SPATIAL STRUCTURE, SPATIAL INTERACTION, AND DISTANCE-DECAY PARAMETERS

© ALEXANDER STEWART FOTHERINGHAM

B.Sc.(Hons.),M.A.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree
Doctor of Philosophy

McMaster University

September 1980

SPATIAL STRUCTURE AND DISTANCE-DECAY PARAMETERS

DOCTOR OF PHILOSOPHY (1980)
(Geography)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Spatial Structure, Spatial Interaction, and Distance-Decay Parameters

AUTHOR: Alexander Stewart Fotheringham, B.Sc. (Hons.) (Aberdeen University, Scotland)
M.A. (McMaster University, Canada)

SUPERVISOR: Professor M.J. Webber

NUMBER OF PAGES: xv , 257

ABSTRACT

One result of calibrating a spatial interaction model using interaction data specific to an origin i is to produce an estimate of the distance-decay parameter specific to that origin ($\hat{\beta}_i$). A correct interpretation of $\hat{\beta}_i$ is important since it is frequently used as a descriptive statistic of the spatial system under investigation and as a parameter for predicting unknown interactions.

The traditional and still commonly-accepted interpretation of $\hat{\beta}_i$ is that it is a purely behavioural measure of the relationship between distance and interaction from an origin, *ceteris paribus*. The distance-decay parameter is assumed to measure the rate of decline of interaction between centres as the distance separating the centres increases. However, when a set of origin-specific, distance-decay parameters, $\{\hat{\beta}_i\}$, is mapped, there is usually a marked spatial pattern to the values: accessible origins have less-negative values while inaccessible origins have more-negative values. As a result, several studies have recently attempted to reinterpret $\hat{\beta}_i$ in terms of the spatial structure of the system under investigation. These attempts have not gained general acceptance, however, and no sound theory has been posited to explain how spatial structure can determine $\hat{\beta}_i$.

This thesis demonstrates that a strong relationship between spatial structure and $\hat{\beta}_i$ exists and that the present interpretation of such parameters is wrong. Spatial structure can determine $\hat{\beta}_i$ in two ways: one is a calibration-specific effect (to which existing literature has unsuccessfully alluded) and the second is a calibration-independent effect

which until now has been unrecognized. The former results from multicollinearity in unconstrained interaction models and from the presence of balancing factors in constrained interaction models. The latter is a result of model mis-specification: the accessibility of a destination is shown to be an important explanatory variable of the volume of interaction into that destination. Theoretical and empirical evidence is given that the more accessible is a destination to other destinations, the less attractive it is for interaction, *ceteris paribus*. This relationship is not measured in present spatial interaction models and when it is, the exaggerated spatial variation of distance-decay parameters disappears and the parameters can be interpreted as purely behavioural measures. When the effect of spatial structure is not taken into account, $\hat{\beta}_i$ is shown to be primarily an index of accessibility and not a behavioural parameter.

Thus, current spatial interaction models are mis-specified. A new type of spatial interaction model is presented which is correctly specified and which removes the spatial structure bias from distance-decay parameter estimates.

ACKNOWLEDGEMENTS

This is probably the most widely read yet hardest to write section of a thesis. The distinction between the acknowledgement of deserved credit and sycophancy is sometimes a very fine one and the temptation to become overly-sentimental is often a very great one. To avoid such criticisms, my acknowledgements are short but they are sincerely given.

Mike Webber deserves thanks for proofreading the thesis and for making many useful suggestions on the contents. I also thank Ross MacKinnon and David Smith of the Department of Geography, State University of New York at Buffalo for providing the interaction data on which most of my empirical work is based.

I thank my wife, Barbara, who deserves much credit for typing an earlier version of the thesis; for proofreading the final version; and for undertaking most of the cartography. The sacrifice of her own career while in Canada cannot be overstated and is deeply appreciated.

Rosanna Wan expertly and cheerfully typed the final version of the thesis and Mrs. Margaret Chan helped with the cartography.

Financial assistance was provided by a C.W. Sherman Science and Engineering Scholarship and a McMaster University Urban Studies Fellowship.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

FIGURE		Page

LIST OF TABLES

PREFACE

A necessity of a society in which there is specialisation is spatial interaction, the movement of people, goods and information from one location to another. This interaction is undertaken at the cost of overcoming the spatial separation of locations which encourages short distance interactions rather than long distance interactions. It is important to measure the perception of spatial separation (or distance, as spatial separation is commonly measured by) as a deterrent to interaction in order to predict interactions where they cannot be measured. For this reason, models of spatial interaction are calibrated and distance-decay parameters are estimated.

A distance-decay parameter measures the relationship between observed interaction patterns and distance when all other determinants of interaction are constant. This relationship is assumed to be an accurate reflection of the perception of distance as a deterrent to interaction. Thus, when a distance-decay parameter is estimated in the calibration of a spatial interaction model, the interpretation of the estimate is that it is purely a behavioural measure of the relationship between distance and interaction, *ceteris paribus*. Chapter 1 defines this behavioural interpretation and then questions it. Evidence is given that distance-decay parameters may be a function of spatial structure. Chapter 2 describes and demonstrates the weaknesses in existing theories relating distance-decay parameters to spatial structure.

There are two ways in which spatial structure can bias an estimated distance-decay parameter such that it no longer solely measures the perception of distance as a deterrent to interaction. The first is that because of the spatial structure of opportunities, observed interaction patterns are not a true reflection of the perception of distance as a barrier to interaction. For example, assume that interaction patterns for grocery shopping are observed in two towns which have identical population distributions and the residents of which have identical perceptions of distance with respect to interaction. In town A there is only one grocery store while in town B there are three stores in different locations. From the observed interaction patterns, the mean trip length to buy groceries would be greater in town A than in town B and it would be assumed that the residents of town A perceive distance to be less of a deterrent to interaction than do the residents of town B. This would be a false assumption since the perception of distance is equal in both towns: the observed interaction patterns and the resulting distance-decay relationships are biased by spatial structure. People travel shorter distances in town B simply because the destinations are closer on average than in town A.

The second way in which spatial structure can bias an estimated distance-decay parameter occurs when the parameter estimated in the calibration of a spatial interaction model is not solely a measure of the relationship between observed interaction patterns and distance but is also a function of spatial structure. Inadvertently, either the calibration procedure or the structure of the spatial interaction model, or both, allows the estimated distance-decay parameter to be a function of spatial structure as well as observed interaction patterns.

This thesis concentrates entirely on the latter effect. The former effect, if it occurs, is very difficult to measure or correct since the only information on people's perception of distance that is usually available is the pattern of observed interactions. There is no standard to decide whether these are accurate data from which to measure the perception of distance. In any case, the former effect is unlikely to bias distance-decay parameter estimates severely. The example given of grocery trips in two towns is only valid when town A has only one opportunity at which to buy groceries. If there were two grocery stores, at different locations, observed interaction patterns would reflect the fact that one store is closer to some residents than is the other and the perceived disutility of distance with respect to interaction could accurately be measured.

Chapter 3 of this thesis proposes that the latter spatial structure effect can occur in three ways depending upon the type of spatial interaction model calibrated and the calibration technique. Chapter 4 describes the effects of multicollinearity between mass and distance on parameter estimates derived from unconstrained spatial interaction models; Chapter 5 describes the relationship between estimated distance-decay parameters and the balancing factors of constrained interaction models; and Chapter 6 describes how present unconstrained and constrained spatial interaction models are mis-specified and derives new interaction models which are correctly specified. Chapter 7 contains a comparison of the calibration results from both sets of interaction models.

Thus, the hypothesis upon which this thesis is based is that actual interaction-distance relationships, *ceteris paribus*, are constant over space but estimated distance-decay parameters are not since the latter are biased by spatial structure. It is assumed that observed interaction patterns are an accurate reflection of the perception of distance as a deterrent to interaction and once the spatial structure bias is removed from the models or the calibration procedure, estimated distance-decay parameters will be constant over space.

The conclusions of this work are far-reaching. The interpretation of present distance-decay parameters is shown to be false. Instead of being indicators of interaction behaviour they are primarily indices of spatial structure. New models are presented, however, from which purely behavioural distance-decay parameters can be derived. These latter distance-decay parameters should aid greatly in predicting interactions when there is a change in spatial structure since they will be invariant to such change - a property present parameters do not have.

CHAPTER ONE

AN INTRODUCTION TO DISTANCE-DECAY PARAMETERS AND SPATIAL INTERACTION MODELS

1.1 Distance-Decay and Distance-Decay Parameters

Spatial interaction can be defined as the movement of people, goods or information resulting from a decision to move. The decision to move is based on the utility derived from the movement.[1] In interaction terms, total utility, U, can be defined as the following benefit-cost equality:

$$U = f(B) + g(C) \qquad (1.1)$$

where f(B) is a function of the total benefit derived from interacting,

g(C) is a function of the total cost incurred from interacting,

and utility is assumed to be an increasing function of benefit and a decreasing function of cost.

The total volume of interaction, I, between any two points will be an increasing function of the total utility derived from such interaction:

$$I = h(U) \qquad (1.2)$$

By substituting (1.1) into (1.2), interaction is seen to be a decreasing

[1] More correctly, this discussion should be in terms of expected utility, expected benefit, and expected cost. For the sake of brevity, it is assumed that all expectations are realised.

function of cost. As the cost of interacting increases, the volume of interaction decreases, *ceteris paribus*.[1] This is the essence of distance-decay since the cost of interacting is assumed to be an increasing function of the distance, D, over which the interaction is taking place:

$$C = i(D) \qquad (1.3)$$

Distance-decay can then be defined as: *the rate at which the volume of interaction decreases as the distance over which the interaction is taking place increases, ceteris paribus.* A distance-decay parameter simply measures this rate. It conveys information about the relationship between interaction and distance or, alternatively, between interaction and cost.

1.2 The Importance of Measuring Distance-Decay

The concept of distance-decay is an extremely important one in urban and economic geography. It is difficult to imagine urban geography existing if distance did not play an important role in human organisation. As Olsson [1970, p. 223] states:

> "Under the umbrella of spatial interaction and distance-decay, it has been possible to accommodate most model work in transportation, migration, commuting and diffusion, as well as significant aspects of location theory."

---

[1]The term *ceteris paribus* is used throughout the text. It is necessary to discuss the relationship between the volume of interaction and the cost of interaction with "everything else being constant" since the true relationship may be disguised otherwise. For example, the volume of ~~interaction~~ between Chicago and New York city may be greater than that between Chicago and Duluth even though the cost of the former may be greater. Obviously there are other factors that affect the volume of interaction besides cost (in this case, population) and these must be controlled for in order to determine the true relationship between interaction and cost.

Such an important concept should be fully understood and accurately measured. There are two main reasons why the accurate measurement of distance-decay in spatial interactions is important: (i) to obtain descriptive statistics about the system under investigation and (ii) in order to predict unknown interactions accurately. These two reasons are now expanded.

