

# AN ENTROPY MAXIMISING ANALYSIS OF JOURNEY-TO-WORK PATTERNS IN COUNTY LIMERICK

JAMES A. WALSH  
Department of Geography,  
Carysfort College

## ABSTRACT

Entropy maximising models have become increasingly popular among geographers and regional scientists over the last decade. Yet there remains a considerable amount of confusion about the rationale for these models. This paper attempts to clarify the main issues and to illustrate the potential of the entropy maximising procedure in the area of journey-to-work modelling. Calibration methods and goodness of fit statistics are discussed. The latter half of the paper analyses recent journey-to-work patterns in County Limerick by the entropy maximising method and concludes by suggesting further applications of the model and areas of research.

## *Introduction*

Entropy maximising models have become increasingly popular among geographers over the last decade or so since the publication in 1970 of Alan Wilson's pioneering monograph, *Entropy in urban and regional modelling*. Since then there have been many advances in the theoretical development and refinement of the Wilson models and a number of new models have been introduced by others (Bussiere and Snickars, 1970; Webber, 1975, 1979; Wilson, 1974; Dacey and Norcliff, 1976). Their qualitative results have thrown new light on a variety of geographical problems, most notably those which involve spatial interaction and residential location patterns. The usefulness of the models has been demonstrated by a number of empirical applications (Wilson *et al.*, 1969; Hathaway, 1975; Openshaw, 1976; Thomas, 1977; Walsh, 1978). Despite this progress, the entropy maximising method is not widely understood, partly because it is rather difficult to explain the notion of entropy in a non-mathematical manner and partly because it is a style of analysis which is radically different from that employed by geographers and other *spatial analysts whose models of behaviour are rooted in concepts from micro-economics and psychology*. The difficulties that result from these other traditions are well illustrated by Beckmann and Golob (1972), Hansen (1972) and Fisch (1977).

The general objective of this paper is to explain briefly the ideas which underlie entropy maximising models by examining the derivation of a particular model. A more specific aim is to present a statistical framework for analysing and interpreting journey-to-work patterns. The kind of results that can be obtained from the approach developed in the paper are illustrated through an examination of recent industrial work-trip patterns in Co. Limerick.

*Entropy maximising method*

The method which is outlined in this section seeks to define the least biased assignment of probabilities of spatial interaction within a specific context. It is assumed that the study area can be subdivided into  $N$  non-overlapping zones. The zonal distributions of persons seeking work and of employment opportunities are known and are designated as  $\{O_i\}_{i=1}^N$  and  $\{D_j\}_{j=1}^N$  respectively (usually  $O_i \neq D_{j, i=j}$ ). The distance or cost of interaction between all pairs of zones is assumed to be known and given by the matrix  $\{c_{ij}\}_{i,j=1}^N$ . A researcher who is given all the information outlined above is still left uncertain about the number of work trips that are made between each pair of zones. Let  $p_{ij}$  denote the probability that a person resides in zone  $i$  and works in zone  $j$ . The probability distribution representing all the work trips that are possible is given by  $\{p_{ij}\}$ . Uncertainty remains about each  $p_{ij}$ . The mean uncertainty associated with the distribution  $\{p_{ij}\}$  is equivalent to its entropy,  $S$  (Walsh and Webber, 1977, p. 398), and is:

$$S = -\sum_{ij} p_{ij} \log p_{ij}.$$

The importance of the entropy concept is due to a logical principle advanced by Jaynes (1957) which states that the least biased assignment of probabilities to choices is that which maximises entropy (uncertainty) subject to some constraints (Webber, 1976, 1977). This logical principle is the foundation for all entropy maximising models. Two major questions arise from the principle. The first concerns the nature of the constraints and their specification while the second concerns the maximisation of the uncertainty or entropy measure.

The constraints play a crucial role in entropy maximising models. They represent whatever useful information is available to the researcher. The more information that is available, the greater will be the number and the complexity of the constraints. At the same time, however, the researcher's uncertainty about the distribution  $\{p_{ij}\}$  will be less. The optimal distribution  $\{p_{ij}^*\}$  in a particular case depends on the number and the nature of the constraints (Sheppard, 1976). Often, therefore, a research problem can be viewed as identifying that set of constraints which reduces the uncertainty about choices to a desired level. In this way the entropy maximising model changes the emphasis in interaction studies from examining the interaction patterns *per se* to explaining the origins of the constraints that give satisfactory solutions. Thus it is seen that entropy maximising is a statistical procedure which examines hypotheses about the constraints which need to be explained if an observation of interest (here the trip distribution) is to be regarded as natural.

The information that is available in a journey-to-work study is usually represented by the following system of constraints:

$$\sum_j t_{ij} = O_i$$

$$\sum_i t_{ij} = D_j$$

$$\sum_{ij} t_{ij} \cdot c_{ij} = C$$

where  $t_{ij}$  is the number of individuals who live in zone  $i$  and work in zone  $j$  and  $C$  is an estimate of the total cost of interaction between all

the zones. The cost of each interzonal trip,  $c_{ij}$ , may be measured in terms of time, money, distance or perhaps in time-space terms where it would be considered as the disruption involved in creating a vacancy in an activity routine (Pirie, 1979). Some models that utilise additional prior information have been suggested by Cesario (1973, 1979), Kirby (1975) and Batty (1974; Batty and March, 1976).

The second question arising from the logical principle advanced by Jaynes concerns the maximisation of entropy. In any investigation one of the most serious problems is to avoid personal bias. Bias is basically the utilisation of private information, private in the sense that the information is not shared with others. To avoid bias, one is legitimately entitled to use only that information which is represented explicitly by the constraints (Jaynes, 1968). This restriction compels the researcher to maximise the entropy subject to the constraints, in order to avoid violating the logical principle. If one did otherwise, one would obtain an assignment  $\{q_{ij}\}$  which would be consistent with the constraints but would not maximise the entropy measure and would leave the researcher more certain about an individual's journey-to-work pattern than he should be on the basis of the stated evidence. By using  $q_{ij}$  instead of  $p_{ij}$  one would be implying that one had more information about the choice an individual would make. To summarise, it can be said that entropy maximising models rely not on a particular model of society (as economic models do, for example) but rather on a logical principle which determines what a researcher is legitimately entitled to conclude from a data set.

*A journey-to-work model*

Following the logical principle a journey-to-work model may be specified as:

Maximise 
$$S = -\sum_{ij} p_{ij} \log p_{ij} \tag{1}$$

subject to 
$$\sum_i p_{ij} = P_{.j}, \quad j \ 1, 2, \dots, N \tag{2}$$

$$\sum_j p_{ij} = P_{i.}, \quad i \ 1, 2, \dots, N \tag{3}$$

$$\sum_{ij} p_{ij} \cdot c_{ij} = \bar{c}, \tag{4}$$

where  $P_{i.}$  is the probability that a trip originates in zone  $i$ ,  $P_{.j}$  is the probability that a trip ends in zone  $j$  and  $\bar{c}$  is the observed average cost of a work journey. The solution to the model is obtained by Lagrange's method of undetermined multipliers (Wilson and Kirby, 1975).

