EMPIRICAL TESTS OF
AN ENTROPY MAXIMISING MODEL

EMPIRICAL TESTS OF AN ENTROPY MAXIMISING MODEL OF RETAIL LOCATION AND CONSUMER BEHAVIOUR

BY

MORTON E. O'KELLY B.A.

A Research Paper
Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree
Master of Arts
McMaster University

July 1978

MASTER OF ARTS (1978)
(Geography)

MCMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Empirical Tests of an Entropy Maximising Model of Retail Location and Consumer Behaviour

AUTHOR: Morton Edward O'Kelly, B.A. (University College, Dublin)

SUPERVISOR: Dr. M.J. Webber

NUMBER OF PAGES: xi, 153

ABSTRACT

The aim of this research paper is to show that entropy maximising models of urban land use are better than first generation large scale land use models. In order to substantiate this claim it will be argued that entropy maximising models are internally consistent, theoretically based, realistic and parsimonious.

First it is shown that an extended family of consistent spatial interaction models can be built using the entropy maximising (information minimising) formalism.

The second step is to develop a detailed model. The model presented represents several aspects of urban spatial interaction. In the model, individuals are assigned to place of residence, place of work and shop trip pattern for an arbitrary distribution of service centres. An entropy statistic, defined over the assignment of individuals is used to determine, endogenously, the entropy maximising distribution of centres. Thus, the model predicts the location of retail facilities as well as residential location, shopping patterns and work trip interchanges. The model also has a theoretical basis in that results using a similar formulation for a linear city show that the entropy maximising distribution of service centres includes downtown.

Having set up a theoretically based, consistent model, the next step is to establish that this model produces a realistic representation of the city. Data from Hamilton, Ontario are used to test the fit of the

model.

Results show that the model reproduces the observed data with some accuracy. Furthermore, the endogenously predicted distribution of service centres includes downtown. This result supports the contention that theoretically based entropy models can produce realistic results, and establishes the argument of the paper.

ACKNOWLEDGEMENTS

This research paper is based on a project carried out during the summer of 1977. This project was initiated and designed by my supervisor Mike Webber. Peter Hall wrote the computer programs to edit the data and also the programs based on the model. The applied part of this paper is therefore, the result of a group effort and I am grateful to Mike and Peter for their guidance. Fred Hall commented on an earlier draft and his advice has improved the development of the paper. Sharon Wright typed the paper with her usual efficiency and good-humour.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

There has been a long tradition in geography and related disciplines of urban model building. These models have developed over the years from schematic representations of land use within the city to operational mathematical models of residential, retail and other land uses. Although these models had become highly developed by the mid 60's, they failed to satisfy critics of their usefulness, and by the late 60's the first phase of urban model building had come to a close.

The aim of this research paper is to show that entropy maximising models of urban structure are better than first generation land use models. Chapter 1 reviews first generation land use models developed in North America in the 60's. This review classifies the models and then focuses on the internal consistency and theoretical basis of four major types of model; simulation, linear, optimising and gravity models. The weaknesses of these models are pointed out and requirements for improved models are outlined. To show that entropy maximising models avoid the faults of first generation models it is required to show that the entropy maximising formalism is a method for deriving consistent, theoretically based, realistic models of urban structure.

Chapter 2 aims to derive consistent models which incorporate origin and destination characteristics. The entropy maximising formalism and an extension based on Kullback's information measure are discussed and these techniques are used to derive a consistent family of spatial interaction models.

Not only does the entropy maximising formalism provide a means of deriving consistent operational models but it has also been used to make theoretical statements about urban areas. The following model is considered in Chapter 3. Assume that there exists a set of service centres and a set of individuals. The individuals are classified by place of residence, place of work and shopping trip pattern. The assignment of individuals to these classes is an entropy maximising assignment subject to constraints on (a) work trip destinations, (b) mean trip length for both work trips and shopping trips, (c) mean number of shopping trips per household, *for a pattern of shopping centres chosen by the modeller.* On each iteration a new pattern of centres is chosen and individuals are again assigned to classes. The object is to find that pattern of centres which maximises the entropy of the assignment of individuals. Webber (1977b) has shown a Hotelling type solution (i.e. firms agglomerate) to be entropy maximising for a linear city if jobs are not locationally fixed. This paper reports the results of tests of the model in more realistic circumstances.

The aim is to show that the pattern of centres predicted by the model corresponds to observed patterns in that it is concentrated near the city centre. Furthermore, the model *should*, if it is realistic, match other aspects of the city: worktrips, shopping trips and shopping behaviour by place of residence. If these distributions can be replicated with high correlations the claim of realism for the model will have been justified. Note that a failure on the part of the model to reproduce real world data will suggest new constraints and will hopefully draw attention to overlooked processes. The models developed in Chapter 4 proceed in this way - taking the errors of one model as a guideline for further model development. Care is taken to heed Batty's (1976) call for parsimony. This requirement of urban models is particularly relevant in entropy models as it is always tempting to continue adding constraints. It is clear however that by adding constraints the modeller is reducing average uncertainty, and ultimately the model will become deterministic.

With this restriction in mind, Chapter 4 develops and tests several entropy maximising models which encompass the real world processes of retail agglomeration economies, employment allocation and land use competition. In each of the models the aim is to find (1) the entropy maximising distribution of centres; and (2) to compare the assignment of individuals with observed patterns.

If it can be shown that a parsimonious, theoretically based model is also realistic the aim of the research will have been achieved.

CHAPTER 1

URBAN LAND USE MODELS: A CLASSIFICATION AND REVIEW OF NORTH AMERICAN EXAMPLES TO 1968

1.1 Introduction

Many models of urban land use have been formulated and tested. In a recent review, Senior (1973; 1974) separated these models into three broad categories - urban ecological, spatial interaction and economic models. This paper focusses on operational models, that is those spatial interaction and economic models which have been applied using real world data. The models reviewed are mainly spatial interaction models but reference is also made to the Herbert and Stevens (1960) operationalization of Alonso's work. Before classifying and discussing the group of operational models some terms must be clarified. This is a relatively simple task since there have been a number of discussions of the role of theory in models (see for example Harris, 1966) and the role of models in planning (see for example Lee, D., 1973). Many different but essentially complementary definitions of modelling have been given. These are now summarized in an increasingly specific order.

At a general level a model is a simplified statement of important elements of a real world situation. Abstract mathematical models, which will be of concern here, represent the real world by symbols and allow the modeller to experiment with the city. A particularly important class of models is that defined by Harris (1966, 8) as *experimental designs based on some theory.* Not all models of land use are based on theory - in fact some rely only on statistical regularity - but it will be argued that models can only perform certain desirable tasks if they are founded on some theoretical notions. Actually the importance of a theoretical background varies with the function of the model (namely description, prediction or prescription; see Lowry, 1965; later goal statements are referred to in Lee, D., 1973, footnotes 3 and 4). Descriptive models are intended to replicate existing urban structure and as such may be purely statistical. Predictive models must specify cause and effect relationships. An important type of predictive model is the conditional forecast in which the effects of future changes in exogenous variables on other variables are examined (Lowry, 1965). Obviously the validity of the cause and effect relationship (and hence the accuracy of the conditional prediction) will depend on the strength of the theoretical foundation of the model (see Lee, D., 1973). Planning or prescriptive models specify a preferred solution from a set of alternatives. This involves the evaluation of policy packages. Difficult decisions have to be made on the costs and benefits of various options and some weight has to be attached to the importance of different social groups. Wilson (1974) and Lichfield *et al.* (1975) have shown that models can be used together with cost-benefit analysis to achieve some rigour in the design and evaluation stages of the planning process. The task of modellers in this area is not made any easier by the conspicuous lack of definite theories for the analysis of urban phenomena.

The purpose of this chapter is to review some "large-scale land use models" (Lee, D., 1973) with reference to their theoretical content. In section 1.2 the early development of land use modelling in North America is traced. Section 1.3 provides a classification of land use models and discusses the decisions facing model builders. (While the main focus here is on early land use models some comparison will be made with important recent methodological developments whenever this is appropriate.) In section 1.4 examples of early models are used to illustrate some of the many possible ways of examining the city. The Lowry model in particular is examined as it provided much of the impetus for later models of spatial interaction.

1.2 The development of land use modelling in North America

In the early post war years transportation planners did not recognize the interdependence between land use and transportation (McLoughlin, 1969; Isard, 1975). As Isard (1975) puts it there was a misunderstanding of the effects of the development of the transportation system on land use. Planners were involved in the design of physical equipment (McLoughlin, 1969, 66). A master plan or a blueprint for the entire city was drawn up and decisions were then made individually using the master plan as a guide. Effectively this meant that the city was treated as a patchwork of isolated problems. McLoughlin (1969, 65) suggests that at this stage land use problems "were dealt with one at a time by improvements conceived in isolation". This then was an era of small scale planning where the concern was with the detailed design of small parts of the city without very much realization of the possible repercussions of these changes. To begin with this method worked - the relatively small scale of investment in urban transportation facilities meant that there was no need for a wider view.

An abrupt change was precipitated by the unprecedented scale of the National Highway Program. Suddenly traffic engineers found themselves with problems on a vast scale for which solutions had to be found rapidly. There was a change from maintaining the road network by small scale piecemeal planning to a planning for "region wide integrated multi-modal systems" (Wingo, 1961). The scale (in dollar terms) of the investment was so great that planners were forced to predict the future trends in population and employment and a greatly increased awareness of the dependence of transportation on land use resulted. At first the idea of the dependence of traffic on land use was developed in the work of Mitchell and Rapkin (1954) and in the transportation studies in Chicago (Hamburg and Creighton, 1959) and Detroit (Kain, 1962a). The main emphasis was on forecasting land use as a means of generating future traffic patterns. Although the idea of interdependence between land use and transportation was initially misunderstood or ignored (McLoughlin, 1969) research was directed towards this important factor by Voorhees (1959, 58) and Hansen (1959). These works represent the beginning of the land use modelling efforts of the 60's.

1.3 A classification of land use models

This section presents a classification of operational urban land use models. Only the main characteristics of these models are mentioned. Detailed descriptions can be found in the original works (referred to below). For critical reviews of these models see James (1974), Putman

(1975), Lee, D. (1973), Lowry (1968) and Brown *et al.* (1972). More general reviews include Lee, C. (1973), Krueckeberg and Silvers (1974) and Masser (1972).

This wide range of land use models and forecasting techniques has been organized under many different classification schemes. For example Reif (1973) proposed a set of three dimensions, Harris (1968) outlined and discussed a number of dichotomies and Wilson (1974) suggested nine characteristics of models. Table 1 is a synopsis of these efforts. It is based on Batty's (1976a) distinction between substantive issues and design characteristics in model building. (This is a distinction that was recognized by Garrison (1966) when he wrote of behavioural and structural decisions in model design; and also by Harris (1961; 1965).) The various solutions to the two problems which face model builders (i.e. what structure to use and how to design the model) provide a convenient method of discussing the dimensions of urban land use models.

(1) Substantive Issues

The major substantive issues facing a model builder are:

(i) The level of explanation
(ii) Positive-normative decision
(iii) Level of aggregation
(iv) The treatment of time

These issues each present the modeller with a set of choices. Each dimension is therefore discussed in turn as a set of choices facing the model builder.

(i) Choice of the level of explanation.

At a very broad level of generalization the choice of the level of explanation is one between a deductive and an inductive approach. More specifically the choice can be crystalized as being between a micro level

(1) Substantive Issues

		CLUG	UNC	EMPIRIC	LOWRY	LAKSHMANAN-HANSEN	SEWRPC	HERBERT-STEVENS	URBAN SPATIAL STRUCTURE MODEL
(i) Level of explanation	micro-behavioural	•					•	•	
	meso-statistical			•					
	macro-growth		•	•	•	•			•
(ii) Normative focus	optimising						•	•	
	non-optimising	•	•	•	•	•			•
(iii) Level of aggregation	spatial micro		•						
	spatial macro	•		•	•	•	•	•	•
Level of aggregation	sectoral partial		•			•		•	
	sectoral general	•		•	•		•		•
(iv) Time	Static		•		•	•	•		•
	Dynamic	•		•				•	

(2) Design Issues

		CLUG	UNC	EMPIRIC	LOWRY	LAKSHMANAN-HANSEN	SEWRPC	HERBERT-STEVENS	URBAN SPATIAL STRUCTURE MODEL
(i) Formulation	linear		•	•			•	•	
	non-linear	•			•	•			•
(ii) Solution strategy	simultaneous		•	•	•	•	•	•	•
	sequential	•							
(iii) Solution techniques	analytical			•					
	numerical				•	•	•	•	•
	simulation	•	•						
(iv) Endogenous variables	residential LU	•	•		•		•	•	•
	industrial LU	•					•		
	retail LU	•			•	•	•		•
	population			•	•		•		
	primary employ.			•					
	secondary employ.			•					•
	tertiary employ.			•	•				

Table 1: Urban Land Use Models

of explanation at which behavioural relationships are specified and a macro approach which is based purely on symmetrical relationships (Karlqvist, 1977).

The micro level of analysis is dominated by a behavioural approach and relies mainly on causal relationships. This presumably allows them to be used in long term predictions and also to be extended to dynamics. Of the models to be discussed below the Herbert and Stevens (1960) model exemplifies this approach. There are real problems associated with the operationalization of this and other models based on consumer behaviour concepts.

At the meso level of explanation there are models based on the entropy maximising formalism. These models consider the characteristics of groups of actors rather than of specific individuals (Karlqvist, 1977, 4). These system wide characteristics are used to assign groups of individuals to classes in a minimally prejudiced way. Models of this type are currently in vogue through the work of Wilson (1967) and will be discussed in Chapter 2 below where reference will be made to the foundations of the approach in statistical mechanics and information theory.

At the macro level of explanation theory deals with the properties of the system as a whole and is not related to the actions of individual decision makers (Karlqvist, 1977, 6). Much geographical research has been at this level. For example the gravity model of interaction is based on properties of places (e.g. populations or attractions) and not on any realized travel pattern. As Karlqvist (1977, 6) puts it "a population centre can act as an attractive force without the presumption that any travel or transportation is actually realized over space".

A fundamental question which arises is the appropriate level of explanation for any given situation. In general the micro level models are faced with an aggregation problem (i.e. how to add up individual preferences to some form of collective choice). At the other extreme macro models should be used only for short run prediction since parameters are unlikely to be stable over longer periods. The meso level methods provide operational formulations without difficulty and the model reported in Chapter 3 will be at this level.

(ii) Normative focus

The normative focus of the model has a different interpretation depending on the place of the model on the description-prescription continuum. In models used for description and prediction normative considerations enter in the following way: we impute to individuals an optimising behaviour that hopefully reproduces their actual behaviour (Harris, 1961, 714). In other words we assume that the actors in the model display optimising behaviour. Although the major focus here will not be on such optimising behaviour there is one exception - the important work of Herbert and Stevens (1960) in which the Alonso bid rent concept is operationalized. (Recent work connecting this model with spatial interaction includes Senior and Wilson (1974a); see section 1.4(3).)

In prescriptive models normative considerations enter via the judgement by the modeller of what is good. In other words the modeller picks the plan which satisfies certain criteria, perhaps by using mathematical programming techniques. The Schlager (1966) model (SEWRPC) to be discussed below exemplifies this approach.

(iii) Scale of analysis

a. The level of sectoral aggregation

This distinction is perhaps more clearly seen than some of the others but presents one of the major problems currently facing model builders. Partial models are obviously those which focus on one or two aspects of the urban area. (Many models examine only the retail sector.) Conversely general models are those which in some way combine the sectors of the urban economy into a much larger model.

Two problems are apparent. Firstly there is the difficulty of linking the subsections of the urban economy in a consistent way; this problem is treated in Wilson (1974). Secondly even when a number of sectors have been completely articulated there is still the problem of incompleteness - even the most general of current models of land use does not account for interrelationships with other aspects of the urban economy such as health and recreation. While there is obviously a need for general urban models the problems posed by the construction of such models mean that one is confined to models which tackle only the major sectors (e.g. Lowry, 3 major sectors). Lowry (1966) has argued for this macro approach but he states that the theoretical basis for macro models should come from micro behavioural studies.

b. The level of spatial aggregation

The choice here is that of the size of zones. The problem is generally one of data availability. For a small zone system one is faced with large input requirements and large travel time matrices. In the modelling efforts of the 1960's there was a tendency to use detailed zonal systems based on specially assembled data sets. More recently Broadbent

(1970) has stressed the necessity of designing zonal systems which meet certain requirements specific to spatial interaction modelling (see below Chapter 2). Openshaw (1976) has shown that zoning has implications for model accuracy - which indicates the importance of the zoning scheme quite apart from data consideration.

(iv) The treatment of time.

The problem of treating time in models of urban land use is a difficult one. The majority of models constructed have begun as static models with no explicit treatment of time, though some have subsequently been converted to dynamic versions (see below 1.4(4)).

Two approaches to constructing dynamic models are possible: firstly there is the treatment of time recursively (i.e. iterations in which the computer distributes increments to the previous total). EMPIRIC (Hill *et al.*, 1965) represents this approach. A second approach to dynamic modelling would be to treat time explicitly. The major problem associated with constructing such a dynamic model is technical (i.e. it is difficult to formulate the necessary equations). Batty (1977) also raises a philosophical issue of wider importance - the difficulty of measuring variables in a dynamic context. (Some recent efforts in this area include Wilson's (1974) quasi-dynamic residential location model and Batty's (1977) work.)

Because of these difficulties most urban land use models are static. A useful compromise is the 'one-shot' or conditional forecast (Lowry, 1965). The modeller uses an exogenous forecast of say basic employment to produce a conditional prediction for some future period. The reliability of this method depends on the ability of planners to predict future parameter values.

(2) Design Issues

Design issues in urban modelling arise as a result of having chosen a model structure. They are consequences of substantive issues. The problem of model design is one involving the translation of the required structure into an inexpensive operational model which will produce, as output, values for certain variables. Depending on the structure of the model this output will be a description, prediction or prescription for the city. The question of how this solution is found is a technical matter and will probably vary for every application but there are some broad issues which are tackled when designing an operational form of any model. The four main areas are:

(i) Model formulation - linear and non-linear
(ii) The solution strategy
(iii) The solution technique
(iv) The quantified endogenous variables.

(i) Model Formulation - linear or non-linear.

It has been suggested (Wilson, 1974, 316) that there is an inherent non-linearity in urban travel behaviour and this has important repercussions for any model based on the interaction of activities. These models are generally non-linear, and log transformations are sometimes used as a method of estimating parameters (in Chapter 2.3(2) below other methods of parameter estimation are mentioned).

(ii) Solution Strategy

The choice here is between simultaneous solution or sequential solution. In simultaneous solutions all the allocations are assumed to adjust to each other at one time. This is obviously a required characteristic of static models. It is also possible to solve dynamic models for discrete points in time in a simultaneous way. However Harris (1968,

380) points out that it is more likely that sectoral location models are solved sequentially for each recursive stage.

(iii) Solution Method

Three solution methods have been used: analytical, numerical and simulation. Analytical solutions usually arise only in regression models. Numerical solutions are inherently required in models which have simultaneous non-linear equations or which are optimising. Simulation as a method of solution involves use of a random number generator and some probability distribution over different spatial locations. This use of simulation arises in situations where there is supposed to be some probabilistic (random) process at work. Examples include Kibel (1972), Chapin and Weiss (1968) and Feldt (1972). Other uses of simulation are discussed in Harris (1961, 712) and include cases where (1) the number of variables is too large for a straight forward analytical solution, (2) the relationships between the variables are non-linear and (3) the model is dynamic.

(iv) The choice of variables

It is difficult to separate clearly the design issue of the endogenous variables from the substantive issue of the "level of sectoral aggregation". In constructing a model the two decisions are probably made together - however it can be seen that the choice of a particular level of sectoral aggregation does to an extent determine the types of variables which must be included.

The choice of variables enters into model design through the role of the modeller as a predictor. Many land use models are connected to much larger transportation studies. The output required from the model is determined by the demands of other stages in the development of the transportation plan.

Some models have been set up as highly disaggregated representations of just one sector. Examples include the San Francisco Housing Renewal Model (see Robinson *et al.*, 1965) and the housing market model of Herbert and Stevens (1960). These models required large amounts of data and predict housing production and rents respectively.

Most land use models attempt to represent only the interrelationships between the major sectors of the urban economy (i.e. residential, service employment and basic employment). Harris (1968) describes some of the models which have been developed to explain each of these sectors. The major effort was that of Lowry (1964) who achieved a simplification of the problem of locating land uses which still forms the basis for many applied models (see below 1.4(4) and 2.3).

1.4 Examples of modelling types

The basic dimensions of urban land use modelling have now been discussed. Some of the models mentioned are now described in more detail. There are four subsections:

(1) Simulation modelling
(2) Linear models
(3) Optimising models
(4) Gravity based models

The aim in developing this section is to present a very brief review of the first three types of model and to concentrate more substantially on the fourth category as it provides a link with later location models.

(1) Simulation Models

Simulation modelling is the term used to describe models which rely on probability distributions to allocate activity over an area in association with some random process. Examples in an urban context include CLUG (Feldt, 1972), Kibel's (1975) use of Monte Carlo Simulation and the UNC models (Chapin and Weiss, 1968).

The UNC and the Kibel (1972) models have a similar mechanism. An attractiveness score is assigned to each unit of undeveloped land within the area and the probability of conversion to residential use during the ensuing forecasting period is proportional to that cell's attractiveness score. Discrete units of development are assigned to cells by random sampling from the resulting probability distribution. The sampling process continues until enough land has been developed to accomodate the given increment of urban population (Lowry, 1968, 132).

It is important to note that these models can only perform an allocation function (i.e. assigning units of development to cells). The models are well suited to distributing activity over space but they are not adaptable to making operational forecasts, and the difficulty of complete market simulation robs these models of usefulness in applied problems. The models do however provide an intuitive understanding of the development of the city through the scoring of attractiveness.

(2) Linear Models

As an example of a linear model the EMPIRIC model will be discussed. The authors of the model (Hill, 1965; Hill *et al.*, 1965) make a distinction between located variables (measures of economic activity or location) and locator variables (explanatory variables). The model is formulated as a set of simultaneous linear equations for each district with one equation for each population or employment variable. The dependent variable is the change during the forecasting interval of the district's share of the regional total for that activity. After the model has been solved the changes-in-shares are added to the shares held by the district at the beginning of the forecast interval and the revised share determines the distribution of independently forecast totals for each activity group (Lowry, 1968, 134). There is some question about the suitability of a change in share formulation. As Lowry (1968, 134) points out, change in share is specified without any reference to change in volume, thus one activity can decrease in one zone while another increases and yet both could have positive changes in share. However the authors of the model did know the overall totals of activity (see Hill's (1968, 154) reply to Lowry).

The model was subjected to extensive testing (Hill *et al.*, 1965; Brand *et al.*, 1967). There was investigation into the model's sensitivity to variations in forecast conditions including the length of the forecast period, the specified regional growth rates, the zoning policies for different suburban communities and transportation facilities. It was found that as the length of the forecast period increased errors in the model compounded. One solution to this is to use the model in a 5-10 year recursive framework in which the outputs from one run form the starting values for the next run. Increasing the specified regional growth rate had the effect of increasing absolute population and employment in all the individual zones. However changes in share in each subregion varied. Population decreased in the Boston area and increased in the suburbs. Employment increased in the core area (Boston City) and decreased in the suburbs. The model predictions were also shown to be sensitive to changes in zoning practice. Tests using an improved radial transportation network predicted increased employment in the outer suburbs and also in the core areas of Boston, Cambridge and Somerville (see Hill *et al.*, 1965; for details).

(3) Optimising Models

Optimising models of urban land use include the Herbert-Stevens (1960) model and Schlager's (1966) land use plan design model.

The Herbert and Stevens (1960) model was designed as a component of the Penn-Jersey Transportation Study; however the model was eventually abandoned by that study in favour of other approaches. The model itself is a residential allocation procedure based on an exogenous amount of land and number of households. The model involves concepts linked with Alonso's (1964) formulation of urban land rent. In the model households choose a housing bundle made up of site and dwelling. Bid rent is based on the fact that the household is prepared to pay a certain amount for the entire bundle and some smaller amount for the bundle excluding site. The difference is defined as bid rent.

The model assigns households to sites so as to maximise the aggregate bid rent of the region's population. The assignment of households to sites which maximises bid rent is mathematically equivalent to the process by which the market clearing solution is found in Lowry's (1968) land market paradigm.

As Wilson (1975) points out, the Herbert-Stevens model only deals with spatial interaction implicitly: journey to work costs are inherent in the price of the site (they were originally explicitly entered into the budget but they were later dropped for technical reasons (see Lowry, 1968, 140)). Recent work by Senior and Wilson (1974a) achieves more explicit spatial interaction terms.

Schlager's (1966) land use plan design model was designed as a component of the South East Wisconsin Regional Planning Commission's regional model. The model is set up as a linear program. The objective function relates to the cost of developing land for a given site.

$$c = c_1x_1 + c_2x_2 + \ldots + c_nx_n$$

where: x represents land use types

c represents costs of development.

The solution to the program must satisfy an equality constraint on the total demand requirement for each land use category. Inequality constraints are imposed to limit the amount of land use within a zone and to constrain interzonal (or intrazonal) land use relationships.

(4) Gravity Based Models

As Batty (1976a) points out, a particularly fruitful area of research in this field has been based on gravity models. The best known example is the Pittsburgh model (Lowry, 1964). Other important papers include those by Hansen (1959) on accessibility, Lakshmanan and Hansen (1965) on retail location and Reilly (1931) on trade areas. Harris (1968, 388) refers to other applications. The gravity model has played an important role in these models and in general gravity based allocation models have been empirically successful.

Later developments in spatial interaction modelling are based on a newer derivation of the gravity model using entropy maximising techniques (see Chapters 2 and 4). (It should be noted that the intervening-opportunities models developed by Schneider (1959) and Harris (1964) are not considered here.)

As was stated above the Lowry (1964) model (hereafter LM) is a well known example of gravity based spatial allocation models and in the following paragraphs the model will be discussed in some detail. The justification for this change in emphasis is that the LM has proved to be a starting point for the construction of much more sophisticated models and it also impinges directly on Wilson's (1974) notion of a general land use model. Furthermore the basic mechanism of the model has some similarities to the mechanism of the model to be described in Chapters 3 and 4.

The LM was set up with the objective of developing "an analytical model capable of assigning urban activities to sub-areas of a bounded region in accordance with those principles of locational interdependence that could be reduced to quantitative form" (Lowry, 1964, 2). In the model urban activities are divided into three broad groups:

> A *basic sector*, including industrial, business and administrative establishments whose clients are predominantly non-local. Because these industries are dependent on events outside the local economy they are treated as exogenous.
>
> A *retail sector*, including those business, administrative and other establishments which deal predominantly and directly with the local residential population. The level of employment in this sector depends on local factors and is treated as endogenous.
>
> A *household sector*, consisting of the residential population. The spatial distribution of the households is determined by the distribution of workplaces.

The mechanisms of the model can be illustrated by a simple diagram (see Figure 1). This mechanism provided a new method with which to view urban

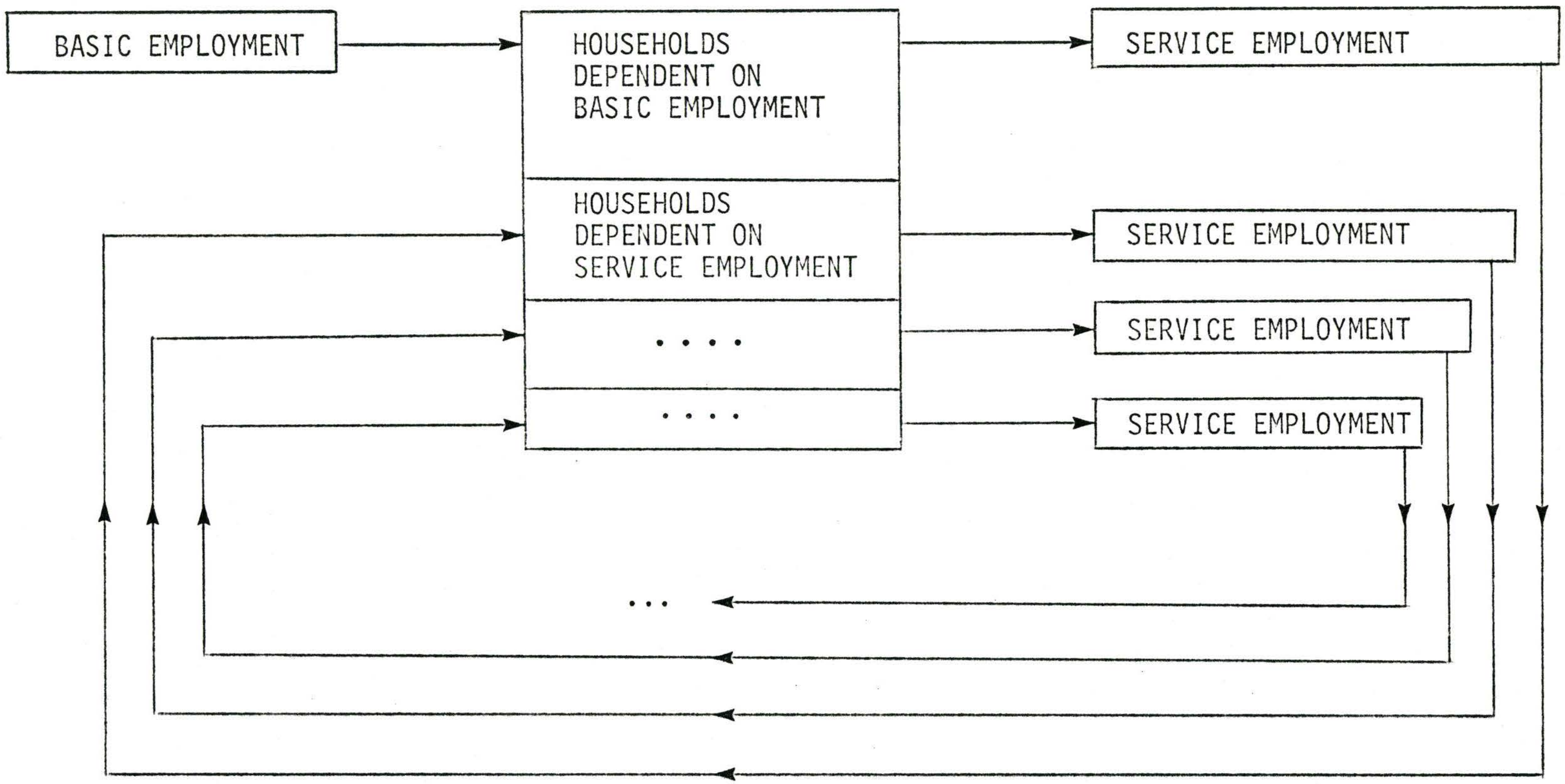

Figure 1: The Mechanism of the Lowry Model (from Goldner, 1971).

areas in an "elementary" way. (Elementary in Wilson's (1969d) sense that it abstracts from individuals preferences and utilities (i.e. a macro level of explanation).) The LM starts by allocating the exogenous level of basic employment to residential areas. An activity rate is then applied which scales employment to population. The next step is to find the amount of service employment required by the population and to allocate these service employees to zones. Obviously this increment of employment in each zone (due to the service sector) will have to be allocated to residences. This procedure is repeated until the additional increments of population and employment become too small to be noticeable. The conversion factors between basic employment and population, and, population and service employment are presumed to be known exogenously (see Batty, 1970, 96). The original application of this model was to data for Pittsburgh and the experience gained in that application enabled Lowry to point out many weaknesses in the model. Three specific weaknesses are now discussed.

