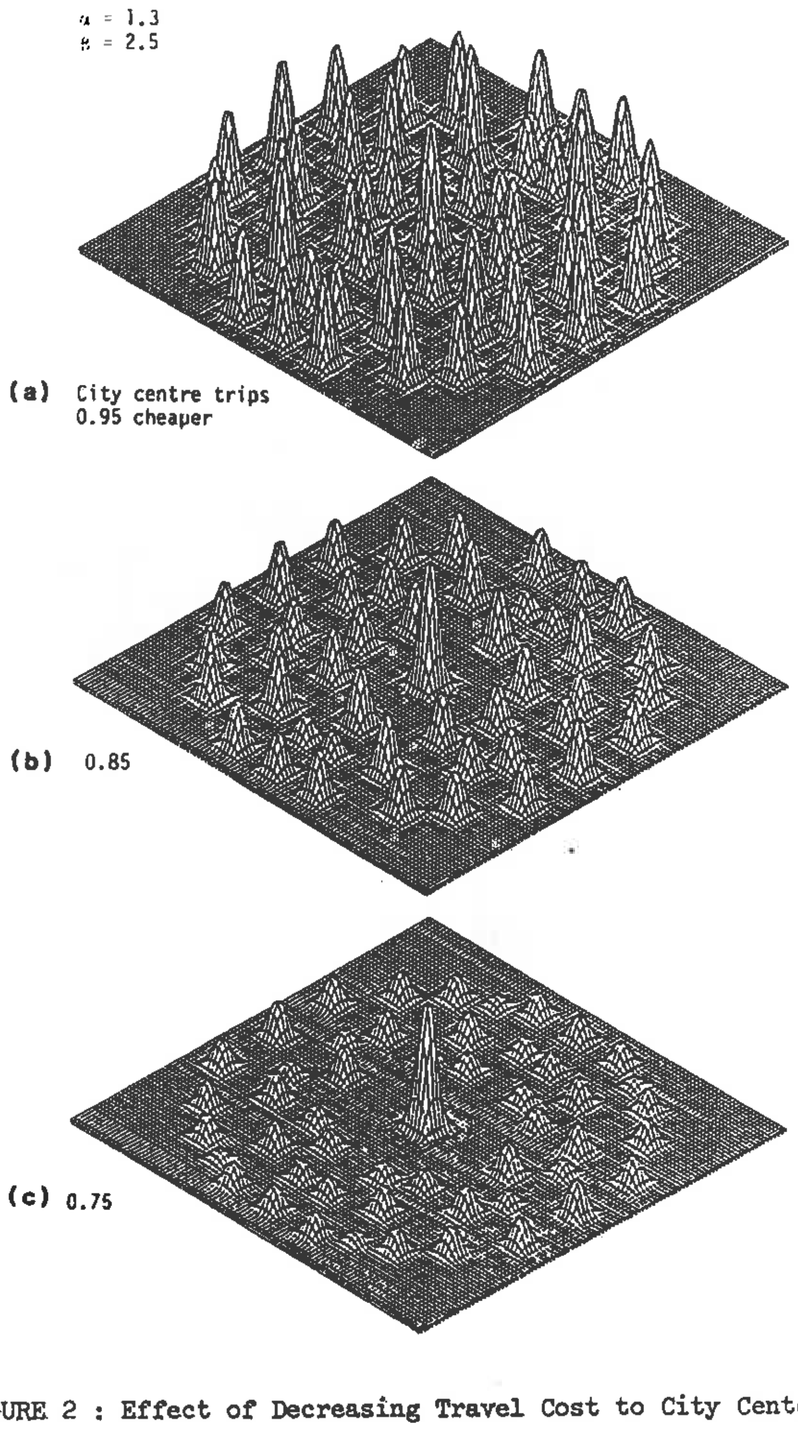

FIGURE 2 : Effect of Decreasing Travel Cost to City Center

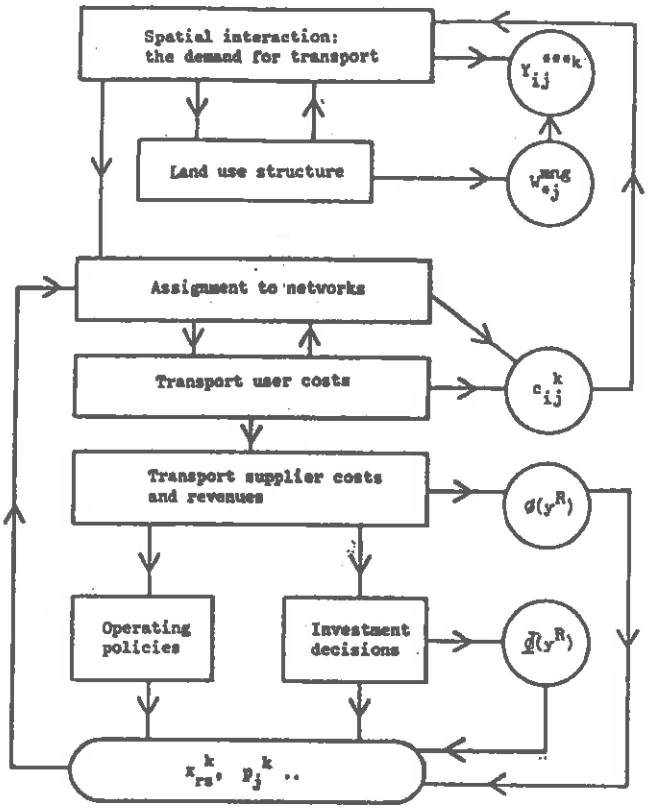

FIGURE 3 : Elements of a general model, emphasising transport

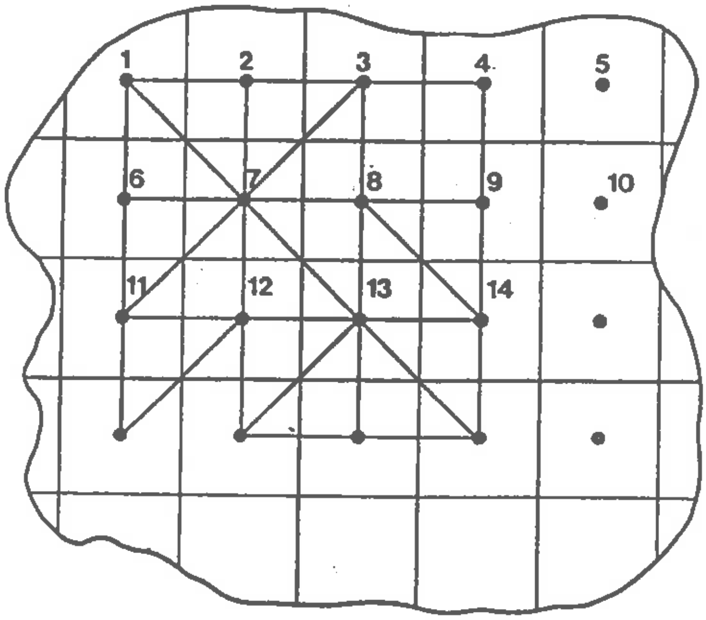