Write the Lagrangian,  $L$ , as

$$L = -\sum_{ij} p_{ij} \log p_{ij} + \sum_i \lambda_i^{(1)} (P_{i.} - \sum_j p_{ij}) + \sum_j \lambda_j^{(2)} (P_{.j} - \sum_i p_{ij}) + \beta(\bar{c} - \sum_{ij} p_{ij} c_{ij}) \tag{5}$$

where  $\lambda_i^{(1)}$ ,  $\lambda_j^{(2)}$  and  $\beta$  are the Lagrangian multipliers. Then

$$\delta L / \delta p_{ij} = -\log p_{ij} - \lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij}$$

which, when set to zero, implies that the least biased estimate of  $p_{ij}$ ,  $p_{ij}^*$ , is

$$p_{ij}^* = \exp [-\lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij}] \tag{6}$$

Equation (6) is the best estimate of the probability that an individual

who chooses to reside in zone  $i$  would work in zone  $j$ . Clearly the probability is not symmetric, i.e. in general,  $p_{ij}^* \neq p_{ji}^*$ . If individuals choose their workplaces independently then the expected number that reside in zone  $i$  and work in zone  $j$  is given by

$$t_{ij}^* = T \cdot p_{ij}^* \quad (7)$$

where  $T$  is the total number of individuals that are at work. Furthermore, it is a condition in any final solution that

$$\sum_j t_{ij}^* = O_i \quad (8)$$

and

$$\sum_i t_{ij}^* = D_j. \quad (9)$$

By combining equations (6) to (9) one obtains the following gravity model

$$t_{ij}^* = \{A_i O_i B_j D_j \exp[-\beta c_{ij}]\}/T \quad (10)$$

where

$$A_i = (T/O_i) \exp[-\lambda_i^{(1)}] \quad (11)$$

and

$$B_j = (T/D_j) \exp[-\lambda_j^{(2)}] \quad (12)$$

This solution is similar to the one obtained by different means by Wilson (1967). However, the interpretation given to  $p_{ij}^*$  here, which is based on Jaynes (1957), Tribus (1969) and Webber (1977), is less restrictive than the usual interpretation given by Wilson and later commentators such as Gould (1972), which was based on statistical mechanics analogies. Furthermore, the conceptual basis for the model given by equation (10) differs fundamentally from that of the classical gravity models introduced by Carrothers (1956) and Isard (1960).

#### *Interpretation of parameters*

There have been a number of interpretations suggested for the Lagrangian multipliers in equation (6) and the parameters  $A_i$  and  $B_j$  in equation (10) by Kirby (1970), Wilson (1970) and Cesario (1974, 1977). It is evident from equation (6) that the larger  $\lambda_i^{(1)}$  and/or  $\lambda_j^{(2)}$  the smaller would be  $p_{ij}^*$ . Thus  $\lambda_i^{(1)}$  and  $\lambda_j^{(2)}$  respectively measure the dispropensity of zone  $i$  to emit trips and of zone  $j$  to attract trips. Similar terminology has been used by Cesario (1974) for his behavioural indices

$$U_i = A_i O_i = \exp[-\lambda_i^{(1)}] \text{ and } V_j = B_j D_j = \exp[-\lambda_j^{(2)}].$$

The parameter  $\beta$  in equation (6) measures the degree of impedance that travel cost exerts upon trip-making behaviour in a region. The smaller the value of  $\beta$  the greater is the average length of the work trip. The Lagrangian multipliers associated with the origin and destination constraints play a role as 'balancing' or 'normalising' factors. When equation (6) is substituted in equation (3) one obtains

$$\sum_j \exp[-\lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij}] = P_i.$$

which implies that

$$\exp[-\lambda_i^{(1)}] = P_i \cdot \sum_j \exp[-\lambda_j^{(2)} - \beta c_{ij}] \quad (13)$$

and each

$$p_{ij}^* = \{ \exp [-\lambda_j^{(2)} - \beta c_{ij}] / \sum_j \exp [-\lambda_j^{(2)} - \beta c_{ij}] \} P_i .$$

Therefore, each  $\exp [-\lambda_i^{(1)}]$  is a normalising factor to ensure that the corresponding origin constraint is met. Similarly, each  $\exp [-\lambda_j^{(2)}]$  guarantees that the corresponding destination constraint is satisfied. The interdependencies between the origin and destination normalising factors become more explicit when one combines equation (6) with equations (11), (12) and (13). Then

$$A_i = \{ \sum_j B_j D_j \exp [-\beta c_{ij}] \}^{-1} \tag{11'}$$

and

$$B_j = \{ \sum_i A_i O_i \exp [-\beta c_{ij}] \}^{-1} \tag{12'}$$

From equation (11') it is evident that  $A_i$  is directly related to the cost of travel from the  $i$ th zone to all the other zones and inversely related to the distribution of employment  $D_j$  and its scalar  $B_j$ . Consequently  $A_i$  can be regarded as a measure of the inaccessibility of zone  $i$  to employment opportunities, given the cost and employment distributions over all  $N$  zones. The parameter decreases in value as the locational potential of a zone as an attractor of work trips increases. Similarly,  $B_j$  measures the inaccessibility of a zone to the workforce distribution. It increases in value as the locational potential of a zone as a source of employees decreases. In this way, each  $A_i$  can be treated as a measure of the effects of inaccessibility on job seekers or employees while each  $B_j$  is a measure of the effects of inaccessibility on employers or job providers. Given that  $A_i$  and  $B_j$  can be viewed as measures of the effects of inaccessibility, their reciprocals  $A_i^{-1}$  and  $B_j^{-1}$  are measures of the effects of variations in accessibility.

The interrelationships between entropy maximising and linear programming interaction models have been explored by, among others, Wilson and Senior (1974) and H. Williams (1976). The maximum entropy associated with the probability distribution  $\{p_{ij}\}$  is obtained by substituting equation (6) into (1) to yield

$$\begin{aligned} S_{Max} &= - \sum_{ij} p_{ij}^* [-\lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij}] \\ &= \sum_i \lambda_i^{(1)} P_i . + \sum_j \lambda_j^{(2)} P_{.j} + \beta \bar{c} \end{aligned} \tag{14}$$

It has been shown by Wilson and Senior (1974) that the maximum value of the entropy function is equivalent to the objective function of the dual,  $S^1$ , of the entropy maximising model given by equations (1) to (4). Therefore each  $\lambda_i^{(1)}$  is a measure of the change in entropy that would result from an increase in  $P_i$ . (Miller, 1972, p. 296). This is the increase in entropy that results when the precise destination of some trip is unknown. Similarly, each  $\lambda_j^{(2)}$  is the increase in entropy when the precise origin of some trip to zone  $j$  is unknown. Following this interpretation  $\lambda_i^{(1)}$  provides the modeller with much useful information about the zones where errors are most likely to occur when the data is incomplete. In summary, the greater the dispropensity of a zone to emit or attract trips the greater the marginal increase in the maximum entropy when the precise destinations or origins of some trips are unknown.

As a result of the duality relationships between entropy maximising and programming models, Wilson and Senior (1974, p. 213) have suggested that the variables  $(-\lambda_i^{(1)}/\beta)$  and  $(-\lambda_j^{(2)}/\beta)$  can be interpreted as measures of comparative advantage. Rewriting equation (14) as

$$S_{Max}/\beta = \sum_i (\lambda_i^{(1)}/\beta) P_i + \sum_j (\lambda_j^{(2)}/\beta) P_j + \bar{c}$$

it is evident that the maximisation of comparative advantage results in a minimisation of entropy.