The distribution of retail employment is based on an equation of the form: (Lowry, 1964, 17)

$$E_j^k = b^k[c^k \sum_{i=1}^{n} \frac{N_i}{T_{ij}^k} + d^k \cdot E_j]$$

where: E_j = basic employment in zone j,

E_j^k = employment in the kth retail type in zone j,

N_i = population in zone i,

T_{ij}^k = a function of distance for the kth retail type,

b^k = a balancing factor to ensure $E^k = \sum_{j=1}^{n} E_j^k$,

c^k = weight of home based shopping trips,

d^k = weight of work based shopping trips.

Clearly retail employment depends on both the population potential at each zone and the amount of basic employment in the zone itself. The assumption being made here is that work based shopping trips end in the same zone as they start in, therefore the model will overestimate retail employment in zones with large amounts of basic employment. It should be noted that the second term in brackets above could be 'distributed' so that retail employment would then depend on both population potential and 'employment' potential.

Specific criticism can also be levelled at the residential location model. As Lowry (1964, 11, equation 8) originally formulated it:

$$N_j = g \sum_{i=1}^{n} \frac{E_i}{T_{ij}}$$

where: g = a balancing factor to ensure $\sum_{j=1}^{n} N_j = N$,

and the other variables are as above.

In this equation N_j (the number of households in each zone) is a function of that zone's accessibility to employment opportunities. As Wilson (1969b) shows, this can be written as:

$$N_j = g \sum_{i=1}^{n} E_i \, f_{res}(c_{ij})$$

where: f_{res} is an increasing function of travel cost c_{ij}.

This is a simple spatial interaction model and Wilson (1974, 222) points out that this equation implicitly estimates journey to work as:

$$T_{ij} = g\, E_i\, f_{res}(c_{ij})$$

and this will not in general satisfy (Senior, 1973, 179):

$$\sum_{i=1}^{n} T_{ij} = E_j$$

Part of the major revision of the LM carried out by Wilson (1974, Chapter 11) and Batty (1976a) has been the invention of suitable algorithms to insure that all the constraints are met while at the same time maintaining consistency.

A third specific criticism of the LM is its lack of ability to incorporate any scale economies in the retail sector. As Lowry (1964, 24) recognized "above minimum efficient size there are neither internal nor external economies of scale". As Lowry states, the errors from this assumption are probably small; however a useful feature of the model to be described in Chapter 4 is that it does encompass agglomeration economies in the retail sector.

To these specific criticisms related to the design characteristics of the model can be added a number of substantive criticisms (see also Fleischer, 1965). The LM is of course open to a question which applies to all modelling efforts - i.e. is the simplification of the urban system that is implied by the model a valid representation of the real world? While this model does have a strong representation of the sectors of the urban economy it is assumed that these sectors adjust in an equilibrium fashion. Equilibrium assumes that there is a balance between a given level of basic employment and transportation facilities and the associated population and retail distributions. Equilibrium is found simultaneously - there is no lag built into the model (Colenutt, 1970, 137). In short the model is atemporal in that it does not consider the process of adjustment of the endogenous sectors to the exogenous sectors (Cordey Hayes, 1972, 366). The use of the model in forecasting situations is consequently purely as a comparative statics device. The iterative sequence used to find the solution to the set of simultaneous equations should not therefore be mistaken for a dynamic representation of the city.

Improvements to the LM have been numerous. Some of these are reviewed in Goldner (1971) and Putman (1975). It is not proposed to dwell on these modifications but more detailed references to the British applied work and the results of these efforts will be discussed in Chapter 2.

1.5 General criticism of large scale models

A characteristic of the large scale modelling efforts of the 1960's was the large amount of constructive criticism offered in review articles (see for example Harris (1965; 1968) and Lowry (1968)). In the late 1960's there was a realization that large scale land use models could not be expected to achieve all the goals which had originally been set for them. One writer (Lee, D., 1973) was led to proclaim the death of large scale models and provided a 'requiem' to mark their passing. Lee, D. (1973) accused large scale models of excessive comprehensiveness, grossness, hungriness and inscrutability. His argument states that these models attempt to replicate systems that are really too complex and yet they fail to produce the level of detail required by policy makers. Furthermore large scale land use models require large amounts of data and in general cost a lot of money. This criticism was delivered at a stage when model building in North America had decreased substantially and before the current growth of British efforts.

It is interesting to note that the preface to Batty's (1976) work contains a list of desiderata for model building which is almost the reverse of Lee's list of 'sins'. Batty (1976a, xxii) requires that urban models have the following characteristics: simplicity, parsimony, clarity and compromise. The use of entropy maximising methods has made it possible to keep to these ground rules. In particular, as will be argued in the next three chapters, the entropy maximising formalism provides a methodology for deriving consistent, theoretically based, parsimonious and realistic models.

CHAPTER 2

A METHODOLOGY FOR SPATIAL INTERACTION MODELLING

2.1 Introduction

The previous chapter demonstrated the weaknesses of several first generation land use models. The purpose of this chapter is to show that there is a methodology (entropy maximisation) available which allows the urban model builder to avoid the weaknesses of the first generation models. Section 2.2 begins with a discussion of the entropy approach to spatial interaction modelling. The following section derives two measures - Shannon's entropy and Kullback's information.

These measures are then used to construct a family of spatial interaction models. A review of applications of spatial interaction models (2.5) provides guidelines for applied model building.

2.2 The use of entropy maximisation in spatial interaction models

The gravity model was developed as a conceptualization of human spatial interaction by Ravenstein (see Carrothers, 1956; Reilly, 1931; Stewart, 1947 and Zipf, 1949). These developments are reviewed in Carrothers (1957). In these works interaction between two centres of population varies directly with some function of the population size of the two centres and inversely with some function of the distance between them. The hypothesis is therefore that the Newtonian gravity model provides a representation of human spatial interaction. The validity of the gravity model has been questioned on both theoretical and empirical grounds. For example Smith (1977) points out that it has yet to be shown how distance aversion behaviour expressible solely in terms of trip probabilities can be equated with gravity type representations of these probabilities. Smith's (1975; 1977) work does however contribute to a more formal basis for the model. On empirical grounds the seemingly good fit of the model could arise from multicollinearity in the variables. It is also possible that error terms in the model are spatially autocorrelated (see also Huff, 1962; Olsson, 1970 and Ewing, 1974). Despite these deficiencies the gravity model has been used extensively in various forms as a model of spatial interaction. Wilson (1967) provides a common framework for these applications and then goes on to show that they can be derived using the methods of statistical mechanics. Wilson (1969h, 231) argues that the statistical mechanics approach provides a method of deriving the gravity model which has two advantages: firstly the desired results can be derived in a natural way and secondly the statistical mechanics analogy suggests fruitful ways of extending gravity models to handle complex situations.

The main features of the statistical mechanics analogy are now outlined (for elementary reviews see Gould, 1972; Cesario, 1975a). The two basic principles from statistical mechanics which are of use are (Georgescu-Roegen, 1971, 142):

> A. The disorder of a microstate is ordinally measured by that of the *associated macrostate* (i.e. the macrostate to which the microstate can be aggregated).

B. The disorder of a macrostate is proportional to the number of the microstates which correspond to that macrostate.

In these principles a "macrostate" is a description of a system which merely gives the numbers of each kind in each state. A "microstate" gives the names of the particles which are of each kind in each state. For example let there be 4 individuals W, X, Y, Z. A typical macrostate description of these individuals would say that two are in category A and two are in category B. A microstate description would name the pairs.

These concepts will now be used to give an intuitive exposition of the notion of entropy and a brief justification for the formalism due to Jaynes (1957) which states that in choosing a distribution one should do so in order to maximise entropy. There are very real difficulties associated with the statistical mechanics approach (see Georgeseu-Roegen, 1971, Chapter 6). These difficulties can be circumvented by using an alternative approach based on information theory.

2.3 The statistical approach to entropy

In the following discussion two cases will be considered: (1) equal categories, (2) unequal categories. In general there will be N individuals and M categories or "boxes".

(1) The equal categories case

Suppose that there are 4 individuals W, X, Y and Z. There are 5 possible ways of placing these individuals into 2 categories:

$\boxed{4 \mid 0} \quad \boxed{3 \mid 1} \quad \boxed{2 \mid 2} \quad \boxed{1 \mid 3} \quad \boxed{0 \mid 4}$

In terms of the definitions given above these are macrostates. The problem is to pick that macrostate which has the highest number of corresponding microstates. The formula for the number of microstates associated with each macrostate is in general

$$R = \frac{N!}{N_1!\, N_2! \ldots N_M!} = \frac{N!}{\prod_{i=1}^{M} N_i!}$$

The principles stated above indicate that this is an ordinal measure of the disorder of any of the microstates as well as of the macrostate (Georgescu-Roegen, 1971, 143), and that the disorder of each of the macrostates is proportional to the number of their corresponding microstates. The maximum entropy formalism states that the distribution which maximises R should be chosen for the positive reason (Jaynes, 1957, 623) that it is the one which is consistent with known information and yet maximally noncommittal with respect to missing information. The question arises as to what exactly is being maximised. It is convenient to take logarithms and to divide by N (this will not affect the maximum):

$$H_1 = \frac{1}{N}\log R = \frac{1}{N}[\log N! - \sum_{i=1}^{M} \log N_i!]$$

where N_i is the number of entries in the category i (see Walsh and Webber, 1977, 402). The measure represents the mean uncertainty about the occurrences of any event in a completely sampled population relative to some known facts about that population and has been deduced from a sampling without replacement process by Walsh (1976). If all N_i are large, use can be made of Stirling's formula, $\log x! \simeq x\log x - x$ to obtain the *approximation* $\tilde{H}_1$

$$\tilde{H}_1 = \frac{1}{N}[N\log N - N - \sum_{i=1}^{M} N_i \log N_i + \sum_i N_i]$$

$$= \log N - \frac{1}{N}\sum_{i=1}^{M} N_i \log N_i$$

$$= -\sum_{i=1}^{M} \frac{N_i}{N} \log\frac{N_i}{N} \equiv \hat{H}_2 \qquad (1)$$

The form of this equation is identical to that used to *estimate* Shannon's (1948) entropy measure,

$$H_2 = -\sum_{i=1}^{M} p_i \log p_i \quad , \qquad (2)$$

where N_i/N is used to *estimate* p_i $(i = 1, \ldots, M)$. Shannon's (1948) measure H_2 is one which satisfies certain reasonable requirements of a measure of uncertainty (see Webber, 1977a; for an outline of these requirements).

The formalism due to Jaynes (1957) states that this quantity should be maximised. It has been found that the principle of maximum entropy is a useful tool for building models of urban phenomena. That there is a need for such a method is unquestionable since as Jaynes puts it "the amount of information available in practical situations is so minute that it alone would never suffice for making reliable predictions" (Jaynes, 1957, 625). Therefore Jaynes states that in choosing a probability distribution one should maximise entropy subject to whatever is known. This formalism produces a probability distribution that is both 'maximally noncommittal' with respect to missing information and which is 'spread out' as uniformly as possible without contradicting the given information (Jaynes, 1968, 231). This formalism has been accepted as a powerful tool (for examples see Tribus, 1969; Dawson and Wragg, 1973) and has found many uses in human geography; Webber (1975, 101) lists several possible geographical applications.

(2) The unequal categories case

It is also possible to work with categories which are not a priori equally likely. As an example begin by placing 4 objects into 2 unequal categories: A (which has two cells) and B (which has one cell). Thus:

A	B	A	B	A	B	A	B	A	B
4	0	3	1	2	2	1	3	0	4

The microstates giving rise to the microstates are:

WXYZ	-	WXY	Z	WX	YZ	W	XYZ	-	WXYZ
		WXZ	Y	WY	XZ	X	WYZ		
		WYZ	X	WZ	ZY	Y	WXZ		
		XYZ	W	XY	WZ	Z	WXY		
				XZ	WY				
				YZ	WX				

As the boxes are drawn there is a difference in size, and they can be further broken down into three equal cells. Then for each of the above microstates there are 2^{N_A} ways of distributing the N_A individuals in category A over its *two* cells, and obviously just one way of placing the individuals in category B into its *single* cell. The first few fine grain microstates are listed here for the above cases:

WXYZ	-	-
WXY	Z	-
WXZ	Y	-
WYZ	X	-
..	..	..

WXY	-	Z
WX	Y	Z
WY	X	Z
XY	W	Z
..	..	..

WX	-	YZ
W	X	YZ
X	W	YZ
-	WX	YZ
..	..	..

W	-	XYZ
-	W	XYZ
X	-	WYZ
-	X	WYZ
..	..	..

-	-	WXYZ

The total number of microstates that correspond to a given macrostate is

$$V = \frac{N!}{N_1!\, N_2! \,\ldots\, N_M!} S_1^{N_1} S_2^{N_2} \ldots S_M^{N_M}$$

where S_i is the number of cells, and N_i is the number of entries, in the ith category. The numbers of microstates which arise in the equal categories case and the number of fine grain microstates which arise in the unequal categories case are summarized in table 2.

Macrostates	4, 0	3, 1	2, 2	1, 3	0, 4
(1) # microstates	1	4	6	4	1
(2) # fine grain microstates	16	32	24	8	1

Table 2: Microstates (equal categories) and fine grain microstates (unequal categories).

Examining this table it is clear that there is a difference in the macrostate which is chosen in each case. Therefore it is important that any information on category size be incorporated into an entropy maximising model. (This point is also raised by Batty (1976b, 3).)

To show what is being maximised in the expression for V, take logs, divide by N and subtract log M.

$$\frac{1}{N}\log V - \log M = \frac{1}{N}[\log N! + \sum_{i=1}^{M} N_i \log S_i - \sum_{i=1}^{M} N_i!] - \log M \equiv H_3.$$

Using Stirling's approximation

$$\tilde{H}_3 = \frac{1}{N}[N\log N + \sum_{i=1}^{M} N_i \log S_i - \sum_{i=1}^{M} N_i \log N_i] - \log M ,$$

$$= +\log\frac{N}{M} - \frac{1}{N}\sum_{i=1}^{M} N_i \log \frac{N_i}{S_i}$$

Rearranging and taking the negative gives

$$-\tilde{H}_3 = \log M + \sum_{i=1}^{M} \frac{N_i}{N} \log \frac{N_i/N}{S_i} \tag{3}$$

Equation (3) has been called inverse spatial entropy by Batty (1976b, 4) and negentropy by Walsh and Webber (1977, 414). Finally combining gives

$$-\tilde{H}_3 = \sum_{i=1}^{M} N_i/N \log \frac{N_i/N}{S_i/M} \equiv \hat{KIG}$$

The form of this equation is identical to that used to calculate Kullback Information Gain (KIG),

$$KIG = \sum_{i=1}^{M} p_i \log \frac{p_i}{q_i} \tag{4}$$

where N_i/N is used to estimate p_i $(i = 1, \ldots, M)$ and S_i/M is used to estimate q_i ($i = 1, \ldots, M$), where q_i is a prior probability.

If the priors are unequal Hobson (1969) has shown that under certain reasonable conditions KIG is a unique (up to a multiplicative constant) measure of the information gained when a distribution p_i ($i = 1, \ldots, M$) replaces a prior q_i ($i = 1, \ldots, M$). Renyi (1970) presents a heuristic argument which illustrates the nature of the change which occurs when a move is made from a prior to a posterior distribution. Implications of this measure are discussed by Hobson (1969) and Hobson and Cheng (1973).

KIG can be accepted as a measure of information gain for cases where the categories (henceforth used interchangeably with 'priors') are unequal. An extension of Jaynes' formalism suggested by Hobson and Cheng (1973) and in a geographical context by Snickars and Weibull (1977) is the so called minimum information principle (see also Evans, 1961). This states that the distribution which minimises information gain (i.e. maximises uncertainty) should be chosen. Formally this means that from the statistical mechanics point of view $\tilde{H}_3$ should be maximised. From an information theory point of view Kullback's uncertainty should be maximised subject to constraints. Kullback uncertainty is defined as the difference between the largest information gain possible and the information gain actually achieved i.e.

$$KU = \sum_{i=1}^{M} p_i^{max} \log \frac{p_i^{max}}{q_i} - \sum_{i=1}^{M} p_i \log \frac{p_i}{q_i}$$

Where p_i^{max} is the distribution representing maximum information consistent with the constraints (i.e. a zero-one distribution). Since the first term on the right does not depend on p_i it produces an irrelevant constant. The modified formalism is therefore to maximise -KIG (or $\tilde{H}_3$). If the prior distribution is uniform then the term being maximised is again Shannon's entropy (H_2) (Hobson and Cheng, 1973, 305; Tribus and Rossi, 1973, 335; Snickars and Weibull, 1977, 146).

There has been some discussion of the modified formalism in a geographical context. March and Batty (1975b) demonstrate that there is a class of minimally prejudiced models of which the Kullback and Shannon based formalisms are special cases. Applications of the modified formalism include March and Batty (1975a), Batty and March (1975) and Snickars and Weibull (1977).

Cesario (1975a) gives a useful numerical example which will serve to illustrate the entropy maximising formalism and the modified formalism based on prior information. Consider the following spatial interaction matrix:

		Destination		
		1	2	T_{i*}
Origins	1	T_{11}	T_{12}	3
	2	T_{21}	T_{22}	3
	T_{*j}	4	2	T

The problem is to find the distribution T_{ij} which is maximally noncommittal with respect to missing information and at the same time consistent with the constraints. The possible microstates which are consistent with the information are:

$$T = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \qquad T = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \qquad T = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$$

(a) (b) (c)

The total number of ways we can select a particular distribution $\{T_{ij}\}$ from T is

$$R = \frac{T!}{T_{11}!\ T_{12}!\ T_{21}!\ T_{22}!}$$

If R is evaluated for each of the above distributions, it is found that (a) = 60, (b) = 180, (c) = 60. The most likely distribution is the one with the most microstates associated with it - hence (b) is chosen.

To extend this example to the unequal categories case suppose that the number of cells in each category is $S_{11} = 2$, $S_{12} = 1$, $S_{21} = 1$, $S_{22} = 2$. Then the number of fine grain microstates consistent with each macrostate is given by

$$V = \frac{T!}{\prod_{i=1}^{n} \prod_{j=1}^{m} T_{ij}!} \prod_{i=1}^{n} \prod_{j=1}^{m} S_{ij}^{T_{ij}}$$

The results for the above values of S_{ij} are (a) 1920, (b) 1440 and (c) 120. Thus (a) is now the macrostate with the greatest number of associated microstates. Snickars and Weibull (1977) suggest that prior information in the trip distribution context could consist of an interaction matrix from another period. Other prior information that could be incorporated includes the number of routes between different zones. In the above case this would mean that there are two routes from zones

1 to 1 and 2 to 2 and only one route from zones 1 to 2 and 2 to 1.

In conclusion the two important measures (2) and (4) from this section are written out in the trip distribution context.

$$H_2 = -\sum_{i=1}^{n} \sum_{j=1}^{m} T_{ij} \log T_{ij} \tag{5}$$

$$-KIG = -\sum_{i=1}^{n} \sum_{j=1}^{m} T_{ij} \log \frac{T_{ij}}{S_{ij}} \tag{6}$$

2.4 A family of spatial interaction models

The various types of models which can be derived using the entropy maximising formalism are distinguished by the constraints which are embodied in them. Four cases can be noted (O = origin, D = destination):

(1) Neither the set of totals O_i nor the set of totals D_j is known,

(2) The set of totals O_i is known,

(3) The set of totals D_j is known,

(4) Both sets of totals O_i and D_j are known.

In addition a constraint on the average cost of trips is added. In the following models zero, one or two of the following constraints may operate.

$$\sum_j T_{ij} = O_i \tag{7}$$

$$\sum_i T_{ij} = D_j \tag{8}$$

and in addition

$$\sum_i \sum_j T_{ij}\, c_{ij} = \bar{c} \tag{9}$$

where (c_{ij}) is a matrix of interzonal costs and $\bar{c}$ is average cost. Wilson (1974) has produced a family of spatial interaction models using these constraints. His results are now generalized using the Kullback based formalism.

(1) Unconstrained case

The result of the maximisation of (5) subject to (9) produces

$$T_{ij} = \exp(-\beta\, c_{ij}) \tag{10}$$

Where β is found so as to satisfy (9). Cordey Hayes and Wilson (1971) introduce the origin and destination attraction factors through the existence of savings in costs associated with certain places. An alternative method for introducing these factors in the unconstrained case is to maximise a function, based on equation (6), of the form:

$$-\sum_i \sum_j T_{ij} \log \frac{T_{ij}}{A_i B_j} \tag{11}$$

where A_i and B_j incorporate some prior knowledge of the attraction of origins and destinations. The result of this maximisation is an expression

$$T_{ij} = A_i B_j \exp(-\beta\, c_{ij}) \tag{12}$$

which is the form of the traditional gravity model.

(2) Production constrained case

When independent estimates of the numbers of flows originating in each zone are known this information must be built into the estimates of T_{ij}. This is achieved by maximising (5) subject to (7) and (9) then the resulting expression is

$$T_{ij} = a_i O_i \exp(-\beta c_{ij}) \tag{13}$$

where $a_i = 1/\sum_{j=1}^{m} \exp(-\beta c_{ij})$ (14)

(this is the well known partition function, see Tribus, 1969, 124). Again it is possible to introduce prior information on the attraction of various destination zones. The problem is to maximise

$$-\sum_i \sum_j T_{ij} \log \frac{T_{ij}}{A_i B_j} \tag{15}$$

where A_i is arbitrary and B_j is non uniform subject to (7) and (9) and the result is

$$T_{ij} = a_i B_j O_i \exp(-\beta c_{ij}) \tag{16}$$

where now

$$a_i = 1/\sum_{j=1}^{m} B_j \exp(-\beta c_{ij}) \tag{17}$$

(3) Attraction constrained

This case is the mirror image of case [2]. D_j is now known independently, and the problem therefore is to maximise (5) subject to (8) and (9). The result of this is

$$T_{ij} = b_j D_j \exp(-\beta\, c_{ij}) \tag{18}$$

where $$b_j = 1/ \sum_i \exp(-\beta\, c_{ij}) \tag{19}$$

Wilson adds in factors to represent the various trip origin zones (based on residential land use) by subtracting a constant for each i from the cost of each flow. This effect can also be achieved by maximising

$$-\sum_i \sum_j T_{ij} \log \frac{T_{ij}}{A_i B_j} \tag{20}$$

where A_i is non uniform and B_j is arbitrary. Subject to (7) and (9) and the result is

$$T_{ij} = A_i b_j D_j \exp(-\beta\, c_{ij}) \tag{21}$$

and now

$$b_j = 1/ \sum_{i=1}^{n} A_i \exp(-\beta\, c_{ij}) \tag{22}$$

(4) Production and attraction constraints

This problem involves maximising (5) subject to (7), (8) and (9). The result of the constrained maximisation is that

$$T_{ij} = a_i b_j O_i D_j \exp(-\beta\ c_{ij}) \tag{23}$$

Where a_i and b_j are interrelated balancing factors which ensure that the constraints on origins and destinations are met. An obvious extension of this result is to incorporate prior information on the distribution T_{ij}. That is maximise (6) where S_{ij} is some existing flow data subject to (7), (8) and (9). The result is an expression of the form

$$T_{ij} = S_{ij} a_i b_j O_i D_j \exp(-\beta\ c_{ij}) \tag{24}$$

and

$$a_i = 1/\sum_{j=1}^{m} S_{ij} b_j D_j \exp(-\beta\ c_{ij}) \tag{25}$$

$$b_j = 1/\sum_{i=1}^{n} S_{ij} a_i D_j \exp(-\beta\ c_{ij}) \tag{26}$$

As a special case let $S_{ij} = A_i B_j$ for all i, j. Then the problem is to maximise

$$-\sum_i \sum_j T_{ij} \log \frac{T_{ij}}{A_i B_j}$$

Subject to (7), (8) and (9). The result is an expression of the form

$$T_{ij} = a_i b_j O_i D_j \exp(-\beta\ c_{ij})$$

and

$$a_i = 1/\sum_j B_j b_j D_j \exp(-\beta\ c_{ij})$$

$$b_j = 1/\sum_i A_i a_i O_i \exp(-\beta\ c_{ij})$$

which is a direct generalization of equation (23) above. (See Snickars and Weibull (1977) for application, and Karlqvist and Marksjö (1971) for discussion of solution methods.) Before describing the results from applied modelling efforts a general discussion of the applicability of different types of spatial interaction model is given.

Firstly production constrained models have enabled the inconsistencies in applications of the gravity model to retail sales to be cleared up. For instance the production constrained model in the later formulation given above (equation 16) is the typical format of early retail gravity models - there are terms for origin zones, for destination zones and for the cost of movement between these zones. In the pre entropy maximising formulation made by Reilly (1931) this format was used. As Wilson (1974, 42) points out the Reilly model was inconsistent in that it was not possible to find a factor K that would ensure that all the origin constraints were met. This inconsistency is overcome in the entropy maximising formulation since the balancing factors are calculated to ensure that the necessary constraints are fulfilled.

It has also been shown that the entropy maximising formalism can be applied to other cases where some prior information on the relative attractiveness (or otherwise) of the zones is available. The major application of attraction constrained models has been to location of residences. The significance of an attraction constrained model is that the modeller can use information on (say) employment as an indicator of work trip attraction together with trip length data to estimate a T_{ij} matrix. Then, summing over the columns of this matrix will give an estimate of trip productions. These can then be interpreted as residential locations. As was shown above the attraction constrained model is of the form

$$T_{ij} = B_j D_j \, f(c_{ij})$$

where B_j is calculated to ensure that the trips ending in zone j match the independent information. Since there is no restriction on $\sum_j T_{ij}$ the model can be said to predict the amount of activity located in i. Thus the simplest modification of the Lowry model (Wilson, 1969g) which allocates workers to zones of residence is

$$P_i = \sum_j B_j D_j \, f(c_{ij})$$

This removes the inconsistency mentioned above in section 1.3. (See Broadbent, 1970, 469-70.) Further developments in residential location modelling are reported in Wilson (1975).

The doubly constrained form of the model has found a wide applicability in trip distribution problems. In these cases it is usually known that certain numbers of trips originate and end each zone. Entropy maximising estimates of the detailed interzonal flows are easily obtained. Extensions have included the combination of the modal split and trip distribution stages of the transportation planning process, by simply adding a subscript for each mode and by using information on average costs for each mode (see Wilson, 1969e, 1970a, 1970c, 1973a, 1974, 1975). Other useful references include Cesario (1973a, 1973c, 1974b), Hathaway (1975).

2.5 The experience gained from applied spatial interaction modelling

A number of Lowry-Wilson type applications have been successfully carried out in regional planning exercises in Britain. It is not proposed to go into detail on these various applications as they have been adequately documented elsewhere (Cripps and Foot, 1969a; 1969b; 1970; Batty, 1969; 1970; Cordey-Hayes, *et al.*, 1970; Echenique, *et al.*, 1969). However a number of points have arisen directly from the experience gained in these applications, and some of these will be noted.

Broadbent (1970), Batty *et al.* (1974) and Barras *et al.* (1971) in particular have paid attention to be operational problems associated with these models. These can conveniently be discussed under a number of headings.

(1) Solution of entropy maximising models

The solution of spatial interaction models which are derived using entropy maximising methods involves finding values of the parameters which ensure that the constraints are met. This problem can sometimes be solved analytically but in general numerical techniques are used (see next subsection).

It is also possible to consider equations such as (12) above as being *given* outside of any entropy maximising framework. In this case the values of the parameters which solve maximum likelihood equations have to be found. It is known (Batty and Mackie, 1972) that the maximum likelihood equations to be solved for the parameters are the constraint equations in the entropy maximising derivation. For instance it can be shown (Hall, 1975, 31-34) that the likelihood function for equation (8) is maximised when the parameter β is the solution of

$$\sum_i \sum_j T_{ij}\, c_{ij} = \sum_i \sum_j N_{ij}\, c_{ij}$$

where N_{ij} is the observed flow from i to j. The right hand side is of course the observed value of $\bar{c}$ used in equation (5) above.