(i) Descriptive Statistic

Consider a spatial system within which interaction is taking place. Knowledge of the distance-decay of interactions within that system can be important in several ways. It can indicate how sensitive particular types of interaction are to distance. Shopping trips, for example, may be more sensitive to distance than are work trips. Trips to consume hamburgers may be more sensitive to distance than are trips to consume steaks. Knowledge of this sensitivity, which is given by the distance-decay parameter, is invaluable in locational analysis. To determine the location for a facility which will maximise its patronage, a good estimate of the relevant distance-decay parameter is necessary. There have been many empirical investigations to obtain this information, *inter alia,* Iklé [1954], Hansen [1959] and Yuill [1967]. A corollary of such knowledge is that distance-decay parameters can aid in determining market areas for various facilities. Huff [1963] and Beaumont [1978] discuss this subject.

Black [1971, 1972], Chisholm and O'Sullivan [1973] and Gordon [1978] have demonstrated that if the interaction under investigation is movement of a particular commodity, then the distance-decay parameter can be used to determine regional market areas for that commodity. They also show how such parameters can be used to determine certain commodity characteristics. A steep distance-decay function, for example, would indicate that the commodity is sensitive to distance and is probably either bulky, perishable or has a low value/weight ratio. A gentle distance-decay function would indicate the opposite characteristics.

Insight into interaction behaviour in different areas is important. It might be determined, through knowledge of distance-decay parameters, that interaction behaviour in one area is very different from interaction behaviour in another area. One can then ask: what causes this difference? Is it differences in the medium through which interaction is taking place or is it differences in distance perceptions? Also, once one is aware of differences in interaction behaviour, one can ask: do we want to change such behaviour? It might be found, for example, that there are many long distance shopping trips from a particular suburb. On investigation, this may be a result of poor shopping facilities within that suburb.

Differences in interaction behaviour over time can be evaluated by comparison of the relevant distance-decay parameters. Hagerstrand [1957], for example, investigated how migration patterns changed over time in Sweden through comparison of distance-decay parameters. Clark and Ballard [1979] similarly studied variations in distance-decay parameters over time using out-migration data for the Appalachian states. They concluded that distance is becoming less of a factor in determining interaction and thus we could expect the "ghettoes" of Appalachians in nearby states to disappear as people migrate further afield.

There have been several other studies in which knowledge of distance-decay parameters has produced insights into human behaviour. MacKay [1958] investigated the effect of boundaries upon interaction through distance-decay parameters. Gould [1975] used distance-decay parameters to analyse spatial variations in the level of information of school children and the transmission of knowledge. Clark [1979] used distance-decay parameters to determine the nature of ancient exchange systems.

Finally, knowledge of distance-decay parameters is useful in a heuristic sense. Besides describing interaction behaviour, patterns of distance-decay parameters can indicate determinants of such behaviour. Olsson [1965] and Young [1975], for example, have suggested from such an analysis that the size of a centre determines, in part, the interaction distance-decay connected with that centre.

Accurate measurement of distance-decay parameters is thus very important in order to describe interaction behaviour. Accurate measurement is a necessity if any inference is to be drawn from such description.

(ii) Prediction

The result of predicting interactions often has major policy implications. Predicted interactions can influence the building of new transit systems and the siting of facilities and thus the accuracy of such predictions is very important. The usual method of prediction is to calibrate a spatial interaction model using interaction data for one particular spatial system in one particular time period.

The parameters derived from the calibration are assumed to remain constant over time or over space, or over both, so that they can be used to predict interactions where data on interactions are not available. Since the distance-decay parameter is an integral component of all major spatial interaction models (see Section 1.3), obtaining accurate estimates of distance-decay parameters is then essential to accurate prediction. These parameters ideally should have little variation over time and space, or if they do vary, the variation should be predictable. Southworth [1980] demonstrates the large errors in predicting interactions as a result of assuming constant distance-decay parameters over time when independent calibrations show that they vary.

1.3 Distance-Decay Parameters and Spatial Interaction Models

It has been shown that interaction is a function of distance:

$$I = f(D) \quad . \tag{1.4}$$

It is clear, however, that interaction is also a function of other variables. The greater the population of a centre, for example, the greater the volume of interaction would be expected to emanate from, and be attracted to, that centre. Interaction between two centres is then a function of the propulsiveness and attractiveness of both centres, as well as the distance between them:

$$I = f(D, P, A) \tag{1.5}$$

where P is a measure of propulsiveness

and A is a measure of attractiveness.

In order to measure accurately the effect of distance upon interaction, all other determinants of interaction must be accounted for. Spatial interaction models permit measurement of the distance-interaction relationship by controlling for varying degrees of propulsiveness and attractiveness. Distance-decay parameters derived from three classes of spatial interaction model will be investigated. The class into which a model is placed is determined by the number of constraints on the predicted interactions operating upon the model. There may be zero, one, or two such constraints producing unconstrained, singly-constrained and doubly-constrained spatial interaction models respectively. The effect of varying the number of constraints produces varying functions of propulsiveness and attractiveness. The distance function, however, is independent of these constraints, and hence remains constant between classes of model. More formally, (1.5) can be rewritten as:

$$I = f(D) \cdot g(P) \cdot h(A) \tag{1.6}$$

where the functions g and h vary between model class while f remains constant. The exact forms of g and h are given below for each type of model. It is useful first, however, to define the function of distance used for each model.

The Distance Function Used

There are many possible functions of distance which could be used in spatial interaction modelling (see, for example, Taylor [1971, 1975], Openshaw and Connolly [1977]). However, three are most commonly used:

A Power function where:

$$f(D) = D^{\beta} , \tag{1.7}$$

β being the distance decay parameter.

An Exponential function:

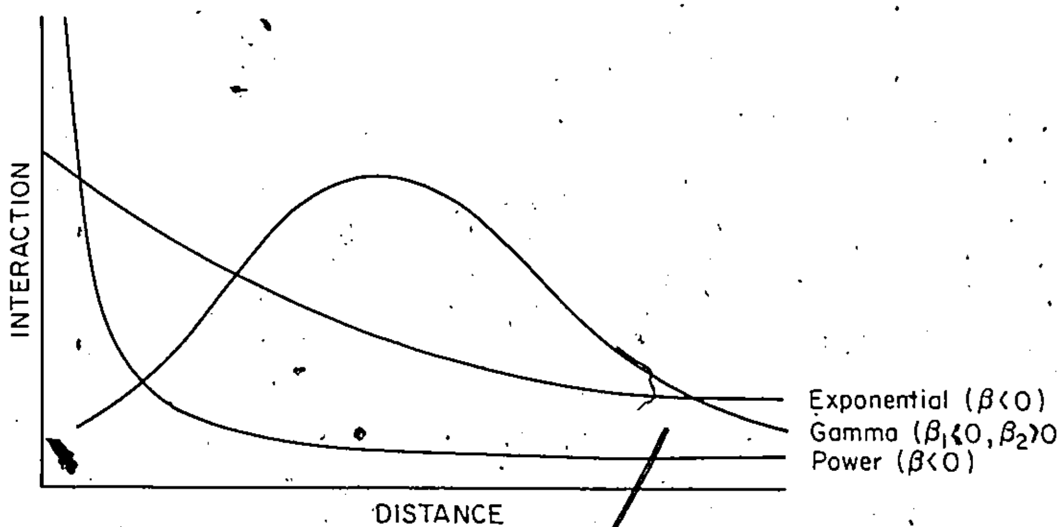
$$f(D) = \exp(\beta D) , \tag{1.8}$$

and a Gamma function:

$$f(D) = \exp(\beta_1 D) . D^{\beta_2} . \tag{1.9}$$

Each of these functions describes how distances are perceived in terms of interaction. Use of the power and exponential functions assumes that a monotonic non-linear relationship exists between distance and interaction. Use of the power function assumes that the relationship can be made linear when both variables are converted to natural logarithms. Use of the exponential function assumes that the relationship can be made linear when only interaction is converted to natural logarithms. The gamma function is more flexible than either the power or exponential function. In (1.9) if $\beta_1 = 0$ and $\beta_2 \neq 0$, the gamma function degenerates to a power function. If $\beta_2 = 0$ and $\beta_2 \neq 0$, the gamma function degenerates to an exponential function. If β_1 and β_2 are both non-zero and have different signs, the gamma function can model relationships that are not monotonic. Examples of the three functions are given in Figure 1.1.

FIGURE 1.1: The Power, Exponential and Gamma Functions of Distance.



The function chosen for analysis in this thesis was the Pareto or power function. There are three reasons for this choice. Firstly, the data analysed in this thesis are inter-urban interactions. It is assumed that the power function is a more accurate description of the perception of distances at this scale than is the exponential function. Support for this assumption is given by Frost [1969], Hyman [1969], Wilson [1970a], Gordon [1976] and Stillwell [1977]. While the gamma function may also be more appropriate than the exponential, the interpretation of its two parameters is not always obvious. Secondly, a primary objective of this thesis is to demonstrate how previously estimated distance-decay parameters are biased. It is then useful to use the same function of distance that has been most commonly used in previous studies. Both Black and Larsen [1972] and Taylor [1975] point out that this is the power function.

While the exponential function has also frequently been used in interaction modelling, the gamma function has not. Thirdly, an advantage of using the exponential function rather than the power function is that the former can model interactions over zero distances while the latter cannot. This is irrelevant in this case, however, since the interaction data to be analysed contain no such interactions.

While a power function of distance is used here, the results obtained would be similar to those obtained using an exponential function. The framework of the analysis would remain the same, as would the interpretation of the distance-decay parameters. Each of the spatial interaction models to be analysed is now briefly described.

(i) An Unconstrained Spatial Interaction Model

Consider a spatial system consisting of m origins i, and n destinations j. An unconstrained spatial interaction model can then be written as:

$$I_{ij} = \alpha \, m_i^{\omega} \, m_j^{\gamma} \, d_{ij}^{\beta} \qquad (1.10)$$

where I_{ij} is the interaction between i and j,

m_i is the measure of the propulsiveness of i,

m_j is the measure of the attractiveness of j,

d_{ij} is the distance between i and j,

and α, ω, γ and β are parameters to be estimated/given data on I_{ij}, m_i, m_j and d_{ij}.

When (1,10) is calibrated, it is rewritten as:

$$\hat{I}_{ij} = \hat{\alpha} \, m_i^{\hat{\omega}} \, m_j^{\hat{\gamma}} \, d_{ij}^{\hat{\beta}} \qquad (1.11)$$

where "^" denotes an estimated value. Hence, $\hat{\beta}$ is the estimated distance-decay parameter. The usual method of calibrating (1.10) is to take the natural logarithms of both sides of the equation and to estimate the parameter values by ordinary least squares regression (OLS). This produces an equation of the form:

$$\hat{I}_{ij}^* = \hat{\alpha}^* + \hat{\omega}\, m_j^* + \hat{\gamma}\, m_j^* + \hat{\beta}\, d_{ij}^* \qquad (1.12)$$

where "*" denotes a natural logarithm. Comprehensive reviews of this model are given in Carrothers [1956], Isard [1960] and Olsson [1965b].