Finally it is interesting to examine the change in the propensity to emit trips of any zone, say zone  $k$ , over other zones due to a unit increase in the impedance exerted by travel cost on trip making. Following equations (3), (6) and (13) the change can be measured as

$$\begin{aligned} -\delta\lambda_k^{(1)}/\delta\beta &= \{\sum_j c_{kj} \exp[-\lambda_j^{(2)} - \beta c_{kj}]\} / \{\sum_j \exp[-\lambda_j^{(2)} - \beta c_{kj}]\} \\ &= \sum_j c_{kj} \cdot p_{kj} = \bar{c}_k \end{aligned} \quad (15)$$

which is the mean cost of a trip originating in zone  $k$ . Therefore the change in the propensity to emit trips when there is an overall increase in transportation costs is greatest in the zones that are furthest from employment opportunities. On the other hand, it can be shown in a similar manner that the change in the propensity of any zone to attract trips when transportation costs increase will be proportional to the current mean length of trips into the zone.

To summarise, it has been shown that the Lagrangian multipliers in equation (6) play a crucial role as normalising factors. More generally they can be interpreted as propensity or comparative advantage measures, while in the context of entropy maximising models they have a further interpretation as measures of the rate of change in the maximum value of entropy associated with the distribution  $\{p_{ij}\}$  when the precise origins or destinations of some trips are unknown.

In the remainder of the paper attention will be confined to the aggregate model given by equation (10), which is recalled here for convenience:

$$t_{ij}^* = \{A_i O_i B_j D_j \exp[-\beta c_{ij}]\} / T$$

which provides the least biased estimate of the number of trips between zones  $i$  and  $j$  subject to the constraints that were set.

#### *Calibration and testing procedures*

The task of finding estimates of the values of the model parameters that provide the best simulation of a real pattern is known as calibration. The calibration of spatial interaction models has increasingly assumed a central role in the design and construction of such models (Ayeni, 1979). The traditional calibration procedure for gravity models is by logarithmic transformation of the equations into linear functions which are then fitted by the method of least squares (Olsson, 1965). However, the model outlined in this paper is intrinsically non-linear and therefore requires some alternative calibration procedures. A number of procedures have been proposed and tested (Batty and Mackie, 1972; Evans and Kirby, 1974; I. Williams, 1976). These include the most commonly used maximum likelihood methods due to Hyman (1969), Evans (1971) and Hathaway (1973) and non-linear least squares methods (Cesario, 1975; Openshaw,

1976). The non-linear least squares method is based on the assumption that the best statistic for measuring the goodness-of-fit of the calibration is some function of the sum of the deviations between the observed and predicted values. However, the parameter estimates that are obtained by this method may not always satisfy the constraints of the model (Openshaw, 1976). In contrast the maximum likelihood method uses the mean observed trip cost as the calibration statistic. The procedure is based on finding the maximum of a likelihood function which expresses the probability of different parameter values producing an observed result. The procedure will yield parameter estimates that satisfy the model constraints if the sampling distribution of trips is Poisson (Kirby, 1974).

In this study Hyman's (1969) iterative procedure is used to find the best estimates of  $A_i$ ,  $B_j$  and  $\beta$ . The procedure involves specifying some initial values of  $\beta$  and  $B_j$  from which an estimate of  $A_i$  is computed. Using this estimate a new  $B_j$  can be computed. If the new  $B_j$  differs from the initial one the procedure is repeated with the new  $B_j$  as input. The iterations continue until  $A_i$  and  $B_j$  stabilise. Then the mean cost of interactions that result from these parameters is compared with an observed mean cost of interaction. This test criterion was derived by Bayesian methods. If there is a difference between the model and the observed mean costs a new series of iterations is initiated to accommodate the effects of variations in  $\beta$ . The procedure continues until a stable solution that satisfies the entropy maximising constraints is achieved. Evans (1971) has proved that in almost all situations the value of  $\beta$  obtained by this method is unique. Some modifications and improvements to the method have been suggested by I. Williams (1976, 1977). The calibration programme used in this study was a modified version of one presented by Baxter (1973).

While calibration is concerned with finding the most probable values of the model parameters, testing involves investigating the overall goodness-of-fit of the model predictions to the observed data. There are a large number of goodness-of-fit statistics available. They include:

- (1) The mean and standard deviation of the residuals.
- (2) The coefficient of determination,  $R^2$ , and the parameters of the regression of the predicted trips on the observed ones. In the case of a good fit the intercept parameter would be close to zero and the slope of the regression line would be close to unity. Great care must be exercised in interpreting the results of this test and, in particular, associated tests of significance if the trip distribution is not approximately normal.
- (3) The chi-square statistic defined as

$$\chi^2 = \sum_i \sum_j (t_{ij} - t_{ij}^*)^2 / t_{ij}^*, \quad \forall t_{ij}^* \geq 6$$

- (4) The dissimilarity index,  $G$ , defined as

$$G = \left\{ \sum_i \sum_j |t_{ij} - t_{ij}^*| \right\} \cdot 1/2T.$$

The index ranges between zero, representing a perfect correspondence, and 100, representing a situation of maximum possible difference between the two distributions. The magnitude of  $G$  represents the percentage of the trips in the predicted matrix that would have to be reallocated in order to replicate the observed trip matrix.

- (5) The mean absolute error,

$$M.E. = M^{-1} \sum_i \sum_j |t_{ij} - t_{ij}^*|$$

where  $M$  is the number of trip pairs.

- (6) The information gain index. The index measures the amount of extra information that is required, relative to what is already known, to alter some prior probability distribution into a posterior distribution. Following Jaynes (1968), the best definition of the prior distribution is the one which has the maximum freedom to vary subject to whatever is known about a problem; that is the set of maximum entropy probabilities. Then the prior distribution may be taken as  $\{p_{ij}^*\}$  and the posterior as  $\{q_{ij}\}$  where  $q_{ij} = t_{ij}/T$ . The index is defined as

$$I = \sum_i \sum_j q_{ij} \log (q_{ij}/p_{ij}^*).$$

It is dimensionless and ranges between zero and  $\log(N^2) - 1$ . For a detailed discussion of the measure and its properties see Walsh and O'Kelly (1979).

It is worth noting that these goodness-of-fit measures differ substantially from the two rather vague criteria used in the Dublin Transportation Study, 1972. The criteria in that study were (a) that 'both curves (of the trip length distribution) should be relatively close to one another' and (b) that the differences between the durations of the average trip lengths should not exceed  $\pm 3\%$  (D.T.S. Technical Report No. 15).