Wilson (1974, 318) points out the connection between the maximum likelihood method and the entropy maximising method:

> "For each parameter of the model the [maximum likelihood] procedure produces an equation to be solved for that parameter. This equation turns out to be the constraint equation which would be used to generate the same model as an entropy-maximising model, and the equivalent parameter is then the Lagrangean multiplier associated with the constraint. Either way, the parameter is obtained by solving the appropriate equation. *It is really a matter of taste and convenience as to whether maximum-likelihood methods or entropy-maximising methods are used to produce the equation* which is to be solved for each parameter."
>
> (Bracketed words and italics mine)

While this suggests that there is nothing to choose between the two approaches there is a strong reason for preferring the entropy maximising derivation of the equations. This is that the entropy measure represents a measure of mean uncertainty and in conjunction with Jaynes formalism it produces a minimally prejudiced distribution. Webber (1977a, 1977b) suggests that this is an appropriate method for use in geographical research. In many applications (for example Batty (1976a) and Openshaw (1976)) equations such as (23) above are used without any recourse to entropy maximising techniques and in these cases maximum likelihood methods can be used to estimate parameters. There may however be "better" estimates (Openshaw (1976) and Stetzer (1976)).

A numerical example illustrating the role of parameters in a trip distribution context is given in Fisk and Brown (1975).

(2) Solution techniques

As a result of the discussion of calibration problems a number of numerical techniques for estimating parameters have been developed. These techniques are reviewed in Batty and Mackie (1972), Batty (1976a) and in an applied context by Batty (1970) and Batty *et al.* (1971). Alternative approaches include Cesario (1973b, 1975b). Recent work using geometric and other forms of programming are discussed in Charnes *et al.* (1972), Dinkel *et al.* (1977), Charnes *et al.* (1977), March and Batty (1975a), Evans (1973; 1976) and Nijkamp (1975).

The major conclusions reached in applied work is that techniques based on second order methods (e.g. Newton-Raphson search) have the advantages of being fast, reliable and versatile. Versatility is important in models such as those to be discussed in Chapter 4 because many changes to the models configuration were made, involving the addition to extra parameters.

As part of the discussion of calibration methods there have been analyses of existence and uniqueness of solutions (Evans, 1973). In practice the problem of non-uniqueness will not arise since for any average cost data greater than the minimum (and less than the maximum) solution to the corresponding transportation problem there will be a unique solution for the parameters (see Evans, 1973). In all the applications referred to above the problem of non-existence (i.e. a failure to converge) was not encountered. However the problem may be the amount of time required to find the solution. Batty (1976a) in particular has been concerned with problems of computation time and has suggested methods for achieving faster solution times.

(3) Dimensionality and zoning schemes (Broadbent, 1970)

Many early spatial interaction models ignored the difficulties associated with varying zone size. The original Lowry application used a 450 cell grid of even sized zones and so avoided the problem.

The crux of the problem can be put as follows in the allocation equation:

$$P_i = \sum_j B_j E_j \exp(-\beta\, c_{ij})$$

population P_i is an extensive variable while the accessibility to employment is intensive - the two should not be mixed (Seidman, 1967; referred to in Broadbent, 1970). In view of the need to have models which are capable of handling variations in zone size, Batty (1976b) has suggested that a prior probability vector giving the weights of each zone should be used. This represents an application of the suggested modification to Jaynes formalism. In the model to be presented below residential acreage will be used to construct an attraction prior.

Broadbent (1970) and Masser *et al.* (1975) have discussed the problems associated with designing an adequate zoning scheme. The model builder is faced with the task of using a scheme that is detailed enough to show interzonal movement - that is the zone must be small so that a large proportion of the trips beginning in that zone end outside it. It has been suggested that 85% of the trips starting in a zone should end outside it. As against this there is the problem of collecting data from various sources - which are unlikely to use the same zoning scheme. In practice one is usually forced to accept some fairly large zones as designed by the City Planning department, in order to be sure of an adequate data base. Openshaw (1976) has discussed the effects of zone size on model accuracy.

2.6 Summary

This chapter has shown that the entropy maximising formalism is a useful methodology for the development of urban land use models. In section 2.3 measures of average uncertainty and information gain were described. This discussion provides a method which allows minimally prejudiced consistent models to be built. In section 2.4 a family of such models was derived using Kullback based measures. It was shown that the format of these models is the same as in Wilson (1974) but with the addition of terms to represent origin and destination characteristics. Section 2.5 referred to applications of spatial interaction models with a view to gathering hints for empirical tests. From a review of research on actual solution techniques it was found that Newton-Raphson methods are the most suitable for the purposes of this paper as they provide fast and reliable parameter estimates. Research on zone schemes in spatial interaction modelling recommends that some account of variations in size be incorporated. The model to be developed below (which is related to the family of models in 2.4) will take these results into account.

In conclusion, this chapter has shown that there exists a methodology, namely the entropy-maximising formalism for developing consistent models of cities. This methodology has been applied by Wilson (1974), Batty (1976b) and others to various aspects of the city - residential location, retail trade and transportation planning. Clearly then one of the major strengths of the entropy maximising formalism is its versatility in deriving consistent urban models. The following chapter continues the argument by developing an entropy maximising model which has theoretical foundations.

CHAPTER 3

A MODEL OF URBAN LAND USE

3.1 Introduction

The preceding chapters have reviewed the early developments of land use modelling in North America and the more recent emergence of entropy maximising models. In particular Wilson's (1974) work in spatial interaction modelling was reviewed and his family of spatial interaction models was generalised. From this general group of models a particular example will now be developed and tested for Hamilton, Ontario.

The model works in the following manner. For any given spatial distribution of retail facilities (and fixed work places), the model assigns individuals to place of residence, shopping trip pattern and place of work. The assignment of individuals is made subject to constraints on mean journey to shop length and mean journey to work length for given distributions of service centres and employment. The solution to the model is in the form of equation (21) in Chapter 2. (Specific solutions will be given below.) Associated with the solution to the model is an entropy score S* (the asterisk denotes values of S and p_{ijk} associated with a *solution* for an arbitrary pattern of centres)

$$S^* = -\sum_i \sum_j \sum_k p^*_{ijk} \log p^*_{ijk}$$

where p^*_{ijk} is the joint distribution which solves the constrained maximisation problem. There is a different S* statistic associated with each distribution of service centres; the objective is to find the distribution for which S* is a maximum. This distribution will be chosen as the endogenously predicted pattern of service centres.

The model is being used in two roles. Firstly, as a spatial interaction model it assigns individuals to a set of categories in a way which is maximally noncommittal with respect to missing information and which at the same time does not contradict given data. Secondly, the model is being used to determine the distribution of service centres which maximises the S* score.

It has been shown (Webber, 1977b) for a simple linear city under restrictive assumptions, that a centralized distribution of service centres maximises the entropy score S* over the assignment of individuals. From this theoretical work it is to be expected that the distribution of service centres that 'wins' will include some central city zones. In order to test this research hypothesis we compare the distribution of service centres predicted by the model to the existing pattern of service centres for Hamilton. Further tests of the 'fit' of the model to observed data can be performed, for example by comparing observed and predicted trip distribution matrices. It must be stressed however that the criterion used to pick the solution is that S* should be a maximum - there is no guarantee that this result will correspond to the highest level of correlation.

The aim of the modelling exercise is to show that the entropy maximising model is a realistic representation of the city. By examining the fit of the model, evaluated at the entropy maximising pattern of facilities, the realism can be assessed. Thus this part of the research paper determines whether or not an entropy maximising model is a realistic representation of the city. If it can be shown that the model "fits" then the claim that entropy models are *realistic* will have some support.

This chapter describes the categories to which the model assigns individuals (3.2), derives a number of endogenous distributions (3.3) and describes the data used to test the model (3.4).

3.2 Categories

The model assigns individuals to place of residence, shopping trip pattern and place of work. Details of these categories are now given.

(1) Place of residence, $(i = 1, \ldots, 14)$

The 14 residential categories correspond to the zones shown in Figure 2. These zones are based on the Planning Districts used by the City of Hamilton Planning Department.

(2) Shopping trip pattern, $(j = 1, 2, \ldots, J)$

These patterns are generated by allowing individuals to make $m = 0, 1, 2$ trips to a given set of service centres. For example if there are service centres in zones 5, 7 and 10 the possible shopping trip patterns are:

	Pattern Number						
	1	2	3	4	5	6	7
Visit Centre	NONE	5	7	10	(5,7)	(5,10)	(7,10)

In general where there are N service centres there will be J shopping trip patterns, where

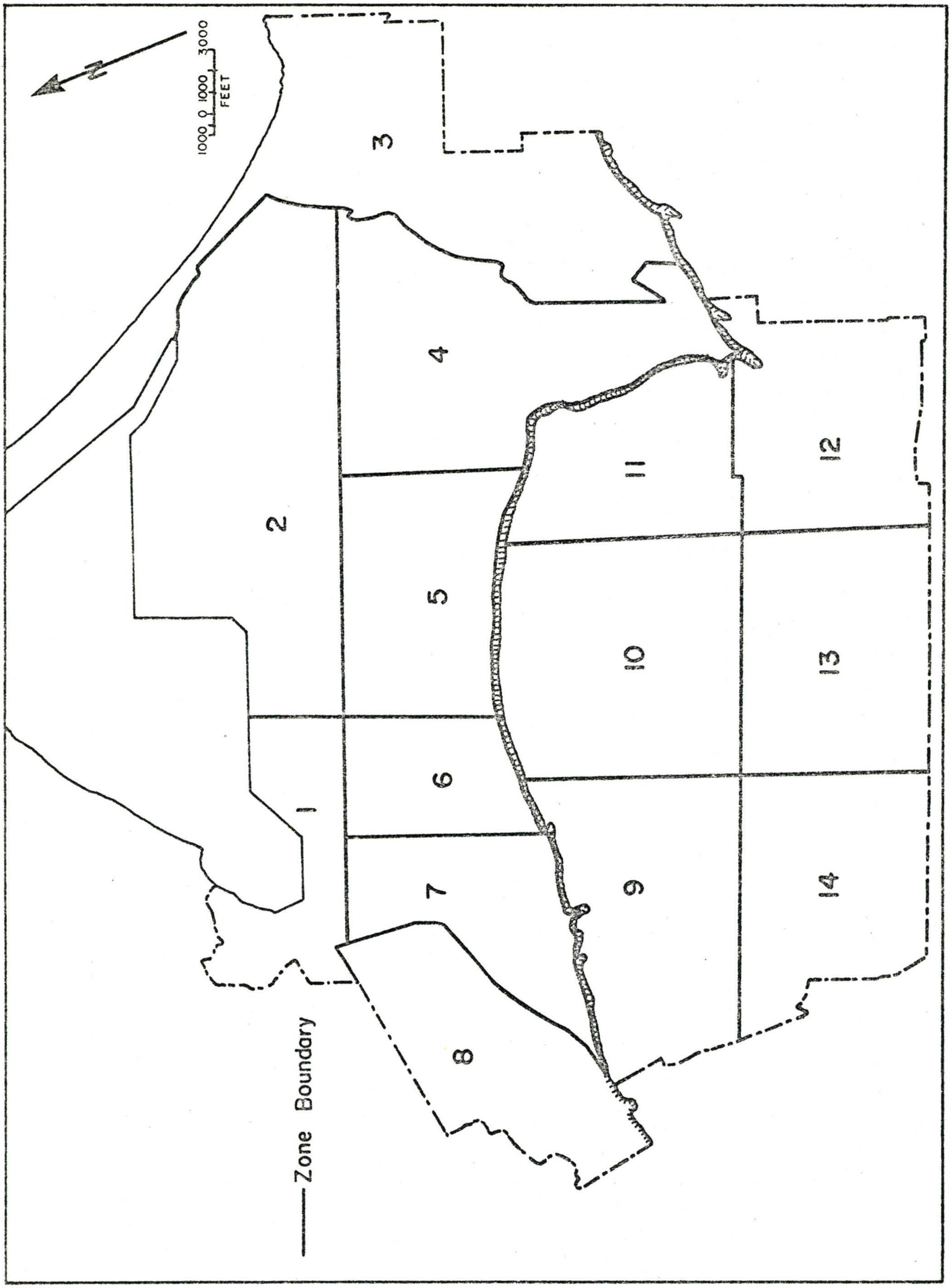

Figure 2 The Zoning Scheme, City of Hamilton

$$J = \binom{N}{0}+\binom{N}{1}+\binom{N}{2} .$$

This method of classification provides a realistic link between the actual distribution of centres and shopping behaviour through the calculation of costs associated with each pattern. (Details of these cost calculations will be given below.)

(3) Work place, ($k = 1, \ldots, 15$)

The 15 work place categories correspond to the zones shown in Figure 2 ($k = 1, \ldots, 14$) and one other category called "no work trip". This last category is included to accomodate those households observed to make no work trip because of unemployment or retirement.

3.3 Details of endogenous distribution

The models to be described in Chapter 4 produce assignments p^*_{ijk}. (The asterisk will be dropped from now on.) These assignments can be manipulated to replicate certain aspects of the urban geography of Hamilton, i.e. residential location, work trip interchanges, shopping patterns and retail sales. Derivations of these distributions follow.

(1) Residential location

The distribution of residents is

$$p_i = \sum_j \sum_k p_{ijk} \qquad (i = 1, \ldots, 14).$$

The joint distribution over zones of residence and trip patterns is

$$p_{ij} = \sum_k p_{ijk} \ .$$

Define a new matrix (f_{im}), m = 0, 1, 2 based on (p_{ij}) where

$$f_{i0} = p_{i1} \qquad (i = 1, \ldots, 14),$$

$$f_{i1} = \sum_{j=2}^{N+1} p_{ij} \qquad (i = 1, \ldots, 14),$$

$$f_{i2} = \sum_{j=N+2}^{J} p_{ij} \qquad (i = 1, \ldots, 14).$$

(f_{im}) is a joint distribution over residential location and shopping trip behaviour where households make m = 0, 1, 2 trips. Note that

$$\sum_{m=0}^{2} f_{im} = \sum_{j=1}^{J} p_{ij} = p_i \qquad (i = 1, \ldots, 14).$$

(2) Employment

Similar manipulations can be performed for the work place distributions. However p_{ijk} is constrained to reproduce the work trip destinations d_k. The model does not therefore "predict" this distribution.

$$d_k \equiv \sum_i \sum_j p_{ijk} \qquad (k = 1, \ldots, 15).$$

Note that

$$d_{ik} = \sum_j p_{ijk} \qquad (i = 1, \ldots, 14;\ k = 1, \ldots, 15),$$

is the familiar trip distribution matrix (T_{ij} in Wilson, 1974). The joint distribution over trip patterns and work places is:

$$d_{jk} = \sum_i p_{ijk} \ .$$

Define a new matrix (e_{km}) based on (d_{jk}) where

$$e_{k0} = d_{1k} \qquad (k = 1, \ldots, 15),$$

$$e_{k1} = \sum_{j=2}^{N+1} d_{jk} \qquad (k = 1, \ldots, 15),$$

$$e_{k2} = \sum_{j=N+2}^{J} d_{jk} \qquad (k = 1, \ldots, 15).$$

This new distribution (e_{km}) is the joint probability of working in zone k and making m shopping trips. Note that

$$\sum_{m=0}^{2} e_{km} = \sum_{j=1}^{J} d_{jk} = d_k \ .$$

(3) Shopping behaviour and retail trade

The distribution (p_{ijk}) can be manipulated to give information on shopping behaviour and retail trade. The proportion of individuals making m shopping trips is

$$f_m = \sum_{i=1}^{14} f_{im} \equiv z_m = \sum_{k=1}^{15} e_{km} \quad (m = 0, 1, 2)$$

The flows between residential zones and shopping centres are calculated as

$$r_{ig} = \frac{\sum\sum_{j\varepsilon g} p_{ij}}{\sum_i \sum_j p_{ij} x_j} = \frac{\sum_{j\varepsilon g} p_{ij}}{\bar{x}}$$

where the notation ($j\varepsilon g$) indicates those shopping trip patterns j which are associated with a particular centre g = 1, ..., N. The denominator ensures that the values of r_{ig} sum to unity, since $\bar{x}$ aggregates all flows made to the centres via single trip patterns and twice the flows made to the centres via two trip patterns; which is precisely the sum of the elements in the numerator.

The proportion of retail sales in each centre is

$$r_g = \sum_{i=1}^{14} r_{ig} = \frac{1}{\bar{x}} \sum_{i=1}^{14} \sum_{j\varepsilon g} p_{ij}$$

The flows between residential zones and shopping centre r_{ig} as well as the share of each centre in retail sales will be mapped for the various models in Chapter 4.

The various distributions which can be calculated have now been listed. This completes the description of the location model. This aspect of the model produces the pattern of residential location, trip interchanges, retail sales etc. *for a given set of centres.* Recall however that the model is being used to find the distribution of service centres which maximises S*. This means that in the results reported in Chapter 4, the primary concern will be with the distribution of service centres and the associated predictions which maximise S*.

3.4 The data

(1) Household characteristics

The data for this study is a travel characteristics study carried out by the Hamilton Wentworth Planning and Development Department (hereafter HWPDD) in 1974.

The survey information collected by telephone, describes all the trips made by each member of each sample household over a given 24 hour period. The survey took place in the period from September to December 1974 and includes an equal sampling of each of the five week days. The sample contained 1634 households, which represents an areally stratified one percent random sample of all households in the study area, drawn from data in local assessment files. A total of 1402 successful interviews were completed. For each household, information was collected on its location, composition, licensed drivers, occupation and industry of members, place of work of household members, number of automobiles and an indicator of residential density; for each individual aged more than five years in the household, data were collected on all the trips made by that person - trip purpose, location of origin and destination, arrival and departure time, travel mode, parking charges and number of persons making that trip. Apart from the absence of data on income and expenditures, the major problem associated with these data are in the fact that only one day's trips are sampled. The travel characteristics report of HWPDD (1975) describes the survey and the travel characteristics of the sample in more detail.

The data contain 10,209 records to describe the trips (or lack of trips) made by the members of the 1402 households. These records have been severely edited to simplify analysis. The first editing process removed all information except that pertaining to home-based work trips, home-based shopping trips and households whose members make no trips. The aggregate trip characteristics of the remaining 2195 records are shown in Appendix B Table B1: in these data one work trip comprises a journey to work (and return) by one individual whereas a shopping trip consists of a single journey made from home to shopping facilities (i.e. several members of a household may travel on a single shop trip). This editing process simplifies the description of trips by reducing the number of possible kinds of trips in the sample. The second stage in editing the data simplified the description of households by ignoring some of the trips made by households. The maximum number of work trips made by each household is constrained to one; if the members of a given household make more than one work trip, then the work trip made by the highest "ranking" family member who works is chosen ("rank" being defined in the survey as head of household, spouse, sons and daughters by age, other relatives by age, and unrelated persons). Although this limit on work trip data was imposed to simplify the analysis, it in effect reflects the hypothesis that the only influence of job location on household location occurs through the job of the highest "ranking" family member who works. Using a similar procedure, the maximum number of distinct shopping trips per household is limited to two. The effect of this editing process is to ignore data on 489 work trips and 21 shopping trips. Table B2 describes the characteristics of the sample households after this edit. The final editing process removes households which reside or work outside the city of Hamilton, in order to reduce the number of zones needed to encompass the study area. The remaining households reside in the city of

Hamilton, do not work in employment outside the home or work within the city, and may make shopping trips to places within or outside the city of Hamilton. Data for the trips made by 769 households remain, the trips of which are described by purpose (work or shopping, but not both), origin (14 zones) and destination (14 zones).

(2) Travel time data

The survey contained departure and arrival times for the trips reported. This information did not however provide an adequate basis for an interzonal travel time matrix since for many routes no data existed. Trial runs with a matrix made up of interzonal travel times from the data and estimates of missing data proved to be unsatisfactory. An alternative source of data was sought.

Travel times computed by the city of Hamilton (1973) are available for a network. Many of the centroids on this network closely approximate the centres of the Planning Districts. A set of 14 centroids corresponding to the 14 zone system illustrated in Figure 2 was chosen and a program MINPATH based on the TNET package (Homburger, 1972) was used to calculate travel times. The times in minutes are given in Table 3. The results are to a large extent dependent on the accuracy of this interzonal travel time matrix which in turn is dependent on the fit of the network and centroids to the zonal system. It should be noted that there are cases where the centroids of the zones do not correspond very well to network nodes. Given these limitations the interzonal travel times were used to calculate journey to work and journey to shop travel times in a straightforward manner, with zero costs being assigned to the no-work and no-shop categories. An improved method of estimating the cost of these patterns

DISTANCE MATRIX

DESTINATION

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	5.25	7.05	32.01	19.39	11.90	12.89	16.07	26.99	23.72	24.77	29.57	40.41	27.77	28.93	24.24
2	7.05	4.63	25.19	12.34	11.07	20.32	24.19	35.11	30.77	31.52	29.06	39.90	36.63	35.98	26.27
3	32.01	25.19	6.30	12.61	18.20	24.84	28.71	39.63	35.09	22.64	17.84	14.75	25.42	29.54	47.69
4	19.39	12.34	12.61	4.07	7.00	16.25	20.12	31.04	34.15	21.70	16.90	27.56	24.29	39.70	25.17
5	11.90	11.07	18.20	7.00	5.20	9.25	13.12	24.04	20.08	22.93	27.73	33.99	25.32	25.29	18.83
6	12.89	20.32	24.84	16.25	9.25	4.38	3.87	14.79	18.39	21.24	26.04	32.30	23.63	23.60	25.24
7	16.07	24.19	28.71	20.12	13.12	3.87	3.82	10.92	15.03	19.97	27.48	32.28	30.48	39.85	30.74
8	26.99	35.11	39.63	31.04	24.04	14.79	10.92	6.32	25.95	38.40	43.20	54.04	41.40	30.89	53.68
9	23.72	30.77	35.09	34.15	20.08	18.39	15.03	25.95	4.23	12.45	17.25	24.12	15.45	4.94	49.71
10	24.77	31.52	22.64	21.70	22.93	21.24	19.97	38.40	12.45	4.29	4.80	11.67	3.00	17.39	53.53
11	29.57	29.06	17.84	16.90	27.73	26.04	27.48	43.20	17.25	4.80	5.02	10.84	7.57	22.19	58.45
12	40.41	39.90	14.75	27.56	33.99	32.30	32.28	54.04	24.12	11.67	10.84	3.70	10.67	14.79	74.62
13	27.77	36.63	25.42	24.29	25.32	23.63	30.48	41.40	15.45	3.00	7.57	10.67	3.45	17.41	61.00
14	28.93	35.98	29.54	39.70	25.29	23.60	39.85	30.89	4.94	17.39	22.19	14.79	17.41	3.73	61.70

Table 3 Interzonal Travel Times (minutes)

is given in Chapter 4.2.

Additional data on zonal land use characteristics and employment are shown in Appendix A, Tables A1 and A2.

3.5 Summary

This chapter has described a specific model which is related to the spatial interaction models in Chapter 2 and which also has a theoretical basis. The model searches for the entropy maximising pattern of retail centres which is expected to include some downtown zones. The model allows inferences to be made about other aspects of the city (residential location etc.). The aim is to show that these distributions predicted by the model fit observed distributions. The final section of the chapter described a data base to test the model.

CHAPTER 4

MODEL DEVELOPMENT AND RESULTS

4.1 Introduction

The previous chapter introduced the model and described the data. It is clear that a consistent, theoretically based model has been set up. The aim now is to test the realism of this model. Specifically the objective is to show that the assignment produced by the model compares with observed patterns. If the model fails to reproduce real world data then obviously the constraint procedure needs to be adjusted and hopefully in the process of making these adjustments insights to the workings of the urban area will be gained. The models developed in this chapter attempt to capture realistically the urban area - therefore a series of revisions, each built on the errors of the previous model, is tested. There is of course the temptation to continue adding constraints until the model almost completely matches observed data. This would however result only in a deterministic model in which uncertainty would be approaching zero. Moreover such a model would contravene Batty's (1976a) requirement of parsimony in urban modelling.

With this restriction in mind the models tested are as follows: (1) a spatial interaction model with two centres; (2) a spatial interaction model with five centres; (3) a model with agglomeration economies in the retail sector; (4) a model with employment allocation; and (5) a model with land use competition.

4.2 Model 1: Singly constrained joint location spatial interaction model with two centres

The aim of the model (as discussed in 3.1 above) is to find that distribution of centres which maximises S*. This means that the model must be run for many different distributions of centres. In this version of the model all possible distributions of two service centres are examined.

(1) The model

The assignment of individuals to categories is such as to maximise

$$-\sum_i \sum_j \sum_k p_{ijk} \log \frac{p_{ijk}}{l_i} \tag{27}$$

subject to

$$\sum_i \sum_j p_{ijk} = d_k \qquad (k = 1, \ldots, 15) \tag{28}$$

and

$$\sum_i \sum_j \sum_k p_{ijk}\, c_{ijk} = \bar{c} \tag{29}$$

where the range of summation here and hereafter is i = 1, ..., 14; j = 1, ..., J; k = 1, ..., 15; and

p_{ijk} = joint probability of living in i and having shopping trip pattern j,

l_i = prior probability based on zone size,

d_k = the proportion of trips observed to end in k,

$c_{ijk} = c_{ij} + c_{ik}$,

c_{ij} = cost of living in i and having shopping pattern j,

c_{ik} = cost of living in i and working in k,

$\bar{c}$ = average trip cost (time in minutes),

J = the number of trip patterns (see 3.1).

The solution to this constrained maximisation problem is found by forming the Lagrangean:

$$L = -\sum_i \sum_j \sum_k p_{ijk} \log \frac{p_{ijk}}{l_i} - \sum_k \lambda_k (\sum_i \sum_j p_{ijk} - d_k)$$

$$- \beta(\sum_i \sum_j \sum_k p_{ijk}\, c_{ijk} - \bar{c})$$

Then

$$\frac{\partial L}{\partial p_{ijk}} = -1 - \log p_{ijk} + \log l_i - \lambda_k - \beta c_{ijk} = 0$$

Rearranging, and absorbing -1 into λ_k we get

$$\log \frac{p_{ijk}}{l_i} = -\lambda_k - \beta c_{ijk}$$

$$p_{ijk} = l_i \exp(-\lambda_k - \beta c_{ijk})$$

Then, since $\sum_i \sum_j p_{ijk} = d_k$

$$\sum_i \sum_j l_i \exp(-\lambda_k - \beta c_{ijk}) = d_k ,$$

$$\exp(-\lambda_k) = \frac{d_k}{\sum_i \sum_j l_i \exp(-\beta c_{ijk})} ,$$

and so

$$p_{ijk} = \frac{l_i\ d_k \exp(-\beta c_{ijk})}{\sum_i \sum_j l_i \exp(-\beta c_{ijk})}$$

and rewriting this:

$$p_{ijk} = l_i b_k d_k \exp(-\beta c_{ijk}) ,$$

where

$$b_k = 1/\sum_i \sum_j l_i \exp(-\beta c_{ijk}) .$$

The solution p_{ijk} is an assignment of individuals to the various categories which satisfies known facts and is at the same time noncommittal with respect to missing information. The distribution p_{ijk} is then manipulated to give the various marginal distributions described in section 3.3. The values of p_{ijk} which solve the constrained maximisation problem (equations 27-29) are found using Newton Raphson methods. (The computer program for the latest version of the model is reproduced in Appendix D.)

(2) Results: both centres at 5

The values of p_{ijk} which solve the constrained maximisation problem were found for all possible distributions of two centres. (That is,

$\binom{14}{2} + \binom{14}{1} = 105$.) The distribution of centres which maximises S* is (5,5). The results associated with this maximum are shown in Tables 4 and 5 and Fig. 3. Correlation coefficients (R^2) between observed and predicted values are given below the relevant distributions. The model has obviously produced low correlations and there are a number of reasons for this. The average cost constraint (equation 29) does not differentiate between work trips and shopping trips. This constraint will be disaggregated in model 2. A major problem with the model is the large proportion of individuals living in the outer city assigned to the category 'no shop trip'. (See Table 4, f(12,0), f(13,0), f(14,0).) This occurs because zero costs are associated with the 'no shop trip' category, thereby allowing the model to spread individuals out and at the same time meeting the average cost constraint. Efforts to correct this will be detailed in the next section.

The small number of service centres makes it impossible for the model to represent the actual distribution of centres. It is however encouraging that the model located the two centres near the CBD (zone 6 in Figure 2). This provides some support for the hypothesis that the pattern that maximises S* contains city centre zones. In later versions of the model the number of centres will be increased. The flows to retail centres predicted by this model (Figure 3) are obviously a very crude representation of the pattern as there is only *one* destination.

4.3 Model 2: Modified cost calculations

(1) Introduction

The changes suggested in 4.2(2) have been incorporated in this model. Data for a disaggregated average trip cost constraint were obtained

Residence Zone	Number of Shopping Trips made by Household (f_{im}) 0		1		2		Total (p_i)	
1	.012	(.017)	.010	(.003)	.002	(.000)	.025	(.020)
2	.024	(.020)	.022	(.008)	.005	(.001)	.051	(.029)
3	.071	(.059)	.040	(.026)	.005	(.005)	.116	(.090)
4	.066	(.107)	.080	(.035)	.025	(.005)	.171	(.147)
5	.046	(.124)	.046	(.038)	.012	(.004)	.105	(.165)
6	.029	(.094)	.030	(.013)	.008	(.000)	.066	(.107)
7	.023	(.038)	.018	(.013)	.004	(.003)	.044	(.053)
8	.033	(.061)	.012	(.016)	.001	(.003)	.046	(.079)
9	.048	(.044)	.022	(.025)	.003	(.000)	.073	(.069)
10	.064	(.098)	.026	(.042)	.002	(.007)	.091	(.146)
11	.030	(.046)	.008	(.018)	.001	(.007)	.039	(.070)
12	.036	(.004)	.006	(.001)	.000	(.000)	.043	(.005)
13	.049	(.005)	.016	(.003)	.001	(.001)	.067	(.009)
14	.047	(.008)	.016	(.003)	.001	(.001)	.063	(.012)
TOTAL (f_m)	.578	(.725)	.352	(.244)	.070	(.037)	1.000	(1.006)

$R^2(f_{im}) = 0.19$ $\qquad$ $R^2(p_i) = 0.24$

$R^2(d_{ik}) = 0.28$

Retail Sales, r_g = (1.000)

Table 4: Model 1, Centres at (5,5); predictions by place of residence.