FIGURE 4 : A 'spider' network

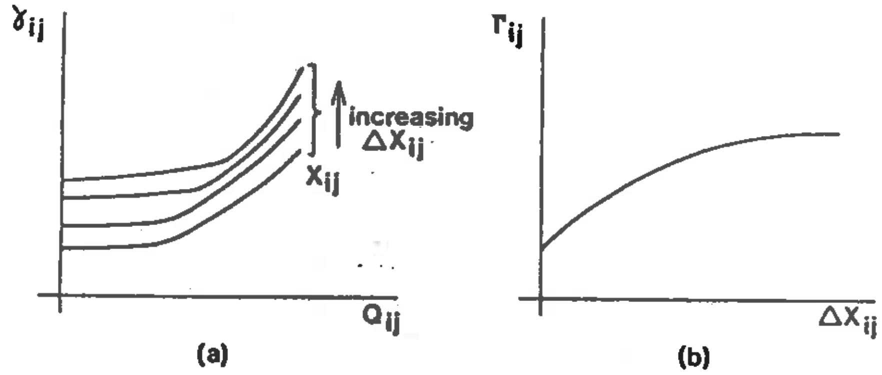

FIGURE 5 : (a) Time-flow relationships (b) Capital costs vs scale of investment

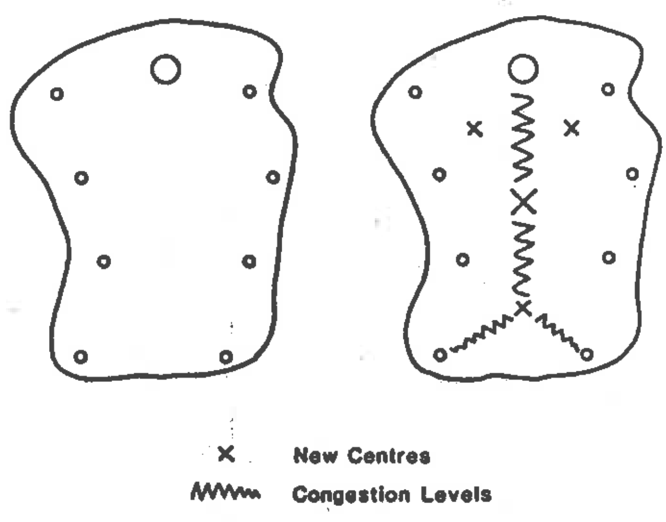

FIGURE 6 : Sources of demand for new transport infrastructure