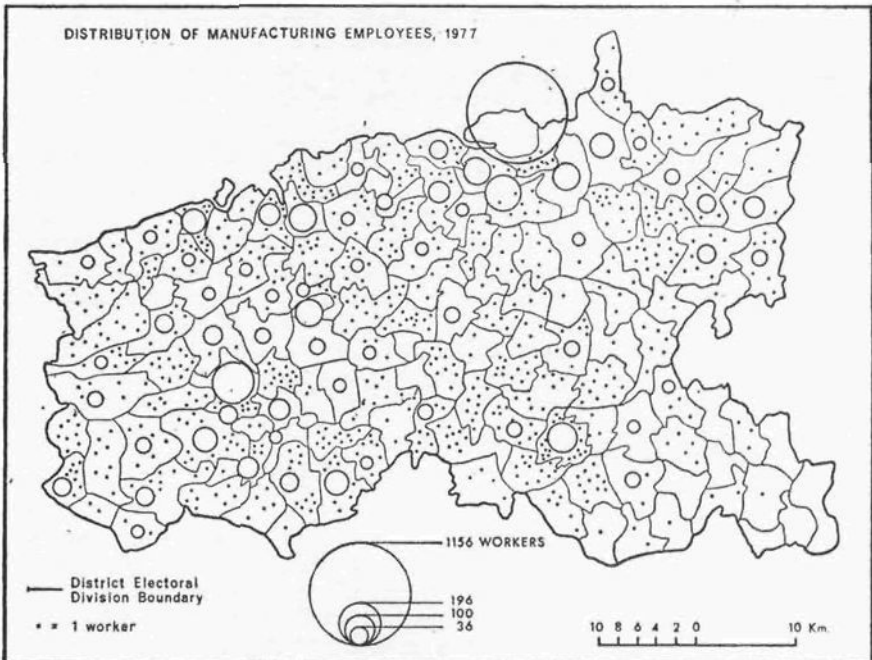


Figure 1



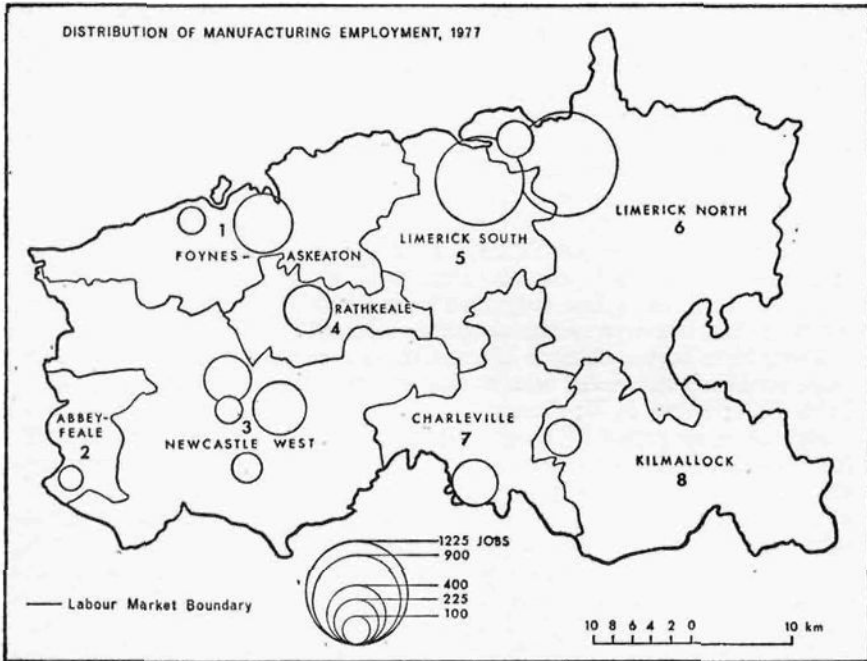


Figure 2

*Application: journey-to-work patterns in Co. Limerick, 1977*

Two interaction matrices are required for calibration and evaluation of the model performance. One consists of the number of work trips that are observed to occur between each origin-destination pair at a particular time while the other is a matrix of the costs of work trips between all pairs of zones at the same time.

*The work-trip matrix.* The data on journey-to-work patterns used in this study was collected during the summer of 1977 by the planning department of Limerick County Council. All manufacturing firms in the county with at least forty employees were asked to supply the council with a list of the home addresses of their workers. This involved nineteen firms with employment figures ranging from 64 to 1552. The median employment level was approximately 130. The total number of employees among the firms was 4466, of whom 3958 were resident in the city or county. The exact addresses of the remainder were unknown and, therefore, they had to be excluded from the study. A further difficulty arose in relation to individuals residing in Limerick but working outside the county. The principal adjacent sources of employment outside the county were in Charleville, Mitchelstown and Tarbert. The leading employers in these towns were contacted, but only the one in Charleville supplied the necessary information. A total of 232 residents of Limerick County were employed there. The treatment of flows across region boundaries has been the subject of a recent paper by Masser (1979).

The 4190 workers resident in the county were distributed over 138 zones that correspond with the District Electoral Divisions (Fig. 1). The

employment opportunities were concentrated in fourteen of these zones (Fig. 2). The observed trip frequencies were summarised in a  $138 \times 14$  matrix. In this matrix only 17.8% of the trips were intrazonal. Therefore, at this level most of the interactions were interzonal and the proportion approximates to Broadbent's rule of thumb that, if patterns of interaction are the major research interest, then at least 85% of all flows should occur between zones (Broadbent, 1970). However, an examination of the trip matrix for this study showed that because of the degree of localisation of employment opportunities in the manufacturing sector approximately 93% of the cells in the matrix had entries of less than five. Because of the sparseness of the matrix and also since there is no logical justification for the use of D.E.Ds as data collection units for this type of analysis, it was felt that an alternative system should be devised.

There have been a number of attempts recently to devise algorithms to solve problems associated with zoning system design (Masser *et al.*, 1975, 1976, 1977; Hirst, 1977; Openshaw, 1977). The procedure of multi-level specification suggested by Masser and Brown (1977) was adopted here. It involves essentially partitioning a large region into weakly interacting subregions and modelling the pattern of interaction at two levels of spatial resolution. Trips between the subregions are modelled using the subregions themselves as zones and trips within each subregion are modelled on a fine zoning system based on the minimisation of intrazonal trips. The Masser and Brown (1975) hierarchical aggregation algorithm was applied to the trip matrix to define subregions which are called labour markets. A number of objective functions were examined with the algorithm. Unfortunately none of them yielded satisfactory solutions since many residential zones were being misallocated in the sense that they were being combined with an employment zone to which they sent fewer workers than they sent to some other zone. In an attempt to overcome these difficulties an effort was made to define local labour markets from an inspection of the trip matrix. The following rules were applied:

1. Each residential zone  $i$  ( $i=1, 2, \dots, 138$ ) was allocated to employment zone  $j$  ( $j=1, 2, \dots, 14$ ) for which  $t_{ij}$  was maximum.
2. When  $t_{ij}$  was equal to  $t_{ik}$  for  $j \neq k$ , zone  $i$  was allocated to the nearer of zones  $j$  and  $k$ , where the nearness of zone  $i$  to either zones  $j$  or  $k$  was defined as the straight line distance between the centroids of the relevant pair of zones.
3. Each labour market consisted of contiguous zones.

The output of this procedure was eight labour markets (Fig. 2) which now become the zones for the remainder of the study. From the resulting trip matrix (Table 1) it can be seen that 32% of the trips now take place between zones. Since the pattern of intrazonal interaction is largely known (given the constraints applied in the definition of the labour markets) the remainder of the analysis, while making use of the total data input of 4190 trips, is concerned primarily with the interzonal pattern.

*The cost matrix.* An  $8 \times 8$  matrix of the costs of interaction between all pairs of labour markets was required for the calibration. Since it was not possible in the survey to obtain information on the time or money spent on travelling to work it was necessary to use distance as a proxy measure of cost. To calculate the interzonal distances the centroid of each labour market was connected to the primary and secondary road networks. All intersections in the road network were identified and the length of the

shortest path between each pair of zones was calculated, following a minimum path algorithm.