Note: The table describes the actual proportion (in brackets) and predicted proportion of households in each category. There are small rounding errors.

Work Zone	Number of Shopping Trips made by Household (e_{km}) 0		1		2		Total (p_i)
1	.004	(.004)	.001	(.004)	.001	(.000)	.008
2	.111	(.153)	.084	(.052)	.019	(.009)	.215
3	.020	(.025)	.012	(.009)	.002	(.000)	.034
4	.026	(.042)	.018	(.005)	.004	(.000)	.047
5	.052	(.073)	.036	(.022)	.008	(.001)	.096
6	.080	(.104)	.050	(.031)	.010	(.005)	.140
7	.022	(.026)	.014	(.010)	.003	(.003)	.039
8	.023	(.034)	.014	(.005)	.003	(.001)	.040
9	.015	(.016)	.006	(.005)	.001	(.003)	.023
10	.018	(.022)	.008	(.005)	.001	(.000)	.027
11	.006	(.005)	.002	(.004)	.000	(.000)	.009
12	.000	(.000)	.000	(.000)	.000)	(.000)	.000
13	.000	(.000)	.000	(.000)	.000	(.000)	.000
14	.002	(.003)	.000	(.000)	.000	(.000)	.003
15	.196	(.216)	.104	(.088)	.018	(.014)	.319

$R^2(e_{km}) = 0.76$

Table 5: Model 1, Centres at (5,5); predictions by place of work.

Note: Small rounding errors account for the discrepancy between f_m and e_m.

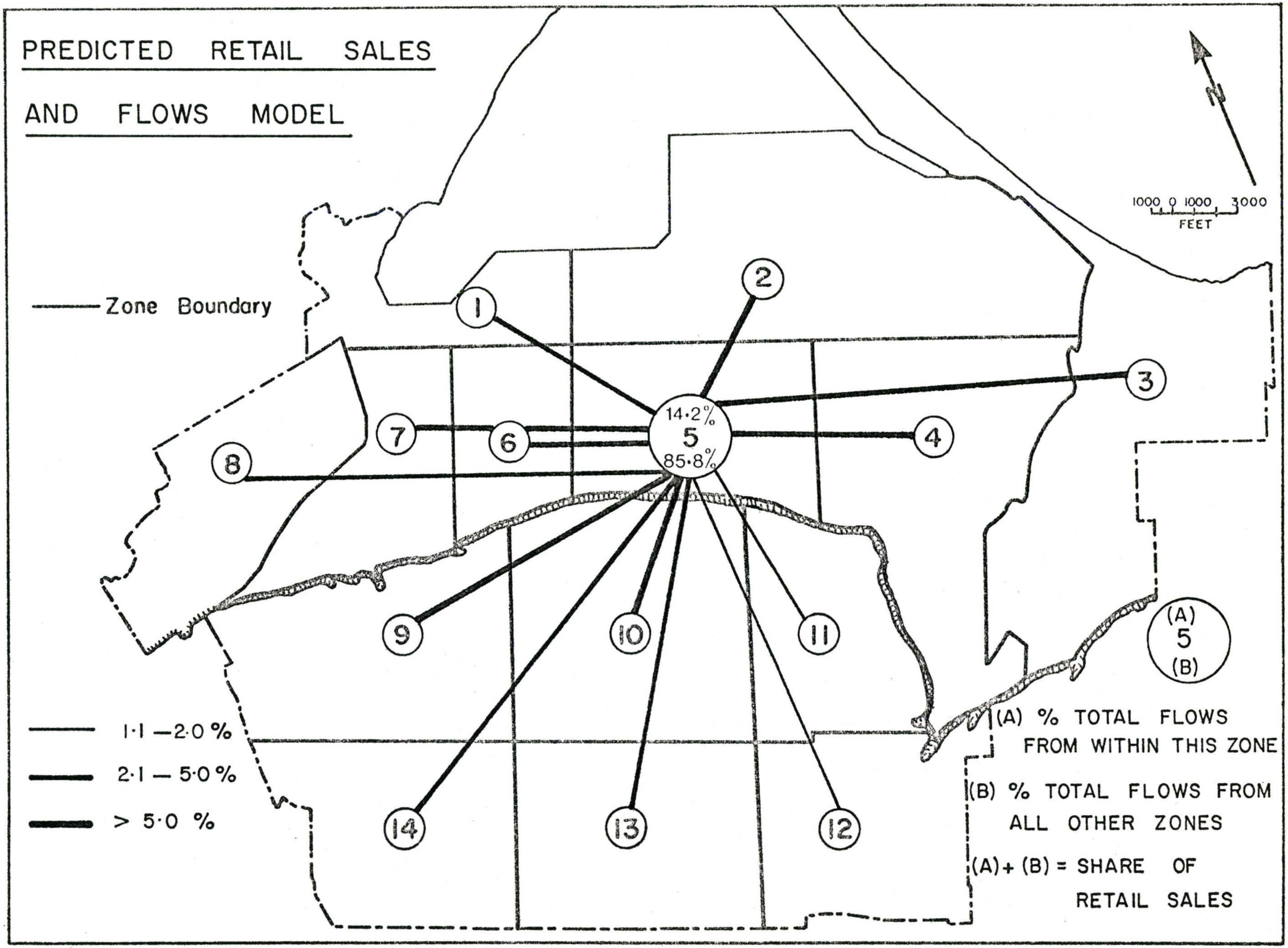

Figure 3 Model 1, Centres at 5; Retail Sales and Flows to Retail Centres

from the survey. The values used are $\bar{c}_S = 5.21$, $\bar{c}_W = 13.33$. An additional constraint on the average number of shopping trips per household ($\bar{x} = .31$) has been included to decrease the proportion assigned to the no shop trip category.

In model 1 no cost was assigned to the "no shop trip" category, and as the results indicated, this produced an unsatisfactory assignment of individuals. (See Table 4.) A modified method of calculating costs was designed whereby $\tilde{c}_{ij}$ is based on the expected number of trips per day for a household which is observed to make m = 0, 1, 2 trips. Specifically for $j = 1$ $(m = 0)$ $\tilde{c}_{ij}$ is calculated as the product of [the time taken to travel to the nearest shopping centre to zone i] and the [expected number of shopping trips] made by households observed to make zero trips. (The calculation of these expectations is described in Appendix C.) For $j = 2, \ldots, M+1$, $(m = 1)$ $\tilde{c}_{ij}$ is calculated as the product of [the time taken to reach the chosen centre] and [the expected number of trips made per day by households observed to make one trip]. The travel costs for shopping patterns involving two facilities are calculated as the product of [the sum of the travel times to the two centres] and [the expected number of trips per day made by a household which was observed to make two trips]. Table C1 (Appendix C) illustrates the effect of this change. These calculations can be summarized as follows

$$\tilde{c}_{ij} = \begin{cases} c_{ign}[\frac{1}{6} \sum\limits_{l=0}^{\infty} l\, P(l|0)] \text{ ,} & j = 1; \\ c_{ij(g)}[\frac{1}{6} \sum\limits_{l=0}^{\infty} l\, P(l|1)] \text{ ,} & j = 2, \ldots, N+1; \\ (\sum\limits_{g\varepsilon j} c_{ij(g)}) \; [\frac{1}{6} \sum\limits_{l=0}^{\infty} l\, P(l|2)] \text{ ,} & j = N+2, \ldots, J. \end{cases}$$

where: $\tilde{c}_{ij}$ is the new cost matrix for shop trips,

c_{ign} is the observed time taken to travel from i to the nearest centre g,

$c_{ij(g)}$ is the observed time taken to travel from i to the centre g associated with pattern j,

$g \varepsilon j$ signifies that the sum is over those centres associated with a pattern j,

$P(l|m)$ m = 0, 1, 2 is defined in Appendix C, and is the conditional probability of making l trips in a week given m.

(2) The model

The model can be stated as follows: Maximise (27) subject to (28) and the following constraints

$$\sum_i \sum_j \sum_k p_{ijk} \tilde{c}_{ij} = \bar{c}_S \tag{30}$$

$$\sum_i \sum_j \sum_k p_{ijk} c_{ik} = \bar{c}_W \tag{31}$$

$$\sum_i \sum_j \sum_k p_{ijk} x_j = \bar{x} \tag{32}$$

where: $\tilde{c}_{ij}$ = the new shopping trip cost matrix,

$\bar{c}_S$ = mean length of shopping trip,

$\bar{c}_W$ = mean length of work trip,

x_j = number of trips associated with trip pattern j (j = 1, 2, ..., J),

$\bar{x}$ = mean number of shopping trips per household.

The solution to this constrained maximisation problem is

$$p_{ijk} = l_i b_k d_k \exp(-\beta_1 \tilde{c}_{ij} - \beta_2 c_{ik} - \gamma x_k)$$

where

$$b_k = 1/\sum_i \sum_j \exp(-\beta_1 \tilde{c}_{ij} - \beta_2 c_{ik} - \gamma x_k) \quad ,$$

and the parameters β_1, β_2 and γ are found to satisfy (30), (31) and (32). The solution p_{ijk} can be manipulated as in model 1 to provide information about residential location, work trip interchanges, shopping behaviour and retail sales.

(3) Results: Centres at (5,10) and (5,7,10)

As can be seen in Tables 6 and 7 the correlations have improved over those in model 1 - especially the residential location correlation. However the model continues to underpredict the lower city population share and to overpredict the mountain's share. Clearly the model can distribute individuals over a wide area and at the same time meet average cost constraints. This means that cost calculations are still not stringent enough, or possibly that the estimates of average cost used are too high. In later versions of the model efforts to improve this are made. The work trip matrix shows a correlation of $R^2 = 0.46$ which represents an improvement over model 1. The predicted distribution of service centres is (5,10). Retail sales in each centre are predicted as (.452, .548) and are illustrated with predicted flows in Figure 4.

A complete search through the distributions of three centres produced a maximum at (5,7,10). The results associated with this distribution of service centres are shown in Table 8 and Figure 5. Retail sales for

Residence Zone	Number of Shopping Trips made by Household (f_{im}) 0		1		2		Total (p_i)	
1	.011	(.017)	.004	(.003)	.000	(.000)	.015	(.020)
2	.023	(.020)	.007	(.008)	.001	(.001)	.031	(.029)
3	.052	(.059)	.022	(.026)	.003	(.005)	.076	(.090)
4	.090	(.107)	.033	(.035)	.004	(.005)	.127	(.147)
5	.068	(.124)	.025	(.038)	.003	(.004)	.095	(.165)
6	.036	(.094)	.014	(.013)	.002	(.000)	.051	(.107)
7	.020	(.038)	.008	(.013)	.001	(.003)	.029	(.053)
8	.016	(.061)	.005	(.016)	.000	(.003)	.021	(.079)
9	.051	(.044)	.020	(.025)	.003	(.000)	.074	(.069)
10	.117	(.098)	.042	(.042)	.004	(.007)	.163	(.146)
11	.053	(.046)	.018	(.018)	.002	(.007)	.073	(.070)
12	.042	(.004)	.013	(.001)	.001	(.000)	.057	(.005)
13	.097	(.005)	.034	(.003)	.003	(.001)	.135	(.009)
14	.037	(.008)	.014	(.003)	.002	(.001)	.053	(.012)
TOTAL (f_m)	.713	(.725)	.259	(.244)	.029	(.037)	1.001	(1.006)

$R^2(f_{im}) = 0.50$ $\qquad$ $R^2(p_i) = 0.25)$

$R^2(d_{ik}) = 0.46$

Retail Sales, $r_g = (0.452, 0.548)$

Table 6: Model 2, Centres at (5,10); predictions by place of residence.

See Note with Table 4.

Work Zone	Number of Shopping Trips made by Household (e_{km}) 0		1		2		Total (d_k)
1	.006	(.004)	.002	(.004)	.000	(.000)	.008
2	.153	(.153)	.055	(.052)	.006	(.009)	.215
3	.024	(.025)	.009	(.009)	.001	(.000)	.034
4	.033	(.042)	.012	(.005)	.001	(.000)	.047
5	.068	(.073)	.025	(.022)	.003	(.001)	.096
6	.100	(.104)	.037	(.031)	.004	(.005)	.140
7	.028	(.026)	.011	(.010)	.001	(.003)	.039
8	.029	(.034)	.010	(.005)	.001	(.001)	.040
9	.017	(.016)	.006	(.005)	.001	(.003)	.023
10	.020	(.022)	.007	(.005)	.001	(.000)	.027
11	.007	(.005)	.003	(.004)	.000	(.000)	.009
12	.000	(.000)	.000	(.000)	.000	(.000)	.000
13	.000	(.000)	.000	(.000)	.000	(.000)	.000
14	.002	(.003)	.000	(.000)	.000	(.000)	.003
15	.228	(.216)	.082	(.088)	.009	(.014)	.319
TOTAL (e_m)	.715	(.723)	.259	(.240)	.028	(.036)	1.002 (.999)

$R^2(e_{km}) = 0.99$

Table 7: Model 2, Centres at (5,10); predictions by place of work.

See Note with Table 5.

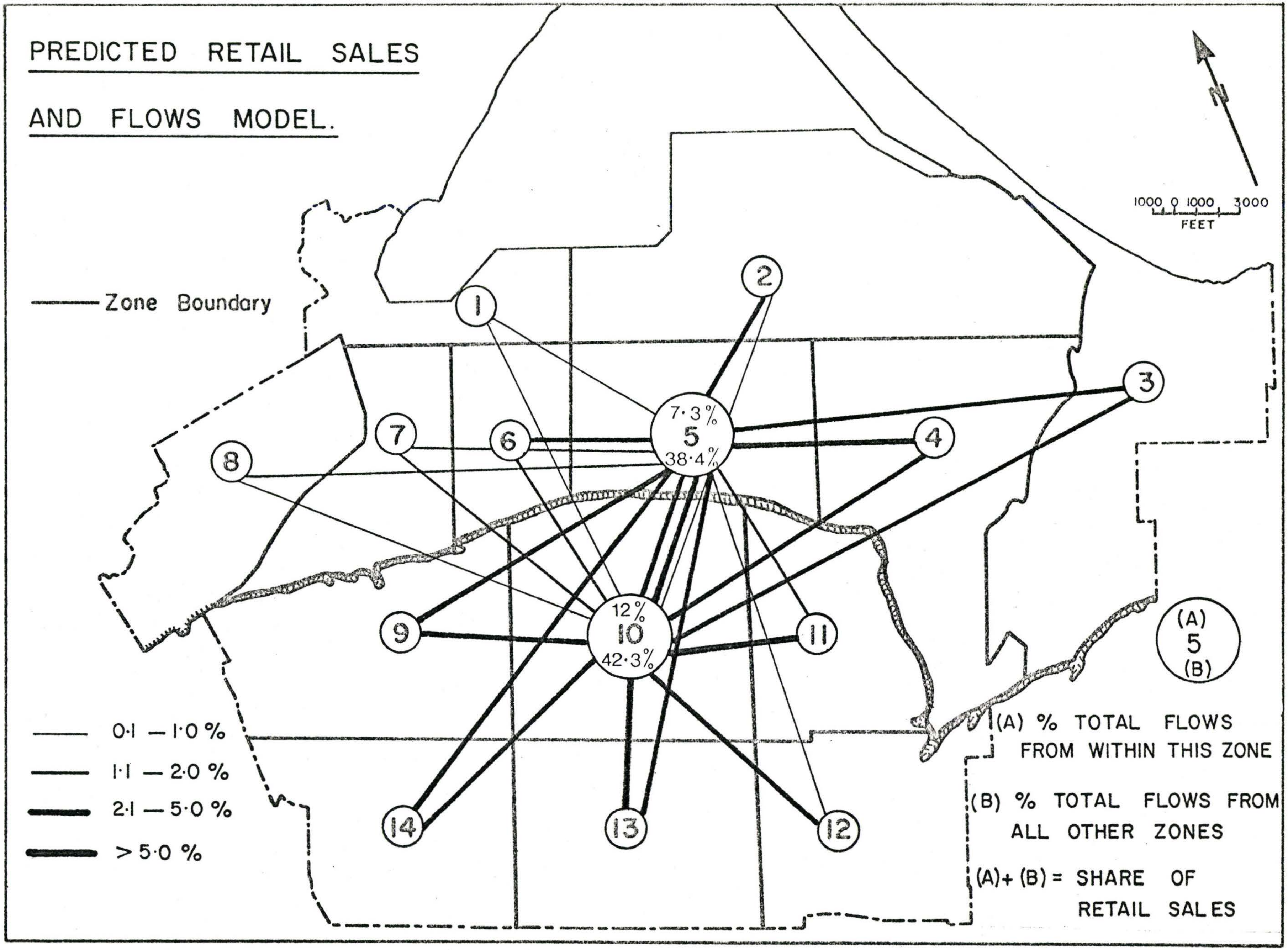

Figure 4 Model 2, Centres at (5,10); Retail Sales and Flows to Retail Centres

Residence Zone	Number of Shopping Trips made by Household (f_{im}) 0		1		2		Total (p_i)	
1	.010	(.017)	.004	(.003)	.000	(.000)	.014	(.020)
2	.022	(.020)	.007	(.008)	.000	(.001)	.029	(.029)
3	.054	(.059)	.020	(.026)	.003	(.005)	.077	(.090)
4	.082	(.107)	.028	(.035)	.005	(.005)	.115	(.147)
5	.060	(.124)	.022	(.038)	.004	(.004)	.086	(.165)
6	.045	(.094)	.018	(.013)	.003	(.000)	.066	(.107)
7	.032	(.038)	.011	(.013)	.002	(.003)	.045	(.053)
8	.036	(.061)	.011	(.016)	.001	(.003)	.048	(.079)
9	.050	(.044)	.022	(.025)	.004	(.000)	.076	(.069)
10	.109	(.098)	.035	(.042)	.005	(.007)	.149	(.146)
11	.050	(.046)	.014	(.018)	.002	(.007)	.066	(.070)
12	.043	(.004)	.012	(.001)	.002	(.000)	.057	(.005)
13	.091	(.005)	.025	(.003)	.004	(.001)	.120	(.009)
14	.039	(.008)	.012	(.003)	.001	(.001)	.052	(.012)
TOTAL (f_m)	.723	(.725)	.241	(.244)	.036	(.037)	1.000	(1.006)

$R^2(f_{im}) = (.7412)^2$ $\quad$ $R^2(p_i) = (.5648)^2$

$R^2(d_{ik}) = (.7283)^2$

Retail Sales, $r_g = (0.327, 0.286, 0.387)$

Table 8: Model 2, Centres at (5,7,10); predictions by place of residence.

See Note with Table 4.

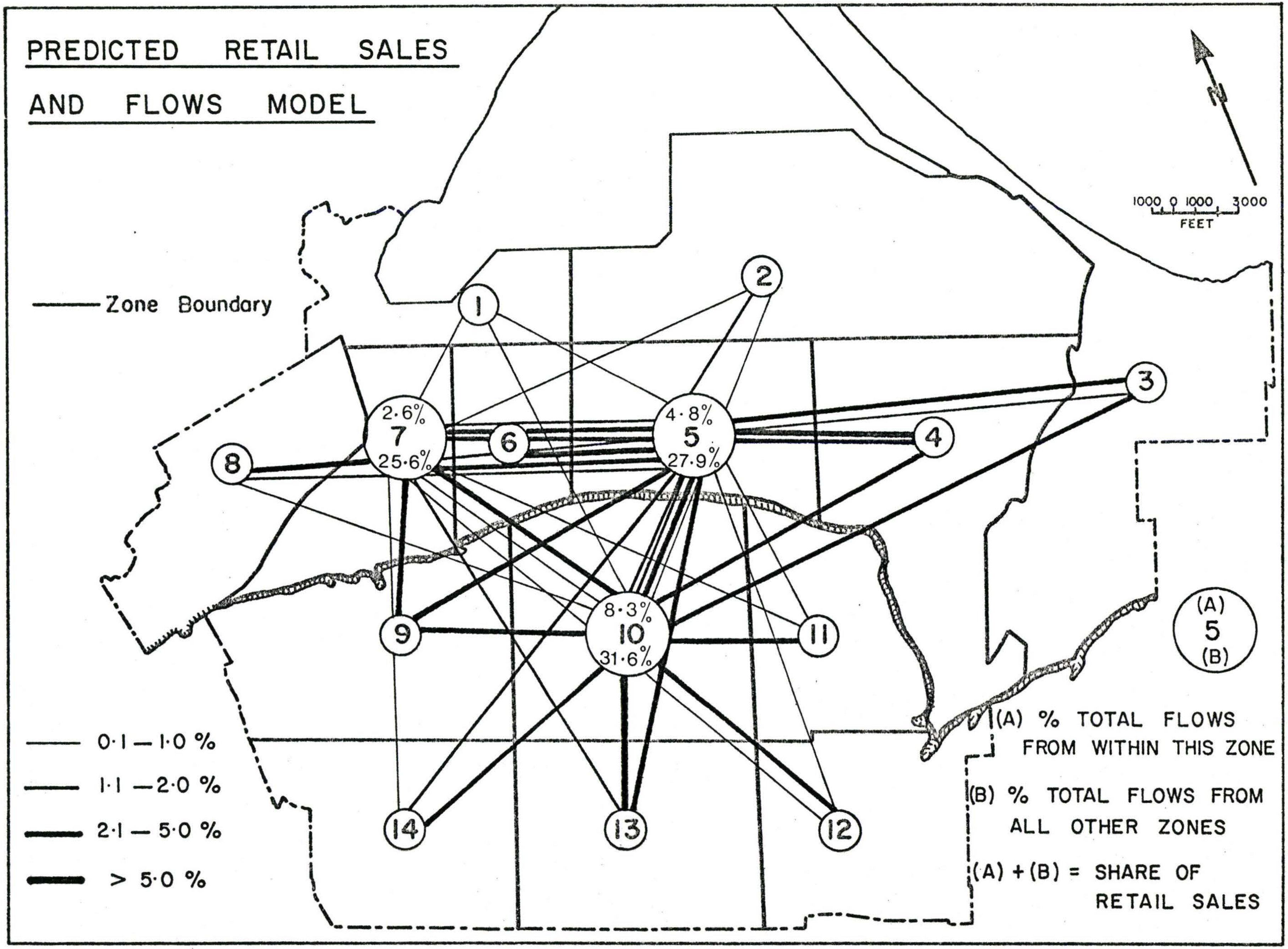

Figure 5 Model 2, Centres at (5,7,10); Retail Sales and Flows to Retail Centres

each centre are (41.8%, 28.6%, 38.7%) which is roughly comparable to the actual retail employment distribution in the upper and lower city.

(4) Modifications to Model 2

The prior probability of zone size l_i is calculated as developable land (i.e. residential, vacant and extractive acreage). The data used for these calculations are shown in Table A1. This prior is now modified to include only residential and vacant acreage. Preliminary tests with this modification on distributions (5,7) and (5,10) indicate that with two centres the residential correlation increased. Comparing Tables 6 and 9 it is clear that the degree of underprediction in the lower city has decreased. Furthermore the magnitude of the errors in zones 12, 13 and 14 has decreased. A test of the modified model using the pattern (5,7,10) produced similar changes in results (see Table 10). Notice that the correlations for the work trip matrix and shopping pattern by place of residence have improved. The modified model predicts higher retail sales in zone 7 and lower retail sales in zone 10.

(5) Algorithm for 5 centre case

The model is designed to run with an arbitrary number of service centres. However in practice the number of centres that can be modelled is 5. Even with 5 centres there are 14 ways of choosing one zone to house all five facilities, $\binom{14}{2}$ ways of choosing one zone for four facilities and one for one facility, another $\binom{14}{2}$ ways of choosing zones to house three facilities and two facilities, $\binom{14}{3}$ ways of locating three facilities in one zone and one in each of two other zones and of locating two facilities in each of two zones and one facility in another zone, $\binom{14}{4}$ ways of locating two facilities in one zone and one in each of three other zones, and $\binom{14}{5}$

Residence Zone	Number of Shopping Trips made by Household (f_{im}) 0		1		2		Total (p_i)	
1	.011	(.017)	.004	(.003)	.000	(.000)	.016	(.020)
2	.024	(.020)	.008	(.008)	.001	(.001)	.033	(.029)
3	.050	(.059)	.021	(.026)	.003	(.005)	.073	(.090)
4	.093	(.107)	.034	(.035)	.004	(.005)	.131	(.147)
5	.070	(.124)	.025	(.038)	.003	(.004)	.098	(.165)
6	.037	(.094)	.014	(.013)	.002	(.000)	.052	(.107)
7	.021	(.038)	.009	(.013)	.001	(.003)	.031	(.053)
8	.017	(.061)	.005	(.016)	.000	(.003)	.023	(.079)
9	.055	(.044)	.022	(.025)	.003	(.000)	.080	(.069)
10	.128	(.098)	.046	(.042)	.005	(.007)	.179	(.146)
11	.058	(.046)	.020	(.018)	.002	(.007)	.080	(.070)
12	.039	(.004)	.012	(.001)	.001	(.000)	.052	(.005)
13	.078	(.005)	.027	(.003)	.003	(.001)	.108	(.009)
14	.032	(.008)	.012	(.003)	.001	(.001)	.045	(.012)
TOTAL (f_m)	.713	(.725)	.259	(.244)	.029	(.037)	1.001	(1.006)

$R^2(f_{im}) = 0.58$ $\qquad$ $R^2(p_i) = 0.37$

$R^2(d_{ik}) = 0.53$

Retail Sales, $r_g = (.459, .541)$

Table 9: Model 2, Centres at (5, 10), modified prior; predictions by place of residence.

See Note with Table 4.

Residence Zone	Number of Shopping Trips made by Household (f_{im}) 0		1		2		Total (p_i)	
1	.011	(.017)	.004	(.003)	.000	(.000)	.016	(.020)
2	.023	(.020)	.007	(.008)	.000	(.001)	.031	(.029)
3	.052	(.059)	.019	(.026)	.003	(.005)	.075	(.090)
4	.085	(.107)	.029	(.035)	.005	(.005)	.118	(.147)
5	.062	(.128)	.023	(.038)	.004	(.004)	.089	(.165)
6	.046	(.094)	.018	(.013)	.003	(.000)	.067	(.107)
7	.034	(.038)	.013	(.013)	.002	(.003)	.048	(.053)
8	.038	(.061)	.011	(.016)	.001	(.003)	.051	(.079)
9	.055	(.044)	.024	(.025)	.004	(.000)	.082	(.069)
10	.118	(.098)	.038	(.042)	.006	(.007)	.162	(.146)
11	.054	(.048)	.014	(.018)	.002	(.007)	.072	(.070)
12	.039	(.004)	.011	(.001)	.002	(.000)	.051	(.005)
13	.072	(.005)	.019	(.003)	.002	(.001)	.094	(.009)
14	.033	(.008)	.010	(.003)	.001	(.001)	.045	(.012)
TOTAL (f_m)								

$R^2(f_{im}) = 0.64$ $\qquad$ $R^2(p_i) = 0.47$

$R^2(d_{ik}) = 0.60$

Retail Sales, $r_g = (0.328, 0.293, 0.379)$

Table 10: Model 2, Centres at (5,7,10), modified prior; predictions by place of residence.

See Note with Table 4.

ways of locating one facility in each of five zones: there are 3,927 different patterns of location of five facilities in 14 zones. Therefore an algorithm was designed, to find the pattern of facility location for which S* is a maximum. Given an initial location pattern, a new set of location patterns of facilities was defined to comprise all location patterns generated from the initial pattern by shifting each facility in turn to each of its neighbouring zones. The location pattern from this set for which S* is a maximum is chosen as the initial location pattern in the next iteration of the algorithm. If S* is maximised by the initial location pattern, the algorithm terminates. Although the algorithm is not guaranteed to discover the location pattern for which S* is maximised, tests on all patterns comprising two facilities and three facilities (beginning from arbitrary patterns) revealed that in those cases, maxima of S* were found.

Model 2 was run through several batches with 5 centres using this algorithm. The results produced by these trial runs reveal several weaknesses in the model. These problems are (i) the model assigns too many individuals to the no work category, (ii) the model assigns too many individuals to the no shop trip category on the mountain, and (iii) there is a lack of aggregation in the retail sector. These problems and the methods to overcome them are discussed in the following section.

4.4 Model 3: Retail sector agglomeration economies

(1) Introduction

In models 1 and 2 zero costs are associated with the no work trip category which makes it too easy to assign individuals to this group. It is necessary therefore to derive some cost for making no work trip. Previous efforts in this area have used balancing factors to 'patch up' errors due to non trip makers (Senior, 1973). This however is not satisfactory for as Kain (1962b) points out an increasing number of households are without a member in the work force. This implies a need to incorporate these individuals in a more appropriate way. Kain (1962b, 138) goes on to suggest that other facilities such as recreational and cultural centres influence location. In this model, though, it is assumed that households which make no work trip are still influenced by the location of workplaces, the rationale being that the household made trips to work in the past and is now retired or unemployed. On this basis the cost of being assigned to the no work trip category is computed as the weighted average of the cost of living in zone i and working in the other zones ($k = 1, \ldots, 14$), with employment d_k used as a weight. Thus:

$$c_{i,15} = \sum_{k=1}^{14} c_{ik}\, d_k/(1 - d_{15})$$

where d_k is the proportion known to be in the work trip category. The value of d_{15} used is .315 which is quite high. However this figure represents those who are unemployed, retired and also those with irregular work schedules so it has been accepted as a reasonable figure.

A vector of prior probabilities s_j ($j = 1, \ldots, J$) is necessary because preliminary tests show that the model fails to produce a downtown cluster without it. The prior probabilities are designed to measure agglomeration economies in the retail sector and thus solve one of the problems avoided by the Lowry Model (see section 1.4(4)). There is obviously some advantage to households in shopping at facilities which are close to other shopping facilities. The growth of suburban shopping plazas is an outcome of the need on the part of retailers to locate together to benefit from the increased patronage attracted by a cluster.

