WORKING PAPER 390

LOCATION THEORY

A.G. WILSON

WORKING PAPER
School of Geography
University of Leeds

WORKING PAPER 390

LOCATION THEORY : A UNIFIED APPROACH

A.G. WILSON

School of Geography
University of Leeds
Leeds LS2 9JT.

September 1985

LOCATION THEORY: A UNIFIED APPROACH

A G Wilson

1. Introduction

There are two directions of approach to location theory (as succinctly and elegantly pointed out by Lowry, 1967). The first focusses on the spatial distribution of activities (type by type) across space; the second on the set of activities, the land-use mix, at a particular place. The customary focus is the first, but both need to be integrated for a location theory to be effective because of the interdependence arising from different activities competing for land. In this paper, we are concerned with the range of theories available for the analysis of various types of activity location and we present a framework within which these can be approached and viewed in a unified way. We also show that interdependencies can be introduced into this framework which provide the basis for solving the problem of the land-use mix and related issues.

Typical fields of location theory, essentially defined by the type of activity being located, are: agricultural location, public utilities, industrial location, private services, public services and housing. Some integration has been attempted, usually at rather coarse levels of resolution, with approaches ranging from central place theory to variants of the Lowry (1964) model. An important distinction should be borne in mind at the outset: that introduced by Paelinck and Nijkamp (1975) between land-consuming activities, of which the main examples in the above list are agriculture and housing, and the rest, for which location can be considered, as a first approximation, to be at a point. (In areas of intense economic activity, like central business districts, this approximation breaks down of course.)

The 'classical' approaches to these problems appear to be very different in style. Consider, for example, von Thunen (1826) on agricultural location, Weber (1909) on industrial location, Burgess (1927) and Hoyt (1939) on residential land use. Or consider later generations of theorists in the same three fields, say Found (1971), Hoover (1967) and Alonso (1960, 1964). These all differ from the early attempts at integrated theory by Christaller (1933) and Losch (1954). As a further contrast, consider the 'modelling' approach of Lowry (1964), much of it based on the concepts of spatial interaction modelling (Lakshmanan and Hansen, 1965, Wilson, 1967, 1974). At a later stage still, approaches to modelling the supply side have been added, as in Harris and Wilson (1978).

The purposes of this paper are: (i) to review briefly a set of rules for model design and to show that this provides a basis for the classification and understanding of alternative approaches; (ii) to present a unified approach based on modelling ideas and to show that many other contributions to location theory can then be seen as special cases; and (iii) to provide the basis for further extensions. This is largely achieved by drawing together ideas which have been presented elsewhere in an attempt to create a new synthesis.

We begin (Section 2) by reviewing a set of rules for model design and then present, in these terms, the approach to location theory to be adopted in the rest of the paper (Section 3). In Section 4, we comment on some earlier theories in relation to this framework and show how they can be incorporated in models which illustrate the unified approach. In Section 5, we make some concluding comments and outline ongoing research tasks.

2. Rules for model design: the basis of a theory-classification system

We begin by reviewing briefly (and, in one case, expanding) the rules for model design which have been presented elsewhere (eg. in Wilson, 1981-A). It is argued that the model builder takes decisions in relation to the following dimensions:

(i) entitation - the enumeration of individuals, organisations and other components of the (sub) system of interest.

(ii) levels of resolution: sectoral (how to categorise components), spatial and temporal.

(iii) partialness/comprehensiveness: whether to model, at one extreme, a single unit taking everything else as a given environment - a kind of marginal analysis; or whether to be comprehensive and handle numbers of (usually competing) units.

(iv) spatial representation - discrete or continuous.

(v) elements of *theory*:

(a) processes shaping the 'environment';
(b) processes associated with consumer behaviour (the 'demand' side);
(c) processes associated with producer/organisation behaviour (the 'supply' side);
(d) any 'whole system' criteria: planning policies, maximisation of consumers' surplus or whatever;
(e) the representation of interdependence;
(f) the development of explanatory concepts.

(vi) development of techniques for building models which represent the theory.

By asking the question, how did X (von Thunen, Weber, etc.) take these decisions, new insights can be gained into old models; alternative approaches can often be seen; and the basis for a classification is provided. We show examples of this in Section 4, but first present the broad outlines of our own approach set against this framework.

3. The general approach advocated

We describe the decisions on model design which constitute the approach to be recommended here in a slightly different order to that used in the previous section. First, we note that it appears to be nearly always more convenient to use a discrete zone system as the basis for characterising location and for modelling. The number of zones determines the fineness of spatial resolution adopted. In general, at least for this author, more powerful mathematical tools seem to be available for building the resulting models. It is also easier to handle land-use constraints. Indeed, at fine levels of resolution at least, discrete zones are more 'natural' since activities are located on units of land of finite size.

Another feature of the representation is determined by sectoral resolution. Typically below, sectors are distinguished by zone totals, not individuals or firms. In other words, we work with variables like 'population in an income group in a zone' (P_i^w, say) or 'retail floor space for furniture in a zone' (W_j^g, say). Again, this turns out to be mathematically convenient and to enable processes based on competition to be represented to a reasonable degree of approximation without having to deal with the geographical equivalent of the n-body problem in physics: a large number of individual competing households or firms in continuous space.

These decisions together also imply a third: to attempt to build comprehensive models, treating as many elements as possible endogenously and seeking explicitly to represent competitive and other interacting processes. It is a combination of the discrete zone system and representing the main entities as 'totals by zone' which allow this to be done more effectively than in other approaches.

We can then turn to the various elements of theory. We take the points in the order listed under (v) in Section 2 above.

(a) The 'environment'. Once the system of interest is defined, we also try to model where appropriate (as distinct submode's) the main exogenous variables which form a backcloth. For many locational models, for example, this involves providing population and macro-economic backcloths.

(b) Consumer behaviour is usually modelled in terms of spatial interaction between residences and point of supply. The most difficult part of this exercise is to characterise the 'attractiveness' of alternative supply points in realistic detail. These models can be made consistent with utility maximising approaches if appropriate, or other alternative theories. An important element of most models deployed here is that consumers do not necessarily go to the *nearest* facility. This is more in accord with reality than the alternative hypotheses of many economic theories. The specification of the attractiveness function can be seen as incorporating the benefits and costs from the point of view of the consumer including the representation of consumer scale economies. A crucial element of this approach is that 'demand' can be allocated to 'service points' *taking account* of competition, and the summation of flows at each destination is a measure of 'revenue' or its equivalent.

(c) On the supply side, it is easiest to model private sector profit maximising behaviour, as we will see. However, since profit is defined as 'revenue minus costs', we can in principle substitute 'benefits' for revenue and define some appropriate public sector hypotheses. Otherwise, different kinds of hypotheses have to be used. As we saw, the model of consumers' behaviour estimates revenue or benefits. There may also be an element of spatial interaction phenomena associated with costs, for example, in the costing of inputs from different places, taking account of transport costs and again allowing for the effects of competition. Each producer may be assumed to seek the cheapest inputs (possibly with some 'dispersion' built in as with the model of consumers spatial interaction).

(d) There are two kinds of examples of system-wide effects (and we distinguish them from the 'environmental backcloth' of (a) above because they are such an integral part of the system). First, a number of market prices, for many goods for example, will be fixed at this level (or bases for prices to which a locational variation, say for transport costs, is added within the locational model). A good deal of economic and social theory needs to be added to specify how these prices are determined: in the market, in relation to supply and demand curves (and, in turn, production functions and utility functions); or through other processes, such as the struggle between capital and labour. Secondly, there is the exercise of the whole range of policy instruments of government in its various forms (expenditure, regulation, fiscal, form of administrative organisation - see Wilson, 1974, Chapter 2) and the way in which these affect the system of interest.

(e) Interdependence is important in many ways. Variables which are exogenous in one subsystem will often be endogenous in another, and this forms an important set of linkages. Another important feature of interdependence is provided by the transport infrastructure of the sytem as a whole. (The way this develops is also an important field of study in its own right - see Wilson, 1983-A).

(f) An important stage is the development and deployment of particular theoretical concepts as the building blocks of theories. The main principles to be advocated here are two. First, as far as possible, define concepts directly in terms of elements of the system. (It is more effective, for example, to work with a variable which is the proportion of those who live in one zone and shop in another than to invoke a concept of non-overlapping market areas.) Secondly, employ Occam's Razor. The second principle in a sense modifies the first where appropriate. If it is *necessary* to use 'higher-level' abstract concepts, then that is OK; but in a critical spirit.