TABLE 1

ACTUAL MATRIX OF JOURNEY TO WORK TRIPS CO. LIMERICK, 1977

Origins \ Destinations	Destinations								$O_i$
	1	2	3	4	5	6	7	8	
1	351	1	24	25	39	12	5	0	457
2	0	49	35	7	1	0	0	0	92
3	35	19	679	67	6	13	56	0	893
4	27	0	30	128	7	10	6	0	208
5	23	0	19	8	253	59	9	0	371
6	45	0	9	3	560	1167	18	12	1814
7	9	1	10	5	10	14	95	16	160
8	1	0	2	0	2	24	43	123	195
$D_j$	491	70	826	243	878	1299	232	151	4190

The estimation of intrazonal distances is a problem encountered in all studies of this nature and has been solved in a variety of ways. Here the intrazonal distances were taken as the mean of the straight line distances between the centroids of residential and employment D.E.Ds within each labour market. The distances obtained in this manner are listed in Table 2. From Tables 1 and 2 it was possible to estimate the mean observed trip length as 7.3 miles. This figure compares well with the figure of 8.8 miles obtained from the census in 1971 as the national average distance travelled to work by non-agricultural residents of rural areas and towns with less than 5,000 inhabitants.

TABLE 2

THE DISTANCE MATRIX

Origins \ Destinations	Destinations							
	1	2	3	4	5	6	7	8
1	4.13	28.30	16.81	6.25	17.50	19.50	22.87	31.44
2	28.30	2.99	17.49	22.07	33.30	43.60	33.50	42.10
3	16.80	17.40	4.04	10.56	21.81	32.11	20.19	28.75
4	6.25	22.07	8.50	1.72	11.25	21.55	16.62	25.19
5	17.50	33.20	21.81	11.25	2.61	10.30	13.60	22.20
6	19.50	43.20	32.11	21.55	10.30	6.39	23.90	27.25
7	22.80	33.50	20.19	16.62	13.68	23.98	5.93	8.56
8	31.40	42.12	28.75	25.19	22.24	27.25	8.56	1.94

*Results.* The entropy maximising solution for equations (10)–(12) was found to occur after ten iterations with a  $\beta$  value of 0.1882. This value predicted a trip matrix satisfying the origin and destination constraints as well as the observed mean travel cost of 7.3 miles per person. The value of  $\beta$  found here differs from those found in other studies (Hathaway, 1975; Openshaw, 1976; Thomas, 1977). For example, the  $\beta$  value that solved the entropy maximising model for Merseyside was 0.567 with an average cost per trip of 2.63 miles. It is interesting to compare the observed mean

trip length with what it would be if there were no deterrence effects arising from distance. In that case  $\beta=0$  and the associated mean trip length is

$$c_0 = (\sum_i \sum_j c_{ij} O_i D_j) / T^2.$$

In Co. Limerick it implies that the mean trip length in the absence of spatial deterrence effects would be 17.9 miles. Clearly then these effects have a considerable impact on interaction patterns. From the distance matrix and the residence and employment distributions one can calculate that the minimum mean trip length that is possible is 5.247 miles which is associated with a  $\beta$  value of 0.45. The relationship between mean trip length and  $\beta$  is shown in Figure 3. It is evident that the entropy maximising solution is much nearer to the cost minimising linear programming one than to a cost maximising solution. This simply means that cost minimisation plays an important role in journey-to-work decisions in Limerick. The theoretical relationships between gravity model and linear programming solutions have been explored in detail by Evans (1973).

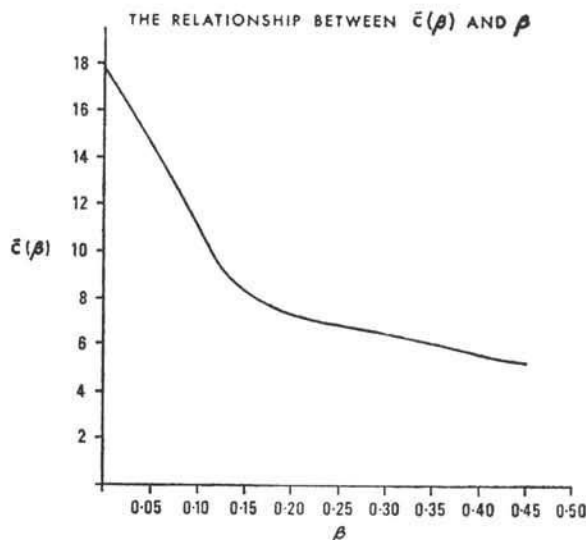


Figure 3

The second set of model parameters that are of interest are the values of  $A_i$  and  $B_j$  which ensure that the trip origin and destination constraints are met. The values of these parameters and the Lagrangian multipliers,  $\lambda_i^{(1)}$  and  $\lambda_j^{(2)}$ , are listed in Table 3. The  $A$  values range from 1.78 in zone 5 to 12.58 in zone 2 (Abbeyfeale). Generally, the values tend to increase with distance from the main centre of employment around Limerick city. The pattern of  $\lambda_i^{(1)}$  is less varied as indicated by its coefficient of variation. Because of their size zones 6 and 3 have a higher propensity to emit high cost work trips than the smaller and more central zones 4, 5 and 7. Despite the inaccessibility of zones 2 and 8 to the primary centres of employment they have relatively low propensities to emit trips. The reason for this is the relatively small discrepancy between the number of workers and the number of jobs available in both zones.

TABLE 3  
PARAMETER VALUES

Zone	$A$	$B$	$-\lambda^{(1)}$	$-\lambda^{(2)}$
1	2.872	4.100	7.18	7.61
2	12.575	5.385	7.05	5.93
3	3.008	2.831	7.90	7.76
4	1.981	3.231	5.99	6.67
5	1.777	3.037	6.49	7.89
6	2.948	2.000	8.58	7.86
7	3.943	7.224	6.44	7.42
8	4.223	5.519	6.71	6.73
Mean	4.166	4.172	7.07	7.23
St. Dev.	3.274	1.627	0.78	0.67
Coeff. of Var.	78.6%	39.0%	11.07%	9.21%

The distribution of  $B_j$  values is similar to the  $A_i$  pattern with the inaccessibility of employment zones to the residences of workers increasing away from the main centres of employment. However, the spatial variation in the  $B_j$  values is substantially less than the variation in the  $A_i$  distribution (Table 3). This, of course, is due to the fact that the workforce residences are more widely dispersed throughout the county than the employment opportunities. A similar situation existed in Merseyside (Thomas, 1977) while the opposite situation has been found in a number of urban areas (Masser, 1977). The higher level of variability of the  $A$  values in the county indicates that overall the inaccessibility effects are more discriminatory among employees than on employers. An optimal solution, perhaps, would be where the variability in the  $A_i$  and  $B_j$  distributions was comparable. Data from a 1969 transportation study in Limerick city suggests that the distributions of residences and employment opportunities were optimal in this sense at the time.