(ii) A Singly-Constrained Spatial Interaction Model

The singly-constrained model to be investigated is a production-constrained model of the form:

$$I_{ij} = Z_i\, O_i\, D_j\, d_{ij}^{\beta} \qquad (1.13)$$

where,

$$Z_i = [\sum_{j=1}^{n} D_j\, d_{ij}^{\beta}]^{-1} . \qquad (1.14)$$

When calibrated, the model is rewritten as:

$$\hat{I}_{ij} = Z_i\, O_i\, D_j\, d_{ij}^{\hat{\beta}} \qquad (1.15)$$

where,

$$Z_i = [\sum_{j=1}^{n} D_j\, d_{ij}^{\hat{\beta}}]^{-1} . \qquad (1.16)$$

O_i is the true outflow[1] from origin i, or $\sum_{j=1}^{n} I_{ij}$, and D_j is

[1] For prediction purposes, of course, O_i must be estimated or measured externally.

a measure of the attractiveness of destination j. Z_i is a balancing factor which ensures that the constraint,

$$\sum_{j=1}^{n} \hat{I}_{ij} = \sum_{j=1}^{n} I_{ij} \qquad \forall i \tag{1.17}$$

is met. In words, (1.17) states that the predicted total outflow is equal to the true total outflow for each origin.

Proof of (1.17)

Summing (1.15) over all j gives:

$$\sum_{j=1}^{n} \hat{I}_{ij} = Z_i \; O_i \sum_{j=1}^{n} D_j \; d_{ij}^{\hat{\beta}} \; . \tag{1.18}$$

Substituting Z_i from (1.14) into (1.18):

$$\sum_{j=1}^{n} \hat{I}_{ij} = [\sum_{j=1}^{n} D_j \; d_{ij}^{\hat{\beta}}]^{-1} \; O_i \sum_{j=1}^{n} D_j \; d_{ij}^{\hat{\beta}} \tag{1.19}$$

$$= O_i = \sum_{j=1}^{n} I_{ij} \; .$$

Q.E.D.

Since (1.13) is non-linear in parameters, it cannot be calibrated using OLS. Instead it is calibrated by maximum likelihood estimation assuming a multinomial likelihood function[1] with the constraint that the estimated mean trip length for the system equals the true mean trip length or,

---

[1] The exact form of this likelihood function and the procedure for the constrained maximum likelihood estimation of the parameters of spatial interaction models is described in detail by Batty and Mackie [1972,212-214].

$$\frac{\sum_{i=1}^{m} \sum_{j=1}^{n} \hat{I}_{ij} \, d_{ij}}{\sum_{i=1}^{m} \sum_{j=1}^{n} \hat{I}_{ij}} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij} \, d_{ij}}{\sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij}} . \tag{1.20}$$

The singly-constrained spatial interaction model is discussed and its derivation from entropy-maximising principles is outlined by Wilson [1971].

(iii) A Doubly-Constrained Spatial Interaction Model

The doubly-constrained model investigated is:

$$I_{ij} = Z_i \, O_i \, B_j \, D_j \, d_{ij}^{\beta} \tag{1.21}$$

where,

$$Z_i = [\sum_{j=1}^{n} B_j \, D_j \, d_{ij}^{\beta}]^{-1} \tag{1.22}$$

and,

$$B_j = [\sum_{i=1}^{m} Z_i \, O_i \, d_{ij}^{\beta}]^{-1} . \tag{1.23}$$

The measurement of j's attractiveness is now equal to the true total inflow into j[1], or $\sum_{i=1}^{m} I_{ij}$. When calibrated, (1.21), (1.22), and (1.23) are rewritten as follows:

$$\hat{I}_{ij} = Z_i \, O_i \, B_j \, D_j \, d_{ij}^{\hat{\beta}} \tag{1.24}$$

where,

$$Z_i = [\sum_{j=1}^{n} B_j \, D_j \, d_{ij}^{\hat{\beta}}]^{-1} \tag{1.25}$$

[1] Again, this has to be estimated or measured independently for prediction.

and,

$$B_j = [\sum_{i=1}^{m} Z_i \; O_i \; d_{ij}^{\hat{\beta}}]^{-1} \tag{1.26}$$

In the above systems, Z_i and B_j are constraints which ensure that:

$$\sum_{j=1}^{n} \hat{I}_{ij} = \sum_{j=1}^{n} I_{ij} \quad \forall i \tag{1.27}$$

and,

$$\sum_{i=1}^{m} \hat{I}_{ij} = \sum_{i=1}^{m} I_{ij} \quad \forall j \tag{1.28}$$

respectively. The interpretation of (1.27) is the same as that of (1.17). Equation (1.28) states that the total predicted inflow is equal to the true total inflow for each destination.

Proof of (1.27)

Summing (1.24) over j gives:

$$\sum_{j=1}^{n} \hat{I}_{ij} = Z_i \; O_i \sum_{j=1}^{n} B_j \; D_j \; d_{ij}^{\hat{\beta}} \; . \tag{1.29}$$

Substituting Z_i from (1.25) into (1.29):

$$\sum_{j=1}^{n} \hat{I}_{ij} = [\sum_{j=1}^{n} B_j \; D_j \; d_{ij}^{\hat{\beta}}]^{-1} \; O_i \sum_{j=1}^{n} B_j \; D_j \; d_{ij}^{\hat{\beta}} \tag{1.30}$$

$$= O_i = \sum_{j=1}^{n} I_{ij} \; .$$

Q.E.D.

Proof of (1.28)

Summing (1.21) over i gives:

$$\sum_{i=1}^{m} \hat{I}_{ij} = B_j D_j \sum_{i=1}^{m} Z_i O_i d_{ij}^{\hat{\beta}} . \qquad (1.31)$$

Substituting B_j from (1.26) into (1.31):

$$\sum_{i=1}^{m} \hat{I}_{ij} = [\sum_{i=1}^{m} Z_i O_i d_{ij}^{\hat{\beta}}]^{-1} D_j \sum_{i=1}^{m} Z_i O_i d_{ij}^{\hat{\beta}} \qquad (1.32)$$

$$= D_j = \sum_{i=1}^{m} I_{ij} .$$

Q.E.D.

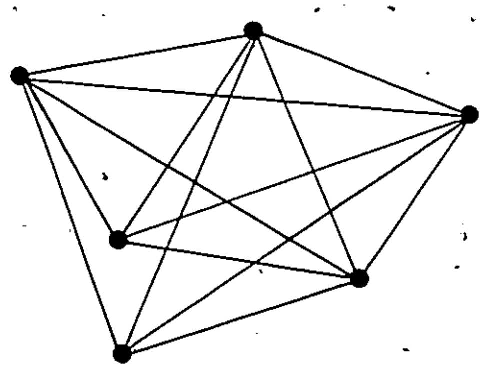
Equation (1.21) is calibrated by maximising the same likelihood function dicussed for the singly-constrained model. This maximisation is again achieved by ensuring that the constraint given in (1.20) is met. The calibration of the doubly-constrained model is more demanding computationally, however, than the singly-constrained model since the balancing factors given in (1.22) and (1.23) must be estimated iteratively. MacGill [1975] gives a proof of the convergence of the balancing factors while Wilson [1970b, 1971] discusses further details of the model.

Origin-Specific Spatial Interaction Models

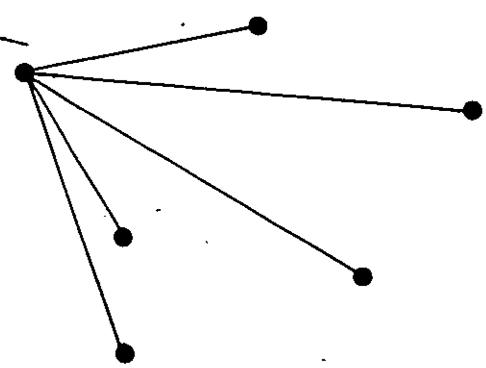
The models given in (1.10), (1.13) and (1.21) describe interactions in a spatial system consisting of m origins and n destinations. Such a system can be considered as an aggregation of m subsystems: each subsystem being the set of interactions from a particular origin. This is shown in Figure 1.2. Calibrating an interaction model for each of these subsystems

FIGURE 1.2: An Interaction System and Subsystem.

(a) Interactions in a Spatial System



(b) Interactions in a Spatial Subsystem (origin-specific)



would thus produce useful information on the interaction patterns specific to each origin. With slight modifications, the models given in (1.10), (1.13) and (1.21) can be used to obtain this information. These origin-specific models are written as follows:

(i) An unconstrained origin-specific model:

$$I_{ij} = \alpha_i \, m_j^{\gamma_i} \, d_{ij}^{\beta_i} \qquad \forall i \, . \tag{1.33}$$

The m_i variable in (1.10) is a constant when i is constant and is subsumed in α_i .

(ii) A singly-constrained origin-specific model:

$$I_{ij} = Z_i \, O_i \, D_j \, d_{ij}^{\beta_i} \qquad \forall i \tag{1.34}$$

where,

$$Z_i = [\sum_{j=1}^{n} D_j \, d_{ij}^{\beta_i}]^{-1} \qquad \forall i \, . \tag{1.35}$$

(iii) A doubly-constrained origin-specific model:

$$I_{ij} = Z_i \, O_i \, B_j \, D_j \, d_{ij}^{\beta_i} \qquad \forall i \tag{1.36}$$

where,

$$Z_i = [\sum_{j=1}^{n} B_j \, D_j \, d_{ij}^{\beta_i}]^{-1} \qquad \forall i \tag{1.37}$$

and,

$$B_j = [\sum_{i=1}^{m} Z_i \, O_i \, d_{ij}^{\beta_i}]^{-1} \qquad \forall j \, . \tag{1.38}$$

In all three models the parameters to be estimated are origin-specific parameters. β_i is then the origin-specific distance-decay parameter and $\hat{\beta}_i$ its estimate. The interpretation of $\hat{\beta}_i$ is the major concern of this thesis. Consequently, the models given in (1.33), (1.34) and (1.36) will be closely investigated in subsequent chapters. However, it is useful to note here three points concerning these models. The first is that the subscripts given in (1.33) appear unbalanced. They are not since there is only one origin i at each calibration of the model. The calibration of (1.33) is entirely independent of data concerning the other origins of the system. This is also true of (1.34).

The second point follows from the first. The model given in (1.36), although calibrated for a specific origin, is not independent of the other origins in the system. The model includes a constraint on the total predicted inflows into each destination and hence (1.38) contains a summation over all origins in the system. By following the proofs of (1.27) and (1.28) it is easily seen that (1.37) and (1.38) ensure that,

$$\sum_{j=1}^{n} \hat{I}_{ij} = \sum_{j=1}^{n} I_{ij} \qquad \forall i$$

and,

$$\sum_{i=1}^{m} \hat{I}_{ij} = \sum_{i=1}^{m} I_{ij} \qquad \forall j$$

respectively.