The final step in the argument is to assemble the necessary techniques, usually mathematical in the case of modelling, but not necessarily so, which can be used to build a formal model of a theory - which is to say, a clearly-formulated theory - from all the definitions, ideas, concepts and hypotheses which have been drawn together using the above procedures. This can often be crucial. It is only relatively recently, for example, that the techniques of catastrophe theory and bifurcation theory have become available from mathematics to enable some fundamental problems in locational theory to be essentially solved. Before that, what can be seen with hindsight is a struggle to solve these problems with inadequate tools.

It is useful to show in a formal way what these kinds of models look like. Let i, j (and, on occasion, k) be locational labels, m, n (and *r*) population groups or economic sectors, and g (and h) a product of some kind - a good or a service. Any of the indices m, n, r or g can be replaced by lists if more resolution is required. Let Y_{ij}^{mng} be the flow of g from sector m in i (an output of that sector) to sector n in j - an input to a sector, or a flow to a sector of final demand. It is convenient to think of the demand term, whether intermediate or final, 'pulling in' quantities from competing suppliers. Let $\hat{X}_j^{ng}$ be demand for g by n at j (either intermediate or final) and let $\dot{W}_i^{mg}$ be the attractiveness of suppliers of g, in sector m, at i. Let c_{ij}^g be the unit cost of transport of g from i to j. Then, formally

$$Y_{ij}^{mng} = Y_{ij}^{mng}(W_i^{mg}, \hat{X}_j^{ng}, c_{ij}^g) \quad (1)$$

Usually, as we will see when we specify an example of the model in more detail below, it will be appropriate to use an attraction-constrained spatial interaction model for (1).

Our main interest is likely to be in a set of variables $\{Z_j^{ng}\}$, the total product of g in sector n at j. To make this focus clear, it is necessary to distinguish intermediate demand and final demand. For simplicity, we will assume that each sector is so defined that it demands either a set of intermediate inputs or final demand. Intermediate inputs are calculated on the basis of input-output coefficients $\{\alpha^{ngh}\}$, so that $\alpha^{ngh}Z_j^{nh}$ is the set of inputs g needed from all sectors m to produce Z_j^{nh} units of h in n at j and $\sum_h \alpha^{ngh}Z_j^{nh}$ is the total demand for g in n at j. Final demand is taken as X_j^{ng}. Then, in (1), we have the two cases:

(i) Intermediate demand flows:

$$W_i^{mg} = W_i^{mg}(Z_i^{mg}, \ldots) \quad (2)$$

$$\hat{X}_j^{ng} = \sum_h \alpha^{ngh} Z_j^{nh} \quad (3)$$

(ii) Final demand

$$W_i^{mg} = W_i^{mg}(Z_i^{mg}, \ldots) \quad (4)$$

$$\hat{X}_j^{ng} = X_j^{ng}. \quad (5)$$

Equations (2) and (4) imply that W_i^{mg} will be a function of availability of products $\{Z_i^{mg}\}$, though of course these quantities will be simultaneously determined by solving the whole system of equations. They may also be functions of other variables, such as price.

It can be seen from Figure 1 that a rather complicated feedback structure is involved. This can be made clearer if we restrict the use of

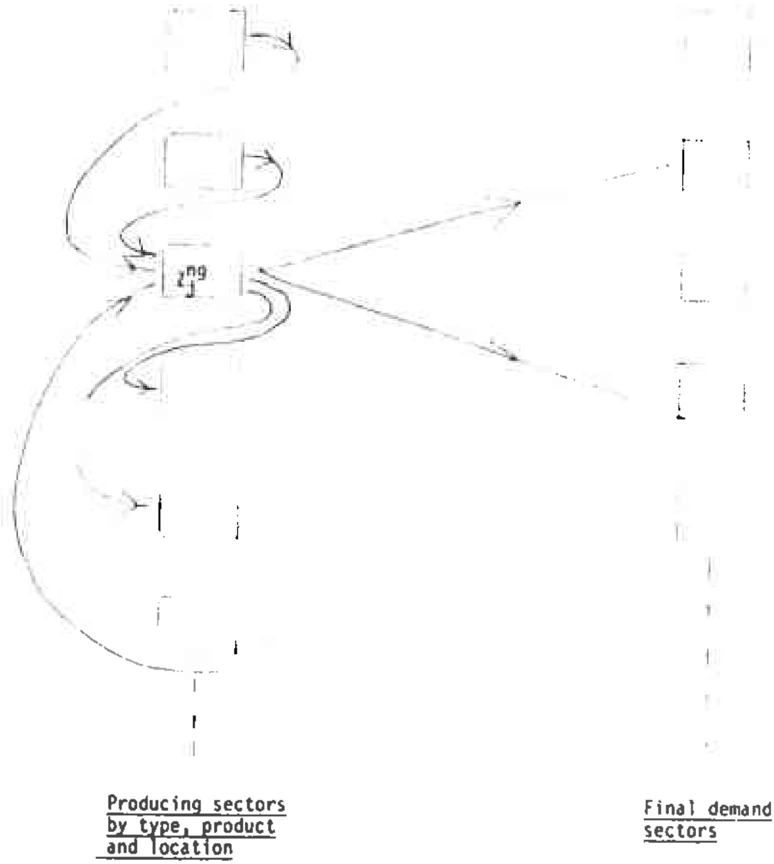

z_j^{ng}
Producing sectors by type, product and location
Final demand sectors

FIGURE 1

i - j spatial labels for input flows to j and use j - k as output flows from j to k. The input-output process centred on Z_j^{ng} can then be demonstrated as follows by rewriting the appropriate elements of equation (1) and the array $\hat{Y}_{ij}^{mnhg}$ defined in relation to equation (6):

$$\hat{Y}_{ij}^{mnhg} = \hat{Y}_{ij}^{mnhg}(W_i^{mh}, \alpha^{nhg}Z_j^{ng}, c_{ij}^h) \tag{6}$$

are the set of inputs, h, to Z_j^{ng}, from a set of sources (i, m).

$$Y_{jk}^{nrg} = Y_{jk}^{nrg}(W_j^{ng}(Z_j^{ng}, \ldots), \hat{X}_k^{rg}, c_{jk}^g) \tag{7}$$

show the destinations (k, r) of the amount of g produced in (j, n).

We will present specific applications of these ideas in Section 4 below. Meanwhile, we can proceed to examine the costs, revenues and dynamics of the production process for g at (j, n).

Assume f.o.b. pricing and let p_i^g be the unit price of i at g. Then total costs for producing Z_j^{ng} are

$$C_j^{ng} = \sum_{ihm} \hat{Y}_{ij}^{mnhg}(p_i^h + c_{ij}^h) + \gamma_j^{ng}Z_j^{ng} \tag{8}$$

where the first term represents the cost of inputs (and $\hat{Y}_{ij}^{mngh}$ is the flow of h to the manufacture of g in n at j and as in (6), disaggregated out of Y_{ij}^{mnh} given by (1) and p_i^g are prices), and the second term represents production costs, assumed proportional to Z_j^{ng} through the unit cost γ_j^{ng}. (It may also be appropriate to include some fixed costs, or more generally to make this part of a more realistic production function).

The revenue collected would be

$$D_j^{ng} = \sum_{kr} Y_{jk}^{nrg}p_j^g. \tag{9}$$

With these forms of costs and revenues, a suitable model of suppliers behaviour in n at j would be:

$$\dot{Z}_j^{ng} = \varepsilon^{ng}[D_j^{ng} - C_j^{ng}]Z_j^{ng} \tag{10}$$

(cf. Harris and Wilson, 1978, Wilson, 1981, for detailed discussion of this in the retailing case). The ε^{ng}s are constant.

The equilibrium state, which may never be achieved but which may influence trajectories, would be the solution of the simultaneous equations

$$D_j^{ng} = C_j^{ng}. \tag{11}$$

Either the set of simultaneous equations (10) can be used (for a fully dynamic model) or the equations (11) (for comparative statics). Models of both types have been extensively explored in the context of retailing (starting with Harris and Wilson, 1978). A key point to be emphasised in this paper is that by appropriate definition and representation of the indices (n, k, j) and of the quantities D_j^{nk} and C_j^{nk}, a much wider range of subsystems becomes susceptible to modelling using these methods and it is this idea which provides the basis for unifying location theory. We will illustrate the argument in detail in Section 4 below.