To model these agglomeration economies the mean distance t_g required to travel from shopping facility g to each of the other N-1 facilities is calculated. Next it is assumed that the probability of visiting shopping facility i is

$$\alpha e^{-\gamma t_i}$$

That is, a negative exponential function of the average distance between centres. (A worthwhile modification, for future research, would be to include a weight for each centre depending on its size.) The prior probability associated with shopping pattern j is then defined as

$$s_j = \begin{cases} \alpha & \text{for } j = 1 \\ \alpha e^{-\gamma t_{j-1}} & \text{for } j = 2, 3, \ldots, N+1 \\ \alpha^2 e^{-\gamma(t_m + t_n)} & \text{for } j = N+2, \ldots, J \end{cases}$$

These three equations correspond to the 0, 1 and 2 trip patterns respectively. In the case of 2 trips the household visits facilities in zones m and n. Since 68.53% of the households made zero shopping trips the value of α was set to .6853. The other parameter γ is chosen to ensure that the prior probabilities sum to 1.

(2) The model

The model incorporating these changes is: Maximise

$$-\sum_i \sum_j \sum_k p_{ijk} \log \frac{p_{ijk}}{m_i s_j} \tag{33}$$

Subject to (28), (30), (32) and

$$\sum_i \sum_j \sum_k p_{ijk} \check{c}_{ik} = \bar{c}_w$$

where s_j is a prior to represent agglomeration economies in the retail sector,

$\check{c}_{ik}$ is the new cost matrix for worktrips.

(3) Results: Centres at (5,5,9,7,10)

Experiments with earlier versions of the model for five centres indicated that the optimal pattern includes zones 5, 7 and 10. These centres were included at the start of the algorithm and the distribution of service centres which maximises S* was found to be (5,5,9,7,10). The results associated with the pattern are shown in Tables 11 and 12. These results show that the model now predicts residential location with much greater accuracy. The work trip matrix correlation is now $R^2 = 0.67$. The retail sales and flow predictions are shown in Figure 6.

Despite the simplifications caused by the limited data and the editing of those data, the model now predicts the location of shopping facilities and the shopping behaviour of households with reasonable accuracy.

Residence Zone	Number of Shopping Trips made by Household (f_{im}) 0		1		2		Total (p_i)	
1	.023	(.017)	.009	(.003)	.001	(.000)	.033	(.020)
2	.055	(.020)	.019	(.008)	.002	(.001)	.077	(.029)
3	.040	(.059)	.015	(.026)	.002	(.005)	.056	(.090)
4	.138	(.107)	.050	(.035)	.006	(.005)	.194	(.147)
5	.125	(.124)	.050	(.038)	.007	(.004)	.181	(.165)
6	.076	(.094)	.030	(.013)	.004	(.000)	.110	(.107)
7	.053	(.038)	.019	(.013)	.003	(.003)	.075	(.053)
8	.033	(.061)	.010	(.016)	.001	(.003)	.044	(.079)
9	.045	(.044)	.015	(.025)	.002	(.000)	.061	(.069)
10	.054	(.098)	.016	(.042)	.002	(.007)	.072	(.146)
11	.023	(.046)	.006	(.018)	.001	(.007)	.029	(.070)
12	.009	(.004)	.002	(.001)	.000	(.000)	.012	(.005)
13	.024	(.005)	.006	(.003)	.001	(.001)	.031	(.009)
14	.020	(.008)	.005	(.003)	.001	(.001)	.026	(.012)
TOTAL (f_m)	.718	(.725)	.252	(.244)	.033	(.037)	1.003	(1.006)

$R^2(f_{im}) = 0.79$ $\qquad R^2(p_i) = 0.64$

$R^2(d_{ik}) = 0.67$

Retail Sales, r_g = (0.26, 0.26, 0.146, 0.225, 0.108)

Table 11: Model 3, Centres at (5,5,9,7,10); predictions by place of residence.

See Note with Table 4.

Work Zone	Number of Shopping Trips made by Household (e_{km}) 0		1		2		Total (d_k)
1	.006	(.004)	.002	(.004)	.000	(.000)	.008
2	.152	(.153)	.055	(.052)	.007	(.009)	.215
3	.025	(.025)	.008	(.009)	.001	(.000)	.034
4	.033	(.042)	.012	(.005)	.001	(.000)	.047
5	.068	(.073)	.025	(.022)	.003	(.001)	.096
6	.100	(.104)	.036	(.031)	.005	(.005)	.140
7	.028	(.026)	.010	(.010)	.001	(.003)	.039
8	.029	(.034)	.010	(.005)	.001	(.001)	.040
9	.018	(.016)	.005	(.005)	.001	(.003)	.023
10	.021	(.022)	.006	(.005)	.001	(.000)	.027
11	.007	(.005)	.002	(.004)	.000	(.000)	.009
12	.000	(.000)	.000	(.000)	.000	(.000)	.000
13	.000	(.000)	.000	(.000)	.000	(.000)	.000
14	.002	(.003)	.001	(.000)	.000	(.000)	.003
15	.228	(.216)	.080	(.088)	.010	(.014)	.319
TOTAL (e_m)	.717	(.723)	.252	(.240)	.031	(.036)	1.000 (.999)

$R^2(e_{km}) = .99$

Table 12: Model 3, Centres at (5,5,9,7,10); predictions by place of work.

See Note with Table 5.

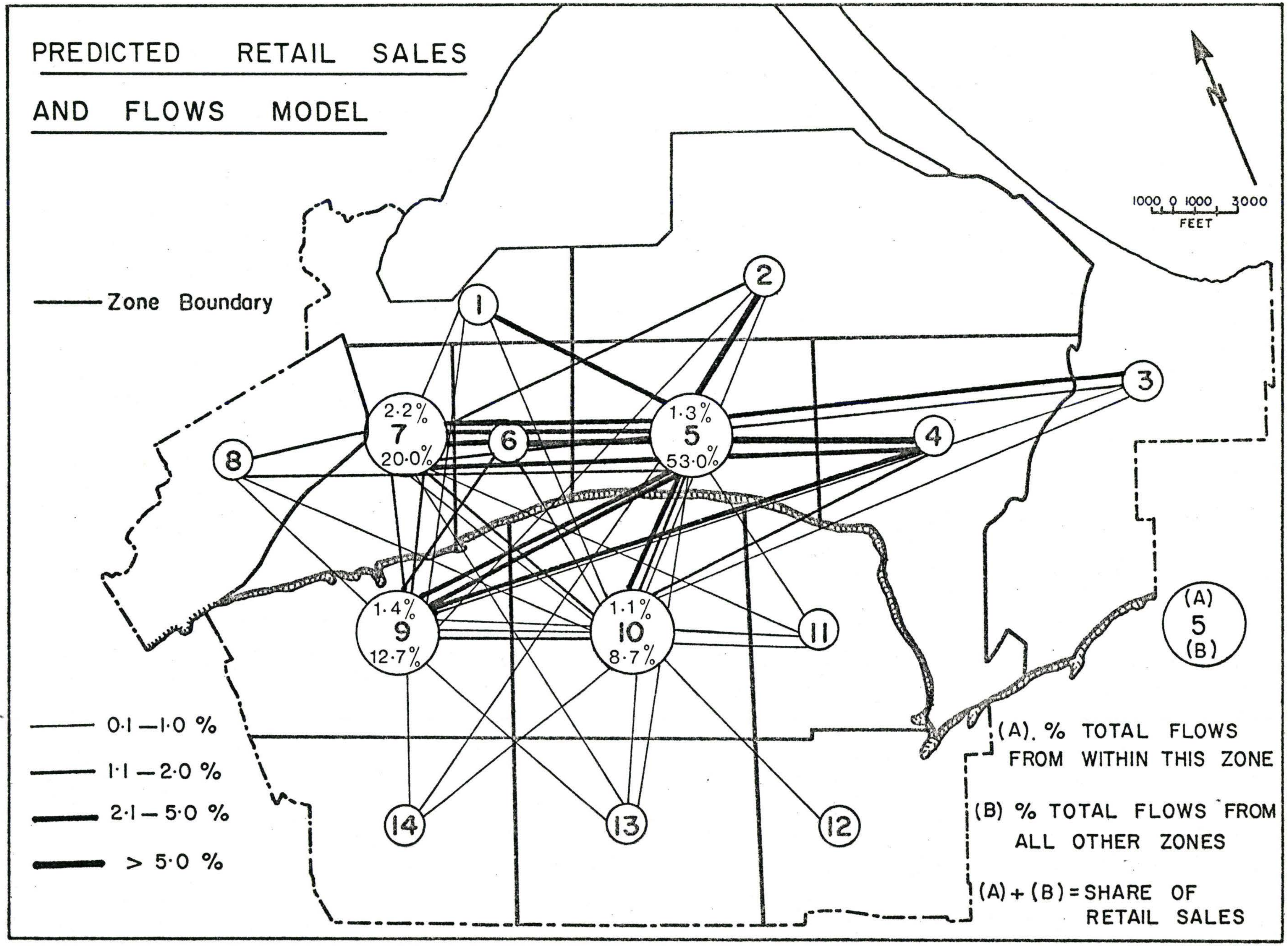

Figure 6 Model 3, Centres at (5,5,9,7,10); Retail Sales and Flows to Retail Centres

4.5 Model 4: Modified employment constraint

(1) Introduction

While the model described has given a good fit to the observed data, a number of modifications are now introduced which help to improve its realism. It is clear that work trip destinations which are observed in the data have not been disaggregated by type of employment. This means that the destination vector d_k (k = 1, ..., 15) to which trips are constrained includes many trips which were made to retail work places. Since the intention is to make retail location endogenous this lack of differentiation presents a problem. The task of rectifying this problem is tackled in two stages.

Firstly a detailed breakdown of employment by sector and by planning district was constructed (see Table A2). These figures are used to calculate basic employment (i.e. Manufacturing and Wholesale) and are used together with a known proportion of unemployed/retired (.319) to make up a new destination constraint, which consists only of basic employment and unemployed, (k = 1, ..., 15). The model is constrained to fit this basic employment distribution. Since we are now using data from two sources it is expected that the correlations will no longer be as high as in model 3. For instance the observed work trip interchange matrix against which the model is tested is obviously not compatible with the basic employment distribution as it contains trips to *all* kinds of work. Thus the results for work places produced by the model are no longer directly comparable (via correlation) with the data observed as the model is no longer being fitted to a complete set of observations. The work pattern results are therefore reported without comparable observed figures.

(2) Preliminary results: centres at (5,4,7,1,10)

The maximum entropy pattern produced by the search was (5,4,7,1,10). The results indicate that the model is now producing improved residential location correlations over those produced by model 2 with the same pattern of centres (Tables 13 and 14). However there are now substantial over-predictions in the *lower* city. Model 4 also produced higher retail sales for all lower city centres, which indicates that the revised employment vector has underestimated the employment opportunities on the mountain. (This can also be seen from the predicted work trip matrix which shows fewer entries for mountain zones.) The allocation of basic jobs used is biased towards the lower city area, with the result that the model is unable to spread residents over a wide area and keep the average cost constraints. A further modification therefore is to include the service employment created by the allocation of a retail centre to a zone in the calculation of the zonal employment figure. This modification is discussed in the following paragraphs.

(3) Partly endogenous employment

Data from the HWPDD (1977) (Table A1) indicate a total employment figure of 135,220 in 1971. Of these 66,050 (48.9%) worked in basic, 51,519 (38.1%) worked in service and 17,618 (13%) worked in retail jobs. These figures are used, together with an observed proportion of 31.9% making no work trips, to construct a new breakdown by employment type as follows:

Employment Sector	%
Basic	33.3
Service, Retail and other	34.8
No work trip	31.9

Residence Zone	Number of Shopping Trips made by Household (f_{im}) 0	1	2	Total (p_i)
1	.044 (.017)	.016 (.003)	.002 (.000)	.062 (.020)
2	.116 (.020)	.043 (.008)	.004 (.001)	.163 (.029)
3	.053 (.059)	.010 (.026)	.002 (.005)	.074 (.090)
4	.167 (.107)	.062 (.035)	.006 (.005)	.235 (.147)
5	.133 (.124)	.050 (.038)	.005 (.004)	.188 (.165)
6	.051 (.094)	.019 (.013)	.002 (.000)	.072 (.107)
7	.031 (.038)	.011 (.013)	.001 (.003)	.043 (.053)
8	.018 (.061)	.006 (.016)	.001 (.003)	.025 (.079)
9	.026 (.044)	.009 (.025)	.001 (.000)	.036 (.069)
10	.031 (.098)	.011 (.042)	.001 (.007)	.043 (.146)
11	.015 (.046)	.005 (.018)	.001 (.007)	.021 (.070)
12	.007 (.004)	.002 (.001)	.000 (.000)	.009 (.005)
13	.011 (.005)	.004 (.003)	.000 (.001)	.015 (.009)
14	.010 (.008)	.004 (.003)	.000 (.001)	.014 (.012)
TOTAL (f_m)	.713 (.725)	.261 (.244)	.026 (.037)	1.000 (1.006)

$R^2(f_{im}) = 0.53$ $\quad$ $R^2(p_i) = 0.33$

$R^2(d_{ik}) = 0.43$

Retail Sales, r_g = (0.324, 0.206, 0.197, 0.178, 0.095)

Table 13: Model 4, Preliminary results, Centres at (5,4,7,1,10); predictions by place of residence.

See note with Table 4.

Work Zone	Number of Shopping Trips made by Household (e_{km}) 0	1	2	Total (d_k^{mod})
1	.005	.002	.000	.007
2	.310	.114	.011	.436
3	.017	.006	.001	.023
4	.013	.005	.000	.018
5	.049	.018	.002	.069
6	.048	.018	.002	.068
7	.031	.012	.001	.044
8	.005	.002	.000	.007
9	.002	.001	.000	.003
10	.002	.001	.000	.003
11	.001	.000	.000	.001
12	.001	.000	.000	.001
13	.001	.000	.000	.001
14	.001	.000	.000	.001
15	.227	.084	.008	.319
TOTAL (e_m)	.713	.263	.025	1.001

Table 14: Model 4, Preliminary results, Centres at (5,4,7,1,10); predictions by place of work.

See text.

Using the proportion in basic employment (33.3%) as a total the observed employment vector ($k = 1, \ldots, 14$ only) was rescaled. This vector is then input to the model and for each new pattern of centres the employment in the zones with centres is incremented by its share of 34.8%. This in effect makes employment partly endogenous within the model. A slight complication arises in that the costs associated with the no work category have to be calculated at every iteration since the weights, d_k^{mod}, change with each new pattern of centres.

Specifically d_k^{mod} is calculated as:

$$d_k^{mod} = \begin{cases} \dfrac{b_k}{E}(1 - U) & k = 1, \ldots, 14 \text{ if } k \text{ has no service centre,} \\[2ex] \dfrac{b_k}{E} + \dfrac{x}{5}\,\dfrac{E - B}{E}\,(1 - U) & k = 1, \ldots, 14 \text{ if } k \text{ has } x \text{ service centres,} \\[2ex] U & k = 15. \end{cases}$$

where b_k = manufacturing and wholesale employment in zone k,

E = city total employment,

U = share of 'no work trip' in employment,

B = city total 'basic' employment.

(4) Results: centres at (5,4,7,1,6)

The maximum entropy pattern produced by the search is (5,4,7,1,6). The results (Table 15) still show substantial overprediction in the lower city (e.g. zones 1, 2, 4). Table 16 shows the "predicted" employment vector. Note that when these results are rescaled to exclude the unemployed category that the model suceeds in matching some aspects of the total

Residence Zone	Number of Shopping Trips made by Household (f_{im}) 0		1		2		Total (p_i)	
1	.046	(.017)	.019	(.003)	.001	(.000)	.066	(.020)
2	.120	(.020)	.057	(.008)	.004	(.001)	.181	(.029)
3	.019	(.059)	.011	(.026)	.001	(.005)	.032	(.090)
4	.122	(.107)	.058	(.035)	.004	(.005)	.184	(.147)
5	.216	(.124)	.066	(.038)	.004	(.004)	.287	(.165)
6	.080	(.094)	.026	(.013)	.002	(.000)	.107	(.107)
7	.047	(.038)	.018	(.013)	.001	(.003)	.067	(.053)
8	.016	(.061)	.008	(.016)	.001	(.003)	.024	(.079)
9	.013	(.044)	.005	(.025)	.000	(.000)	.019	(.069)
10	.012	(.098)	.004	(.042)	.000	(.007)	.016	(.146)
11	.004	(.046)	.002	(.018)	.000	(.007)	.006	(.070)
12	.002	(.004)	.001	(.001)	.000	(.000)	.005	(.005)
13	.004	(.005)	.001	(.003)	.000	(.001)	.003	(.009)
14	.003	(.008)	.001	(.003)	.000	(.001)	.004	(.012)
TOTAL (f_m)	.704	(.725)	.277	(.244)	.018	(.037)	.999	(1.006)

$R^2(f_{im}) = 0.47$ $\qquad$ $R^2(p_i) = 0.29$

$R^2(d_{ik}) = 0.31$

Retail Sales, r_g = (0.244, 0.108, 0.216, 0.137, 0.295)

Table 15: Model 4, Centres at (5,4,7,1,6); predictions by place of residence.

Work Zone	Number of Shopping Trips made by Household (e_{km}) 0	1	2	Total (d_k^{mod})
1	.051	.021	.001	.073
2	.145	.064	.004	.213
3	.007	.004	.000	.012
4	.053	.024	.002	.078
5	.073	.029	.002	.104
6	.073	.027	.002	.102
7	.065	.025	.002	.091
8	.002	.001	.000	.004
9	.001	.000	.000	.001
10	.001	.000	.000	.002
11	.000	.000	.000	.001
12	.000	.000	.000	.001
13	.000	.000	.000	.000
14	.000	.000	.000	.000
15	.231	.083	.005	.319
TOTAL (e_m)	.702	.278	.018	.998

Table 16: Model 4, Centres at (5,4,7,1,6); predictions by place of work.

employment distribution (Table A2). That is, it allocates most of the employment to the lower city. However the model overpredicts the shares of 4, 5 and 7 and underpredicts the share of 6. This indicates that the model is still not able to predict correctly the high rate of activity in zone 6.

4.6 Model 5: Land use accounting

(1) Introduction

The models developed so far have had no explicit accounting mechanism for the amount of land used in each zone. An implication of this omission is that the retail centres take up no space and therefore residential land can be allocated to a retail centre. As a partial solution to this problem the amount of land used by various categories is used as a control on the allocation procedure. The details of this method follow. Using the data in Table A1 the total amount of land used by residential, vacant, service and office activity (L_i) in each zone is obtained. Next these figures are scaled by the total amount of residential and vacant land. This produces a vector ($L_i/(R+V)$) which obviously sums to a figure greater than one. From this vector residential priors are calculated in the following way

$$\tilde{l}_i = \begin{cases} \dfrac{L_i}{R+V} & \text{if } i \text{ has no centre,} \\[2ex] \dfrac{L_i}{R+V} - x\dfrac{S}{N} & \text{if } i \text{ has } x \text{ centres.} \end{cases}$$

where $\tilde{l}_i$ = the new residential prior,

L_i = sum of residential, vacant, service and office land use in i,

R = total residential land,

V = total vacant land,

S = total service and office land,

N = number of centres ($\leq$ 5).

The effect of this method is to reduce the prior probability for residential land use as soon as a centre is allocated to the zone. This makes it difficult for the model to allocate residential land to a zone with facilities and therefore builds in a type of land use competition.

(2) Results: centres at (5,4,5,6,4)

The entropy maximising distribution of service centres is found to be (5,4,5,6,4). Since the model is no longer being fitted directly to observed data not much significance can be attached to the correlations recorded in Table 17. The pattern of retail sales is however quite close to the actual downtown configuration, although there seems to be an overprediction for zone 5. The predicted flows to the various centres are shown in Figure 7.

This completes the development of the model. The next subsection summarizes the model in its most complete form and notes the major results.

4.7 Summary of model

The model has been developed through five versions and it is necessary to gather together the entire structure. The results produced by the model are also summarized briefly.

Residence Zone	Number of Shopping Trips made by Households (f_{im}) 0		1		2		Total (p_i)	
1	.045	(.017)	.018	(.003)	.002	(.000)	.064	(.020)
2	.112	(.020)	.039	(.008)	.003	(.001)	.155	(.029)
3	.036	(.059)	.015	(.026)	.001	(.005)	.052	(.090)
4	.197	(.107)	.070	(.035)	.006	(.005)	.272	(.147)
5	.184	(.124)	.061	(.038)	.005	(.004)	.250	(.165)
6	.043	(.094)	.020	(.013)	.002	(.000)	.065	(.107)
7	.030	(.038)	.017	(.013)	.001	(.003)	.049	(.053)
8	.011	(.061)	.007	(.016)	.001	(.003)	.019	(.079)
9	.013	(.044)	.007	(.025)	.001	(.000)	.021	(.069)
10	.015	(.048)	.006	(.042)	.001	(.007)	.022	(.146)
11	.007	(.046)	.004	(.018)	.000	(.007)	.011	(.070)
12	.004	(.004)	.002	(.001)	.000	(.000)	.005	(.005)
13	.006	(.005)	.003	(.003)	.000	(.001)	.009	(.009)
14	.005	(.008)	.003	(.003)	.000	(.001)	.007	(.012)
TOTAL (f_m)	.708	(.725)	.272	(.244)	.023	(.037)	1.003	(1.006)

$R^2(f_{im}) = 0.48$ $\qquad$ $R^2(p_i) = 0.34$

$R^2(d_{ik}) = 0.37$

Retail sales, r_g = (0.273, 0.194, 0.273, 0.067, 0.194)

Table 17: Model 5, Centres at (5,4,5,6,4); predictions by place of residence.

Work Zone	Number of Shopping Trips made by Household (e_{km}) 0	1	2	Total (d_k^{mod})
1	.003	.001	.000	.004
2	.153	.056	.005	.213
3	.008	.003	.000	.012
4	.105	.039	.003	.148
5	.122	.047	.004	.173
6	.067	.032	.003	.102
7	.014	.007	.001	.021
8	.002	.001	.000	.004
9	.001	.000	.000	.001
10	.001	.000	.000	.002
11	.000	.000	.000	.001
12	.000	.000	.000	.001
13	.000	.000	.000	.000
14	.000	.000	.000	.000
15	.231	.081	.007	.319
TOTAL (e_m)	.707	.267	.023	.997

Table 18: Model 5, Centres at (5,4,5,6,4); predictions by place of work.

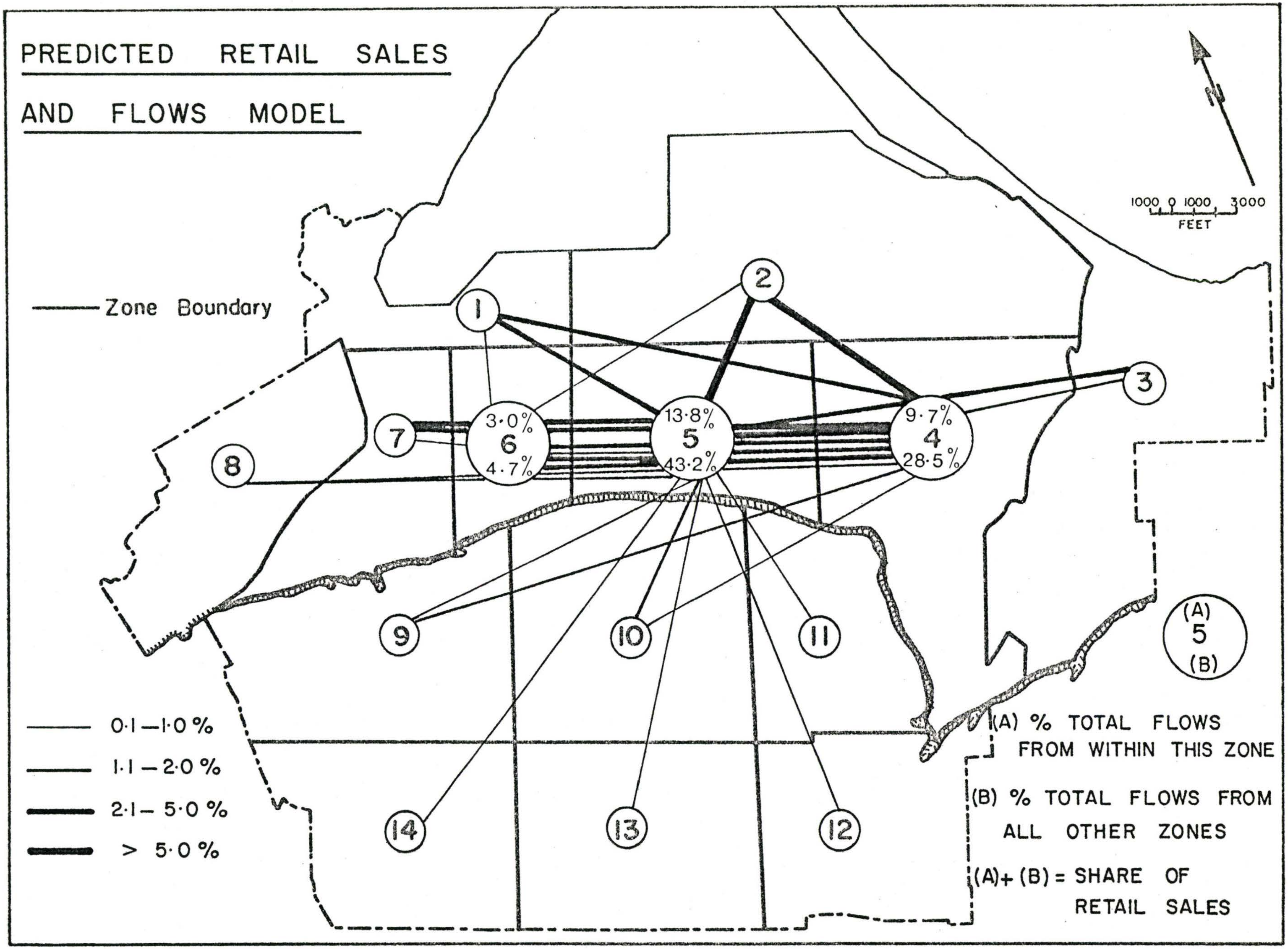

Figure 7 Model 5, Centres at (5,4,5,6,4); Retail Sales and Flows to Retail Centres

(1) The model

Maximise

$$-\sum_i \sum_j \sum_k p_{ijk} \log \frac{p_{ijk}}{\tilde{l}_i s_j}$$

Subject to

$$\sum_i \sum_j p_{ijk} = d_k^{mod} \qquad (k = 1, .., 15).$$

$$\sum_i \sum_j \sum_k p_{ijk} \tilde{c}_{ij} = \tilde{c}_s ,$$

$$\sum_i \sum_j \sum_k p_{ijk} \check{c}_{ik} = \bar{c}_w ,$$

$$\sum_i \sum_j \sum_k p_{ijk} x_j = \bar{x} .$$

The terms in these equations are now defined.

Residential prior

$$\tilde{l}_i = \begin{cases} \dfrac{L_i}{R+V} & \text{if } i \text{ has no centre,} \\[2ex] \dfrac{L_i}{R+V} - x\dfrac{S}{N} & \text{if } i \text{ has } x \text{ centres,} \end{cases}$$

where $\tilde{l}_i$ = a prior probability for residential zones,

L_i = sum of residential, vacant, service and office land use in i,

R = total residential land use,

V = total vacant land use,

S = total service and office land use,

N = number of service centres ($\leq$ 5).

Retail agglomeration prior

$$s_j = \begin{cases} \alpha & j = 1 \\ \alpha \, e^{-\gamma t_i} & j = 2, \ldots, N+1 \\ \alpha^2 \, e^{-\gamma(t_n + t_m)} & j = N+2, \ldots, J \end{cases}$$

where α = probability that a household makes no shopping trip, from data α = .6853.

γ = parameter to ensure $\sum_{j=1}^{J} s_j = 1$.

t_i = mean distance from a service centre at i to all other service centres. (Varies with distribution of service centres.)

Work trip constraint

$$d_k^{mod} = \begin{cases} \frac{b_k}{E}(1 - U) & k = 1, \ldots, 14 \text{ no service centre,} \\ \frac{b_k}{E} + x\frac{E - B}{E} \quad (1 - U) & k = 1, \ldots, 14 \text{ if } k \text{ has } x \text{ centres,} \\ U & k = 15. \end{cases}$$

where b_k = manufacturing and wholesale employment zone k, E = city total employment,

U = share of 'no work trip' in trip ends,

B = city total basic employment.

(Data in Appendix A, Table A2.)

Shopping trip costs

$\bar{c}_s$ is the average shopping trip time.

$$\tilde{c}_{ij} = \begin{cases} c_{ign}[\frac{1}{6}\sum_{l=0}^{\infty} l\, P(l|0)] , & j = 1; \\ c_{ij(g)}[\frac{1}{6}\sum_{l=0}^{\infty} l\, P(l|1)] , & j = 2, \ldots, N+1; \\ (\sum_{g\varepsilon j} c_{ij(g)})[\frac{1}{6}\sum_{l=0}^{\infty} l\, P(l|2)], & j = N+2, \ldots, J, \end{cases}$$

where $\tilde{c}_{ij}$ is the cost matrix for shop trips,

c_{ign} is the observed time taken to travel from i to the nearest centre g,

$c_{ij(g)}$ is the observed time taken to travel from i to the centre g associated with pattern j,

$g\varepsilon j$ signifies that the sum is over all those centres associated with pattern j.

$P(l|m)$ m = 0,1,2, is described in Appendix C.

Work trip costs

$$\overset{v}{c}_{ik} = \begin{cases} \text{observed travel time, } c_{ik} & (k = 1, \ldots, 14) \\ \sum_{k=1}^{14} (d_k^{mod}\, c_{ik}) / \sum_{k=1}^{14} d_k^{mod} & (k = 15) \end{cases}$$

$\bar{c}_w$ = observed average work trip length

Number of shopping trips

x_j = the number of shopping trips associated with trip pattern j ($j = 1, 2, \ldots, J$)

$\bar{x}$ = mean number of shop trips per day per household

The terms in the combined model have now been defined. The details of these terms and their significance are discussed above in the main body of this chapter. It remains to summarize the results. Detailed results have already been given for each of the versions of the model. Of these Model 3 clearly provides the most encouraging results in terms of correlations. The map which illustrates the results for model 3 (Fig. 6) summarizes our best effort at fitting the observed data (see Fig. 8). This crude representation of the service centres is a fair approximation to the existing configuration.