The third point concerns the constraint necessary for the calibration of (1.34) and (1.36) by maximum-likelihood estimation. For the complete spatial system, this was given by (1.20). However, for origin-specific calibration, it is necessary to constrain the estimated mean trip length for each origin so that it is equal to its true value. Formally, the constraint is written as:

$$\frac{\sum_{j=1}^{n} \hat{I}_{ij} d_{ij}}{\sum_{j=1}^{n} \hat{I}_{ij}} = \frac{\sum_{j=1}^{n} I_{ij} d_{ij}}{\sum_{j=1}^{n} I_{ij}} \qquad \forall i \; . \tag{1.39}$$

1.4 The Interpretation of $\hat{\beta}_i$

Traditionally, the sign and magnitude of $\hat{\beta}_i$, the estimated origin-specific distance-decay parameter, are assumed to be solely functions of interaction behaviour. The estimate of $\hat{\beta}_i$ is assumed to reflect underlying interaction behaviour with respect to distance: the underlying behaviour being a result of many individual decisions. Under this assumption of $\hat{\beta}_i$, if the underlying decision-making process was constant over space, $\hat{\beta}_i$ would be constant for all origins. Variations in $\hat{\beta}_i$ are ascribed to variations in interaction decision-making. However, this thesis will demonstrate that $\hat{\beta}_i$ is also a function of spatial structure so that even if the underlying interaction behaviour is constant over space, $\hat{\beta}_i$ can vary due to variations in spatial structure. An example is useful here to demonstrate the meaning of spatial structure and behaviour as determinants of estimated distance-decay parameters. Consider airline passenger interaction between the hundred largest cities in the United States. Such data are used to obtain estimated distance-decay parameters for each city. Suppose the parameter value for Los Angeles was -2.0 while that for Chicago was -0.5. Then suppose that there was a complete exchange of residents between Los Angeles and Chicago (their 1980 populations are approximately equal) and interaction patterns were again measured and origin-specific distance-decay parameters estimated. If the new estimated distance-decay parameter for Chicago was -2.0 and the new parameter value for Los Angeles was -0.5, then the determinant of $\hat{\beta}_i$ would be entirely behavioural, where behavioural refers to the discriminating characteristics of people interacting. $\hat{\beta}_i$ is determined solely by people's perceptions of the utility to be derived from interacting over various distances. However, if the resulting values after the exchange of residents were still -2.0 for Los Angeles and -0.5 for Chicago, the determinant of $\hat{\beta}_i$ would be entirely a function of spatial structure: $\hat{\beta}_i$ is determined solely by the configuration and size of destinations faced by an origin. Obviously, $\hat{\beta}_i$ can be determined by both behaviour and spatial structure and the resulting estimated parameter values for Los Angeles and Chicago would then be somewhere between -2.0 and -0.5. At this point, since the nature of the relationship between $\hat{\beta}_i$ and spatial structure has not been given, the interpretation of $\hat{\beta}_i$ continues under the assumption that $\hat{\beta}_i$ is solely a function of underlying decision-making behaviour.

It is evident that in each of the models discussed above, the expected sign of $\hat{\beta}_i$ is negative. As the distance increases between centres, interaction decreases, *ceteris paribus*. It would be very difficult to accept as accurate a positive $\hat{\beta}_i$ value since it would imply that the greater the separation between two centres, the greater is the interaction between them, *ceteris paribus*. There is no rationale for such behaviour. Thus, the expected sign of $\hat{\beta}_i$ is negative. The more negative is $\hat{\beta}_i$, then the greater is the "friction of distance" effect or the distance-decay effect: interaction declines more steeply as distance increases and there are proportionally more short distance interactions and proportionally fewer long distance interactions.

The interpretations given above can easily be demonstrated using (1.33). Assume the following:

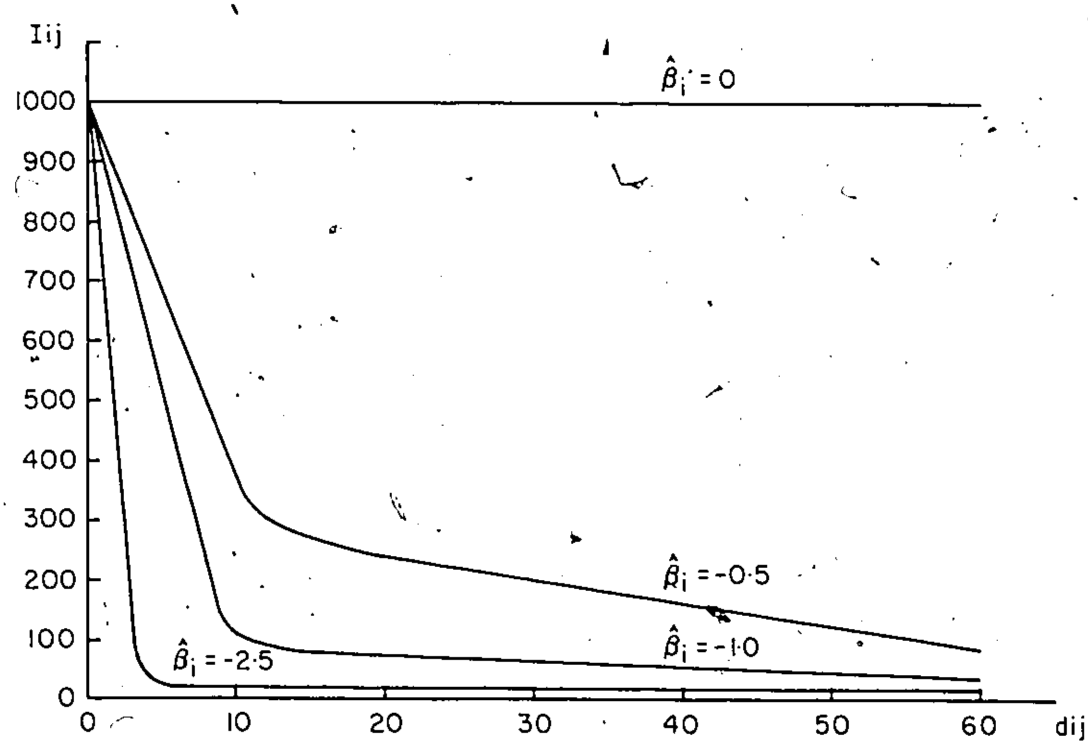
$$\alpha_i = 0.1$$

$$\gamma_i = 1.0$$

$$m_j = 10{,}000 \quad \forall j .$$

Then Figure 1.3 describes the different distance-interaction relationships measured by different values of $\hat{\beta}_i$.

FIGURE 1.3: $\hat{\beta}_i$ and the Distance-Interaction Relationship



A large negative value of $\hat{\beta}_i$ indicates a very steep interaction distance-decay while a small negative value of $\hat{\beta}_i$ indicates a shallow distance-decay. If $\hat{\beta}_i = 0$, then distance is not a factor in explaining interactions. If $\hat{\beta}_i$ is positive, then the volume of interaction increases from a minimum of 1,000 as distance increases. The steepness of this increase is indicated by the magnitude of $\hat{\beta}_i$. A very rapid increase would produce a large positive $\hat{\beta}_i$ while a very gentle increase would produce a small positive $\hat{\beta}_i$.

Thus, the estimation of $\hat{\beta}_i$ is very important in measuring interaction behaviour with respect to distance.

Origin-specific distance-decay parameters have now been defined and the models from which they are commonly estimated have been described. The remaining section of this introductory chapter reviews some currently accepted empirical results of the calibration of these models. Particular attention is paid to the spatial variation of $\hat{\beta}_i$.

1.5 Existing Empirical Evidence Relating $\hat{\beta}_i$ to Spatial Structure

The spatial structure of an interaction system is defined here as the spatial configuration of origins and destinations between which interaction takes place. Spatial configuration is taken to encompass not only the relative locations of centres but also the relative sizes of centres. This thesis investigates effects of spatial structure which are unaccounted for in the estimation of distance-decay parameters from present spatial interaction models. This "spatial structure effect" is thus assumed to be a function of the configuration of origins and destinations within the spatial system under investigation. If there is such an effect, it should be evident in the patterns of estimated origin-specific distance-decay parameters. There should be a systematic pattern of parameter estimates since only the configuration of destinations can affect $\hat{\beta}_i$, and adjacent origins will have similar configurations of destinations. Seven empirical studies are briefly reviewed to investigate whether or not such a pattern is evident. These studies are identified by their respective author(s).

(i) Chisholm and O'Sullivan [1973]

Chisholm and O'Sullivan calibrated a singly-constrained model very similar[1] to (1.34) for 78 zones in Britain using freight flow data. Their map of the resulting distance-decay parameters is reproduced in Figure 1.4. A clear spatial pattern emerges. Generally, more accessible origins have less negative $\hat{\beta}_i$ values, while less accessible origins have more negative $\hat{\beta}_i$ values. This pattern was recognised by Chisholm and O'Sullivan [p.72] and the natural logarithm of potential accessibility[2] was found to explain 48% of the variance of $\hat{\beta}_i$. Their explanation for this pattern was that more remote zones attempt to avoid the penalties of long hauls to central zones and therefore have proportionately more short distance interactions, while central areas distribute their traffic with less regard to distance. Chisholm and O'Sullivan also noted that because there was such a large spatial variation in parameter values, a single model applied to the whole country may result in serious over- or under-estimates of interaction between certain zones. This point is returned to later. The range of Chisholm and O'Sullivan's parameter estimates is given in Table 1.1.