Equations (1) - (11) look relatively straightforward when presented in this formal way. The formality, however, masks enormous complexity, and we can begin to see this by spelling out the issues related to the next two items on the theory-building list. These are the related topics of prices and interdependence and we start with the broader second topic. We take each equation in turn and show how key variables have to be made functions of other variables in order to capture potentially important interrelations.

In equation (5), final demand, X_j^{ng}, will be a function of offered price, the prices of substitutes, and the transport cost to sources of supply (so this will not simply be c_{ij}^k but a composite measure such as accessibility). So we might write

$$X_i^{mg} = X_i^{mg}(\underline{p}, \underline{V}, \ldots) \tag{12}$$

where $\underline{p}$ is a composite vector of prices, $\underline{V}$ a vector of accessibilities and so on. (Other variables would include, for example, the relative wealth of the m-group.)

The characteristics of attractiveness in (2) and (4) will obviously include terms like Z_j^{ng} as shown but also measures of the competition of other n's and j's - so in this sense the balancing factor in a singly-constrained spatial-interaction model, which can be taken to represent competition, can be considered formally to play a role in the measurement of attractiveness. The price vector, $\underline{p}$, is relevant again, now in terms of choice of destination rather than scale of demand. The accessibility vector, $\underline{V}$, will also be relevant. For instance, a visit may be made to get k from n at j if j is also highly accessible to other k's. So we extend equations (2) and (4) formally as

$$W_i^{mg} = W_i^{mg}(Z_j^{ng}, \text{competition}, \underline{p}, \underline{V}, \ldots) \tag{13}$$

There are many potential complexities associated with the inputs to Z_j^{ng}, and their prices. The inputs may well be non-linear functions of the amount produced, to represent scale economies for example, and this would have to be incorporated in a production function via prices, or possibly through more elaborate input-output coefficients. The prices are functions of the prices at input sources and the cost of transporting them, as shown in equation (6), and again, in some cases, it may be important to replace those simple assumptions by those which incorporated, for example, economies of scale.

Note that the inputs should include capital, land and labour as well as goods, services and materials and this can be achieved by appropriate definition of sectors. The element of $\underline{p}$ which is the price of land is particularly important in relation to interdependence. Many activities compete for land and this competitive process (with other influences such as government policy, the history of past ownership and so on) will determine the land price element. If, in effect, no land is available for (n, k, j) activity, then this will show itself as a very high price. This will transmit itself to C_j^{nk} via the production costs in (8), and then into (10) Because the right hand side of (10) will obviously be negative in such as instance, we will have $Z_j^{ng} = 0$. An example is given of this kind of mechanism for residential location and housing in Clarke and Wilson (1983-A).

It is worth trying to pursue the issue of handling land a little further. There are two fundamental difficulties to be dealt with and then a number of consequences follow in relation to the approaches which solve these. The first involves the investigation of the concept of rent, the second, some problems arising from the size of zones in the discrete zone representation. We discuss each of these in turn.

The theory of land rent has its origins in the work of von Thunen (1826) and in its modern form has been much developed by Alonso (1960, 1964) and others. The key idea of relevance here is that of bid rent: that different potential users of a unit of land notionally bid a rent which is the surplus (if normal profits are included in costs) to be gained from economic activity or which is related to utility in the case of housing. The rent which can be gained by the land owner is then the highest bid rent at a particular site. The difficulty posed by this formulation for the present framework is that land rent is not obviously a smooth function of other system variables. If what are specified are schedules of possible profits or consumers' surpluses, then these have to be searched and the market cleared in order to construct the rent surface. In modelling terms, this is an algorithmic process rather than an analytical function of other obvious variables. There are at least two ways of dealing with this in principle. First, to develop an analytic function for rent as an approximation to the process - for example, by making land prices a function of various accessibilities. Secondly, a model of the rent-bidding algorithm could be built in explicitly. This may not be as fearsome as it at first appears since, in a run of the model through time, there would always be the previous period's prices to act as a base; and, initially, empirically-measured prices could be used.

The second problem is concerned with the size of zones. Rent theories work in terms of either price per unit area (and this may in principle vary across a zone) or price for plots of land for particular purposes (which may be parts of zones or may straddle zones). If the first kind of approach is adopted, then an average would have to be assumed for the whole zone; the second is more intractable unless the zone system can continually be redefined in relation to plots (which is impossible in practice). So we can assume a zonal average. The remaining problem is then how to handle a mix of land uses within a zone and to build in appropriate constraints on the total land available in that zone. In a run of the model through time, this is further complicated, as noted above, by the problem of 'converting' old land uses to new ones. We will offer specific mechanisms in the context of examples in Section 4 and Section 5 below.

Arrays like the accessibilities, $\underline{V}$, or the transport costs, $\underline{c}$, are specified in terms of other variables which are already endogenous to the problem. They do introduce new kinds of nonlinearities. What is more difficult to follow through at present is the nature of the elements of the price vector $\underline{p}$. At a macro level, some of these elements can be considered to be determined by the intersection of supply and demand curves in some 'backcloth' aggregate model. Others might be fixed by government policy. Yet others will be functions of the spatial configurations of activities and the associated transport system. Above all, many prices of goods at locations will be functions of the quantities produced, because of scale economies and other nonlinearities. This, of course, adds a major feedback loop to the model system. Formally, we can write

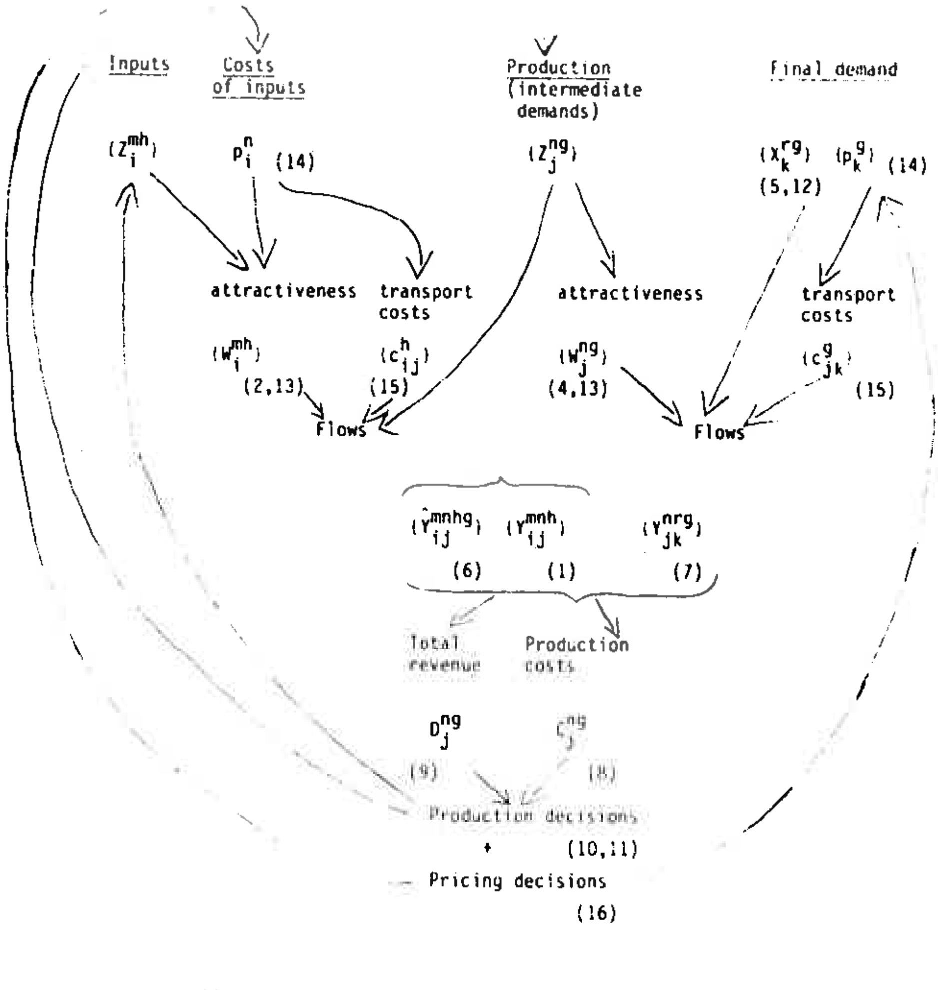
$$\underline{p} = \underline{p}(\underline{X}, \underline{Z}, \underline{c}, \underline{V}, \ldots) \tag{14}$$

and to complete the equation system show transport costs as determined by infrastructure supply which will be a function of demands, government allocations and so on:

$$\underline{c} = \underline{c}(\underline{X}, \underline{Z}, \text{government allocation to transport} \ldots) \tag{15}$$

The whole system is shown in Figure 2 which makes clear many of the important feedbacks which have been built into even a formally-specified model. Note that these are not only feedbacks within the (n, k, j) system shown, but between systems, for all the reasons outlined above and sketched on the figure.