The  $-\lambda_j^{(2)}$  parameter measures the propensity of zones to attract trips relative to other zones. The zones with the highest levels of attractiveness are 5 and 6. These are followed by zones 3 and 1. On the other hand, the zone with the smallest propensity to attract trips is number 2, Abbeysfeale. Clearly, the  $\lambda_j^{(2)}$  values are largely determined by the levels of zonal employment. However, there are some other interesting features in the  $\lambda_j^{(2)}$  distribution. Firstly, the relative attractiveness indices of zones 4 and 8 do not differ very much even though zone 8 is twice the distance from zone 5 that zone 4 is. The low attractiveness index for zone 4 is due to its relatively small number of employment opportunities while for zone 8 it is mostly due to its relative inaccessibility. Secondly the attractiveness index for zone 7 is more than twice that for zone 4 even though employment opportunities in zone 7 are slightly less than in zone 4. This situation arises because employers in zone 4 have to compete for workers against employers in zones 1, 3 and 5. On the other hand, the propensity of Charleville in zone 7 to attract work trips is heightened by its relative inaccessibility, away from competitors for labour.

Using the  $A_i$  and  $B_j$  values in Table 3 and the  $\beta$  value of 0.1882 the entropy maximising matrix of predicted trips was computed. The predictions are listed in Table 4. The goodness-of-fit statistics for the model are summarised in Table 5. For comparison similar statistics are included for the more conventional *ad hoc* gravity model of the form

$$p_{ij}^* = K O_i D_j \exp[-\beta c_{ij}].$$

The latter model is called model *B* and it has a  $\beta$  value of 0.1811.

TABLE 4  
THE PREDICTED TRIP MATRIX

Origins \ Destinations	Destinations								$O_t$
	1	2	3	4	5	6	7	8	
1	290	1	31	76	31	21	7	1	458
2	3	59	24	3	1	0	1	0	91
3	55	9	701	69	28	4	24	2	892
4	61	1	46	56	32	4	7	1	208
5	12	0	6	15	257	59	20	2	371
6	66	0	7	17	490	1203	24	6	1813
7	4	0	8	5	31	4	83	25	160
8	1	0	2	1	8	3	66	114	195
$D_j$	492	70	825	242	878	1298	232	151	4188

TABLE 5  
THE GOODNESS-OF-FIT STATISTICS

	Model A	Model B
Residual $\bar{x}$	-0.031	0.000
" S.D.	19.90	129.48
$R^2$	0.988	0.862
Intercept coeff.	0.675	-34.90
Slope "	0.989	1.423
$\chi^2$	406.12	3393.4
$G$	8.90	26.44
Mean error	11.67	34.62
$I/I_{\max}$	0.077	0.3454

Clearly the solution obtained from the entropy maximising model *A* is vastly superior to the one from model *B*. The residual mean from model *A* is very close to the expected value of zero and the standard deviation is relatively small. Furthermore, the total residual distribution does not depart significantly from a normal distribution. Almost 78% of the residuals are within plus or minus one standard deviation and 93% are within two standard deviations of the mean. The coefficient of determination between the predicted and observed number of trip interchanges is  $R^2 = 0.988$ , while the intercept and slope coefficients of the regression line are 0.6769 and 0.9890 respectively. The chi square statistic is more difficult to interpret for a single model. Its principal use seems to be for comparison of performances of competing models. The smaller the value of chi square the better the performance of a model. Here the statistic strongly favours model *A*. The  $G$  statistic is probably the easiest to interpret in terms of overall model performance. Its magnitude indicates that model *A* correctly allocates 91% of the journey to work trips. On the other hand, with model *B* over one quarter of the trips in the predicted matrix would have to be reallocated in order to replicate the observed trip matrix. The mean absolute error in the trip frequencies is 11.67. Finally, the information gain index,  $I$ , also highlights the relatively good performance of model *A*. Since the maximum value for  $I$  is 0.806 only 7.7% more information would be required relative to what is known already to change the predicted matrix to the actual one. In summary, the goodness-of-fit statistics clearly demonstrate that the entropy maximising model given by equations (12)

to (14) provides a very good solution to the journey-to-work problem in Co. Limerick. Similar results have been obtained by the author in a study of work travel patterns in Limerick city. The results also highlight the considerably better performance of the theoretically derived model over the rather simple *ad hoc* formulation represented by model *B*, which, unfortunately, is still used quite often.

The results obtained here differ substantially from those reported for similar studies in England (Hathaway, 1975; Openshaw, 1976; Thomas, 1977). The goodness-of-fit statistics obtained here are much better than those obtained in any of the other studies. In fact, the goodness-of-fit statistics reported by Thomas for Merseyside are not even as good as those for the *ad hoc* unconstrained model *B* in this study. The principal difference between this study and the others is in relation to the zoning system.

It must therefore be emphasised that a trade-off exists between the number of zones and the level of model performance. At present, over-elaborate zoning systems that ignore much of the information available to an investigator tend to yield poor model performances. The information used in designing the zoning system for this study was contained in the interaction matrix and, therefore, the procedure was consistent with entropy maximising. Of course, the approach taken here may not always be possible, according to the purpose of the study. For instance, it may not be very suitable for a transportation study where one of the objectives would be to make recommendations about the alterations that may be required to the network. On the other hand, the approach outlined here may be quite suitable for regional impact studies, particularly in areas that are still largely rural. In any investigation it is important that the zoning system be related to the purpose of the study.

TABLE 6  
THE RESIDUAL MATRIX

Destinations Origins	Destinations							
	1	2	3	4	5	6	7	8
1	-61	0	7	51	-8	9	2	1
2	3	10	-11	-4	0	0	1	0
3	20	-10	4	2	22	-9	-32	2
4	34	1	16	-72	25	-6	1	1
5	-11	0	-13	7	4	0	11	2
6	21	0	-2	14	-70	36	6	-6
7	-5	-1	-2	0	21	-10	-12	9
8	0	0	0	1	6	-21	23	-9

Despite the comparatively good performance of the model there were still some errors in the predictions (Table 6). These require further investigation. The relative importance of errors associated with particular zones can be assessed by the following indices,

$$I_i = \sum_j q_{ij} \log (q_{ij}/p_{ij}^*)$$

$$I_j = \sum_i q_{ij} \log (q_{ij}/p_{ij}^*)$$

The index  $I_i$  is termed the origin information gain and measures the

TABLE 7

## ORIGIN AND DESTINATION INFORMATION GAIN INDICES

Zone	1	2	3	4	5	6	7	8
$I_i$	0.008	0.002	0.011	0.016	0.005	0.005	0.005	0.009
$I_j$	0.008	0.001	0.004	0.017	0.011	0.012	0.006	0.003

relative importance of the errors on all trips leaving zone  $i$ ;  $I_j$  is termed the destination information gain and measures the relative importance of errors on all trips entering zone  $j$ . The values of  $I_i$  and  $I_j$  are contained in Table 7. Clearly zone 4 (Rathkeale) is the one for which it is most difficult to predict work-travel patterns. The principal errors in the model solution were a gross underprediction of the number of intrazonal trips and an overprediction of the numbers travelling from zone 4 to zones 1 and 5. At the same time the number of trips from zone 1 was highly overpredicted. This pattern of errors suggests that in the distance matrix  $d(4,1)$ ,  $d(1,4)$ ,  $d(4,5)$  and  $d(5,4)$  have been underestimated. The route between zones 1 and 4 is a county road of poor quality. Furthermore, given the shape of zone 1 it was decided to double the distance to zone 4. The distance between zones 4 and 5 was increased by 50% to take account of the actual location of employment opportunities, rather than the centroid, in zone 5. Since there was a relatively large underprediction of trips between zones 6 and 5 (Table 6),  $d(6,5)$  was reduced to 9.5. The reduction can be justified because many of the workers in zone 5 reside in part of the city environs which is included in zone 6. After these adjustments were made to the distance matrix the model was recalibrated. The goodness-of-fit statistics are summarised in Table 8.