It is interesting to note that Frost [1969] calibrated (1.33), the unconstrained model, using the same data as Chisholm and O'Sullivan. His results were similar: the logarithm of potential accessibility explained 31% of the variance in $\hat{\beta}_i$.[3]

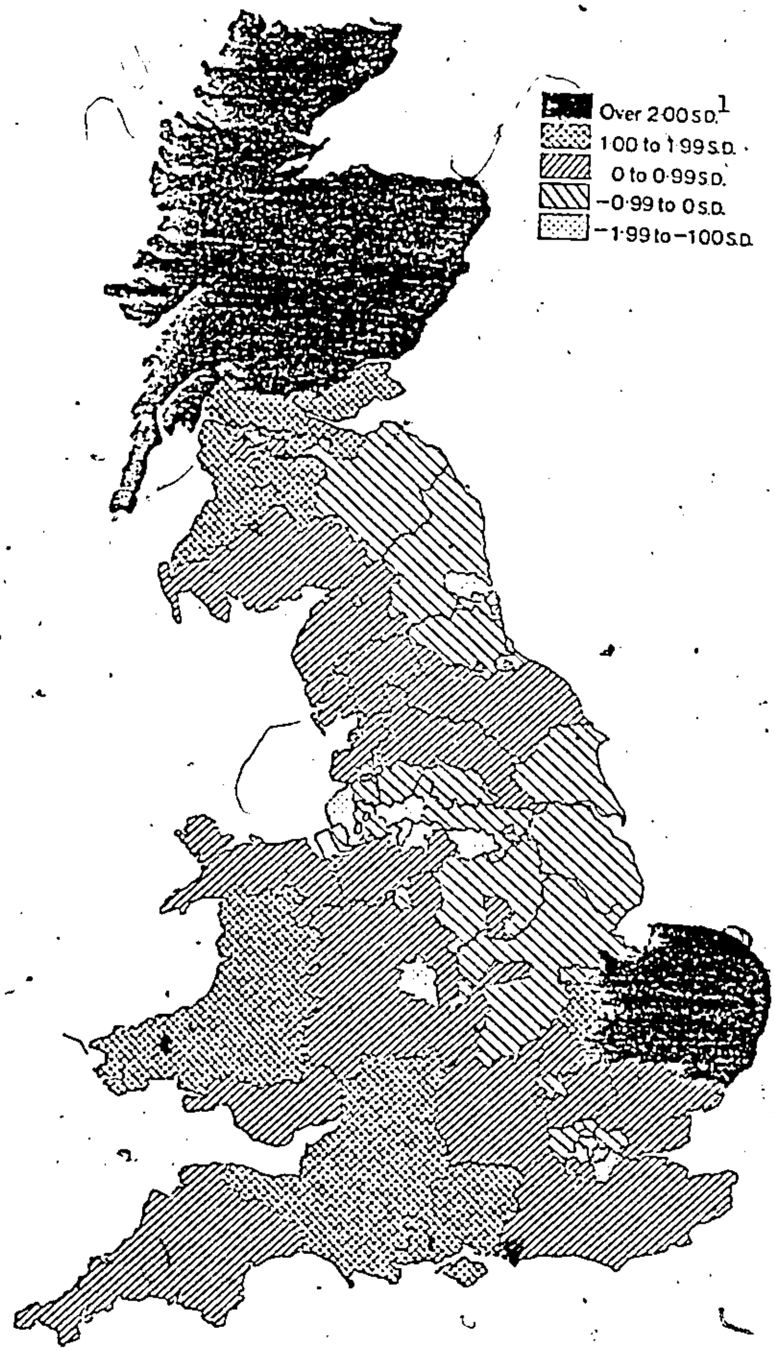
---

[1] The only difference ~~is~~ that Chisholm and O'Sullivan's model contained an exponent on the D_j variable. This in no way alters the interpretation of $\hat{\beta}_i$, however.

[2] Potential accessibility was defined by $\sum_{j=1}^{n} \frac{D_j}{d_{ij}^{\delta}}$, where δ was assumed to to equal 1.0.

[3] This is my calculation. The potential accessibilities used are those reported by Chisholm and O'Sullivan.

FIGURE 1.4: The Spatial Variation of $\hat{\beta}_i$ Estimated from British Freight Flow Data (Reproduced from Chisholm and O'Sullivan [1973, p. 71]).



[1] The $\hat{\beta}_i$ values were assumed to be negative in Chisholm and O'Sullivan's study. A positive Z-score therefore indicates a $\hat{\beta}_i$ value more negative than the mean.

TABLE 1.1

RANGE OF $\hat{\beta}_i$'s IN NINE EMPIRICAL STUDIES

Study	No. of $\hat{\beta}_i$'s	Range of Values: Least Negative $\hat{\beta}_i$	Range of Values: Most Negative $\hat{\beta}_i$
Chisholm and O'Sullivan	78	-1.3	-4.8
Frost	78	+0.3	-5.2
Gould	33	-0.2	-1.0
Stillwell	44	-0.8	-2.9
Leinbach	16	-0.4	-1.0
Greenwood and Sweetland	50	-0.4	-2.5
Smith	49	0.0	-2.7
Linneman	75	> 0[1]	-2.3
Griffith and Jones[2]	24	-0.1	-0.4

[1] Exact value not given.

[2] Their model used an exponential function of distance. The other studies used power functions.

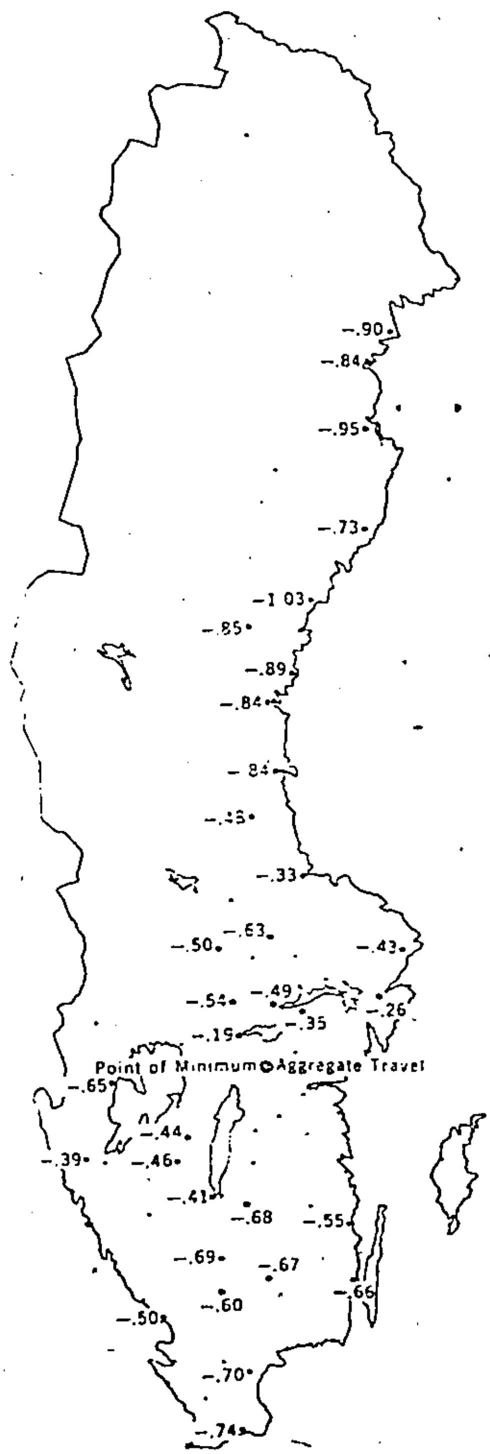
(ii) Gould [1975]

Gould calibrated an unconstrained origin-specific interaction model (1.33) for 33 Swedish towns. The interaction data were measures of children's spatial awareness and the $\hat{\beta}_i$ values measured the distance-decay rate of the acquisition of knowledge at each town. The spatial pattern of the distance-decay parameter is reproduced in Figure 1.5. With reference to the point of minimum aggregate travel, it is clearly seen that, in general, the further a town is from this point (i.e. the more inaccessible it is), the more negative is the parameter value for that town. Gould calculated that the square root of distance to the ~~point of~~ minimum aggregate travel explained 64% of the variance of $\hat{\beta}_i$. He also noted that the variance in $\hat{\beta}_i$ was much greater than in $\hat{\gamma}_i$, the mass parameter. Gould could not give a behavioural rationale for the large variance in $\hat{\beta}_i$, nor for its spatial pattern. He proposed that in egalitarian societies, such as Sweden, interaction behaviour with respect to distance should be reasonably constant. Thus, to explain his findings, Gould suggested that $\hat{\beta}_i$ is not only a function of behaviour but also of spatial structure. On this subject, he referred to the theoretical work of Curry [1972] which is discussed in Chapter 2.

(iii) Stillwell [1978]

Using British interregional migration data for 44 regions, Stillwell calibrated a doubly-constrained origin-specific interaction model (1.36). The spatial pattern of the distance-decay parameters indicates the duality between accessible and inaccessible origins observed in the above studies. More accessible origins have less negative parameter estimates and less accessible origins have more negative parameter estimates. One apparent anomaly is that Scotland, an inaccessible origin, has a very low negative parameter estimate. However, an exaplanation for this, in

FIGURE 1.5: The Spatial Variation of $\hat{\beta}_i$ Estimated from Swedish Interaction Data (Reproduced from Gould [1975, p. 93]).



terms of spatial structure, is given in Section 5.6.

Another interesting finding from this study was that there was no significant relationship between $\hat{\beta}_i$ and the mean trip length from origin i. A strong positive relationship would be expected if $\hat{\beta}_i$ were purely a behavioural measure: the greater the mean trip length from an origin, the less negative should be $\hat{\beta}_i$, and vice versa. This expected relationship was only found in Stillwell's results when spatial structure was controlled for by holding a region constant and then disaggregating the interaction data for that region by age and sex. These results were not noted by Stillwell who interpreted the pattern of distance-decay parameters as solely a result of varying migration behaviour over space.

(iv) Leinbach [*1973*]

In a study on modernisation, Leinbach analysed telephone message flows between exchanges in West Malaysia. Equation (1.33) was calibrated and a strong relationship between the location of an exchange and its estimated distance-decay parameter was found. Leinbach reports that the Spearman Rank correlation coefficient between distance from the modernisation core (a central location on the island) and the magnitude of the distance coefficient was .75. Since the coefficients were all negative, this indicates that the more inaccessible an origin, the more negative was it parameter estimate and vice versa. Leinbach assumed that this pattern was a result of varying behaviour.

(v) Greenwood and Sweetland [*1972*]

In a study of migration between 50 major cities in the U.S., Greenwood and Sweetland calibrated an unconstrained interaction model for each city. As they note, there is a strong relationship between the accessibility of an origin and its estimated distance-decay parameter. The smallest

$|\hat{\beta}_i|$ reported is for Chicago while the largest is for Salt Lake City. For the 10 cities in the sample that are located in the North-East (accessible centres), Greenwood and Sweetland report that the average $\hat{\beta}_i$ is -0.87 while that for the 11 cities located in the West (inaccessible centres), the average $\hat{\beta}_i$ is -1.74. It is interesting to note that for the same cities, the parameters reported for migration are very similar to those reported for 1960 airline passenger interaction by Smith [1973] who interpreted the pattern as a result of varying travel behaviour. Both sets of parameter estimates appear to be determined primarily by the accessibility of the origins.

(vi) Linneman [1966]

Linneman's origin-specific econometric model of trade flows is:

$$X_{ij}^* = \phi_0 + \phi_1 Y_j^* + \phi_2 N_j^* + \phi_3 D_{ij}^* + \phi_4 P_{ij}^{*1} + \phi_5 P_{ij}^{*2} + \phi_6 P_{ij}^{*3} \qquad (1.40)$$

where X_{ij} is the trade flow between i and j,

Y_j is the GNP of j,

N_j is the population of j,

D_{ij} is the distance between i and j,

P_{ij}^1, P_{ij}^2 and P_{ij}^3 are dummy variables related to trade blocs,

and "*" denotes a natural logarithm.