One final observation is in order before we leave this form of the general model. As we have shown elsewhere (Clarke and Wilson, 1983-B), this model system can be run in what is essentially a comparative static mode or as a fully dynamic model, recognising that the corresponding real system will



[Appropriate Section 3 equations numbers are shown in brackets near to each set of variables]

FIGURE 2

be in disequilibrium most of the time though on trajectories which will often be governed by underlying equilibria. The key difference in the model as presented is whether equation (10) or equation (11) is used and these should be seen as the dynamic and static alternatives respectively.

There is an implicit assumption now that adjustments when the model is run in its dynamic mode relate wholly to the quantities produced, as described by equation (10), and that any price adjustments take place instantaneously as in equation (14). An obvious extension would be to add an equation corresponding to (10) for prices, say

$$\dot{p}_j^{ng} = \hat{\varepsilon}^{ng}[D_j^{ng}(p_j^{ng}) - C_j^{ng}(p_j^{ng})]p_j^{ng} \tag{16}$$

where $\hat{\varepsilon}^{ng}$ is a suitable set of constants and we have shown D_j^{ng} and C_j^{ng} as functions of p_j^{ng}. The relative sizes of the constants ε^{ng} in equation (10) and $\hat{\varepsilon}^{ng}$ in equation (16) will then determine the relative importance of adjustment in quantity or price respectively.

The model can now be run in dynamic mode using equations (10) and (16) or in comparative static mode using (11) and (14).

In the next section, we interpret this general model in the terms of various classical contributions to location theory and show how to extend these to fit in with the framework of this section. Meanwhile, we briefly examine the equivalent framework for a public sector system where there are no market prices and so the mechanism implied by equations (10) and (11) is not - at least at first sight - possible. Equation (1) might still describe consumers behaviour and equation (2) attractiveness (though the units of flow will probably be trips and not money). Equation (9) will count total trips with the price factor omitted. Equation (8) still represents costs of supply. Versions of equations (12) - (15) still hold.

The substitute mechanism in this case, at least notionally and possibly in practice, would be to find $\{Z_j^{ng}\}$ by some optimisation method which reflected the policies involved. This involves specifying a function which represents the benefit side of the policies, say $B_j^{ng}(\{Z_j^{ng}\}, \ldots)$. Then we can write

$$\underset{(Z_j^{ng})}{\text{Max}}\ L = \sum_{jng} (B_j^{ng} - C_j^{ng}) \qquad (17)$$

subject to a version of equation (1) as constraint. In this case, we would usually wish to assume a simple input structure and assume that final demand represented a level of service provision which had to be met. Equation (1) would therefore take the form shown in equation (7) and that would be used as a constraint, with a given distribution of demands X_k^{rg}. The r-superscript could be dropped - or used to reflect the different demands of different population groups. It is also likely that other constraints would be added. In particular, following Leonardi (1981), it can be argued that a public authority will maximise net benefits while spending up to its total budget, B, say.

$$\sum_{jnk} C_j^{ng} < B \qquad (18)$$

is therefore an additional constraint (with C_j^{ng} then removed from the objective function). There may also be further constraints to represent the operations of this kind of planning system.

Because B_j^{ng}, C_j^{ng} and the various terms in the constraints are functions of $\underline{W}$, $\underline{p}$, $\underline{q}$, $\underline{V}$ and $\underline{c}$, the equations (12) - (15) also, in effect, function as constraints. It can then be seen that the whole system is a highly nonlinear one. In some cases, the nonlinearities can be removed from the constraints and added into the objective function using so-called embedding theorems (Coelho and Wilson, 1977; Coelho, Williams and Wilson, 1978). For example, if (17) is replaced by

$$\underset{(Z_j^{ng}, Y_j^{mng})}{\text{Max}}\ L = \sum_{jng} (B_j^{ng} - C_j^{ng})$$

$$- \sum_{jkmng} Y_{jk}^{mng} \log Y_{jk}^{mng}$$

$$- \beta \sum_{jkmng} Y_{jk}^{mng} c_{jk}^{g}$$

$$+ \alpha \sum_{jkmng} Y_{jk}^{mng} \log W_j^{mg} \qquad (19)$$

and the constraint

$$\sum_{jnk} Y_{ij}^{mng} = X_i^{mg} \tag{20}$$

added, then because of the addition of entropy and other relevant terms into the objective function, it can be shown that Y_{jk}^{mng}, which is now a variable in the optimisation problem, satisfies a suitable spatial interaction model equation.

It remains an open and interesting research question as to whether the two formats, differential equations and optimisation problem, can be linked - perhaps through the iterative procedures which are commonly used to find solutions in each case and by using shadow prices or some other quasi prices (implicit in the benefit measures) in the second. In the private system case, the suppliers budgets are, in effect, determined by the X_k^{rg} terms. So another possibility is to 'transmit' suppliers cost, notionally, to users through an accounting system and then to use these as quasi-budget totals. This can be done through the implicit accounting system provided by the interaction terms. The public model could then possibly be run in the same way as the private model.

4. Examples: from old to new

In this section, we consider four examples of 'classical' theory and show how the ideas can be recast in the framework of Section 3. We take, in turn, von Thunen's (1826) model of agricultural location, Weber's (1909) model of industrial location, Alonso's (1960, 1964) model of residential land use and Christaller's (1933) central place theory.

von Thunen's theory is based on continuous space and, as is common in such cases, and as we will see again with central place theory, the geographical analysis task is seen as one of charting boundaries, in this case to demarcate different agricultural land uses. The resolution levels are relatively coarse. The approach appears at first sight to be comprehensive because all land is dealt with. But in practice this is not the case because of the single centre and inelastic demand assumptions: the process of farmers competing to serve different centres is missing. The most striking success of von Thunen's model, however, is in introducing the theoretical concept of bid rent whose uses, as we will see later, extend beyond agricultural location. Meanwhile, we explore these central ideas, but recast them in discrete zone format. We work broadly in the notation of Section 3 but with appropriate simplifications. Here, for example, we do not distinguish between sectors and goods - m and n or g, and so we use g as type of crop and let it serve as a label for both. We use k as the location of final demand.

Let Z_j^g be the yield of crop g at location j, p^g its unit price, γ_j^g the unit production costs at j, r^g the unit transport costs assumed paid by the former and d_{jk} the distance from j to the market at k. Z_j^g can be considered to be a function of the intensity of farming measured by the unit cost of inputs, γ_j^g and so could be written $Z_j^g(\gamma_j^g)$. We can avoid a difficulty later if we work with $\hat{Z}_j^g$, the yield from a unit area of land at j from crop g. von Thunen's key idea was then to take

$$E_j^g = \hat{Z}_j^g(\gamma_j^g)(p^g - \gamma_j^g - r^g d_{jk}) \tag{21}$$

as the profit to be derived per unit area from cultivation of g at j. This can also be interpreted as the maximum rent a farmer would be prepared to bid for p at j. In circumstances of perfect competition, E_j^g would have to be paid as rent to the owner of the land.

von Thunen's analysis can now be implemented. Essentially it involves finding the g (and γ_j^g) which maximises E_j^g at each location. He, of course, worked with continuous space. Here, we have added the zone label j. His analysis produces the well known 'rings' of agricultural land use. With the discrete zone system introduced here, such rings would be reproduced in approximate form (if there was one market centre, k): that is, if g (and γ_j^g) were assigned to each j such that (21) was maximised in each case. Indeed, Stevens (1968), by shifting to discrete zones, shows how the model can be represented in mathematical programming terms.

This constitutes a route into location theory via rent and land use. The rent function incorporates the main revenues and costs. We can use the approach offered by equations (8) and (9) to collect these together in a different way and then first reproduce von Thunen's results and secondly generalise them. The crucial first step is to return to the discussion of land rent in Section 3. Define

$$\rho_j = \underset{g}{\mathrm{Max}}\ E_j^g \tag{22}$$

where E_j^g is given by (21) and this can be taken to apply to an area of land L_j. The procedure implied by (22) is one of land owners seeking the highest bidder and being able to charge that as rent. Note that ρ_j is independent of g. The bid rents in (21) would have to be calculated on the basis of a known yield function $Z_j^g(\gamma_j^g)$, varying over j and g. We also need to define a_j^g as the amount of land needed to produce a unit of g at j - and this also could be a function of γ_j^g: $a_j^g(\gamma_j^g)$. For the time being, we also assume a fixed set of cost inputs, $\{\gamma_j^g\}$. One problem should be noted. With discrete zones, there is no reason why this rent surface should even approximate continuity. It may be a good idea, therefore, to fit a function through the ρ_j's produced by (22).