Models 4 and 5 adopt a new approach which attempts to model more accurately real world processes of employment allocation and land use competition. This is achieved by making the distribution of employment d_k

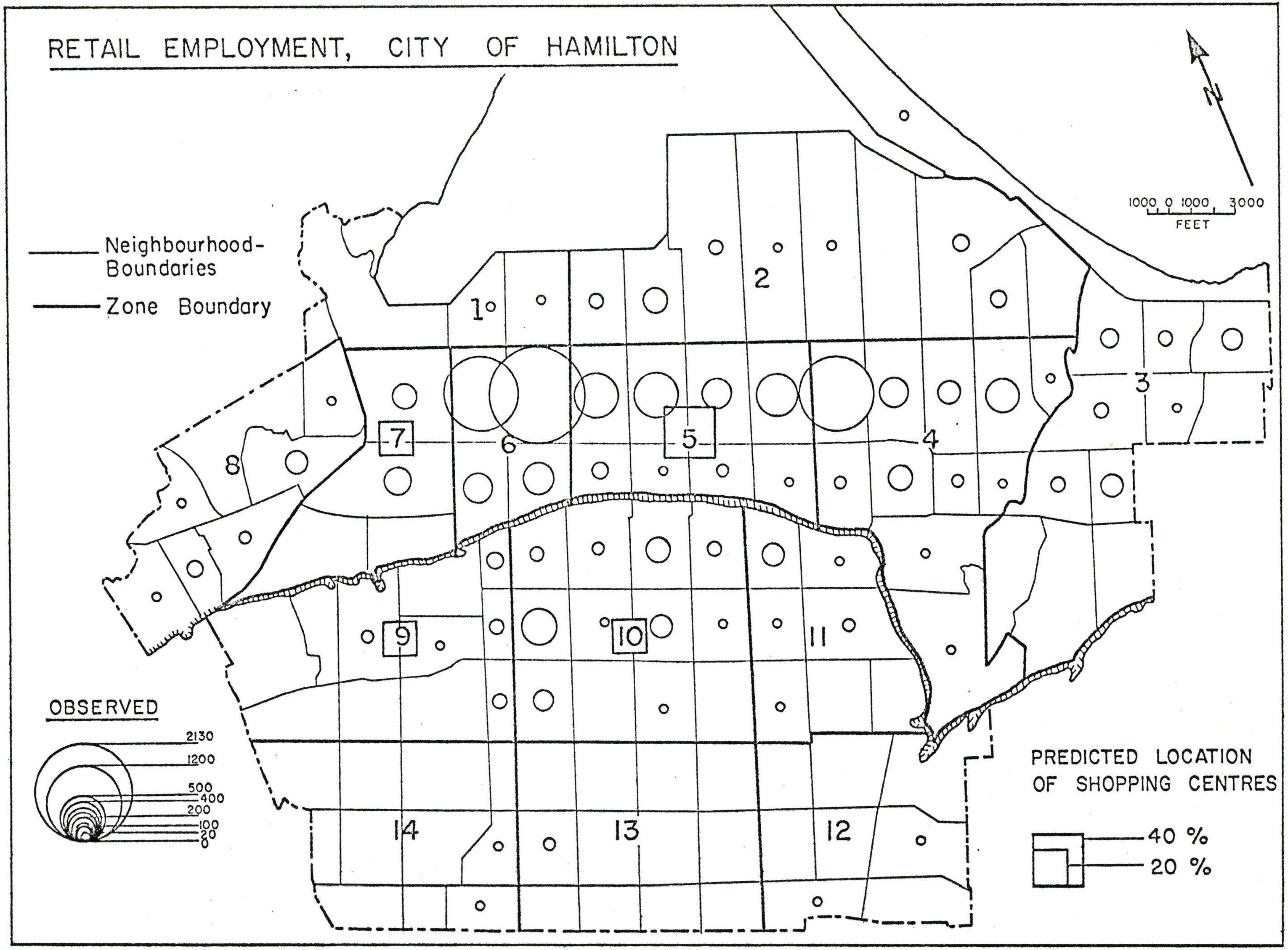

Figure 8 Observed distribution of retail employment

partially endogenous (model 4) and by incorporating a land use account (model 5). The results in these last two models are less encouraging in terms of correlations. It is felt however that models 4 and 5 could also achieve high correlations (as did model 3) using a more integrated data base.

The conclusion reached on the basis of these models is that the pattern of centres which maximises S* is indeed a centralized one and therefore matches observed retail locations. Furthermore the results associated with the entropy maximising assignment are in the case of model 3 highly correlated with other aspects of the city. This shows therefore that by pursuing a theoretical question using an entropy model we have also produced a realistic representation of the city.

CHAPTER 5

SUMMARY AND CONCLUSIONS

5.1 Summary

This paper began with a demonstration of the weaknesses of early land use models developed in North America between 1955 and the mid 1960's. The criticism by D. Lee (1973) pointed clearly to the failure of these efforts and forced urban analysts to face the problem of designing consistent, theoretically based, parsimonious and realistic models. The main focus of this paper is to attempt to show that entropy maximising models satisfy the above criteria. The research proceeded in three stages.

Chapter 2 discussed a methodology for building consistent parsimonious and realistic models; that is, the entropy maximising formalism. An extension of this formalism based on Kullbacks measure was described and this extended formalism was used to generalize Wilson's family of spatial interaction models. Applications of spatial interaction modelling were then reviewed to gather guidelines for empirical model building. Thus Chapter 2 set up the general methodology and reviewed the necessary techniques to apply this method.

Chapter 3 then described a specific model which is not only related to the family of spatial interaction models, but also has a theoretical basis. The model searches for the entropy maximising pattern of retail centres, which is expected to include downtown centres. The model is sufficiently detailed to allow inferences to be made about other aspects of the city. These other endogenous distributions include residential location, work trips and shopping trips. The realism of entropy maximising solutions was then assessed in Chapter 4 by comparing the fit of these predicted distributions to observed patterns.

The following models were tested:

Model 1, spatial interaction - retail location

This model was run with two centres. It produced poor results because of the unrealistic calculation of costs associated with the shopping patterns and with the no work-trip category.

Model 2(a), modified cost calculations

This model produced improved correlations. It was run for both 2 and 3 centre distributions.

Model 2(b), modified residential prior

The residential prior was recalculated to include only residential and vacant acreage. Correlations improved. At this stage an algorithm was developed to handle higher numbers of service centres. Tests on this model with 5 centres failed to produce a downtown core.

Model 3, retail sector agglomeration

A prior probability distribution was included to account for agglomeration economies in the retail sector. This distribution was based on the proximity of shopping centres to each other. Non-zero costs were also assigned to the 'no work trip' category. The model was run with 5 centres and produced substantial improvements over the previous models. The predicted pattern of service centres was (5,5,9,7,10). This pattern places some service activity in the area south of the city centre (the so called "Mountain").

Model 4(a), new employment constraint

The work trip destination vector was recalculated using data from HWPDD to provide a more accurate representation of the distribution of basic employment. The correlations disimproved over model 3 as the model was being constrained to fit one set of data, and was tested against another set (i.e. the original survey). The maximum entropy distribution of service centres was (5,4,7,1,10) which is more dispersed than the distribution for model 3.

Model 4(b), partially endogenous employment

Retail centres obviously employ people, so it is necessary to include these employers in the work trip constraint. In this version of the model the work trip constraint is adjusted to include a share of the service employment in the city (from HWPDD, 1977). The work trip constraint becomes partly endogenous. The results for this model again showed poor correlations, however the pattern of retail centres associated with the maximum was (5,4,7,1,6) which clearly covers the downtown area of the city.

Model 5, land use accounting

Retail centres consume space, so it is necessary to take this use of space into account in the calculation of the residential prior. Up to this point the residential prior has been based on developable land - now this amount of developable land is decremented in accordance with the pattern of retail centres. The correlations in this model improved over those in model 4. The maximum entropy distribution of service centres was (5,4,5,6,4) which covers the actual downtown in Hamilton, but misses the observed service activity on the mountain.

The best results in terms of correlations were obtained in model 3. The revisions to the model made in versions 4 and 5 attempt to represent more realistically the real world processes of employment allocation and land use competition. The results of these models are less encouraging in terms of correlations as some of the data used as an exogenous input to the model have themselves become the object of prediction.

5.2 Conclusions

The aim of this research paper was to show that entropy maximising models are better than first generation models of urban land use.

The conclusions reached are that the entropy maximising formalism is a powerful methodology for developing models which are better than other models. The strength of these models is apparent on a number of grounds.

Firstly from the point of view of the applied model builder, the maximum entropy (minimum information) formalism is a powerful tool. It can be shown that the formalism provides a means of consistently deriving existing land use models. For example, the Lowry model discussed in Chapter 1 has been improved and derived consistently using entropy maximising method by Wilson (1974) and Batty (1976a). This alone would seem to ensure the use of the formalism in practical planning situations, and indeed the references cited in Chapter 2 attest to the widespread acceptance of this method.

A second, and possibly stronger conclusion, is with regard to the link between theoretical models and applied urban models. Entropy maximising methods have been used by Webber (1977b; 1977c) to examine the structure of the city. It has been shown that many of the results derived by economists in the neoclassical tradition can be found using the formalism. Thus the Alonso (1964) - Herbert-Stevens (1960) tradition is faced by a powerful alternative methodology. The empirical model in Chapter 3 is a simplified version of the theoretical model in Webber (1977b). Thus the *theoretical* model provides the basis for analysis in a specific real world situation. This is an advantage which gains in significance when the difficulty of constructing operational models in the neoclassical tradition is recalled. Furthermore the model developed in this paper is only a simplified version of the full model - the scope for further development can quite legitimately be said to be immense. Thus the second conclusion is that the entropy maximising formalism allows the modeller to generate theoretical models of the real world which are readily applied to real world data.

This conclusion can of course be viewed in the opposite direction. Thus the empirical model generated and tested in this paper (and other models of spatial interaction) gain support from the development of a theoretical background. This ability to trace the models to a theoretical foundation is a feature that was absent from many of the ad hoc models reviewed in the first chapter.

A third conclusion concerns the flexibility of the model in an applied context. The great advantage of the formalism is that a failure on the part of an entropy maximising model to represent real world data will always provide suggestions for improved realism. Care must however be exercised to keep the mechanism of the model as parsimonious as possible. Otherwise there is a temptation to continue adding constraints (i.e. information) up to the point where the model is almost fully constrained. At this stage, average uncertainty approaches zero and the model becomes deterministic (or tautological). It is felt that the models developed in Chapter 4 avoided this sin, and maintained parsimony (as suggested by Batty, 1976a).

The conclusions reached from the model itself are that the maximum entropy formalism allows quite complex models of some aspects of the urban geography of Hamilton to be developed. Clearly with better data, more retail facilities and more realistic representation of multi-purpose trips the model could be greatly improved. Nevertheless, the models in Chapter 4 (summarized above, 5.1) do manage to reproduce some of the features of retail location and consumer behaviour in Hamilton.

APPENDIX A

DATA SOURCES

City of Hamilton Planning Department 1972, "The Year 2001, Trends Forecast?", City of Hamilton, December 1972.

City of Hamilton Planning Department 1973, "1971 Employment Survey Statistics", City of Hamilton, April 1973.

City of Hamilton Planning Department 1977, Unpublished data, Land Use Summary Sheets for Planning Divisions, 61-76.

Hamilton-Wentworth Regional Planning and Development Department, 1975, *Transportation: A Substudy of the Regional Official Plan*, Report Number 2, HWPDD, Hamilton, October 1975.

Hamilton-Wentworth Planning the Development Department, 1977a, "Selecting a preferred Regional Development Pattern", HWPDD, Hamilton, February 1977.

Hamilton-Wentworth Planning and Development Department, 1977b, Unpublished Data, "Land Use Data: Allocation to Traffic Zones", HWPDD, Hamilton, October, 1977.

Homburger, W.S., 1972, Traffic Estimation Computer Programs for Educational Purposes, 4th ed., Institute of Transport and Traffic Engineering, University of California.

Zone	Residential Land Use	Vacant Land Use	Unusable Land	Office Land Use	Service Land Use	Basic Land Use	Other Land Use	Total
1	113.0	87.2	351.0	1.9	3.9	6.9	121.5	685.4
2	187.5	214.9	3045.2	42.1	63.8	2299.9	927.0	6780.4
3	977.1	771.6	714.4	5.5	146.4	124.6	676.0	3415.6
4	1150.5	47.9	861.7	32.8	146.3	58.7	316.7	2614.6
5	734.7	11.6	103.2	42.5	84.7	77.1	254.4	1308.2
6	295.9	250.1	8.1	43.4	67.9	44.4	240.9	950.7
7	370.2	44.1	437.8	2.9	31.8	90.5	329.4	1306.7
8	523.5	344.5	396.8	9.1	29.4	28.8	325.2	1657.3
9	1034.3	314.7	55.0	3.5	50.3	0.8	384.8	1843.4
10	1435.6	471.4	98.5	7.5	103.0	13.5	221.5	2351.0
11	735.6	185.4	22.1	1.0	32.3	0.0	76.7	1053.1
12	359.9	788.8	275.9	0.5	17.8	17.4	346.8	1807.1
13	582.1	600.6	53.6	16.4	34.3	8.1	499.5	1794.6
14	634.3	588.2	85.1	1.1	33.7	0.0	444.6	1787.0
	9134.2	4721.0	6508.4	210.2	845.6	2770.7	5165.0	29355.1

Table A1: Land Use Data, by Planning District.

Source: City of Hamilton Planning Department, Unpublished Data, 1977.

Note:

Residential	Property Code 01-19	Service	Property Code 81-89
Vacant	Property Code 28	Basic	Property Code 71-73
Unuseable	Property Code 21-27	Other	Property Code 31-37, 41-49, 51-59, 76-79, 91-99.
Office	Property Code 61-66		

Where this code refers to the Planning Department's Classification.

Zone	Population	B‡ Manufacturing and Wholesale Employment	Retail Employment	Service Employment	Other Employment	Total Employment
1	7355	699	89	1087	16	1891
2	11325	42274	659	1933	4	44870
3	17730	2313	1022	1226	0	4561
4	49085	1728	4355	3623	3	9709
5	55755	6755	2314	9245	0	18314
6	27310	6429	5466	18237	1	30133
7	20485	4230	557	1376	0	6163
8	18160	721	561	4533	0	5815
9	22205	188	503	3880	0	4571
10	51290	334	1391	5120	0	6845
11	21975	106	514	822	0	1442
12	840	133	40	130	0	303
13	2685	88	81	206	9	384
14	2610	52	66	101	0	219
CITY TOTAL	308810	66050	17618	51519	33	135220

Table A2: Population and Employment Data, by Planning District.

Source: Aggregated from "Land Use Data: Allocation of Traffic Zones", Unpublished data provided by the Planning and Development Department of the Regional Municipality of Hamilton-Wentworth.

‡ abbreviations used in Chapter 4.

APPENDIX B

TABLE B1

Trip Characteristics of Sample Households After First Edit

	NUMBER OF WORK TRIPS (X)							
	0	1	2	3	4	5	6	
No. of Households making X trips	344	606	288	61	12	7	3	Total No. of Households 1321

	NUMBER OF SHOPPING TRIPS (Y)					
	0	1	2	3	4	
No. of Households making Y trips	940	282	82	13	4	Total No. of Households 1321

	MEAN PER HOUSEHOLD	STANDARD DEVIATION
Work Trips	1.11	0.83
Shop Trips	0.38	0.94

Source: see text.

TABLE B2

Trip Characteristics of Sample Households After Second Edit

Number of Households Classified by Work Trip and Shop Trip Frequency

Shop Trips	0	1	2	TOTAL
Work Trips				
0	228	95	21	344
1	712	187	78	977
TOTAL	940	282	99	1,321

	Mean per Household	Standard Deviation
Work trips	0.74	0.51
Shop trips	0.36	1.03

APPENDIX C

CALCULATION OF EXPECTED NUMBER OF SHOPPING TRIPS

If there are N shopping facilities to be located in the city (N = 5) then the possible trip patterns of a household are:

j = 1: no shopping trips were made by that household;

j = 2, 3, ..., N+1: one shopping trip is made, to any one of the N facilities in the city;

j = N+2, ..., J: the household visits two different shopping facilities. Evidently this list assumes that no household visits the same facility twice in one day.

Now the survey data record a single day's travel for each household. This data collection method poses serious analytical problems, for some 70% of sampled households made no shopping trip on the day on which they were interviewed, which clearly does not mean that 70% of Hamilton households make an average of zero shopping trips per day over a week or a month. Presumably, location decisions of households reflect average long run shopping trip frequencies rather than the vagaries of trip making behaviour on an individual day, and so a means must be found of estimating average shopping trip frequencies from the data.

The results of Table B1 describe the mean and variance of shopping trip frequencies for the sample households. Following Tribus (1969) and Webber (1976), the entropy maximising probability distribution with given mean and variance is a discrete approximation to a normal distribution, which may be interpreted as describing the probability, computed over the entire sample of households, that a household makes 0, 1, 2, ... shopping trips per day. If households choose a number of trips on one day which is independently (of previous days' trips) drawn from this normal distribution, the probability of a household making 0, 1, 2, ... trips per week (of six days) can be calculated. Call this distribution $\underline{\underline{G}} = (g_0, g_1, g_2, \ldots)$: denote by $P(l|m)$ the conditional probability that a household makes l shopping trips in a week if it made m trips on the day of the interview; and let $\underline{\underline{F}} = (f_0, f_1, f_2)$ be the observed proportion of households making 0, 1 or 2 trips on the day of the interview. Then

$$P(l|2) = \begin{cases} 0 & \text{if } l < 2 \\ (f_2 \times g_l)/(\sum_{l=2}^{\infty} g_l) & \text{if } l \geq 2; \end{cases}$$

is an estimate of the probability that a household makes l trips in a week if it made two trips on the day of the interview. If

$$\underline{\underline{G}}^1 = (g_0^1, g_1^1, g_2^1, \ldots) = (g_0, g_1, g_2 - P(2|2), g_3 - P(3|2), \ldots),$$

then

$$P(l|1) = \begin{cases} 0 & \text{if } l < 1 \\ (f_1 \times g_l^1)/(\sum_{l=1}^{\infty} g_l^1) & \text{if } l \geq 1; \end{cases}$$

a similar estimate of $P(l|0)$ is

$$P(l|0) \;=\; g_l - P(l|1) - P(l|2) \;.$$

If a household makes m shopping trips on the day it is interviewed, the expected number of shopping trips made per day in that week is therefore

$$\frac{1}{6}\sum_{l=0}^{\infty} l \times P(l|m), \qquad\qquad m = 0,\, 1,\, 2 \;.$$

TABLE C1

The Effect of Chanaging the Calculation of

Cost Matrix for Centres at (5,7)

Residence	TRIP PATTERN [travel times in minutes] 1	2	3	4
1	0.00	11.90	16.07	27.97
2	0.00	11.07	24.19	35.26
3	0.00	18.20	28.71	46.91
4	0.00	7.00	20.12	27.12
5	0.00	9.62	13.12	22.74
6	0.00	9.25	3.87	13.12
7	0.00	13.12	8.33	21.45
8	0.00	24.04	10.92	34.96
9	0.00	20.08	15.03	35.11
10	0.00	22.93	19.97	42.90
11	0.00	27.73	27.48	55.21
12	0.00	33.99	32.28	66.27
13	0.00	25.32	30.48	55.80
14	0.00	25.29	39.85	65.14

Residence	TRIP PATTERN [travel times in minutes] 1	2	3	4
1	5.91	7.20	9.72	9.58
2	5.50	6.70	14.63	12.08
3	9.05	11.01	17.37	16.07
4	3.48	4.24	12.17	9.29
5	2.58	3.15	7.94	6.27
6	1.92	5.60	2.34	4.49
7	1.90	7.94	2.31	5.80
8	5.43	14.54	6.61	11.97
9	7.47	12.15	9.09	12.03
10	9.93	13.87	12.08	14.69
11	13.66	16.78	16.63	18.91
12	16.04	20.56	19.53	22.70
13	12.58	15.32	18.44	19.11
14	12.57	15.30	24.11	22.31

(This change effective from Model 2 onwards.)