Define,

$$W_j \equiv Y_j \cdot N_j \qquad (1.41)$$

and,

$$\alpha \equiv \phi_0 + \phi_4 P_{ij}^{*1} + \phi_5 P_{ij}^{*2} + \phi_6 P_{ij}^{*3} \qquad (1.42)$$

and then (1.40) can be rewritten as:

$$X_{ij}^* = \alpha + (\phi_1 + \phi_2) W_j^* + \phi_3 D_{ij}^* \qquad (1.43)$$

which is of the same form as (1.33) and where ϕ_3 is the origin-specific distance-decay parameter and is henceforth referred to as $\hat{\beta}_i$.

Linneman calibrated this unconstrained spatial interaction model for 75 countries. Two of the resulting $\hat{\beta}_i$ values were positive but Linneman dismissed these as "nonsensical" [p. 91]. Eight extremely large negative $\hat{\beta}_i$ values and eight extremely small negative $\hat{\beta}_i$ values were given. The former are associated with inaccessible countries such as South Africa, Australia and Japan: the latter are associated with accessible countries such as Cuba, Cyprus and the U.K.. Linneman demonstrates that this relationship is supported by calculating $\hat{\beta}_i$ values for different continents and sub-continents. The smallest negative value is for Central America while the largest negative value is for Africa. He interprets these findings as indicative that relatively "unfavourably located" countries direct a greater part of their exports to their closest neighbours than do "more favourably" located countries. However, Linneman states that it is not easy to understand why this should be so.

(vii) Griffith and Jones [*1980*]

Using journey - to - work data for 24 Canadian cities, Griffith and Jones calibrated a doubly-constrained interaction model for each origin. In a detailed analysis, they showed how the estimated distance-decay parameters were related to spatial structure. Various measures of the spatial autocorrelation of destinations accounted for 41% of the variation in $\hat{\beta}_i$. In absolute terms, smaller distance-decay parameters were associated with geometric patterns in which similar numbers of jobs tended to cluster, whereas larger distance-decay parameters were associated with geometric patterns in which dissimilar numbers of jobs clustered. No significant relationship was found between $\hat{\beta}_i$ and various economic characteristics of urban areas. Griffith and Jones thus concluded that spatial structure and levels of distance-decay do covary and that [p.199]:

> "distance decay exponents for spatial interaction models measure both the influence of map pattern and the true friction of distance."

Thus, from the above studies it appears evident that there is a relationship between $\hat{\beta}_i$ and the location of i and consequently $\hat{\beta}_i$ may be a function of spatial structure. The evidence for the latter statement falls into four categories.

(i) There is an empirical regularity of less accessible origins having more negative distance-decay parameter estimates and of more accessible origins having less negative parameter estimates. This regularity is unexplained by *a priori* behavioural reasoning.

(ii) Following Gould's reasoning, interaction behaviour within a relatively homogeneous society should be fairly constant. Thus, the variation in estimated distance-decay parameters should be minimal. Table 1.1 indicates that this is not so and there are large differences between parameter estimates in all of the studies described above.

(iii) Near-zero negative values of $\hat{\beta}_i$ and even positive values of $\hat{\beta}_i$ occur. Such results are contrary to all intuitive understanding of interaction behaviour.

(iv) If $\hat{\beta}_i$ is solely a function of behaviour, a strong positive relationship between $\hat{\beta}_i$ and the mean trip length from origin i, $\bar{d}_i$, would be expected. As $\bar{d}_i$ increases, $\hat{\beta}_i$ should become less negative and vice versa. This relationship is rarely found, however, and often $\hat{\beta}_i$ and $\bar{d}_i$ are negatively related.

The investigation of a link between spatial structure and $\hat{\beta}_i$ is important since, if $\hat{\beta}_i$ is a function of spatial structure, then much of the variation in $\hat{\beta}_i$ would be a result of variations in spatial structure. Interaction behaviour could be constant over space and yet $\hat{\beta}_i$ would vary. This has important implications for prediction since a problem in predicting interactions has long been that parameter values vary over space.

While there are several researchers who are aware of the empirical findings presented above, and who believe that $\hat{\beta}_i$ is a function of spatial structure, a problem has been to demonstrate this relationship theoretically. How is $\hat{\beta}_i$ related to spatial structure? Demonstrating a theoretical relationship between $\hat{\beta}_i$ and spatial structure is the concern of this thesis. It is an important concern since, if such a relationship exists, it would be necessary to reinterpret the results of calibrating present spatial interaction models. Demonstrating a theoretical relationship between $\hat{\beta}_i$ and spatial structure is also important as a necessary step in obtaining estimates of $\hat{\beta}_i$ which are unbiased by spatial structure and which can be interpreted as solely behavioural measures. Existing approaches to this topic are now described and their weaknesses highlighted.

CHAPTER TWO

EXISTING THEORIES RELATING DISTANCE-DECAY PARAMETERS TO SPATIAL STRUCTURE

2.1 Introduction

While it has long been hypothesised (for example, Porter [1956]) that estimated distance-decay parameters and spatial structure are related, it is only within the past ten years that this relationship has been investigated theoretically. However, these investigations have only studied unconstrained models and there exists no theory which relates spatial structure to distance-decay parameters in constrained interaction models.

Three separate ideas have emerged. The first links $\hat{\beta}$ to the range of distances which separate origins and destinations and is attributable to Johnston. The second links $\hat{\beta}$ to the spatial autocorrelation of populations within a spatial system and is the result of work by Curry, Griffith and Sheppard. Thirdly, Fotheringham and Webber describe the unconstrained spatial interaction equation as one equation of a simultaneous equation system.

While the Fotheringham and Webber study is very recent and its general acceptance cannot be commented on, the studies by Johnston and Curry, Griffith and Sheppard do not appear to have gained universal acceptance. It is still generally assumed that the variance of $\{\hat{\beta}_i\}$ is a function solely of varying behaviour. There are two reasons for this.

The first is that the work of Johnston contains a flaw in logic while that of Curry, Sheppard and Griffith contains unnecessary and misleading complications. The second is that both approaches are only applicable to unconstrained spatial interaction models. Since empirical studies have shown that the variation in $\{\hat{\beta}_i\}$ is similar in constrained and unconstrained models, the explanation may be common to both types of model. While any theory that is peculiar to one type of model may explain some of the variance of $\{\hat{\beta}_i\}$, it is unlikely to describe the underlying relationship between $\hat{\beta}_i$ and spatial structure unless, of course, the similarity between $\{\hat{\beta}_i\}$'s estimated from unconstrained and constrained models is purely coincidental. The Fotheringham and Webber study is guilty of this latter criticism but not of the former.

The three theories which have attempted to demonstrate a link between estimated distance-decay parameters and spatial structure are now reviewed and their weaknesses highlighted. The chapter concludes by briefly reviewing several studies which have proposed methods to eliminate the effect of spatial structure in interaction models.

2.2 The Work of Johnston

Johnston [1973, 1975, 1976] attributed the spatial variation of $\{\hat{\beta}_i\}$ to variations in the range of logarithmic distances from an origin to its destinations. For example, an accessible origin is likely to have very near destinations as well as very distant destinations; hence, the range of logarithmic distances to its destinations will be large. An inaccessible origin, on the other hand, is unlikely to have nearby destinations and hence, in logarithms, the range of distances will be small.

This variation in the range of logarithmic distances is then related to the variation in $\{\hat{\beta}_i\}$ as follows. Assume that in (1.33) $T_{ij} \equiv I_{ij}/m_j^{\gamma_i}$ and that $\gamma_i = 1.0$ for all i. Then the OLS estimator for $\hat{\beta}_i$ is:

$$\hat{\beta}_i = \frac{\text{covariance}_i\ \{T_{ij}^*,\ d_{ij}^*\}}{\text{variance}_i\ \{d_{ij}^*\}} \qquad (2.1)$$

Johnston assumes that $\{T_{ij}^*\}$ is constant for all i so that as the variance of logarithmic distances varies, $\hat{\beta}_i$ will vary. Under these assumptions, there can be no corresponding change in the covariance between $\{T_{ij}^*\}$ and $\{d_{ij}^*\}$ for origin i as the variance of $\{d_{ij}^*\}$ for origin i varies. Thus, from (2.1), as the variance of logarithmic distances increases, $\hat{\beta}_i$ will become less negative (the numerator of (2.1) is assumed to be negative). As the variance decreases, $\hat{\beta}_i$ will become more negative. This is shown graphically by Johnston [1975, p. 282]. Hence, if Johnston's assumptions are correct, the estimated distance-decay parameter will become less negative as the accessibility of an origin increases. Johnston attributes this variation solely to variations in spatial structure where spatial structure is measured by the range of logarithmic distances.

An immediate weakness in Johnston's argument is that it only applies to an unconstrained spatial interaction model. The distance-decay parameter in singly and doubly-constrained interaction models is estimated by MLE and the estimator contains no term similar to the denominator of (2.1). Yet, as already mentioned, the spatial variation of distance-decay parameters derived from constrained models is similar to that derived from an unconstrained model.

Another weakness of Johnston's argument is that there is a very large discrepancy between his assumption of a constant set of interactions over varying sets of distance, and reality. As Sheppard [1979] indicates, Johnston implies that interaction behaviour is solely a function of intervening opportunities and not of distance. Hence, Johnston is discussing a case of model misspecification whereby a gravity-type model is used to describe intervening opportunity behaviour. His argument is irrelevant when interaction behaviour is a function of distance and not of intervening opportunities. This can be shown as follows. Consider the interaction model:

$$T_{ij}^* = \alpha_i^* + \beta_i d_{ij}^* + \varepsilon_{ij}^* \qquad (2.2)$$

where,

$$T_{ij}^* = I_{ij}^* - \gamma_i m_j^*$$

and,

$$\gamma_i = 1.0 \qquad \forall i.$$

The OLS estimator of β_i from (2.2) is given in (2.1). Assume that the variance of $\{d_{ij}^*\}$ increases and estimates of β_i are derived for three situations. In the first, the variance of $\{d_{ij}^*\}$ increases and $\{T_{ij}^*\}$ remains constant. This is the situation envisaged by Johnston. In the second, as the variance of $\{d_{ij}^*\}$ increases, $\{T_{ij}^*\}$ varies as a result of T_{ij}^* being a deterministic function of d_{ij}^*. In the third, as the variance of $\{d_{ij}^*\}$ increases, $\{T_{ij}^*\}$ varies as a result of T_{ij}^* being a stochastic function of d_{ij}^*. The latter two situations are inconsistent with Johnston's discussion but are more representative of reality than is the first situation. The latter two situations assume interaction is a function of distance: the former assumes interaction is a function of intervening opportunities. An estimate of $\hat{\beta}_i$ is derived for each of the three situations and these estimates are compared to the original estimate given in (2.1). The original estimate is denoted by $\hat{\beta}_i$ (OLD) and each new estimate is denoted by $\hat{\beta}_i$ (NEW).