We can now assume that the rent is paid and write down revenue and costs for farmers in zone j who are considering crop g. Recall that there is one market at k and 'final demand' at this point is assumed to be equal to what can be profitably produced (given fixed and exogenous prices).

$$D_j^g = Z_j^g p^g \tag{23}$$

$$C_j^g = Z_j^g(\gamma_j^g + a_j^g \cdot \rho_j + d_{jk} r^g). \tag{24}$$

The growth of $\{Z_j^g\}$ will be given by (cf. (10))

$$\dot{Z}_j^g = \varepsilon^g [D_j^g - C_j^g] Z_j^g \tag{25}$$

and the equilibrium condition will be (cf. (11))

$$D_j^g = C_j^g \tag{26}$$

We will only have $D_j^g > C_j^g$ (for reasonably large values of Z_j^g) when g is the optimal use in the von Thunen sense and (26) will also only be satisfied on the same basis. These results follow from the definition of ρ_j in (22). A further modification is needed: to add a notional increase to the rent if the amount of land available at j is exceeded, and so to prevent such unrealistic increases. The mechanism for this is given in relation to the discussion around equations (32) - (34) below.

At present, we are assuming fixed and given $\{Y_j^g\}$, and we are also assuming that everything which is produced can be sold at the market. The simplest assumption is that Y_j^g is fixed so that E_j^g in (21) is a maximum with respect to it. In this case, the results we derive from an application of the general framework of Section 3 simply reproduce von Thunen's results without adding to them (except that solving (25) from a position of disequilibrium would demonstrate something about the *path* to equilibrium). The next step in the argument is to generalise the von Thunen model.

The most obvious extension is to introduce several markets, k, and prices, p_k^g which vary by location of these (perhaps in relation to local demand). Let X_k^g be demand at k, initially fixed, and let Y_{jk}^g be the flow of g from j to k. We might hypothesise that, say

$$Y_{jk}^g = B_k^g Z_j^g X_k^g e^{-\beta^g c_{jk}} \tag{27}$$

where

$$B_k^g = 1/\sum_j Z_j^g e^{-\beta^g c_{jk}}. \tag{28}$$

The revenue and cost equations become

$$D_j^g = \sum_k Y_{jk}^g p_k^g \tag{29}$$

$$C_j^g = Z_j^g(Y_j^g)(\gamma_j^g + \rho_j) + \sum_k d_{jk} r^g Y_{jk}^g \tag{30}$$

The von Thunen surplus at j from g per unit is

$$E_j^g = \hat{Z}_j^g(Y_j^g)(\sum_k Y_{jk}^g p_j^g - \gamma_j^g - \sum_k d_{jk} r^g Y_{jk}^g) \tag{31}$$

and ρ_j is still defined by (22). $\{Y_j^g\}$ would also be fixed, though ideally optimised, in relation to (21) as before.

The problem as a whole, however, is now more difficult (and interesting) than before, essentially because of the influence of possible markets and the fact that Y_{jk}^g depends on Z_j^g, and on all the other Z_j^g's (in effect representing a competitive process). Thus, solving the equilibrium equation (26) for $\{Z_j^g\}$

is now non-trivial; and, worse, E^g_j in (31), and hence ρ_j, will now depend on the $\{Z^g_j\}$ pattern in what is likely to be quite a sensitive way. The problem in this form can be made obviously more complicated by making X^g_k a function of p^g_k. This is also a matter for further exploration.

There remains the second kind of difficulty (with land use models) mentioned in Section 3 and above in relation to the (23) - (26) model: the problem of taking land area into account and, in the case of a system of relatively large zones, of handling multiple land use in a zone. We have to find a way of letting costs rise very steeply when the land runs out in a zone. In other words, when $Z^g_j(\gamma^g_j)$ reaches a certain level such that the land used exceeds L_j, then an additional 'cost', $\hat{\gamma}^g_j$, has to be imposed. Let $a^g(\gamma^g_j)$ be the land-used per unit yield at level of intensity γ^g_j. Then the additional cost should be

$$\hat{\gamma}^g_j = \begin{cases} b^g & \text{if } a^g(\gamma^g_j)Z^g_j(\gamma^g_j) > L_j \\ 0 & \text{otherwise.} \end{cases} \qquad (32)$$

$a^g(\gamma^g_j)$ is the amount of land used to generate a unit of crop g at j when the unit intensity of input is γ^g_j. b^g is a suitable constant. If necessary, b^g could be replaced by $b^g Z^g_j$ which would ensure not only a substantial immediate increment, but also rapid further increases. Where multiple uses are possible, the condition would have to relate to these:

$$\hat{\gamma}^g_j = \begin{cases} b^g & \text{if } \sum_g a^g(\gamma^g_j)Z^g_j(\gamma^g_j) > L_j \\ 0 & \text{otherwise.} \end{cases} \qquad (33)$$

$\hat{\gamma}^g_j$ should then be added to the γ^g_j costs in (30):

$$C^g_j = Z^g_j(\gamma^g_j)(\gamma^g_j + \hat{\gamma}^g_j + \rho_j) + \sum_k d_{jk} r^g Y^g_{jk} \qquad (34)$$

and this would prevent further development via (25).

What is now needed is a set of numerical experiments to investigate the range of patterns produced and these and further model developments are reported elsewhere (Wilson and Birkin, 1984).

As a second example, we can examine Weber's approach to industrial location. What is commonly referred to as the Weber model deals with the location of a single firm in continuous space. The approach is therefore obviously highly partial (albeit at a fine level of resolution). Weber was well aware of the limitations of his approach and his 1909 book contains illuminating passages which form the basis for something more comprehensive. Unfortunately, he did not have the mathematical tools available to implement his ideas. Our aim now is to show how the methods of Section 3 begin to provide the basis for a solution to many of the problems he formulated and lead the way to a comprehensive model of industrial location. We first present a version of the original Weber problem and then a more modern version of the same problem. We can then extend the latter into a more comprehensive model.

The original model relates a firm to two material sources M_1 and M_2 and a single consumption point C. Let (x,y,z) be the coordinates of the firm and let $a_1(x,y,z)$ be the distance from M_1, $a_2(x,y,z)$ the distance from M_2 and $a_3(x,y,z)$ the distance from C. Then if w_1 is the weight of materials needed from M_1 and w_2 from M_2 to produce w_3 for C, and if transport costs vary with weight and distance only, (x,y,z) is to be found to minimise total transport costs:

$$\underset{(x,y,z)}{\text{Min}} \quad z = \sum_i a_i(x,y,z)w_i. \tag{35}$$

There are algorithms for finding (x,y,z) - see Scott (1971) for a review.

The first task is to examine a multi-firm discrete zone version of the model - which is known in the literature as the 'Weber problem on a network' since the discreteness is taken as a set of nodes which, for our purposes, can be taken as zone centroids. The model, which is usually referred to in the literature as the p-median problem, turns out to be more appropriate for certain kinds of public facility location problems rather than industrial location. It can be written:

$$\underset{\{\lambda_{ij},x_j\}}{\text{Min}} \quad z = \sum_{ij} O_i\lambda_{ij}x_jc_{ij} \tag{36}$$

subject to

$$\sum_j \lambda_{ij} = 1 \tag{37}$$

$$\sum_j x_j = p \tag{38}$$

$$\lambda_{ij}, x_j = 0,1 \tag{39}$$

x_j is 1 if a facility is assigned to j, 0 otherwise; λ_{ij} assigns origin i to facility j. This locates p facilities to minimise total transport costs. Sources O_i are assigned to the nearest facility. As noted above, this problem is more appropriate for certain kinds of services for example when a set of 'needs', $\{O_i\}$ have to be met though in the notation of Section 3 we would then transpose it from an (i,j) problem to a (j,k) problem. (If some of the O_i are taken as consumption points, then it is more like an industrial location model, but in that case connections between inputs and outputs need to be established.) To extend Weber's problem we need to ensure that the correct (if fixed) demands reach consumption points. Let S_ℓ be a set of nodes which supply input material ℓ and let V be a set of consumption points with X_j of product demanded at the node j from this set. Suppose p factories are to be located and there is no constraint on the amount of material to be taken from any material source. Let w_ℓ be the weight of material ℓ needed to produce a *unit* weight of final product. Then consider the following problem:

$$\underset{\{\lambda^\ell_{ij}, \lambda^p_{ij}, x_j, w^p_j\}}{\text{Min}} \quad Z = \sum_{\ell i \in S_\ell, j} w_\ell \lambda^\ell_{ij} Z_j x_j c_{ij} + \sum_{jk \in V} Z_j x_j \mu_{jk} c_{jk} \tag{40}$$

$$\sum_j \lambda^\ell_{ij} = 1, \; i \in S_\ell \tag{41}$$

$$\sum_j \mu_{jk} = 1, \; k \in V \tag{42}$$

$$\sum_j x_j = p \tag{43}$$

$$\sum_j Z_j x_j \mu_{jk} = X_k, \; k \in V \tag{44}$$

$$\lambda_{ij}, \mu_{ij}, x_j = 0, 1 \tag{45}$$

where Z_j is the total weight of product at j, λ_{ij}^{ℓ} assigns sources to factories and μ_{jk} assigns factories to consumption points. The x_j's are as before. This is an obvious extension of the p-median problem. Sources still come from the nearest facility, factories send to the nearest consumption point, transport costs are minimised. It would be a straightforward further addition to add inequality constraints on material supply points if there were limits.

It is clear that the model can be easily extended by a shift to a discrete zone system, but, as presented, it is still limited in important respects. The spatial flow patterns are likely to be more dispersed than the 0-1 $\underline{\lambda}$ and $\underline{\mu}$ variables suggest. But above all, the focus on minimising transport costs is restrictive. It relies on assumptions about fixed demand and spatially invariant prices. By switching to the Section 3 framework, we can easily make all these extensions. The procedure for doing it will now be obvious.

The flow arrays $\{Y_{ij}^{mng}\}$, $\{\hat{Y}_{ij}^{mnhg}\}$ and $\{Y_{jk}^{nrg}\}$ of Section 3 can be used to describe inputs and outputs in industrial location. The 'sectors' m and n can be industrial sectors, but also labour (for inputs), final demand (for outputs) and so on. A decision has to be made about how to model prices, and if we assume that all pricing is fob, so that purchasers pay transport costs, then we can specify costs and revenues for the production of k in n at j as follows:

$$C_j^{ng} = \sum_{imh} \hat{Y}_{ij}^{mnhg}(p_i^h + c_{ij}^h) \tag{46}$$

where the summation is over all inputs h, and

$$D_j^{ng} = \sum_{rk} Y_{jk}^{nrg} p_j^g \tag{47}$$

for the sale of g to all r in all k. Equations (10) and (11) can then be applied directly:

$$\dot{Z}_j^{ng} = \varepsilon^{nk}(D_j^{nk} - C_j^{nk})Z_j^{nk} \tag{48}$$

with equilibrium conditions

$$D_j^{ng} = C_j^{ng}. \tag{49}$$

The next step is the spelling out of the spatial interaction arrays $(\hat{Y}_{ij}^{mnhg})$ and (Y_{jk}^{nrg}). Flows to intermediate demand can be represented in total by

$$Y_{ij}^{mnh} = B_j^h W_i^{mh} (\sum_g \alpha^{mhg} Z_j^{ng}) e^{-\beta^h c_{ij}^h} \tag{50}$$

with

$$B_j^h = 1/\sum_{im} W_i^{mh} e^{-\beta^h c_{ij}^h}. \tag{51}$$

The α^{mhg} are input output coefficients defined as the amount of h needed to produce a unit of g (and coming from m - which could perhaps be dropped). This can be split into $\hat{Y}_{ij}^{mnhg}$ by removing the summation sign in (50). Also

$$Y_{jk}^{nrg} = \hat{B}_k^g W_j^{ng} X_k^{rg} e^{-\beta^g c_{jk}^g} \tag{52}$$

where

$$\hat{B}_k^g = 1/\sum_{jn} W_j^{ng} e^{-\beta^g c_{jk}^g}. \tag{53}$$

For the present, we need simply to note that the elements will be nonlinear functions of the production totals (Z_j^{ng}) and should be constructed so that known input-output relations are preserved. Scale economies can now be taken into account because the prices (p_j^g) can be made functions of production levels of g in j (though, in this representation, this can only be done in an approximate way because we are dealing with sectors in a zone rather than with individual firms). Agglomeration economies can also be handled through the role of the c_{ij}^g terms in (46) (and hence in (48) and (49)) and through the spatial interaction parameters in the various sub-models. Thus two key features which could not be incorporated in the Weber approach can now be tackled. The models which result are of immense complexity and the first stage is understanding their capabilities is to carry out numerical experiments. For this example, these and further model developments are reported elsewhere together with a more extensive review of alternative theories of industrial location when viewed against the

Section 3 framework (Birkin and Wilson, 1984; Wilson and Birkin, 1983).

As a third example, we turn to residential location and housing supply. The 'classical' models are those of the Chicago sociological 'ecologists' school - based on concepts of invasion and succession of territory. Residential structure in cities is then seen as having a ring structure (Burgess, 1927) with a sectoral structure superposed (Hoyt, 1939) and the possibility of multi-nucleated structures (Harris and Ullman, 1945). These theories are essentially descriptive but they do give some idea of the patterns we might seek to reproduce. These patterns have been recorded and reviewed using the techniques of social area analysis which also provides a useful background (cf. Rees, 1979, for a good review). The ecologists approach was based on continuous space and, as in von Thunen's work, focussed on demarcating land uses and identifying boundaries. It operates at a coarse level of resolution, however, and so does not really allow for mixing of different types of residential land use. The approach is, however, comprehensive. The important breakthrough in theory came with Alonso's (1960, 1964) development of von Thunen's ideas in this context: he used the theory of consumer's behaviour and the theory of the firm to calculate the bid rents which householders or firms could make for land to achieve given levels of utility or profit respectively. The consumer utility functions can be seen as representations of preferences for housing.

Alonso's achievement was not only in developing the bid rent concepts in this context, but also in showing that the market could be cleared: highest bid rents ' accepted' to determine a land price surface; utilities and profits maximised. The approach was also based on continuous space and a fine level of resolution, and this made the theory very difficult to implement. An operational version was, however, skilfully constructed by Herbert and Stevens (1960), essentially because they were able to switch to a discrete zone representation and to consider people in groups. They were then able to use a mathematical programming formulation to represent market clearing. Let T_i^{gn} be the number of type n people in type g housing in zone i. Let b^{gn} be the bid rents (or preferences) for type n people for type g houses and let p_i^{gn} be the price to be paid *exclusive of site cost* (but *including* transport costs). Then, if s^{gn} is the amount of land, on average, used by a type n household in a type g house, if there are P^n type n people, and if L_i is the land area in zone i, their model is:

$$\underset{\{T_i^{gn}\}}{\mathrm{Max}}\ Z = \sum_{ign} T_i^{gn}(b^{gn} - p_i^{gn}) \tag{54}$$

subject to

$$\sum_{gn} s_{gn} T_i^{gn} < L_i \tag{55}$$

$$\sum_{ig} T_i^{gn} = P^n. \tag{56}$$

At the end of the 60's, disaggregated spatial interaction models of residential location became available (Wilson, 1970) and these were then integrated with the Alonso-Herbert-Stevens approach soon afterwards (Senior and Wilson, 1974). The key new theoretical idea which permitted greater realism to be introduced into economic models was the use of the entropy concepts to represent *dispersion* within perfect markets. For present purposes, we simply need to know that, for appropriately defined variables, a spatial interaction model of the form (1) can be built and that this could be given an economic interpretation.

In order to build a spatial interaction model related to equation (1) and the other concepts of Section 3 above, it is useful, as with the earlier examples, to use a slightly unconventional notation and to amend that used to present the Herbert-Stevens model above. As a special case of a Z_j^{ng}-variable, let H_j^g be the amount of type-g housing produced in zone j, and let this be related to a 'final demand' for labour in zone k by person-type groups, w. If this 'final demand' is E_k^w, then the appropriate interaction model is, in effect, allocating workers to housing in the form

$$Y_{jk}^{gw} = B_k^w W_j^g E_k^w e^{-\beta^w c_{ij}} \tag{57}$$

where g is used to label different housing 'sectors' and so it appears as a sector superscript and the type-of-good superscript is redundant and so is omitted. B_k^w is the usual balancing factor and β^w the usual impedance parameters. It turns out that $\log W_j^g$ can be interpreted as the benefit derived from living in a type g house in zone j and this is the basis for reconciling the economic and spatial interaction forms of model. With a suitable specific definition of W_j^g, and letting $\beta^w \rightarrow \infty$, the Herbert-Stevens model can then be derived as a special case of (57).