TABLE 8

## GOODNESS-OF-FIT STATISTICS FOR MODEL WITH ADJUSTED DISTANCES

Residual $\bar{x}$	0.031
" S.D.	13.746
$R^2$	0.994
Intercept coeff.	0.444
Slope "	0.992
$\chi^2$	398.57
$G$	7.00
Mean Error	9.17
$  I_{\max}$	0.053

All these statistics represent an improvement in the model performance (cf. Table 5). This model is able to allocate correctly 93% of the total work trips in the county. The  $\beta$  parameter associated with this model is 0.1892. The other parameters and the emissiveness and attractiveness indices are contained in Table 9. Since the distribution patterns of these parameters and indices are similar to those for Table 3 no further comment is required.

*Discussion*

The distinctive feature of the model described in this paper is that it is the one which provides the least biased *a priori* estimates of the elements of an interaction pattern. This feature distinguishes it from regression



TABLE 9  
PARAMETER VALUES FOR MODEL WITH ADJUSTED DISTANCES

Zone	<i>A</i>	<i>B</i>	$-\lambda^{(1)}$	$-\lambda^{(2)}$
1	3.148	4.247	7.27	7.64
2	12.673	5.362	7.06	5.93
3	3.013	2.785	7.90	7.74
4	2.636	3.953	6.31	6.87
5	1.840	2.958	6.53	7.86
6	2.727	2.155	8.51	7.94
7	3.963	7.233	6.45	7.43
8	4.232	5.539	6.72	6.73
Mean	4.2789	4.2789	7.09	7.27
St. Dev.	3.064	1.5844	0.72	0.66
Coeff. of Var.	0.716	0.370	0.10	0.09

models of interaction which fit estimated distributions to observed ones. When the model was calibrated and tested against data on recent work-travel patterns in Limerick the degree of correspondence between the model and reality was close, after an appropriate zoning system had been identified. In this final section two issues are considered: (1) applications of the model and (2) areas for further research.

The model presented in this paper has at least two types of application. One is concerned with the forecasting and assessment of impacts of alterations in the distribution patterns of the workforce, in sources of employment, or in travel cost change. The second is concerned with providing an interpretation of an observed pattern of interactions.

In the forecasting area there are a number of possibilities and difficulties. For example, in the context of preparing five-year development plans one may have projections of the workforce  $\{O_i\}$  and employment  $\{D_j\}$  distribution for a particular area. If

$$\sum_i O_i = \sum_j D_j$$

and an appropriate assumption is made about the mean trip length then the model presented in this paper could be used to predict the future pattern of interactions.

An alternative application for the model is in the area of impact studies. In that situation the research problem can be defined in the following terms: given a distribution of employment opportunities what is the least biased estimate of the distribution of workers for each industry? The problem is solved by obtaining the probability distribution

$$\{p_{ji}\} \text{ where } p_{ji} = p_{ij}^* / \sum_i p_{ij}^*$$

is the probability that someone works in zone  $j$  and resides in zone  $i$ . From this distribution one can estimate the extent of the spread effects (in terms of employment generation) that may result from the establishment of a new industry. This type of analysis has been carried out in the hinterland of the Shannon Estuary to estimate the spatial effects on employment generation that are likely to be associated with the development of the Alcan project at Aughinish Island.

The value of the model for interpreting a given interaction pattern has

already been demonstrated in this paper. Increased understanding of existing patterns may be achieved through analyses of the *A* and *B* patterns. Analyses of this type can often provide useful information for evaluating planning policies in an area (Thomas, 1977). A more general type of application would be to use the interaction model as a component of a more comprehensive macro-model of landuse patterns (Wilson, 1974).

There are a number of questions arising from the model presented in this paper that require further research. On the one hand, more attention needs to be given to measurement of the costs of interaction and the definition of appropriate zoning systems. The effect of the zoning system design and the measurement of costs on model performance will be examined in a forthcoming paper. On the other hand, attention needs to be directed towards explaining the origin of the constraints that generate the observed patterns of interaction. It is important to remember that since the entropy maximising model is based on a logical principle rather than on a model of behaviour, it does not provide an explanation for an observed trip pattern. In fact it could be argued that the relationship identified by the model is essentially tautological (Sheppard, 1978). Nevertheless, the entropy maximising model makes an important contribution to our understanding of interaction patterns by identifying the constraints that generate the observed pattern. This means that in the Limerick area the questions that need to be answered to provide an explanation for observed travel patterns are (1) why should the value of  $\bar{c}$  be 7.3 miles, given that any value between 5.47 and 17.9 would result in a model which satisfied the origin-destination constraints, and (2) what forces in society determine the employment and workforce distributions? When the answers to these questions are known we will have a better understanding of the organisation of the economic landscape and we will be better able to predict changes in patterns of spatial interaction.

### Conclusion

The principal conclusion to be drawn from this study is that the entropy maximising method provides a sound basis for modelling spatial interaction. Much more research is required on the zoning and cost measurement problems encountered in applying models of this type. Since many of the problems can be solved only by empirical applications it is important that the required data be collected and that models of the type outlined in this paper be tested in a variety of economic landscapes.

### REFERENCES

- |                               |      |                                                                                                                                  |
|-------------------------------|------|----------------------------------------------------------------------------------------------------------------------------------|
| Ayeni, B.                     | 1979 | <i>Concepts and techniques in urban analysis</i> , London.                                                                       |
| Batty, M.                     | 1974 | 'Spatial entropy', <i>Geographical Analysis</i> , 5, 1-33.                                                                       |
| Batty, M. and Mackie, S.      | 1972 | 'The calibration of gravity, entropy and related models of spatial interaction', <i>Environment and Planning</i> , 4, 205-33.    |
| Batty, M. and March, L.       | 1976 | 'The method of residues in urban modelling', <i>Environment and Planning</i> , 8, 189-214.                                       |
| Baxter, R.                    | 1973 | 'Entropy maximising models of spatial interaction', <i>Computer Applications</i> , 1, 57-83.                                     |
| Beckmann, M. and Golob, T. F. | 1972 | 'A critique of entropy and gravity in travel forecasting', in <i>Traffic flow and transportation</i> , ed. Newell, G., New York. |