For each of the three situations, the variance of $\{d_{ij}*\}$, is increased by increasing $d_{ij}*$ by an amount proportionate to $d_{ij}*$. Let the increased value of $d_{ij}*$ be denoted by $c_{ij}*$ and define,

$$c_{ij} \equiv \log_n (d_{ij}^{\,z}) \qquad z > 1.0 \tag{2.3}$$

which is equivalent to,

$$c_{ij}* \equiv z\, d_{ij}* . \tag{2.4}$$

Let $\mathrm{Var}_i\{x\}$ and $\mathrm{Cov}_i\{x, y\}$ denote the variance of $\{x\}$ for origin i and the covariance of $\{x\}$ and $\{y\}$ for origin i respectively. Then,

$$\mathrm{Var}_i\{c_{ij}*\} = \frac{1}{n}\sum_{j=1}^{n}(c_{ij}* - \bar{c}_{ij}*)^2 \tag{2.5}$$

$$= \frac{1}{n}\sum_{j=1}^{n}(zd_{ij}* - z\bar{d}_{ij}*)^2$$

$$= z^2\,\mathrm{Var}_i\{d_{ij}*\} . \tag{2.6}$$

The variance of $\{d_{ij}*\}$ has then been increased by a factor of z^2. Consider the new estimate of β_i, given this increase, for each of the three situations described above.

(i) $\{T_{ij}^*\}$ *remains constant*

$$\mathrm{Cov}_i\{T_{ij}^*, c_{ij}^*\} = \frac{1}{n}\sum_{j=1}^{n}(T_{ij}^* - \bar{T}_{ij}^*)(c_{ij}^* - \bar{c}_{ij}^*) \tag{2.7}$$

$$= \frac{1}{n}\sum_{j=1}^{n}(T_{ij}^* - \bar{T}_{ij}^*)(z\ d_{ij}^* - z\ \bar{d}_{ij}^*)$$

$$= z\ \mathrm{Cov}_i\{T_{ij}^*, d_{ij}^*\} \tag{2.8}$$

From (2.1),

$$\hat{\beta}_i\,(\mathrm{NEW}) = \frac{\mathrm{Cov}_i\{T_{ij}^*, c_{ij}^*\}}{\mathrm{Var}_i\{d_{ij}^*\}} \tag{2.9}$$

$$= \frac{z\ \mathrm{Cov}_i\{T_{ij}^*, d_{ij}^*\}}{z^2\ \mathrm{Var}_i\{d_{ij}^*\}} \tag{2.10}$$

$$= \frac{\hat{\beta}_i\,(\mathrm{OLD})}{z} \tag{2.11}$$

Equation (2.11) states that as the $\mathrm{Var}_i\{d_{ij}^*\}$ increases by a factor of z^2, where $z > 1.0$, $|\hat{\beta}_i|$ decreases by a factor of z. This is consistent with the empirical findings described in Section 1.5 since, as mentioned, the variation of logarithmic distances is greater for accessible origins than for inaccessible origins. However, the result depends on the assumption that $\{T_{ij}^*\}$ remains constant even when $\{d_{ij}^*\}$ varies. Since the model being calibrated describes T_{ij}^* as a function of d_{ij}^*, the assumption is a false one.

(ii) There is a deterministic relationship between T_{ij}^* and d_{ij}^*.

Let,

$$M_{ij}^* = \alpha_i^* + \beta_i\, c_{ij}^* \tag{2.12}$$

where M_{ij} represents the new value of T_{ij} when the distance between i and j is increased from d_{ij} to c_{ij}.

Then,

$$\mathrm{Cov}_i\{M_{ij}^*, c_{ij}^*\} = \frac{1}{n}\sum_{j=1}^{n}(M_{ij}^* - \bar{M}_{ij}^*)(c_{ij}^* - \bar{c}_{ij}^*) \tag{2.13}$$

$$= \frac{1}{n}\sum_{j=1}^{n}(\alpha_i^* + \beta_i z d_{ij}^* - (\alpha_i^* + \beta_i z\bar{d}_{ij}^*))(zd_{ij}^* - z\bar{d}_{ij}^*)$$

$$= z^2\,\frac{1}{n}\sum_{j=1}^{n}\left(\frac{\alpha_i^*}{z} + \beta_i d_{ij}^* - \left(\frac{\alpha_i^*}{z} + \beta_i\bar{d}_{ij}^*\right)\right)(d_{ij}^* - \bar{d}_{ij}^*)\ . \tag{2.14}$$

$$= z^2\,\frac{1}{n}\sum_{j=1}^{n}(\alpha_i^* + \beta_i d_{ij}^* - (\alpha_i^* + \beta_i\bar{d}_{ij}^*))(d_{ij}^* - \bar{d}_{ij}^*) \tag{2.15}$$

$$= z^2\,\frac{1}{n}\sum_{j=1}^{n}(T_{ij}^* - \bar{T}_{ij}^*)(d_{ij}^* - \bar{d}_{ij}^*)$$

$$= z^2\,\mathrm{Cov}_i\{T_{ij}^*, d_{ij}^*\}\ . \tag{2.16}$$

From (2.12),

$$\hat{\beta}_i\ (\mathrm{NEW}) = \frac{\mathrm{Cov}_i\{M_{ij}^*, c_{ij}^*\}}{\mathrm{Var}_i\{c_{ij}\}} \tag{2.17}$$

$$= \frac{z^2\ \mathrm{Cov}_i\{T_{ij}^*, d_{ij}^*\}}{z^2\cdot \mathrm{Var}_i\{d_{ij}^*\}} \tag{2.18}$$

$$= \hat{\beta}_i\ (\mathrm{OLD}) \tag{2.19}$$

Thus, when $Var_i\{d_{ij}{}^*\}$ varies and $T_{ij}{}^*$ is a deterministic function of $d_{ij}{}^*$, there is no variation in $\hat{\beta}_i$.

(iii) There is a stochastic relationship between $T_{ij}{}^*$ and $d_{ij}{}^*$.

Let,

$$M_{ij}{}^* = \alpha_i{}^* + \beta_i\, c_{ij}{}^* + e_{ij}{}^* , \tag{2.20}$$

and then,

$$Cov_i\{M_{ij}{}^*, c_{ij}{}^*\} = \frac{1}{n}\sum_{j=1}^{n}(M_{ij}{}^* - \bar{M}_{ij}{}^*)(c_{ij}{}^* - \bar{c}_{ij}{}^*) \tag{2.21}$$

$$= z^2 \frac{1}{n}\sum_{j=1}^{n}\left(\frac{\alpha_i{}^*}{z} + \beta_i\, d_{ij}{}^* + \frac{e_{ij}{}^*}{z} - \left(\frac{\alpha_i{}^*}{z} + \beta_i \bar{d}_{ij}{}^*\right)\right)(d_{ij}{}^* - \bar{d}_{ij}{}^*)$$

$$= z^2 \frac{1}{n}\sum_{j=1}^{n}\left(T_{ij}{}^* - \bar{T}_{ij}{}^* + \frac{e_{ij}{}^*}{z}\right)(d_{ij}{}^* - \bar{d}_{ij}{}^*)$$

$$= z^2\, Cov_i\{T_{ij}{}^*, d_{ij}{}^*\} + z\,\frac{1}{n}\sum_{j=1}^{n} e_{ij}{}^*\,(d_{ij}{}^* - \bar{d}_{ij}{}^*). \tag{2.22}$$

From (2.20), assuming $e_{ij}{}^*$ is a random variable with zero mean and zero covariance with $c_{ij}{}^*$,

$$\hat{\beta}_i\ (\text{NEW}) = \frac{Cov_i\{M_{ij}{}^*, c_{ij}{}^*\}}{Var_i\{c_{ij}{}^*\}} \tag{2.23}$$

$$= \frac{z^2\, Cov_i\{T_{ij}{}^*, d_{ij}{}^*\} + z\,\frac{1}{n}\sum_{j=1}^{n} e_{ij}{}^*\,(d_{ij}{}^* - \bar{d}_{ij}{}^*)}{z^2\, Cov_i\{d_{ij}{}^*\}}$$

$$= \hat{\beta}_i\ (\text{OLD}) + \frac{\frac{1}{n}\sum_{j=1}^{n} e_{ij}{}^*\,(d_{ij}{}^* - \bar{d}_{ij}{}^*)}{z\, Var_i\{d_{ij}{}^*\}} \tag{2.24}$$

The term $\dfrac{\frac{1}{n}\sum_{j=1}^{n} e_{ij}^{*}\,(d_{ij}^{*} - \bar{d}_{ij}^{*})}{z\ \mathrm{Var}_i\{d_{ij}^{*}\}}$ is a random variable and hence, $\hat{\beta}_i$(NEW) is not related to $\mathrm{Var}_i\{d_{ij}^{*}\}$.

Thus, Johnston's explanation for the spatial variation of $\hat{\beta}_i$ is only true when T_{ij}^{*} is not a function of d_{ij}^{*}. It is an irrelevant argument when T_{ij}^{*} is either a deterministic or stochastic function of d_{ij}^{*} — the usual occurrence. Sheppard [1979a] shows, for example, that Johnston's ideas explain none of the variance in the estimated distance-decay parameters reported by Gould and Leinbach (see Section 1.5) because the interactions in each case were a function of distance and not intervening opportunity.

Since it is commonly accepted that interactions are usually a function of distance, Johnston's ideas can be dismissed as an explanation of the effect of spatial structure upon $\hat{\beta}_i$. There is likely to be some other spatial structure effect that determines $\hat{\beta}_i$ and if there is, this has a further implication for the relevance of Johnston's work. Even where interactions are a function of intervening opportunities, the variance of $\{d_{ij}^{*}\}$ is unlikely to be as strong a determinant of $\hat{\beta}_i$ as Johnston imagines since the spatial structure effect that occurs in gravity models is also likely to occur in intervening opportunity models. In terms of (2.1) such a result may occur by the numerator varying as the denominator varies. Johnston does not discuss this possibility although it could nullify the effects of varying $\{d_{ij}^{*}\}$.

2.3 The Work of Curry, Griffith and Sheppard

Curry [1972] hypothesised that in an unconstrained spatial interaction model the estimated distance-decay parameter depends on the spatial autocorrelation of the mass terms present in the model. Mass, or population size, is a commonly-used measure of attractiveness and propulsiveness and the spatial autocorrelation of masses is a description of the spatial structure of the system under investigation. Curry's ideas were contested in an ensuing discussion (Curry, Griffith and Sheppard [1975], Sheppard, Griffith and Curry [1976], and Cliff, Martin and Ord [1974, 1975, 1976]). Subsequently, separate studies were undertaken by Griffith and Jones [1980] and by Sheppard [1979a]. Briefly, the hypothesised relationship between the spatial autocorrelation of masses and the estimation of β can be stated as follows.