The next step in the argument is to use the Section 3 procedures to derive equations for H_j^g. One way to do this is to assume q^w as the average amount a w-type household spends on housing and p_j^g as the (annualised, say) cost of developing a type-g house in zone j, so that

$$C_j^g = p_j^g H_j^g \tag{58}$$

$$D_j^g = \sum_{kw} Y_{jk}^{gw} q^w \tag{59}$$

and then

$$\dot{H}_j^k = \varepsilon^w \left(\sum_{kw} Y_{jk}^{gw} q^w - p_j^g H_j^g \right) H_j^g \tag{60}$$

with

$$\sum_{kw} Y_{jk}^{gw} q^w = p_j^g H_j^g \tag{61}$$

as the equilibrium conditions.

The real situation is immensely more complicated than this. Housing is long-life stock, used many times over. Equation (60) will apply only to new stock. (The inertia of the old will be enough to cause the ring-like growth described by Burgess.) There are many problems of matching expenditure and 'price' of new and old housing. There are problems in defining attractiveness functions, and so on. It is some of the terms in the attractiveness function, for example, representing property for social classes to group together, which can produce Hoyt's sectoral differentiation. These are all addressed in a separate paper (Clarke and Wilson, 1983). For the present, we will assume that some kind of working model can be built which will incorporate the 'old' theory.

As a fourth and final example at this stage we briefly consider Christaller's (1933) central place theory. This provides an early approach to the modelling of service sectors and also differs from most of the early approaches in that it is comprehensive (though at a relatively coarse level of resolution); but like most other early approaches, it uses mainly a continuous space representation. It can be applied (as originally) to settlement structures or to service systems within towns. The main theoretical idea is concerned with market areas. The assumption is made that market areas of service centres do not overlap and that there is a hierarchy of such areas. The concept of the range of a good can be used to calculate the spacing of centres, though once this has been fixed for centres of a given order, it determines, through the market area geometry of the whole system, the spacing of centres in all other orders. The theory is built on other restrictive assumptions which are difficult to relax: a uniform distribution of population, for example, and uniform plain transport networks.

Again, a switch to a discrete zone system and the framework of Section 3 immediately produces a workable and much more flexible model. Equations (1) - (11) can be interpreted quite directly for service systems (and indeed it is in this field that there has been most experience of running this kind of model).

Again, we develop a model in a somewhat unconventional notation with respect to subscripts used, and this connects us firmly to the framework of Section 3 and also provides the possibility of new extensions of the model. Let P_k be the population of zone k (and this could easily be disaggregated if appropriate) and e_k^g the per capita expenditure of this population on goods of type g. $e_k^g P_k$ then form an estimate of 'final demand'. Let W_j^g be the provision of retail facilities (say measured by floorspace) and assume that, raised to a power α^g, it provides a measure of attractiveness. (Other factors can easily be added to the notion of attractiveness but we keep to this simple assumption for ease of exposition.) W_j^g is then again a special kind of Z_j^{mg} variable with g being used to label a sector. The appropriate spatial interaction model is then

$$Y_{jk}^g = B_k (W_j^g)^{\alpha^g} e_k^g P_k e^{-\beta^g c_{jk}} \qquad (62)$$

with

$$B_k = 1/\sum_j (W_j^g)^{\alpha^g} e^{-\beta^g c_{jk}}. \qquad (63)$$

The revenue attracted to g at j is

$$D_j^g = \sum_k Y_{jk}^g \qquad (64)$$

and the cost is taken as a linear function of floorspace:

$$C_j^g = k_j^g W_j^g \qquad (65)$$

for suitable constants k_j^g.

Harris and Wilson (1978) showed how to solve the resulting equilibrium equation

$$D_j^g = C_j^g \qquad (66)$$

for $\{W_j^g\}$. The resulting patterns are functions of the parameters α^g, β^g and k_j^g (and, of course, arrays which are exogenous to this problem, but which might be provided by other models - like e_i^g and P_k). These sets of patterns provide a richer alternative to central place theory.

A dynamic model has also been explored (see, for example, Clarke and Wilson, 1983-B):

$$\dot{W}_j^g = \varepsilon^g [D_j^g - C_j^g] W_j^g \qquad (67)$$

for suitable constants ε^g.

These models can be extended in various ways - see Wilson (1983-B) for some detailed suggestions. In particular, in the format of Section 3, interaction flows of the form Y_{ij}^{mgh} could be explored of the inputs h from sectors m in zone i to (j,g) activity, and this would then provide the basis for a more realistic cost function, C_j^g.

We have already noted that the $\{W_j^g\}$ variables are *particular* Z_j^{ng} variables, and that in this model they are functions of other such variables, like the P_k's. It is an obvious extension of the argument to note that it is really the full $\{Z_j^{ng}\}$ model whose framework was presented in Section 3 which constitutes the appropriate replacement for central place theory - in other words, a set of linked submodels within a comprehensive framework. The development of such a model is very much a matter for future research, but we can gain some useful insights into the exploitation of the framework in this context by investigating its application to the Lowry (1964) model.

In the Lowry model, basic employment is a quasi 'final demand' for labour and is the basis (in the first iteration) of residential location. The population thus generated then have a 'final demand' for services. This implies the use of interaction arrays $\{Y_{jk}^{hw}\}$ and $\{Y_{ij}^{gw}\}$ as the flow of people from type h-houses in j to type w-jobs in k (with w also serving as a person-type label) and the flow of goods or services g from shops or facilities in i to type-w people in j. The appropriate models would be like those in the discussions on residential location and retailing above.

The models would be linked first through the services creating employment to be added to the final demand for labour on later iterations; and secondly through population being a common factor in the two models. This type of linkage reveals an important feature which needs to be built into the Section 3 framework when it is used for the development of comprehensive models. That is: the results of distinct submodels must sometimes be combined to provide inputs to other submodels. A particular illustration here is the addition of employment from services to basic employment to give a total which is then the basis of the residential location model. Generally, because of such features, a comprehensive model is more than simply a sum of submodels.

In all four cases, we have shown how a shift of perspective - to discrete zones, to a comprehensive picture or whatever - can create a model which is more powerful and flexible but which nonetheless fits into the framework of Section 3. In most cases, the theoretical content can be incorporated within the framework - the obvious exceptions being the use of non-overlapping areas or whatever which do not accord with any reality. In the final section of the paper, we review the research tasks involved in building on this foundation.

5. Concluding comments: towards further research

In this final section, we connect the argument of this paper to a number of other relevant ideas and attempt to point the way towards further productive research. First, it is important to recognise, of course, that there are many alternative approaches to location theory. Some, such as random utility theory (cf. Williams, 1977), provide alternative theories of consumer's behaviour and can easily be incorporated into the approaches of Section 3. Others can be seen as more direct descendents of much of the classical theory which we used as a basis for comparison in Section 4. These include, for example, what is often termed the 'new urban economics' (cf. Richardson, 1977, for a good review). These mostly use continuous space representations and though they sometimes operate at fine levels of resolution, the price is usually paid in terms of restrictive assumptions and intractable mathematics. However, as might be expected, more progress has been made with price mechanisms and an important avenue of research is the incorporation of these into the kinds of models sketched in this paper.

The second point to make is a comment about the nature of the models - and this comment will apply to the alternatives too when they are properly formulated. It is clear, even in a formally-specified model, that there will be many nonlinearities and a high degree of interdependence. These properties guarantee the existence of multiple solutions (or even types of solution) for particular values of parameters and exogenous variables. This takes us into the realms of catastrophe theory and bifurcation theory (cf. Wilson, 1981-B) and a concern with critical parameter values at which jumps or other transitions can take place. The evolution of an urban or regional system can then be seen as episodic: periods of smooth change interspersed with more rapid structural change. An important field of research becomes the identification, both theoretically and empirically, of this kind of change.

It is relatively difficult to account for structural change analytically in these models. Significant progress has been made for service systems (cf. Harris and Wilson, 1978; Wilson and Clarke, 1979; Harris, Choukroun and Wilson, 1982), but much more work is needed in fields like residential and industrial location where the nonlinearities are different and more complicated. There are some particularly difficult mathematical problems to solve - particularly the so-called 'backcloth' problem (cf. Wilson and Clarke, 1979). When an analysis is carried out for a particular Z_j^{ng}, the results are dependent on values of all other Z_j^{ng}'s, each of which is potentially subject to a similar analysis and for which there are possible jumps.