- Broadbent, T. A. 1970 'Notes on the design of operational models', *Environment and Planning*, 2, 469-76.
- Bussiere, P. and Snickars, F. 1970 'Derivation of the negative exponential model by an entropy maximising method', *Environment and Planning*, 2, 295-301.
- Carrothers, G. A. 1956 'An historical review of the gravity and potential concepts of human interaction', *Journal of American Institute of Planners*, 22, 94-102.
- Cesario, F. J. 1973 'A generalized trip distribution model', *Journal of Regional Science*, 13, 233-47.
- Cesario, F. J. 1974 'The interpretation and calculation of gravity model zone-to-zone adjustment factors', *Environment and Planning*, 6, 247-57.
- Cesario, F. J. 1975 'Least squares estimation of trip distribution parameters', *Transportation Research*, 9, 13-18.
- Cesario, F. J. 1977 'A new interpretation of the "normalizing" or "balancing" factors of gravity-type spatial models', *Socio-Economic Planning Sciences*, 11, 131-6.
- Cesario, F. J. 1979 'Much ado about entropy', *Geographical Analysis*, 11, 189-96.
- Dacey, M. P. and Norcliff, A. 1976 'New entropy models in the social sciences: elementary residential location models', *Environment and Planning A*, 8, 299-310.
- Dublin Transportation Study 1972 *Technical reports*, An Foras Forbartha.
- Evans, A. W. 1971 'The calibration of trip distribution models with exponential or similar cost functions', *Transportation Research*, 5, 15-38.
- Evans, S. P. 1973 'A relationship between the gravity model for trip distribution and the transportation problem in linear programming', *Transportation Research*, 7, 39-61.
- Evans, S. P. and Kirby, H. R. 1974 'A three-dimensional Furness procedure for calibrating gravity models', *Transportation Research*, 8, 105-22.
- Fisch, O. 1977 'On the utility of entropy maximisation', *Geographical Analysis*, 9, 79-84.
- Gould, P. 1972 'Pedagogic review', *Ann. Ass. Am. Geogr.*, 62, 689-700.
- Hansen, S. 1972 'Utility, accessibility and entropy in spatial modelling', *Swedish Journal of Economics*, 74, 35-44.
- Hathaway, P. J. 1973 *Some statistical problems associated with trip distribution models*, Proceedings of PTRC Summer Annual Meeting, University of Sussex, PTRC, London.
- Hathaway, P. J. 1975 'Trip distribution and disaggregation', *Environment and Planning* 7, 71-97.
- Hirst, M. A. 1977 'Hierarchical aggregation procedures for interaction data: a comment', *Environment and Planning A*, 9, 99-103.
- Hyman, G. M. 1969 'The calibration of trip distribution models', *Environment and Planning*, 1, 105-12.
- Isard, W. 1960 *Methods of regional analysis*, Cambridge, Mass.
- Jaynes, E. T. 1957 'Information Theory and Statistical Mechanics I', *Physical Review*, 106, 620-30.
- Jaynes, E. T. 1968 'Prior probabilities', *IEEE Transactions on Systems Science and Cybernetics*, SSC-4, 227-41.
- Kirby, H. R. 1970 'Normalizing factors of the gravity model—an interpretation', *Transportation Research*, 4, 34-50.
- Kirby, H. R. 1974 'Theoretical requirements for calibrating gravity models', *Transportation Research*, 8, 97-104.

- Kirby, H. R. 1975 'A comment on Cesario's note on trip distribution', *Transportation Research*, 9, 83-4.
- Masser, I. and Brown, P. J. B. 1975 'Hierarchical aggregation procedures for interaction data', *Environment and Planning A*, 5, 81-112.
- Masser, I. 1976 'The design of spatial systems for internal migration analysis', *Regional Studies*, 10, 39-52.
- Masser, I. and Brown, P. J. B. 1977 'Spatial representation and spatial interaction', *Papers Regional Science Association*, 38, 71-92.
- Masser, I. 1977 'A comparative analysis of spatial representation in doubly constrained interaction models', *Environment and Planning A*, 9, 759-69.
- Masser, I. 1979 'A note on the treatment of flows across system boundaries in spatial interaction models', *Environment and Planning A*, 11, 447-54.
- Miller, R. E. 1972 *Modern mathematical methods for economics and business*, New York.
- Olsson, G. 1965 'Distance and human interaction: a migration study', *Geografiska Annaler*, 74B, 3-43.
- Openshaw, S. 1976 'An empirical study of some spatial interaction models', *Environment and Planning A*, 8, 23-41.
- Openshaw, S. 1977 'A geographical solution to scale and aggregation problems in region building, partitioning and spatial modelling', *Trans. Inst. Brit. Geogr.*, new series, 2(4), 459-72.
- Openshaw, S. 1977 'Optimal zoning systems for spatial interaction models', *Environment and Planning A*, 9, 169-84.
- Pirie, G. H. 1979 'Measuring accessibility: a review and proposal', *Environment and Planning A*, 11, 299-312.
- Sheppard, E. S. 1976 'Entropy, theory construction and spatial analysis', *Environment and Planning A*, 8, 741-52.
- Sheppard, E. S. 1978 'Theoretical underpinnings of the gravity hypothesis', *Geographical Analysis*, 10, 386-402.
- Thomas, R. W. 1977 'An interpretation of the journey-to-work on Merseyside using entropy maximising methods', *Environment and Planning A*, 9, 817-34.
- Tribus, M. 1969 *Rational description, decisions and designs*, Oxford.
- Walsh, J. A. 1976 'Information theory and classification in geography', *DP 4, Dept. of Geography, McMaster University*, Hamilton, Ontario.
- Walsh, J. A. and Webber, M. J. 1977 'Information theory: some concepts and measures', *Environment and Planning A*, 9, 395-417.
- Walsh, J. A. 1978 'An analysis of intra-urban residential ownership patterns', *Environment and Planning A*, 10, 17-28.
- Walsh, J. A. and O'Kelly, M. 1979 'An information theoretic approach to measurement of spatial inequality', *Econ. & Soc. Review*, 10(4), 267-86.
- Webber, M. J. 1975 'Entropy, maximising location models for non-independent events', *Environment and Planning*, 7, 99-108.
- Webber, M. J. 1976 'The meaning of entropy maximising models', in G. J. Papageorgiou (ed.), *Mathematical land-use theory*, Lexington.
- Webber, M. J. 1977 'Pedagogy again: what is entropy?' *Ann. Ass. Am. Geogr.*, 67, 254-66.
- Webber, M. J. 1979 *Information theory and urban spatial structure*, London.

- Williams, H. C. W. L. 1976 'Travel demand models, duality, relations and user-benefit analysis', *Journal of Regional Science*, 16, 147-66.
- Williams, I. N. 1976 'A comparison of some calibration techniques for doubly constrained models with an exponential cost function', *Transportation Research*, 10, 91-104.
- Williams, I. N. 1977 'Algorithm 1: three-point rational function interpolation for calibrating gravity models', *Environment and Planning*, 9, 215-21.
- Wilson A. G. 1967 'A statistical theory of spatial distribution models', *Transportation Research*, 1, 253-69.
- Wilson, A. G. 1969 'Developments of some elementary residential location models', *Journal of Regional Science*, 9, 377-85.
- Wilson, A. G., Hawkins, A. F., 1969 'Calibrating and testing the SELNEC transport  
Hill, G. J. and Wagon, D. J. model', *Regional Studies*, 3, 337-50.
- Wilson, A. G. 1970 *Entropy in urban and regional modelling*, London.
- Wilson, A. G. 1974 *Urban and regional models in geography and planning*, London.
- Wilson, A. G. and Kirby, M. J. 1975 *Mathematics for geographers and planners*, Oxford.
- Wilson, A. G. and Senior, M. L. 1974 'Some relationships between entropy maximising models, mathematical programming models and their duals', *Journal of Regional Science*, 14, 207-15.

## ACKNOWLEDGMENTS

The author gratefully acknowledges the assistance of Richard Tobin, Chief Planning Officer, Limerick Corporation, who initiated the original research on which this paper is based and who provided invaluable technical expertise with the programming of the model. Thanks are also extended to Limerick County Council for providing the survey data used in the study.