APPENDIX D

Computer Program (Model 5)

```
      PROGRAM CALBTN(INPUT,OUTPUT,TAPE7,TAPE8,TAPE5=INPUT,TAPE6=OUTPUT)
C
C     THIS IS A CALIBRATION PROGRAM TO SOLVE FOR THE PARAMETERS
C     OF A LOCATION MODEL USING THE NEWTON-RAPHSON OPTIMISATION
C     METHOD.  THE NON-LINEAR EQUATIONS FOR THIS PROGRAM MUST BE
C     CONSISTENTLY DERIVED FROM THE MAXIMUM LIKELIHOOD METHOD.
C
C     NOTE .. THIS VERSION OF THE MODEL IS A THREE PARAMETER FUNCTION
C     CONSTRAINED TO MEAN TRIP LENGTH FOR WORKTRIPS, MEAN TRIP
C     LENGTH FOR SHOPPING TRIPS, AND THE MEAN NUMBER OF SHOPPING
C     TRIPS PER HOUSEHOLD
C
C     THIS PROGRAM WAS WRITTEN BY PETER HALL.
C     PROGRAM DESCRIPTION .. MODIFIED EMPLOYMENT AND RESIDENTIAL PRIOR
C     CALCULATION BASED ON PLANNING DEPT DATA
C     VERSION 5  DECEMBER 14 1977        MORTON O"KELLY
C
      DIMENSION SOBS(5),ERROR(5),SPRED(5),PRED(5),PARA(5),F(5),DERIV(10,
     110),STAT(5),ISHOP(10),H(50),BARA(5),HORSE(4),IDUCK(4),IN1(4,10),RS
     2Q(50,4),COW(4),TIT(4),AB(10),SAL(50,10),LCON(50,10),IZONE1(10),WGH
     3(3)
      DIMENSION DIRTY(9),LINEN(6),YOUR(6),ZNUM(5)
      DIMENSION D(15),O(15),O2(15),INDIC(15)
      COMMON/INDEX/IND(20)
      COMMON/COST/DIS(15,15),CPAT(15,20),X(20),ORD(51),Q(20)
      COMMON/MODE/D2(15),O3(15),A(15),B(15),ORIGIN(15),DENS(15)
      COMMON/PROB/PR(15,15,20)
      COMMON/TRIP/PIS(15,15)
      COMMON/CNTROL/INFO(10)
      COMMON/PAT/XPAT(2,15,3)
      COMMON/EXPNUM/WEIGHT(20)
      COMMON/BEER/IOPNT(10,10)
      LOGICAL CHECK
      INTEGER W
      REAL LINEN
      DATA H/50*0./
      DATA HORSE,RSQ,SAL/4*0.,200*0.,500*0./
      DATA TIT/8HWORKTRIP,8HLIVE/PAT,8HWORK/PAT,8HLOCATION/
      DATA INDIC/15*0/
      READ(5,97) IDX
C
C     CONTROL CARD 1
C     EACH FIELD (COLUMN) CONTAINS EITHER 0 (OR BLANK) OR 1.
C     (PRINT 1 IF YOU WANT THE OPTION)
C        COLUMN                  OPTION
C           1*                SUMMARY STATISTICS
C           2                 FREQUENCY DISTRIBUTION OF TRIP LENGTH
C           3                 SUMMARY MATRICES
C           4                 ENTROPY MEASURES
C           5                 INTERMEDIATE OUTPUT
C
      READ(5,99)(INFO(I),I=1,10)
      DO 41 I=1,10
      INFO(I)=INFO(I)+1
   41 CONTINUE
      W=INFO(5)+5
C
C     CONTROL CARD 2
C     STARTING AT COLUMN 1, WRITE TRUE IF PRINTOUT OF COMPOSITE
C     COST MATRIX NOT WANTED, FALSE OTHERWISE.
C
      READ(5,98) CHECK
C
C     INPUT..NO. OF ORIGINS = N
C            NO. OF DESTINATIONS = M
```

```
C              NO. OF PARAMETERS = K
C              NO. OF ITERATIONS = LIMIT
C              NO. OF POSSIBLE PATTERNS = NPAT
C              NUMBER OF SHOPS TO BE LOCATED = NSHOP
C
      READ(5,100) N,M,K,LIMIT,NPAT,NSHOP
C
      IF(NSHOP .LE. 2) GO TO 48
      NTIMES=6*(NSHOP-2)+2
      GO TO 49
   48 CONTINUE
      NTIMES=2
   49 CONTINUE
      ENCODE(52,138,YOUR) NTIMES
      ENCODE(88,135,DIRTY) NTIMES
      ENCODE(52,136,LINEN) NSHOP,NSHOP
      ENCODE(43,132,ZNUM) NSHOP
C     READ IN EXPECTED NUMBER OF TRIPS PER DAY FOR EACH TRIP
C     FREQUENCY OBSERVED
C
      READ(5,140)(WGH(I),I=1,3)
      CALL WTDIST(NSHOP,WGH,NPAT)
C
C     INPUT..OBSERVED STATISTICS (SOBS) - FORMAT - 10F6.0
C
      READ(5,101)(SOBS(I),I=1,K)
C
C     INPUT..STARTING VALUES FOR PARAMETERS (BARA) - FORMAT 10F6.0
C
      READ(5,101)(BARA(I),I=1,K)
C
C     INPUT TRAVEL COST MATRIX (DIS) - THIS IS A 2-DIMENSIONAL
C     ARRAY USED TO CALCULATE A COMPOSITE COST MATRIX
C
      MAX=MAX0(M,N)
      NN=N-1
      DO 1 I=1,NN
      LT=I+1
      READ(5,103)(DIS(I,J),J=LT,N)
    1 CONTINUE
      DO 2 I=2,N
      L=I-1
      DO 3 J=1,L
      DIS(I,J)=DIS(J,I)
    3 CONTINUE
    2 CONTINUE
      READ(5,126)(DIS(I,I),I=1,N)
      WRITE(W,105)
      LAX=MAX0(N,M,NPAT)
      DO 25 I=1,LAX
      IND(I)=I
   25 CONTINUE
      X(1)=0.
      DO 62 I=1,NSHOP
      J=I+1
      X(J)=1.
   62 CONTINUE
      J=NSHOP+2
      DO 63 I=J,NPAT
      X(I)=2.
   63 CONTINUE
      READ(5,104)(O(I),I=1,N)
      READ (5,104) (O2(I),I=1,N)
      READ (5,1041)(D(I),I=1,M)
C
```

```
C     INPUT OBSERVED TRIP ORIGIN VECTOR ORIGIN(I) - FORMAT 14F5.0
C
      READ(5,122) (ORIGIN(I),I=1,N)
      TOTAL=0.
      DO 20 I=1,N
      TOTAL=TOTAL+ORIGIN(I)
 20   CONTINUE
      DO 8 I=1,N
      ORIGIN(I)=ORIGIN(I)/TOTAL
      DENS(I)=ORIGIN(I)/O(I)
 8    CONTINUE
      WRITE(W,113)
      WRITE(W,115)
      DO 61 I=1,N
      WRITE(W,114) I,C(I),ORIGIN(I),DENS(I)
 61   CONTINUE
C
      WRITE (W,116) TOTAL
C     READ IN WORKTRIP MATRIX AND PRINT IT OUT - MATRIX IS CALLED
C     PIS - FORMAT 15F5.0
C
      DO 9 I=1,N
      READ(5,120)(PIS(I,J),J=1,M)
      DO 9 J=1,M
      PIS(I,J)=PIS(I,J)/TOTAL
 9    CONTINUE
      WRITE(W,105)
      WRITE(W,117)
      WRITE(W,139)
      WRITE(W,127)
      WRITE(W,118)(IND(I),I=1,M)
      DO 16 I=1,N
      IF (I .EQ. 7) GO TO 17
      WRITE(W,119) I,(PIS(I,J),J=1,M)
      GO TO 16
 17   CONTINUE
      WRITE(W,121) I,(PIS(I,J),J=1,M)
 16   CONTINUE
C
C     READ IN OBSERVED TRIP PATTERN MATRICES FOR LIVE IN I, SHOP IN J,
C     AND FOR WORK IN K, SHOP IN J
C
      IJKL=14
      DO 18 I=1,2
      DO 19 J=1,IJKL
      READ(5,123)(XPAT(I,J,L),L=1,3)
      DO 24 L=1,3
      XPAT(I,J,L)=XPAT(I,J,L)/TOTAL
 24   CONTINUE
 19   CONTINUE
      IJKL=IJKL+1
 18   CONTINUE
      READ(5,130)(ORD(I),I=1,51)
C
C     INPUT THE LOCATIONS OF SHOPPING CENTRES.
C
      HSTART=0.
      KOUNT=0
      DO 400 IDIX=1,IDX
      KOUNT=KOUNT+1
      READ(5,142)(ISHOP(I),I=1,NSHOP)
C
C     CALCULATE A NEW EMPLOYMENT VECTOR USING THE LOCATIONS
C     OF THE SERVICE CENTRES TO INCREMENT THE BASIC EMPLOYMENT
C
```

```
      DO 600 I=1,M
      D2(I)=D(I)
600   CONTINUE
      DO 602 I=1,NSHOP
      NX=ISHOP(I)
      D2(NX)=D2(NX)+(.348/FLOAT(NSHOP))
      INDIC(NX)=INDIC(NX)+1
602   CONTINUE
C
C      COUNT THE NUMBER OF ZERO ELEMENTS
C
      NZER=0
      DO 603 I=1,N
      IF (INDIC(I).EQ.0) NZER=NZER+1
  603 CONTINUE
C
C     INCREMENT ZONES THAT ARE ZERO BY THEIR SHARE OF SERVICE LAND
C
      DO 605 I=1,N
      IF (INDIC(I).NE.0) GO TO 604
      O3(I)=O2(I)+(.069/FLOAT(NZER))
      GO TO 605
604   O3(I)=O2(I)
605   CONTINUE
      TAD=.319
      DO 81 I=1,N
      DIS(I,15)=0.
      DO 80 J=1,14
      DIS(I,15)=DIS(I,15)+DIS(I,J)*D2(J)
80    CONTINUE
      DIS(I,15)=DIS(I,15)/TAD
81    CONTINUE
      DO 23 I=1,K
      PARA(I)=BARA(I)
23    CONTINUE
      ERCK=999.
      WRITE(W,105)
      WRITE(W,125)
      WRITE(W,127)
      WRITE(W,128)(IND(I),I=1,M)
      DO 60 I=1,N
      WRITE(W,102) I,(DIS(I,J),J=1,M)
60    CONTINUE
      WRITE(W,105)
      WRITE(W,129) KOUNT
      WRITE(W,124)(ISHOP(I),I=1,NSHOP)
      CALL COSCAL(N,M,NPAT,ISHOP,CHECK,W,NSHOP)
C
C     CALCULATE THE PRIOR PROBABILITIES FOR THE POSSIBLE PATTERNS
C
      CALL PRIOR(NSHOP,ISHOP,W,NPAT)
      WRITE(W,1051)
      WRITE(W,1052) (O3(I),I=1,N)
      AINC=1./EXP(20.)
      KK=0
      WRITE(W,105)
      WRITE(W,106)
4     CONTINUE
      KK=KK+1
C
C     COMPUTE THE PREDICTED STATISTICS
C
      CALL MODEL(N,M,NPAT,K,PARA,SPRED,W)
      DO 5 I=1,K
      P(I)=PARA(I)
```

```
      STAT(I)=SOBS(I)-SPRED(I)
    5 CONTINUE
C
C     COMPUTE FIRST ORDER PARTIAL DERIVATIVES NUMERICALLY
C
      DO 6 I=1,K
      DO 6 J=1,K
      DO 7 L=1,K
      P(L)=PARA(L)
    7 CONTINUE
      P(J)=P(J)+AINC
      CALL MODEL(N,M,NPAT,K,P,PRED,W)
      DERIV(I,J)=(SPRED(I)-PRED(I))/AINC
    6 CONTINUE
C
C     INVERT MATRIX OF DERIVATIVES
C
      CALL MATINV(K,DET,DERIV)
C
C     COMPUTE INCREMENTS TO THE PARAMETERS CALLED ERROR(I)
C
      DO 10 I=1,K
      ERROR(I)=0.
      DO 11 J=1,K
      ERROR(I)=ERROR(I)+DERIV(I,J)*STAT(J)
   11 CONTINUE
   10 CONTINUE
C
C     WRITE OUT THE RESULTS OF ITERATION
C
      WRITE(W,107) KK
      DO 12 I=1,K
      WRITE(W,108) PARA(I),ERROR(I),STAT(I)
   12 CONTINUE
      DO 28 I=1,K
      IF (ABS(STAT(I)) .GT. .0000001) GO TO 29
   28 CONTINUE
      GO TO 14
   29 CONTINUE
      IF(KK.LT.10) GO TO 38
      SUMSQ=0.
      DO 39 I=1,K
      SUMSQ=SUMSQ+STAT(I)**2.
   39 CONTINUE
      IF (ABS(SUMSQ) .GT. ERCK) GO TO 37
      ERCK=SUMSQ
      GO TO 38
   37 CONTINUE
      DO 36 J=1,NSHOP
      LOCN(KOUNT,J)=ISHOP(J)
   36 CONTINUE
      GO TO 400
   38 CONTINUE
C
C     COMPUTE NEW PARAMETERS = OLD PARAMETERS + INCREMENTS
C
      DO 13 I=1,K
      PARA(I)=PARA(I)-ERROR(I)
   13 CONTINUE
      IF (KK .LT. LIMIT) GO TO 4
      WRITE(W,109) LIMIT
      GO TO 400
   14 CONTINUE
      WRITE(W,110)
      WRITE(W,111)
```

```
      DO 15 I=1,M
      WRITE(W,112) B(I)
 15   CONTINUE
C
C     SUMMARY MATRICES PREDICTED BY MODEL (OPTIONAL)
C
      IF (INFO(3) .LT. 2) GO TO 45
      CALL SUMMAR(N,M,K,NPAT,FRED,MAX,TOTAL,COW,ISHOP,W,NSHOP)
 45   CONTINUE
C
C     ENTROPY MEASURES (OPTIONAL)
C
      IF (INFO(4) .LT. 2) GO TO 43
      CALL ENTROP(N,M,NPAT,HMAX,W)
 43   CONTINUE
C
C     FREQUENCY DISTRIBUTION OF TRAVEL TIMES (OPTIONAL)
C
      IF (INFO(2) .LT. 2) GO TO 30
      CALL FREQ(N,M,NPAT,W)
 30   CONTINUE
      CALL PATFIX(NSHOP,ISHOP,W)
      CALL SALES(N,M,NPAT,AB,NSHOP)
      DO 33 I=1,NSHOP
      SAL(KOUNT,I)=AB(I)
 33   CONTINUE
      IF (HMAX .LE. HSTART) GO TO 399
      HSTART=HMAX
      ITER=KOUNT
      DO 47 I=1,NSHOP
      IZONE1(I)=ISHOP(I)
 47   CONTINUE
 399  CONTINUE
      H(KOUNT)=HMAX
      DO 26 I=1,4
      IF (COW(I) .LE. HORSE(I)) GO TO 27
      HORSE(I)=COW(I)
      IDUCK(I)=KOUNT
      DO 46 J=1,NSHOP
      IN1(I,J)=ISHOP(J)
 46   CONTINUE
 27   CONTINUE
      RSQ(KOUNT,I)=COW(I)
      DO 34 J=1,NSHOP
      LOCN(KOUNT,J)=ISHOP(J)
 34   CONTINUE
 26   CONTINUE
      I=KOUNT
      WRITE(8,LINEN)I,(LOCN(I,J),J=1,NSHOP),H(I),(RSQ(I,J),J=1,4),(SAL(I
     1,J),J=1,NSHOP)
 400  CONTINUE
      WRITE(6,105)
      WRITE(6,131) HSTART,ITER
      WRITE(6,ZNUM)(IZONE1(I),I=1,NSHOP)
      DO 32 I=1,4
      WRITE(6,117)
      WRITE(6,137) TIT(I),HORSE(I),IDUCK(I)
      WRITE(6,ZNUM)(IN1(I,J),J=1,NSHOP)
 32   CONTINUE
      WRITE(6,105)
      WRITE(6,134)
      WRITE(6,117)
      WRITE(6,YOUR)
      WRITE(6,DIRTY)(TIT(I),I=1,4)
      DO 21 I=1,KOUNT
```

```
      WRITE(6,LINEN)I,(LOCN(I,J),J=1,NSHOP),H(I),(RSQ(I,J),J=1,4),(SAL(I
     1,J),J=1,NSHOP)
   21 CONTINUE
      STOP
C
C     INPUT OUTPUT FORMAT STATEMENTS
C
   97  FORMAT(I3)
   98 FORMAT(L5)
   99 FORMAT(10I1)
  100 FORMAT(6I3)
  101 FORMAT(10F6.0)
  102 FORMAT(8X,I3,4X,15(F5.2,2X))
  103 FORMAT(13F5.2)
  104 FORMAT(15F5.0)
 1041 FORMAT (8F10.4)
  105 FORMAT("1")
 1051 FORMAT (//,5X,"CALCULATED RESIDENTIAL PRIOR",/)
 1052 FORMAT (5X,20F5.3)
  106 FORMAT(//,4X,*ITERATION*,3X,*PARAMETER ESTIMATE*,5X,* INCREMENT *,
     112X,*ERROR*,//)
  107 FORMAT(3X,I7)
  108 FORMAT(16X,F13.5,7X,F14.10,5X,F14.8)
  109 FORMAT(15X,* NOTE .. PARAMETER ESTIMATES DID NOT CONVERGE WITHIN *
     1,I5,* ITERATIONS*)
  110 FORMAT(//,5X,*BALANCING FACTORS *,/)
  111 FORMAT(12X,*B(J)*,/)
  112 FORMAT(5X,F13.5)
  113 FORMAT(//,33X,*PROPORTION OF RESIDENTIAL LAND AREA AND DENSITY IN
     1 EACH ZONE*,/)
  114 FORMAT(40X,I3,6X,F7.3,7X,F8.3,7X,F6.3)
  115 FORMAT(40X,*ZONE*,5X,*LAND AREA*,5X,*POPULATION*,5X,*DENSITY*,/)
  116 FORMAT(////,40X,*TOTAL NUMBER OF TRIPS IN SAMPLE = *,F7.0)
  117 FORMAT(//)
  118 FORMAT(15X,15(I5,2X),/)
  119 FORMAT(11X,I3,1X,15(F5.3,2X),/)
  120 FORMAT(15F5.0)
  121 FORMAT(4X,*ORIGIN*,1X,I3,1X,15(F5.3,2X),/)
  122 FORMAT(14F5.0)
  123 FORMAT(3F5.0)
  124 FORMAT(5X,*LOCATION MODEL CALIBRATION FOR SHOPPING CENTRES IN *,10
     1I3,/)
  125 FORMAT(//,55X,*DISTANCE MATRIX*,//)
  126 FORMAT(14F5.2)
  127 FORMAT(55X,*DESTINATION*,//)
  128 FORMAT(15X,15(I3,4X),/)
  129 FORMAT(5X,*ITERATION*,I6,/)
  130 FORMAT(51A1)
  131 FORMAT (//,5X,*MAXIMUM ENTROPY OF*,F7.4,* OCCURS ON ITERATION*,I6,
     1/)
  132 FORMAT(35H(6X,*LOCATIONS OF SHOPS ARE ZONES*,,I2,6HI4,//))
  133 FORMAT(2I3)
  134 FORMAT (6X,*SUMMARY OF STATISTICS*)
  135 FORMAT(40H(/,2X,*ITERATION*,1X,*SHOPPING LOCATIONS,2H*,,I2,6HX,*EN
     1T,34HROPY*,2X,4A10,7X,*RETAIL SALES*,/))
  136 FORMAT(10H(2X,I6,6X,,I1,9H(I3,3X),6,10HX,F7.4,2X,,10H4(F7.4,3X),1H
     1,,I1,7H(F5.3)))
  137 FORMAT (6X,*MAXIMUM *,A8,* CORRELATION OF*,F7.4,* OCCURS ON ITERAT
     1ION*,I6,/)
  138 FORMAT(5H(56X,,I2,45HX,*CORRELATION*,15X,*SALES AT EACH LOCATION*)
     1)
  139 FORMAT(42X,*OBSERVED WORKTRIP MATRIX FOR HAMILTON*,/)
  140 FORMAT(10F5.0)
  141 FORMAT(3I10)
  142 FORMAT(7I2)
```

```
      END
      SUBROUTINE COSCAL(N,M,NFAT,ISHOP,CHECK,W,NSHOP)
      COMMON/COST/C(15,15),CPAT(15,20),X(20),ORD(51),O(20)
      COMMON/INDEX/IND(20)
      COMMON/EXPNUM/WT(20)
      DIMENSION ISHOP(10),CT(20)
      LOGICAL CHECK
      INTEGER W
C
C     CALCULATE COSTS FOR A GIVEN TRIP PATTERN
C
      DO 1 I=1,N
C
C     A - FIND COSTS FOR HOUSEHOLDS OBSERVED TO MAKE NO SHOPTRIPS
C
      A1=99.
      KL=1
      DO 5 J=1,NSHOP
      L2=ISHOP(J)
      A1=AMIN1(A1,C(I,L2))
  5   CONTINUE
      CPAT(I,KL)=WT(KL)*A1
C
C     B - FIND COSTS FOR HOUSEHOLDS OBSERVED TO MAKE ONE SHOPTRIP
C
      MONE=NSHOP+1
      DO 2 J=2,MONE
      KL=KL+1
      K=J-1
      L=ISHOP(K)
      CT(KL)=C(I,L)
      CPAT(I,KL)=C(I,L)*WT(KL)
  2   CONTINUE
C
C     C - FIND COSTS FOR HOUSEHOLDS OBSERVED TOM AKE TWO SHOPTRIPS
C
      DO 6 J=2,NSHOP
      L=J+1
      DO 6 K=L,MONE
      KL=KL+1
      CPAT(I,KL)=((CT(J)+CT(K))/2.)*WT(KL)
  6   CONTINUE
  1   CONTINUE
C
C     WRITE OUT PATTERN-COST MATRIX FOR VERIFICATION
C
      WRITE(W,101)
      WRITE(W,102)(ISHOP(I),I=1,NSHOP)
      WRITE(W,103)
      WRITE(W,104)(IND(I),I=1,NFAT)
      DO 3 I=1,N
      WRITE(W,105) I,(CPAT(I,J),J=1,NPAT)
  3   CONTINUE
      RETURN
 101  FORMAT(//)
 102  FORMAT(5X,*COST MATRIX FOR TRIP PATTERNS - SHOPPING DESTINATIONS A
     1RE *,10I3,//)
 103  FORMAT(20X,* PATTERN *,/)
 104  FORMAT(10X,16(I4,3X),/)
 105  FORMAT(2X,I4,4X,16(F6.2,1X))
      END
      SUBROUTINE SUMMAR(N,M,KL,NPAT,PRED,MAX,TOTAL,COW,ISHCP,W,NSHOP)
      COMMON/TRIP/F(15,15)
      COMMON/PROB/P(15,15,20)
      COMMON/MODE/DD(15),OO(15),A(15),B(15),OR(15),DENS(15)
```

```
      COMMON/PAT/Y(2,15,3)
      COMMON/INDEX/IND(20)
      COMMON/CNTROL/INFO(10)
      DIMENSION PRED(5),O(15),D(15),ISHOP(10),ALAB(2),T(15,20),DEN(15),
     1COW(4)
      INTEGER W
      DATA ALAB/4HLIVE,4HWORK/
      NOTE=INFO(1)
      DO 19 I=1,M
      O(I)=D(I)=0.
   19 CONTINUE
C
C     COMPUTE ORIGIN AND DESTINATION AGGREGATE TOTALS
C
      DO 17 I=1,N
      DO 1 J=1,M
      DO 1 K=1,NPAT
      O(I)=O(I)+P(I,J,K)
      D(J)=D(J)+P(I,J,K)
    1 CONTINUE
      DEN(I)=O(I)/OO(I)
   17 CONTINUE
      WRITE(W,100)
      WRITE(W,101)
      WRITE(W,102)(ISHOP(I),I=1,NSHOP)
      WRITE(W,103)
      WRITE(W,122)
      DO 2 I=1,N
      WRITE(W,104) I,O(I),DEN(I),OR(I),DENS(I)
    2 CONTINUE
      IF (NOTE .LT. 2) GO TO 28
      Z=FLOAT(N)
      XIYI=XI=YI=XISQ=YISQ=0.
      DO 30 I=1,N
      XIYI=XIYI+O(I)*OR(I)
      XI=XI+O(I)
      YI=YI+OR(I)
      XISQ=XISQ+O(I)**2.
      YISQ=YISQ+OR(I)**2.
   30 CONTINUE
      RED=(Z*XIYI-XI*YI)/SQRT(ABS((Z*XISQ-XI**2.)*(Z*YISQ-YI**2.)))
      COW(4)=RED
      WRITE(W,101)
      WRITE(W,130) RED
      WRITE(W,101)
      WRITE(W,105)
      WRITE(W,103)
      DO 3 I=1,M
      WRITE(W,107) I,D(I),DD(I)
    3 CONTINUE
   28 CONTINUE
C
C     GENERATE WORKTRIP INTERCHANGE MATRIX
C
      DO 4 J=1,M
      DO 4 I=1,N
      T(I,J)=0.
      DO 5 K=1,NPAT
      T(I,J)=T(I,J)+P(I,J,K)
    5 CONTINUE
    4 CONTINUE
      WRITE(W,100)
      WRITE(W,101)
      WRITE(W,108)
      WRITE(W,109)
```

```
      WRITE(W,110)  (IND(I),I=1,M)
      DO 6 I=1,N
      IF (I .EQ. 7) GO TO 7
      WRITE(W,111) I,(T(I,J),J=1,M)
      GO TO 6
 7    CONTINUE
      WRITE(W,112) I,(T(I,J),J=1,M)
 6    CONTINUE
      NU=N
      IF (NOTE .LT. 2) GO TO 24
      MU=M
      CALL STAT(T,NU,MU,COW(1))
 24   CONTINUE
C
C     LOCATION OF HOUSEHOLDS WITH A GIVEN SHOPPING PATTERN
C
      LMAX=N
      L=1
      DO 8 J=1,NPAT
      DO 8 I=1,N
      T(I,J)=0.
      DO 9 K=1,M
      T(I,J)=T(I,J)+P(I,K,J)
 9    CONTINUE
 8    CONTINUE
      WRITE(W,130) COW(1)
 15   CONTINUE
      WRITE(W,100)
      WRITE(W,101)
      WRITE(W,113)
      WRITE(W,114)
      WRITE(W,115)(IND(I),I=1,NPAT)
      DO 10 I=1,LMAX
      IF (I .EQ. 7) GO TO 11
      WRITE(W,116) I,(T(I,J),J=1,NPAT)
      GO TO 10
 11   CONTINUE
      WRITE(W,117) ALAB(L),I,(T(I,J),J=1,NPAT)
 10   CONTINUE
      CALL MATPRM(LMAX,T,NSHOP,NPAT)
      CALL LITPRT(LMAX,ALAB,L,W,T)
      MU=3
      LINT=L+1
      CALL STAT2(T,NU,MU,L,COW(LINT),W)
      IF (L .EQ. 2) GO TO 12
      LMAX=M
      L=L+1
      DO 13 J=1,NPAT
      DO 13 I=1,MAX
      T(I,J)=0.
      DO 14 K=1,N
      T(I,J)=T(I,J)+P(K,I,J)
 14   CONTINUE
 13   CONTINUE
      GO TO 15
 12   CONTINUE
      WRITE(W,100)
      WRITE(W,101)
      WRITE(W,118)
      DO 16 I=1,KL
      WRITE(W,119) PRED(I)
 16   CONTINUE
      RETURN
 100  FORMAT("1")
 101  FORMAT(//)
```

```
 102   FORMAT (5X,*LOCATION OF HOUSEHOLDS GIVEN SHOPPING LOCATIONS IN ZON
      1ES*,10I3)
 103   FORMAT(19X,*PREDICTED*,10X,*OBSERVED*,/)
 104   FORMAT(5X,I3,5X,2(F7.3,3X,F7.3,1X))
 105   FORMAT (5X,*OBSERVED AND PREDICTED DESTINATION TOTALS*,/)
 107   FORMAT(5X,I3,11X,F6.3,9X,F8.3)
 108   FORMAT(30X,*PREDICTED WORKTRIP INTERCHANGE MATRIX FOR HAMILTON*,/)
 109   FORMAT(45X,*DESTINATION*,/)
 110   FORMAT(15X,15(I5,2X),/)
 111   FORMAT(11X,I3,1X,15(F5.3,2X),/)
 112   FORMAT(4X,*ORIGIN*,1X,I3,1X,15(F5.3,2X),/)
 113   FORMAT(5X,*PREDICTED SHOPPING PATTERN MATRIX FOR HAMILTON*,/)
 114   FORMAT(10X,*PATTERN/LOCATION*,/)
 115   FORMAT(10X,20(I4,2X))
 116   FORMAT(5X,I3,2X,20(F5.3,1X))
 117   FORMAT(1X,A4,I3,2X,20(F5.3,1X))
 118   FORMAT(30X,* PREDICTED STATISTICS*,/)
 119   FORMAT(35X,F8.2)
 121   FORMAT(42X,*OBSERVED WORKTRIP INTERCHANGE MATRIX FOR HAMILTON*,//)
 122   FORMAT(/,5X,*ZONE*,5X,2(*LOCATION*,3X,*DENSITY*,1X),/)
 130   FORMAT (//,5X,*CORRELATION COEFFICIENT = *,F8.4,//)
       END
       SUBROUTINE ENTROP(N,M,NPAT,SPR,W)
C
C      CALCULATION OF MAXIMUM ENTROPY, THE ENTROPY OF THE DISTRIBUTION
C      PREDICTED BY THE MODEL AND A MEASURE OF REDUNDANCY
C
       COMMON/PROB/P(15,15,20)
       COMMON/CNTROL/INFO(10)
       INTEGER W
       SOB=0.
       SPR=0.
       Z=FLOAT(N*M*NPAT)
       DO 1 K=1,NPAT
       DO 1 J=1,M
       DO 1 I=1,N
       SOB=SOB+P(I,J,K)
       IF (P(I,J,K) .EQ. 0) GO TO 1
       SPR=SPR-P(I,J,K)*ALOG(P(I,J,K))
   1   CONTINUE
       SMAX=ALOG(Z)
       RED=1.-(SPR/SMAX)
       WRITE(W,102)
       WRITE(W,103)
       WRITE(W,104) SPR,SMAX
       WRITE(W,105) RED
       WRITE(W,106) SOB
       RETURN
 100   FORMAT("1")
 102   FORMAT(//,25X,*ENTROPY STATISTICS FOR MODEL*,//)
 103   FORMAT(30X,*S (MODEL)*,3X,*S (MAX)*,/)
 104   FORMAT(30X,F7.4,5X,F7.4,//)
 105   FORMAT(30X,*REDUNDANCY =*,F7.4)
 106   FORMAT(//,25X,*PROBABILITY SUMS TO *,F7.4)
       END
       SUBROUTINE MODEL(N,M,NPAT,KIT,PARA,PRED,W)
       COMMON/PROB/P(15,15,20)
       COMMON/COST/C(15,15),CPAT(15,20),X(20),ORD(51),Q(20)
       COMMON/MODE/D(15),O(15),A(15),B(15),OR(15),DENS(15)
       DIMENSION PARA(5),PRED(5),XP(15,15,20)
       INTEGER W
       DO 14 K=1,NPAT
       DO 14 J=1,M
       DO 14 I=1,N
       XPON=-PARA(1)*C(I,J)-PARA(2)*CPAT(I,K)-PARA(3)*X(K)
```

```
      XP(I,J,K)=EXP(XPON)
 14   CONTINUE
      DO 1 J=1,M
      B(J)=0.
      DO 2 I=1,N
      DO 2 K=1,NPAT
      B(J)=B(J)+O(I)*Q(K)*XP(I,J,K)
 2    CONTINUE
      B(J)=1./B(J)
 1    CONTINUE
      SUM1=0.
      SUM2=0.
      SUM3=0.
      DO 3 K=1,NPAT
      DO 3 J=1,M
      DO 3 I=1,N
      P(I,J,K)=O(I)*Q(K)*B(J)*D(J)*XP(I,J,K)
      SUM1=SUM1+P(I,J,K)*C(I,J)
      SUM2=SUM2+P(I,J,K)*CPAT(I,K)
      SUM3=SUM3+P(I,J,K)*X(K)
 3    CONTINUE
      PRED(1)=SUM1
      PRED(2)=SUM2
      PRED(3)=SUM3
      RETURN
      END
      SUBROUTINE FREQ(N,M,NPAT,W)
      COMMON/PROB/P(15,15,20)
      COMMON/COST/C(15,15),CPAT(15,20),XXX(20),ORD(51),Q(20)
      DIMENSION T(25),MO(26),X(51,101)
      INTEGER G,W
      DATA BLANK,DASH,STAR,PLUS,EYE/1H ,1H-,1H*,1H+,1HI/
      AMAX=0.
      DO 1 I=1,25
      T(I)=0.
 1    CONTINUE
      DO 2 I=1,N
      DO 2 J=1,M
      DO 2 G=1,NPAT
      L=0
      DO 3 KK=5,125,5
      K=KK-1
      TIME=FLOAT(K)
      L=L+1
      IF ((C(I,J)+CPAT(I,G)) .LE. TIME) GO TO 4
 3    CONTINUE
 4    CONTINUE
      T(L)=T(L)+P(I,J,G)
 2    CONTINUE
      DO 17 I=1,25
      AMAX=AMAX1(AMAX,T(I))
 17   CONTINUE
      DO 20 I=5,100,5
      AK=FLOAT(I)/100.
      IF (AMAX .LT. AK) GO TO 21
 20   CONTINUE
 21   CONTINUE
      AMAX=AK
C
C     INITIALIZE AND BORDER THE GRAPH
C
      DO 6 I=1,50
      X(I,1)=EYE
      DO 7 J=2,101
      X(I,J)=BLANK
```

```
      X(51,J)=DASH
    7 CONTINUE
    6 CONTINUE
      DO 8 I=1,51,10
      X(I,1)=PLUS
    8 CONTINUE
      DO 9 J=1,101,4
      X(51,J)=PLUS
    9 CONTINUE
C
C     MAP FREQUENCY DISTRIBUTIONS ON GRAPH
C
      L=2
      DO 10 I=1,25
      MM=4*I+1
      MO(I)=IFIX((T(I)/AMAX)*50.+.05)
      IS=51-MO(I)
      DO 11 J=L,MM
      DO 12 K=IS,50
      IF (K .EQ. 51) GO TO 12
      X(K,J)=STAR
   12 CONTINUE
   11 CONTINUE
      L=MM+1
   10 CONTINUE
C
C     WRITE OUT FREQUENCY DIAGRAM
C
      WRITE(W,100)
      WRITE(W,101)
      L=1
      DO 13 I=1,51
      IF (I .EQ. L) GO TO 14
      WRITE(W,102) ORD(I),(X(I,J),J=1,101)
      GO TO 13
   14 CONTINUE
      XY=FLOAT(51-I)*AMAX/50.
      L=L+10
      WRITE(W,103) ORD(I),XY,(X(I,J),J=1,101)
   13 CONTINUE
      DO 15 I=1,26
      K=I
      MO(I)=5*(K-1)
   15 CONTINUE
      WRITE(W,104)(MO(I),I=1,26)
      WRITE(W,105)
C
C     WRITE OUT FREQUENCY DISTRIBUTION ON TABLE
C
      WRITE(W,100)
      WRITE(W,101)
      WRITE(W,106)
      L=0
      DO 16 I=1,25
      K=5*I-1
      WRITE(W,107) L,K,T(I)
      L=K+1
   16 CONTINUE
      RETURN
  100 FORMAT("1")
  101 FORMAT(/,37X,*FREQUENCY DISTRIBUTION OF TRAVEL TIMES PREDICTED BY
     1 MODEL*,/)
  102 FORMAT(10X,A1,7X,101A1)
  103 FORMAT(10X,A1,2X,F5.2,101A1)
  104 FORMAT(16X,26(I3,1X))
```

```
 105    FORMAT(57X,*TRAVEL TIMES (MIN.)*)
 106    FORMAT(42X,*TIME INTERVAL*,10X,*PROPORTION OF TRIPS*,/)
 107    FORMAT(44X,I4,*-*,I4,18X,F7.4)
        END
        SUBROUTINE MATINV(N,DET,Z)
C
C       INVERSION BY THE PIVOTAL CONDENSATION METHOD
C
        DIMENSION Z(10,10)
        DET=1.
        DO 1 J=1,N
        PV=Z(J,J)
        DET=DET*PV
        Z(J,J)=1.
        DO 2 K=1,N
        Z(J,K)=Z(J,K)/PV
 2      CONTINUE
        DO 1 K=1,N
        IF (K-J) 500,1,500
 500    TT=Z(K,J)
        Z(K,J)=0.
        DO 400 L=1,N
        Z(K,L)=Z(K,L)-Z(J,L)*TT
 400    CONTINUE
 1      CONTINUE
        RETURN
        END
        SUBROUTINE STAT(T,N,M,RED)
        DIMENSION T(15,20)
        COMMON/TRIP/F(15,15)
C
C       THIS SUBROUTINE CALCULATES A CORRELATION COEFFICIENT ON THE
C       INPUT MATRICES.  FOR DETAILS OF THE STATISTIC, SEE WILSON (1974),
C       P.317, AND BATTY (1976).
C
        Z=FLOAT(N*M)
        XIYI=XI=YI=XISQ=YISQ=0.
        DO 1 J=1,M
        DO 1 I=1,N
        XIYI=XIYI+T(I,J)*F(I,J)
        XI=XI+T(I,J)
        YI=YI+F(I,J)
        XISQ=XISQ+T(I,J)**2.
        YISQ=YISQ+F(I,J)**2.
 1      CONTINUE
        RED=(Z*XIYI-XI*YI)/SQRT(ABS((Z*XISQ-XI**2.)*(Z*YISQ-YI**2.)))
        RETURN
        END
        SUBROUTINE LITPRT(MAX,ALAB,L,W,F)
        DIMENSION ALAB(2),F(15,20)
        COMMON/INDEX/IND(20)
        COMMON/FAT/Y(2,15,3)
        INTEGER W
        WRITE(W,100)
        WRITE(W,124)
        WRITE(W,125)
        WRITE(W,126)(IND(I),I=1,3),(IND(I),I=1,3)
        DO 1 I=1,MAX
        IF (I .EQ. 7) GO TO 2
        WRITE(W,127) I,(Y(L,I,J),J=1,3),I,(F(I,J),J=1,3)
        GO TO 1
 2      CONTINUE
        WRITE(W,128) ALAB(L),I,(Y(L,I,J),J=1,3),ALAB(L),I,(F(I,J),J=1,3)
 1      CONTINUE
        RETURN
```

```
100   FORMAT(//)
125   FORMAT(28X,*PATTERN*,50X,*PATTERN*,//)
124   FORMAT (5X,*OBSERVED SHOPPING PATTERN MATRIX FOR HAMILTCN*,10X,*AG
     1GREGATED PREDICTED SHOPPING PATTERN MATRIX FOR HAMILTCN*,/)
126   FORMAT(19X,3(I5,3X),33X,3(I5,3X),/)
127   FORMAT (9X,I5,5X,3(F7.3,1X),23X,I5,5X,3(F7.3,1X))
128   FORMAT(2X,A4,3X,I5,5X,3(F7.3,1X),16X,A4,3X,I5,5X,3(F7.3,1X))
      END
      SUBROUTINE MATFRM(LMAX,F,NSHOP,NPAT)
      DIMENSION F(15,20)
      DO 1 I=1,LMAX
      T=0.
      MONE=NSHOP+1
      DO 2 J=2,MONE
      T=T+F(I,J)
2     CONTINUE
      F(I,2)=T
      T=0.
      MONE=MONE+1
      DO 3 J=MONE,NPAT
      T=T+F(I,J)
3     CONTINUE
      F(I,3)=T
1     CONTINUE
      RETURN
      END
      SUBROUTINE SALES(N,M,NPAT,S,NSHOP)
      COMMON/PROB/P(15,15,20)
      COMMON/BEER/IPOINT(10,10)
      DIMENSION S(10)
      DO 1 K=1,NSHOP
      S(K)=0.
      DO 1 L=1,NSHOP
      KK=IPOINT(K,L)
      DO 1 J=1,M
      DO 1 I=1,N
      S(K)=S(K)+P(I,J,KK)
1     CONTINUE
      RETURN
      END
      SUBROUTINE STAT2(T,N,M,L,RED,W)
      COMMON/PAT/X(2,15,3)
      DIMENSION T(15,20)
      COMMON/CNTROL/INFO(10)
      INTEGER W
      Z=FLOAT(N*M)
      XIYI=XI=YI=XISQ=YISQ=0.
      DO 1 J=1,M
      DO 1 I=1,N
      XIYI=XIYI+T(I,J)*X(L,I,J)
      XI=XI+T(I,J)
      YI=YI+X(L,I,J)
      XISQ=XISQ+T(I,J)**2.
      YISQ=YISQ+X(L,I,J)**2.
1     CONTINUE
      RED=(Z*XIYI-XI*YI)/SQRT(ABS((Z*XISQ-XI**2.)*(Z*YISQ-YI**2.)))
      WRITE(W,100) RED
      RETURN
100   FORMAT(////,38X,*CORRELATION COEFFICIENT = *,F8.4,//)
      END
      SUBROUTINE PATFIX(NSHCP,ISHOP,W)
      COMMON/BEER/IPOINT(10,10)
      DIMENSION KRUD(10),ISHOP(10)
      INTEGER W
      WRITE(W,100)
```

```
      WRITE(W,101)
      WRITE(W,102)
      DO 1 I=1,NSHOP
      KRUD(I)=0
      DO 1 J=1,NSHOP
      IPOINT(I,J)=0
    1 CONTINUE
      KL=1
      MONE=NSHOP+1
      DO 2 I=2,MONE
      KL=KL+1
      K=I-1
      L=ISHOP(K)
      WRITE(W,103) KL,L
      KRUD(K)=KRUD(K)+1
      KING=KRUD(K)
      IPOINT(K,KING)=KL
    2 CONTINUE
      DO 6 J=2,NSHOP
      L=J+1
      DO 6 K=L,MONE
      KL=KL+1
      IC=J-1
      JC=ISHOP(IC)
      WRITE(W,104) KL,JC
      KRUD(IC)=KRUD(IC)+1
      KING=KRUD(IC)
      IPOINT(IC,KING)=KL
      IC=K-1
      JC=ISHOP(IC)
      WRITE(W,105) JC
      KRUD(IC)=KRUD(IC)+1
      KING=KRUD(IC)
      IPOINT(IC,KING)=KL
    6 CONTINUE
      RETURN
  100 FORMAT("1",//,5X,*CORRESPONDENCE BETWEEN SHOPPING PATTERNS AND SHO
     1PPING CENTRES BEING PATRONIZED*,/)
  101 FORMAT(20X,*PATTERN*,10X,*SHOPPING LOCATION*,/)
  102 FORMAT(22X,* 1 *,19X,* 0 *,/)
  103 FORMAT(21X,I3,19X,I3,/)
  104 FORMAT(21X,I3,19X,I3)
  105 FORMAT(43X,I3,/)
      END
      SUBROUTINE WTDIST(NSHOP,WGH,NPAT)
      COMMON/EXPNUM/W(20)
      DIMENSION WGH(3)
      M=1
      W(M)=WGH(M)
      M=M+1
      MONE=NSHOP+1
      DO 1 I=2,MONE
      W(I)=WGH(M)
    1 CONTINUE
      MONE=MONE+1
      M=M+1
      DO 2 I=MONE,NPAT
      W(I)=WGH(M)
    2 CONTINUE
      RETURN
      END
      SUBROUTINE PRIOR(NSHOP,ISHOP,L,NPAT),RETURNS(NOGOOD)
      COMMON/COST/C(15,15),CP(15,20),ZAP(20),ORD(51),Q(20)
      COMMON/INDEX/INC(20)
      DIMENSION ISHOP(10),D(10),P(10)
```

```
      S=FLOAT(NSHOP-1)
      ALPHA=.6853
       GAMMA=0.1
      DO 1 I=1,NSHOP
      P(I)=0.
      M=ISHOP(I)
      DO 2 J=1,NSHOP
      IF (J .EQ. I) GO TO 2
      N=ISHOP(J)
      P(I)=P(I)+C(M,N)
2     CONTINUE
      P(I)=P(I)/S
1     CONTINUE
      KK=0
      WRITE(L,100)
19    CONTINUE
      KK=KK+1
      S=V=0.
      DO 3 I=1,NSHOP
      D(I)=EXP(-P(I)*GAMMA)
      S=S+D(I)
      V=V-P(I)*D(I)
3     CONTINUE
      R=Y=0.
      MSHOP=NSHOP-1
      DO 4 I=1,MSHOP
      K=I+1
      T=X=0.
      DO 5 J=K,NSHOP
      T=T+D(J)
      X=X+(P(I)+P(J))*D(J)
5     CONTINUE
      R=R+T*D(I)
      Y=Y+X*D(I)
4     CONTINUE
      FUNC=((ALPHA-1.)/ALPHA)+S+ALPHA*R
      IF (ABS(FUNC) .LT. .001) GO TO 7
      DERIV=V-ALPHA*Y
      ERROR=-FUNC/DERIV
      WRITE(L,105) KK,GAMMA,FUNC,ERROR
      GAMMA=GAMMA+ERROR
      IF (KK .LE. 10) GO TO 19
      KK=KK-1
      WRITE(L,106) KK
      GO TO 8
7     CONTINUE
      KL=1
      Q(KL)=ALPHA
      DO 10 I=1,NSHOP
      KL=KL+1
      Q(KL)=ALPHA*D(I)
10    CONTINUE
      ALPHA=ALPHA**2.
      DO 11 I=1,MSHOP
      K=I+1
      DO 12 J=K,NSHOP
      KL=KL+1
      Q(KL)=ALPHA*D(I)*D(J)
12    CONTINUE
11    CONTINUE
      WRITE(L,101)
      WRITE(L,102)(IND(I),I=1,NPAT)
      WRITE(L,103)(Q(I),I=1,NPAT)
      S=0.
      DO 13 I=1,NPAT
```

```
      S=S+Q(I)
13    CONTINUE
      WRITE(L,104) S
      RETURN
8     CONTINUE
      RETURN NOGOOD
100   FORMAT(//,5X,*ITERATION*,2X,*GAMMA*,5X,*FUNCTION*,5X,*ERROR*,/)
101   FORMAT(//,5X,*CALCULATED PRIORS FOR SHOPPING CENTRES*,/,20X,*PATTE
     1RN*,/)
102   FORMAT(6X,20(I3,2X))
103   FORMAT(5X,20F5.3)
104   FORMAT(//,5X,*PRIOR PROBABILITIES SUM TO *,F6.3)
105   FORMAT(5X,I6,3X,F7.3,5X,F6.3,3X,F7.3)
106   FORMAT(/,5X,*NOTE..NON-CONVERGENCE AFTER*,I3,* ITERATIONS.RUN ABOR
     1TED*,/)
      END
```