These difficulties give rise to the third point: that the notion of 'numerical experiments' has become particularly valuable. It is very difficult, looking at a set of algebraic equations like (1) - (16), to imagine the range of spatial configurations which can be generated by them and the ways in which these are transformed at critical parameter values. By running different versions of these models over ranges of parameter values, a more systematic understanding of these possibilities can be achieved. An example for service sectors is provided by Clarke and Wilson (1983-B) and for residential location and housing, as noted in another context earlier, by Clarke and Wilson (1983-A). There is an advantage to understanding in conducting these experiments in relation to simple hypothetical systems, which, for example, are symmetric in various ways; but it is also important to understand the nature of change in real systems. What does become clear, even from this very brief account, is that for any particular system there is a large number of alternative paths of development and evolution. Historical 'accidents' will play a large part in determining which actual path is chosen. Two interesting types of research problem result from this observation. First, a new perspective is added to historical geography: to chart the particular path which is chosen, with reasons why this is the case. Secondly, it is interesting to ask whether bundles of alternative paths can be grouped together in such a way that they constitute a *type* of city, and that, for a certain region in parameter space, a city will be of this type whatever perturbations occur.

Fourthly, we note that some technical mathematical points (simpler than the 'backcloth' problem) need to be solved. In particular, the relationship of programming models to the differential equations and equilibrium equations (and the relationship of the iterative solutions of the last to the other two) need to be articulated.

The fifth point is concerned with the application of the methods to the full variety of major subsystems and then the exploration of the major linkages between them. In Section 1 we listed the major subsystems: agriculture, public utilities, industrial location, private services, public services and housing. Public utilities have some of the same character as public services, and so we can say that we have given an illustrative approach to each in Section 3 or Section 4. These are two main points then to be emphasised at this stage of the argument. First, although we have argued that Section 3 offers a unified approach, it is important to note that, for each major subsystem, the particular features demand much research, and these particularities may be the dominant features in the results which emerge. Section 3 provides a framework only. Secondly, it is important to remember that most major systems will be linked, first through the competition for land and secondly through the transport system. It remains important to develop methods of land use accounting which are appropriate for comprehensive dynamic models and to achieve an effective representation of transport demand and supply.

Sixthly and finally, we recognise that the most important ongoing problem is the articulation of mechanisms of price determination and change whether in the neo-classical mode or another. It is at this stage, more broadly, that alternative theories of social processes can be incorporated. There is an implicit contention in the whole argument which is an important one, and provides an appropriate note on which to end: whatever the underpinning social theory, the linked accounts implied by the array definitions will still be relevant concepts and in most cases, the general framework of theory offered here will still be relevant.

Acknowledgement

I am grateful to Mark Birkin for valuable comments on an earlier draft of this paper.

References

Alonso, W. (1960) A theory of the urban land market, *Papers, Regional Science Association, 6*, 149-57.

Alonso, W. (1964) *Location and land use*, Harvard University Press, Cambridge, Mass.

Birkin, M. and Wilson, A.G. (1984) Models of industrial location: a review, forthcoming Working Paper, School of Geography, University of Leeds.

Burgess, E.W. (1927) The determination of gradients in the growth of the city, *Publications, American Sociological Society, 21*, 178-84.

Christaller, W. (1933) *Die centralen orte in Suddeutschland*, Jena; English translation by C.W. Baskin, *Central places in southern Germany*, Prentice-Hall, Englewood Cliffs.

Clarke, M. and Wilson, A.G. (1983-A) Exploring the dynamics of urban housing structure: a 56 parameter residential location and housing model, Working Paper 363, School of Geography, University of Leeds.

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

Coelho, J.D. and Wilson, A.G. (1977) Some equivalence theorems to integrate entropy maximising submodels within overall mathematical programming frameworks, *Geographical Analysis, 9*, 160-73.

Coelho, J.D., Williams, H.C.W.L. and Wilson, A.G. (1978) Entropy maximising submodels with overall mathematical programming frameworks: a correction, *Geographical Analysis, 9*, 195-201.

Found, W.C. (1971) *A theoretical approach to rural land use patterns*, Arnold, London.

Harris, B., Choukroun, J-M. and Wilson, A.G. (1982) Economies of scale and the existence of supply-side equilibria in a production-constrained spatial-interaction model, *Environment and Planning A, 14*(6), 823-837.

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*, 371-88.

Harris, C.D. and Ullman, E.L. (1945) The nature of cities, *Annals of the American Academy of Political and Social Science, 242*, 7-17.

Herbert, D.J. and Stevens, B.H. (1960) A model for the distribution of residential activity in urban areas, *Journal of Regional Science, 2*, 21-36.

Hoover, E.M. (1967) Some programmed models of industry location, *Land Economics, 43*, 303-11.

Hoyt, H. (1939) *The structure and growth of residential neighbourhoods in American cities*, Federal Housing Administration, Washington DC.

Lakshmanan, T.R. and Hansen, W.G. (1965) A retail market potential model, *Journal of the American Institute of Planners, 31*, 134-43.

Leonardi, G. (1981) A unifying framework for public facility location problems, *Environment and Planning A, 13*, 1001-28; 1085-1108.

Losch, A. (1954) *The economics of location*, Yale University Press, New Haven.

Lowry, I.S. (1964) *A model of metropolis*, RM-4035-RC, Rand Corporation, Santa Monica.

Lowry, I.S. (1967) Seven models of urban development: a structural comparison, *Special Report 97*, Highway Research Board, Washington DC, 121-45.

Paelinck, J.H. and Nijkamp, P. (1975) *Operational theory and method in regional economics*, Saxon House, Farnborough.

Rees, P.H. (1979) *Residential patterns in American cities*, Research Paper 189, Department of Geography, University of Chicago.

Richardson, H.W. (1977) *The new urban economics: and alternatives*, Pion, London.

Scott, A.J. (1971) *Combinatorial programming, spatial analysis and planning*, Methuen, London.

Senior, M.L. and Wilson, A.G. (1973) Disaggregated residential location models: some tests and further theoretical developments, *Geographical Analysis, 6*, 209-38.

Stevens, B.H. (1968) Location theory and programming models: the von Thunen case, *Papers, Regional Science Association, 21*, 19-34.

von Thunen, J.H. (1826) *Der isolierte staat in beziehung auf land wirstchaft und national ökonomie*, Hamburg; English translation, *The isolated state*, C.M. Wartenburg with a translation by P. Hall, Oxford, 1966.

Weber, A. (1909) *Uber den standort der industrien*, Tubingen; translation by Freidrich, C.J. (1929) *Alfred Weber's theory of location of industries*, University of Chicago Press, Chicago.

Williams, H.C.W.L. (1977) On the formation of travel demand models and economic evaluation measures of user benefit, *Environment and Planning A, 9*, 285-344.

Wilson, A.G. (1967) A statistical theory of spatial distribution models, *Transportation Research, 1*, 253-69.

Wilson, A.G. (1970) Disaggregating elementary residential location models, *Papers, Regional Science Association, 24*, 103-25.

Wilson, A.G. (1974) *Urban and regional models in geography and planning*, John Wiley, Chichester and New York.

Wilson, A.G. (1981-A) The evolution of urban spatial structure: a review of progress and research problems using SIA models, *Bulletin of the Institute of Mathematics and its Applications, 18*, 90-100.

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

Wilson, A.G. (1983-A) Transport and the evolution of urban spatial structure, in [illegible], Guida Editori, Naples, pp. 17-27.

Wilson, A.G. (1983-B) A generalised and unified approach to the modelling of service-supply structures, Working Paper 362, School of Geography, University of Leeds.

Wilson, A.G. and Birkin, M. (1983) Industrial location theory: explorations of a new approach, Working Paper 361, School of Geography, University of Leeds.

Wilson, A.G. and Birkin, M. (1984) Models of agricultural land use: some numerical experiments, forthcoming Working Paper, School of Geography, University of Leeds.

Wilson, A.G. and Clarke, M. (1979) Some illustrations of catastrophe theory applied to urban retailing structures, in M.J. Breheny (ed.) *Developments in urban and regional analysis*, Pion, London, 5-27.

37

Produced by
School of Geography
University of Leeds
Leeds LS2 9JT
From whom copies may be obtained