BIBLIOGRAPHY

Alonso, W., 1964, *Location and Land Use* (Harvard Univ. Press, Cambridge, Mass.).

Barras, R., et al., 1971, "An operational urban development model of Chesire", *Environment and Planning*, *3*, 115-234.

Batty, M., 1969, "Impact of a New Town: An Application of the Garin-Lowry Model", *Journal of the Town Planning Institute*, *55*, 428-35.

Batty, M., 1970, "Some Problem of Calibrating the Lowry Model", *Environment and Planning*, *2*, 95-114.

Batty, M., 1977a, "Design and Construction of a Subregional Land Use Model", *Socio-Economic Planning Sciences*, *5*, 97-124.

Batty, M., 1971b, "A Note on the Calibration of a Retail Location Model using Hyman's Method", Working Paper, Urban Systems Research Unit, Department of Geography, University of Redding.

Batty, M., 1972, "Recent Developments in Land Use Modelling - a review of British Research", *Urban Studies*, *9*, 151-77.

Batty, M., 1976a, *Urban Modelling* (Cambridge University Press, London).

Batty, M., 1976b, "Entropy in Spatial Aggregation", *Geographical Analysis*, *8*, 1-21.

Batty, M., 1977, "Operational Urban Models Incorporating Dynamics in a Static Framework", Paper presented to 'Spatial Interaction Theory and Planning Models', Båstad, Sweden.

Batty, M., and Mackie, S., 1972, "The calibration of gravity entropy and related nodels of spatial interaction", *Environment and Planning*, *4*, 205-233.

Batty, M., et al., 1973, "Spatial System Design and Fast Calibration of Activity Interaction-Allocation Models", *Regional Studies*, *7*, 351-366.

Batty, M., et al., 1974, "Experiments in urban modelling for county structure planning: the Area 8 pilot model", *Environment and Planning A*, *6*, 455-478.

Batty, M. and March, L., 1976, "The method of residues in urban modelling", *Environment and Planning A*, *8*,

Baxter, R. and Williams, I., 1975, "An automatically calibrated urban model", *Environment and Planning A*, *7*(1), 1-123.

Beardwood, J.E. and Kirby, H.R., 1975, "Zone Definition and the Gravity Model", *Transportation Research*, *9*, 363-369.

Blumenfeld, H., 1959, "Are Land Use Patterns Predictable?", *JAIP*, *25*(2), 61-66.

Brand, D., et al., 1967, "Techniques for relating transportation improvement and urban development patterns", *HRR*, *207*, 53-57.

Broadbent, T.A., 1970, "Notes on the design of operational models", *Environment and Planning*, *2*, 469-76.

Brown, H.J., et al., 1972, *Empirical Models of Urban Land Use* (National Bureau Economic Research: New York).

Carrothers, G.A.P., 1956, "An Historical Review of the Gravity and Potential Concepts of Human Interaction", *JAIP*, *22*, 94-102.

Cesario, F., 1973a, "A Note on the Entropy Model of Trip Distribution", *Transportation Research*, *7*, 331-333.

Cesario, F., 1973b, "Parameter Estimation in Spatial Interaction Modelling", *Environment and Planning*, *5*, 503-18.

Cesario, F., 1973c, "A generalized trip distribution model", *JRS*, *13*, 233-248.

Cesario, F., 1974a, "The interpretation and calculation of gravity nodel zone-to-zone adjustment factors", *Environment and Planning A*, *6*, 247-

Cesario, F., 1974b, "More on the generalized trip distribution model", *Transportation Research*, *8*,

Cesario, F., 1975a, "A Primer on Entropy Modeling", *JAIP*, *41*, 40-48.

Cesario, F., 1975b, "Least-Squares Estimation of Trip Distribution Parameters", *Transportation Research*, *9*, 13-18.

Chapin, S.F. and Weiss, S.F., 1968, "A Probabilistic model for residential growth", *Transportation Research*, *2*, 375-90.

Charnes, A., Raikes, W.M. and Bettinger, C.O., 1972, "An External and Information-Theoretic Characterization of Some Interzonal Transfer Models", *Socio-Economic Planning Sciences*, 6, 531-537.

Charnes, A., Haynes, K.E., Phillips, F. and White, G.M., 1977, "Dual Extended Geometric Programming Problems and the Gravity Model", *Journal of Regional Science*, *17*, 71-76.

Chisholm, M., et al., eds., 1971, *Regional Forecasting* (Butterworths, London).

Coelho, J.D. and Wilson, A.G., 1977, "An Equivalence Theorem to Integrate Entropy-Maximising Submodels within overall Mathematical-Programming Frameworks", *Geographical Analysis*, *9*, 160-173.

Colenutt, R.J., 1970, "Building models of urban growth and spatial structure", in *Progress in Geography*, *2*, Ed. C. Board et al. (Arnold, London), 109-52.

Cordey-Hayes, M., 1972, "Dynamic frameworks for spatial models", *Socio-Economic Planning Sciences*, 6, 365-85.

Cordey-Hayes, M. and Wilson, A.G., 1971, "Spatial Interaction", *Socio-Economic Planning Sciences*, *5*, 73-95.

Cordey-Hayes, M., et al., 1970, "An operational urban development model of Chesire", CES-WP-64; Centre for Environmental Studies, London.

Cripps, E.L. and Foot, D.H.S., 1969a, "The empirical development of an elementary residential location model for use in sub-regional planning", *Environment and Planning*, *1*, 81-90.

Cripps, E.L. and Foot, D.H.S., 1969b, "A land use model for subregional planning", *Regional Studies*, *3*, 243-68.

Cripps, E.L. and Foot, D.H.S., 1970, "The Urbanization Effects of a Third London Airport", *Environment and Planning*, *2*, 153-92.

Dawson, D.C. and Wragg, A., 1973, "Maximum-entropy Distributions Having Prescribed First and Second Moments", IEEE, *Transactions on Information Theory*, IT19, 689-93.

Dinkel, J.J., Kochenberger, G.A. and Wong, S-N, 1977, "Entropy maximization and geometric programming", *Environment and Planning A*, *9*, 419-427.

Echenique, M., 1969, "Urban Systems: Towards an Explorative Model", CES-WP-2, Centre for Environment Studies, London.

Echenique, M., Crowther, D. and Lindsay, W., 1969, "Development of a Model of a Town", *Land Use and Built Form Studies*, WP26.

Echenique, M., et al., 1974, "A disaggregated model of urban spatial structure: theoretical framework", *Environment and Planning A*, *6*, 33-63.

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., 1976, "Derivation and Analysis of Some Models for Combining Trip Distribution and Assignment", *Transportation Research*, *10*, 37-57.

Evans, R.A., 1961, "The Principle of Minimum Information", IEEE, *Transactions on Reliability*, R-18, 87-90.

Ewing, G.O., 1974, "Gravity and Linear Regression Models of Spatial Interaction: A Cautionary Note", *Economic Geography*, *50*, 83-88.

Feldt, A.G., 1972, *CLUG: Community Land Use Game: Player's Manual* (Free Press: New York).

Fisk, C. and Brown, G.R., 1975, "The Role of Model Parameters in Trip Distribution Models", *Transportation Research*, *9*, 143-148.

Fleischer, A., 1965, "Review of a Model of Metropolis", *JAIP*, *31*, 175-176.

Foot, D., 1974, "A comparison of some land-use allocation/interaction models", Reading Geographical Papers, Number 31, Department of Geography, Reading Univ.

Garin, R.A., 1966, "A matrix formulation of the Lowry Model for Intra-metropolitan Activity Location", *JAIP*, *32*, 361-64.

Garrison, W.L., 1966, "Difficult Decisions in Land Use Model Construction", *HRR*, *120*, 17-23.

Garrison, W.L., 1965, "Urban Transportation Planning Models in 1975", *JAIP*, *31*, 156-58.

Georgescu-Roegen, 1971, *The Entropy Law and the Economic Process* (Harvard U.P., Cambridge).

Goldner, W., 1971, "The Lowry model heritage", *JAIP*, *37*, 100-110.

Gould, P., 1972, "Pedagogic Review", *AAAG*, *62*, 689-700.

Hall, P.D., 1975, "An Assessment of the Calibration of Spatial Interaction Models", Unpublished M.A. Thesis, Dept. of Geography, McMaster Univ., Hamilton, Ontario.

Hamburg, J.R. and Creighton, R.I., 1959, "Predicting Chicago's Land Use Pattern", *JAIP*, *25*(2), 67-72.

Hansen, W.G., 1959, "How Accessibility Shapes Land Use", *JAIP*, *25*, 73-76.

Hansen, W.B., 1961, "An Approach to the analysis of metropolitan residential extension", *Journal of Regional Science*, *3*, 37-56.

Harris, B., 1961, "Some Problems in the Theory of Intra-Urban Location", *Operations Research*, *9*, 695-721.

Harris, B., 1964, "A note on the probability of interaction at a distance", *Journal of Regional Science*, *5*, 31-5.

Harris, B., 1965, "Urban Development Models: New Tools for Planning", *JAIP*, *31*, 90-95.

Harris, B., 1966, "The Uses of Theory in the Simulation of Urban Phenomena", *HRR*, *126*, 1-16.

Harris, B., 1968, "Quantitative models of urban development: their role in Metropolitan Policy-Making", in *Issues in Urban Economics*, Ed. H.S. Perloff and L. Wingo, JR. (Johns Hopkins, Baltimore), 363-412.

Hathaway, P.J., 1975, "Trip distribution and disaggregation", *Environment and Planning A*, *7*(1), 71-97.

Hemmens, G.C., ed., 1968, *Urban Development Models*, Special Report 97, (Highway Research Board, Washington, D.C.).

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

Hill, D.M., 1965, "A growth allocation model for the Boston region", *JAIP, 31*, 111-20.

Hill, D.M., 1968, Discussion of Lowry's "Models of Urban Development", in Hemmens, G., ed., 1968.

Hill, D.M., et al., 1965, "Prototype Development of Statistical Land Use Prediction Model for Greater Boston Region", *HRR*, *114*, 51-70.

Hobson, A., 1969, "A New Theorem of Information Theory", *Journal of Statistical Physics*, *1*, 3, 383-391.

Hobson, A., and Cheng, B.K., 1973, "A Comparison of the Shannon and Kullback Information Measures", *Journal of Statistical Physics*, *7*(4), 301-310.

Huff, D.L., 1962, "A Note on the Limitations of the Intra-Urban Gravity Model", *Land Economics*, *38*, 64-6.

Hyman, G.M., 1969, "The calibration of trip distribution models", *Environment and Planning*, *1*, 105-12.

Isard, W., 1975, "Transportation Research: some reflections on its development", *Regional Science and Urban Economics*, *5*, 133-135.

James, F.J., 1974, *Models of Employment and Residence Location* (Centre for Urban Policy Research, New Jersey).

Jaynes, E.T., 1957, "Information theory and statistical mechanics", *Physical Review*, *106*(4), 171-190.

Jaynes, E.T., 1968, "Prior probabilities", IEEE, *Transactions on Systems Science and Cybernetics*, SSC-4, 227-241.

Kain, J.F., 1962a, "A multiple equation model of household location and trip making behaviour", Rand Corporation, RM-3086-FF; Santa Monica, Calif.

Kain, J.F., 1962b, "The journey to work as a determinant of residential location", *PPRSA*, *9*, 137-60.

Karlquist, A. and Marksjo, B., 1971, "Statistical Urban Models", *Environment and Planning*, *3*, 83-98.

Karlquist, A., 1977, "Spatial Interaction Theory and Planning Models", Paper presented to 'Spatial Interaction Theory and Planning Models', Båstad, Sweden.

Kibel, B.M., 1972, "Simulation of the Urban Environment", AAG, Commission on College Geography Technical Paper 5.

Kilbridge, M.D., O'Block, R.P. and Teplitz, P.V., 1969, "A Conceptual Framework for Urban Planning Models", *Management Science*, *15*(6), 246-264.

Kirby, H.R., 1970, "Normalizing Factors of the Gravity Model: An Interpretation", *Transportation Research*, *4*, 37-50.

Kirby, H.R., 1974, "Theoretical requirements for calibrating gravity models", *Transportation Research*, *8*, 97-104.

Krueckenberg, D.A. and Silvers, A.L., 1974, *Urban Planning Analysis: Methods and Models* (Wiley: New York).

Lakschmanan, T.R., 1964, "An approach to the analysis of intraurban location", *Economic Geography*, *40*, 348-70.

Lakschmanan, T.R., and Hansen, W.G., 1965, "A Retail Market Potential Model", *JAIP*, *31*, 134-43.

Lee, C., 1973, *Models in Planning* (Pergamon Press, Oxford).

Lee, D.B., 1973, "Requiem for Large Scale Models", *JAIP*, *39*, 163-178.

Lichfield, N., Kettle, P. and Whitbread, M., 1975, *Evaluation in the Planning Process* (Pergamon Press, Oxford).

Lowry, I.S., 1964, "A Model of Metropolis", Rand Corporation, RM-4125-RC; Santa Monica, Calif.

Lowry, I.S., 1965, "A short course in model design", *JAIP*, *31*, 158-165.

Lowry, I.S., 1966, "Discussion of Land Use Forecasting Concepts", *HRR*, *126*, 32-37.

Lowry, I.S., 1968, "Sevens models of urban development: a structural comparison", Special Report 97, (Highway Research Board, Washington, D.C.), 121-145.

March, L., Batty, M., 1975a, "Information-Minimizing Formalism and the Derivation of Non-Parametric Forms for Population and Transportation Models", Technical Report, 6-S, Department of Systems Design, University of Waterloo, Ontario, Canada.

March, L. and Batty, M., 1975b, "Generalized measures of information, Bayes' likelihood ratio and Jaynes' formalism", *Environment and Planning B*, *2*, 99-105.

Masser, I., 1972, *Analytical Models for Urban and Regional Planning* (David and Charles, Newton Abbott).

Masser, I., Batey, P.W.J. and Brown, P.J.B., 1975, "The Design of Zoning Systems for Interaction Models", in Cripps, E.L., ed., 1975.

McLoughlin, B., 1969, *Urban and Regional Planning: a systems approach* (Faber, London).

Mitchell, R. and Rapkin, C., 1954, *Urban Traffic: A Function of Land Use* (Columbia U.P.: New York).

Nijkamp, P., 1975, "Reflections on gravity and entropy models", *Regional Science and Urban Economics*, *5*, 203-255.

Olsson, G., 1970, "Explanation, Prediction and Meaning-Variance: An Assessment of Distance Interaction Models", *Economic Geography*, *46*, 223-233.

Openshaw, S., 1976, "An empirical study of some spatial interaction models", *Environment and Planning A*, *8*(1), 1-122.

Putman, S., 1975, "Urban Land Use and Transportation Models: A State-of-the-Art Summary", *Transportation Research*, *9*, 187-202.

Reif, B., 1973, *Models in Urban and Regional Planning* (Intertext, New York).

Reilly, W.J., 1931, *The Law of Retail Gravitation* (Putman, New York).

Renyi, A., 1970, *Probability Theory* (American Elsevier, New York).

Robinson, I.M., Wolfe, H.B. and Barringer, R.L., 1965, "A Simulation Model for Renewal Programming", *JAIP*, *31*, 126-34.

Rogers, A., 1966, "A Note on the Garin-Lowry Model", *JAIP*, *32*(6), 364-66.

Schlager, K.J., 1966, "A Recursive Programming Theory of the Residential Land Development Process", *HRR*, *126*, 24-32.

Schneider, M., 1959, "Gravity models and trip distribution theory", *PPRSA*, *5*, 51-6.

Scott, A.J., ed., 1969, *London Papers in Regional Science*, volume 1, *Studies in Regional Science*, Proceedings of the Second Annual Conference of The British Section of the Regional Science Association, Pion, London.

Scott, A.J., 1970, "Transportation and Distribution of Population : Some Entropy Maximising Models", Centre for Urban and Community Studies, RP 32, University of Toronto.

Senior, H.L., 1973, "Approaches to Residential Location Modelling 1: Urban Ecological and Spatial Interaction Models", *Environment and Planning*, *5*, 165-197.

Senior, M.L., 1974, "Approaches to Residential Location Modelling 2: Urban economic models and some recent developments", *Environment and Planning A*, *6*, 369-409.

Senior, M.L. and Wilson, A.G., 1974a, "Explorations and syntheses of linear programming and spatial interaction models of residential location", *Geographical Analysis*, *6*, 209-237.

Shannon, C.F., 1948, "A mathematical theory of communication", *Bell System Technical Journal*, *27*, 379-423, 623-656.

Smith, T.E., 1975, "A Choice Theory of Spatial Interaction", *Regional Science and Urban Economics*, *5*, 137-176.

Smith, T.E., 1977, "A general efficiency principle of spatial interaction", Paper presented to 'Spatial Interaction Theory and Planning Models', Båstad, Sweden.

Snickars, F. and Weibull, J.W., 1977, "A Minimum Information Principle Theory and Practice", *Regional Science and Urban Economics*, *7*, 137-168.

Stetzer, F., 1976, "Parameter estimation for the constrained gravity model", *Environment and Planning A*, *8*, 673-683.

Stewart, J.Q., 1947, "Empirical rules concerning the distribution and equilibrium of population", *Geographical Review*, *38*, 461-85.

Swerdloff, C.N. and Stowers, J.R., 1966, "A test of some first generation land use models", *HRR*, *126*, 38-59.

Tribus, M., 1969, *Rational Descriptions, Decisions and Designs* (Pergammon, New York).

Tribus, M. and Rossi, R., 1973, "On the Kullback Information measure as a Bases for Information Theory: Comments on a Proposal by Holsson and Chang", *Journal of Statistical Physics*, *9*(4), 331-338.

Voorhees, A.M., 1959, "The Nature and Uses of Models in City Planning", *JAIP*, *25*(2), 57-60.

Walsh, J.A., 1976, "Information Theory and Classification in Geography", DP-4, Department 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.

Webber, M.J., 1975, "Entropy Maximising Location Models for Non-independent Events", *Environment and Planning A*, *7*, 99-108.

Webber, M.J., 1977a, "Pedagogy again: What is Entropy?", *AAAG*, 67(2), 254-266.

Webber, M.J., 1977b, "Spatial Interaction and the Form of the City", Paper presented to 'Spatial Interaction Theory and Planning Models', Båstad, Sweden.

Webber, M.J., 1977c, "An Elementary Entropy-Maximizing Model of Urban Consumers", *PPRSA*, 39, 251-271.

Williams, I., 1974, "A comment on the derivation of calibration criteria by Batty and Mackie", *Environment and Planning A*, *6*, 603-

Wilson, A.G., 1967, "A Statistical theory of spatial distribution models", *Transportation Research*, *1*, 253-69.

Wilson, A.G., 1968, "Models in Urban Planning: a synoptic review of recent literature", *Urban Studies*, *5*, 249-276.

Wilson, A.G., 1969a, "Notes on some concepts in social physics", *PPRSA*, *22*, 159-193.

Wilson, A.G., 1969b, "Disaggregating elementary residential location models", CES-WP37, Centre for Environmental Studies, London.

Wilson, A.G., 1969c, "The integration of accounting and location theory in urban modelling", in A.J. Scott (ed.), 1969, 89-104.

Wilson, A.G., 1969d, "Development of some elementary residential location models", *Journal of Regional Science*, *9*, 377-385.

Wilson, A.G., 1969e, "The use of entropy maximising models in the theory of trip distribution, mode split and route split", *Journal of Transport Economics and Policy*, *3*, 79-85.

Wilson, A.G., 1969f, "Metropolitan growth models", CES-WP55, Centre for Environmental Studies, London.

Wilson, A.G., 1969g, "Generalizing the Lowry Model", CES-WP56, Centre for Environmental Studies, London.

Wilson, A.G., 1969h, "The use of Analogies in Geography", *Geographical Analysis*, *1*, 225-233.

Wilson, A.G., 1970a, "Advances and problems in distribution modelling", *Transportation Research*, *4*, 1-18.

Wilson, A.G., 1970b, "Interregional Commodity Flows: entropy maximising approaches", *Geographical Analysis*, *2*, 255-282.

Wilson, A.G., 1970c, *Entropy in Urban and Regional Modelling* (Pion, London).

Wilson, A.G., 1970d, "The use of the concept of entropy in system modelling", *Operational Research Quarterly*, *21*, 247-265.

Wilson, A.G., 1970e, "On some problems in urban and regional modelling", in M. Chisholm et al., eds., 1971.

Wilson, A.G., 1971a, "A family of spatial interaction models and associated developments", *Environment and Planning*, 3, 1-32.

Wilson, A.G., ed., 1971b, *London Papers in Regional Science*, volume 2, *Urban and Regional Planning*, Proceedings of the Third Annual Conference of the British Section of the Regional Science Association,

Wilson, A.G., ed., 1972, *London Papers on Regional Science*, volume 3, *Patterns and Processes in Urban and Regional Systems*, Proceedings of the Annual Conference of the British Section of the Regional Science Association,

Wilson, A.G., 1973a, "Further developments of entropy maximizing transport models", *Transportation Planning and Technology*, *1*, 183-193.

Wilson, A.G., 1973b, "Learning and control mechanisms for urban modelling", in Cripps, E.L., ed., *Regional Science*, (Pion, London).

Wilson, A.G., 1974, *Urban and Regional Models in Geography and Planning* (Toronto, Wiley).

Wilson, A.G., 1975, "Some New Forms of Spatial Interaction Model: A review", *Transportation Research*, *9*, 167-179.

Wilson, A.G., et al., 1969, "Calibrating and testing the SELNEC transport model", *Regional Studies*, *3*, 337-350.

Wilson, A.G. and Senior, M.L., 1974, "Some relationships between entropy and maximizing models, mathematical programming models and their duals", *Journal of Regional Science*, *14*, 207-215.

Wingo, L., 1961, *Transportation and Urban Land* (Resources for the Future Inc., Washington).

Wingo, L. and H.S. Perloff, 1961, "The Washington transportation plan: technics or politics", *PPRSA*, *I*, 249-262.

Zipf, G.K., 1949, *Human Behaviour and the Principle of Least Effort: an introduction to human ecology.* (Addison-Wesley, Cambridge, Mass).