Consider the unconstrained spatial interaction model given in (1.10). If the masses of the spatial system under investigation are spatially autocorrelated, and there are k origins and n destinations, then the following equations hold:

$$m_i = \tilde{m}_i + \sum_{\substack{j=1 \\ j\neq i}}^{n} \phi_{ij}(s)\, w_{ij}\, m_j \tag{2.25}$$

and

$$m_j = \tilde{m}_j + \sum_{\substack{i=1 \\ i\neq j}}^{k} \phi_{ij}(s)\, w_{ij}\, m_i \tag{2.26}$$

where m_i is the mass of centre i belonging to the vector M whose elements are spatially autocorrelated,

$\tilde{m}_i$ is the mass of centre i belonging to the vector $\tilde{M}$ whose elements are not spatially autocorrelated,[1]

---

[1]In matrix notation, $\tilde{M} = (I - A)\, M$

where I is the identity matrix, the typical elements of $\tilde{M}$ and M are $\tilde{m}_i$ and m_i respectively and the typical element of A, a_{ij} is $\phi_{ij}(s)\, w_{ij}$.

$\phi_{ij}(s)$ is the spatial autocorrelation statistic of the mass terms at a lag of distance s where s is the distance between i and j,

and w_{ij} is a connectivity measure.

Equations (2.25) and (2.26) can be substituted into (1.10) to produce:

$$I_{ij} = \alpha(\tilde{m}_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} \phi_{ij}(s)\, w_{ij}\, m_j)^{\omega}(\tilde{m}_j + \sum_{\substack{i=1 \\ i \neq j}}^{k} \phi_{ij}(s)\, w_{ij}\, m_i)^{\gamma} d_{ij}^{\beta}. \quad (2.27)$$

The spatial autocorrelation statistic, $\phi_{ij}(s)$, is a function of distance and Curry [1972, p. 132] states that it can often be represented by a negative power function of distance. Hence, the mass terms in (1.10) are, in part, a function of distance and the calibration of this model by OLS regression would produce an estimate of β which includes this mass-distance relationship and estimates of ω and γ which include the distance-interaction relationship. Curry also notes that it is only when masses are not spatially autocorrelated that the distance-decay parameter in an unconstrained spatial interaction model is wholly a measure of behaviour. This, he states, explains the large spatial variation of $\hat{\beta}_i$ since mass-distance relationship vary widely over space even if behaviour is constant.

The work of Curry, Griffith and Sheppard can be interpreted as investigating the effects of multicollinearity in OLS regression. In (2.27), for example, it is assumed that the autocorrelation statistic is a negative power function of distance so that the total effect of the explicit distance function is hidden. The true relationship between distance and interaction is not given by $\hat{\beta}$ but it is included within $\hat{\beta}$,

$\tilde{\omega}$, $\tilde{\gamma}$. To demonstrate this, the intermediate step of introducing spatial autocorrelation is unnecessary. If there is a relationship between mass and distance, (1.10) can be rewritten as:

$$I_{ij} = \alpha(\tilde{m}_i d_{ij}^{\lambda_i})^{\omega} (\tilde{m}_j d_{ij}^{\lambda_j})^{\gamma} d_{ij}^{\beta} , \tag{2.28}$$

or in natural logarithms:

$$I_{ij}^* = \alpha^* + \omega(\tilde{m}_i^* + \lambda_i d_{ij}^*) + \gamma(\tilde{m}^* + \lambda_j d_{ij}^*) + \beta d_{ij}^* , \tag{2.29}$$

where $\tilde{m}_i^*$ is the mass of centre i in logarithms which is not a function of distance, and λ_i is the elasticity of mass with respect to distance. It is clear from (2.29) that β may not be an accurate measure of the relationship between distance and interaction when $\lambda_i \neq 0$ or when $\lambda_j \neq 0$. This is an effect of multicollinearity in OLS regression and this is the essence of the Curry, Griffith and Sheppard papers. Their work is confusing, however, because it introduces and emphasises an intermediate step, spatial autocorrelation, which need not be discussed. What is important in determining $\hat{\beta}$ is the linear correlation between mass and distance. As Cliff, Martin and Ord [1974, p. 282] note:

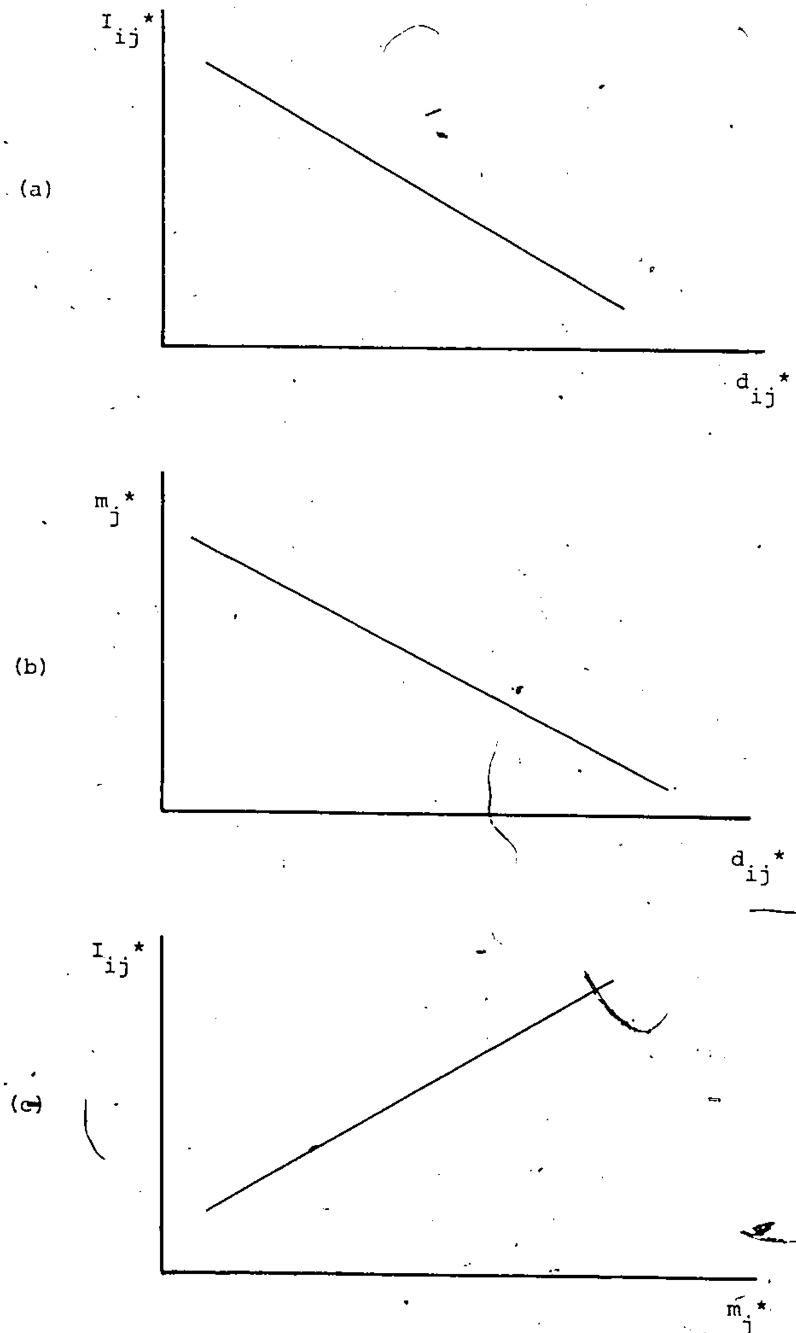
> In such a situation (multicollinearity)[1], the precision of the coefficient estimates would be small so that it would be difficult to disentangle the relative influence of X_1 and X_2, the population and distance variables.

As a simple example, consider the situation given in Figure 2.1. It is impossible here to determine the true effects of mass and distance upon interaction. On one hand, the apparent relationship between mass and distance in 2.1a may be entirely spurious. It may simply be a result of mass decreasing as distance increases, (2.1b), and interaction increasing

---

[1] The wording inside the brackets is mine.

FIGURE 2.1: Relationships between Interaction, Distance and Mass.



as mass increases (2.1c). On the other hand, the apparent relationship between mass and interaction may be entirely spurious. It may simply be a result of interaction decreasing as distance increases and mass decreasing as distance increases. This subject is expanded and discussed in detail in Chapter 4.

The intermediary step of spatial autocorrelation used by Curry, Griffith and Sheppard is defended by Sheppard [1979a] as being a more general measure of the relationship between mass and distance than the specific linear relationship implied by a linear correlation coefficient. However, the OLS estimator of β is not a function of any spatial autocorrelation measure but it is a function of the linear correlation coefficient between the logarithm of mass and the logarithm of distance. Chapter 4 demonstrates this explicitly. It is not variations in the spatial autocorrelation of mass terms that are responsible, wholly or in part, for the spatial variation in origin-specific distance-decay parameters; rather, it is the degree of multicollinearity between mass, in logarithms and distance in logarithms that can produce biased estimates of β. Spatial variations in multicollinearity would then produce spatial variations in estimated distance-decay parameters. The use of a spatial autocorrelation statistic is simply a surrogate measure for the linear correlation coefficient between the logarithm of mass and the logarithm of distance. There is no theoretical reason for the use of the former as a determinant of the spatial variation of $\hat{\beta}_i$. The theory given in Chapter 4 indicates that the relationship between multicollinearity and $\hat{\beta}_i$ is very complex and the use of a spatial autocorrelation statistic to measure this relationship would be a gross oversimplification.

Thus, the discussion relating variations in spatial autocorrelation of masses to variations in $\hat{\beta}_i$ is misleading. Spatial autocorrelation is a surrogate for the degree of multicollinearity between m_j^* and d_{ij}^* present in the data. The problem which Curry, Griffith and Sheppard address should then be restated in three parts. One, how does the degree of multicollinearity between m_j^* and d_{ij}^* affect $\hat{\beta}_i$? Two, if there is a systematic relationship between $\hat{\beta}_i$ and multicollinearity, is the degree of multicollinearity likely to vary over space? It should do if multicollinearity is to explain the spatial pattern of $\hat{\beta}_i$ values commonly found. This problem is not investigated thoroughly by Curry, Griffith and Sheppard but Cliff, Martin and Ord [1974] suggest that multicollinearity is likely to be a problem in intra-urban interaction but not in inter-urban interaction . Three, the analysis of Curry, Griffith and Sheppard is specific to an unconstrained interaction model yet there is empirical evidence that distance-decay parameters estimated from constrained models have similar spatial patterns to those estimated from unconstrained models. The problem is then to determine whether there is some other spatial structure effect that is independent of the method of calibration. If there is such an effect, how much of the variation in $\hat{\beta}_i$ is due to this effect and how much can be attributed to variations in multicollinearity?

Each of these problems is investigated in later sections of this thesis.

2.4 The Work of Fotheringham and Webber

Assume that in a region there are k small centres and n large centres. Interaction occurs between each small centre i and each large centre j in a manner described by the unconstrained stochastic version of (1.12):

$$I_{ij}^* = \alpha^* + \omega\, m_i^* + \gamma\, m_j^* + \beta\, d_{ij}^* + \varepsilon_{ij}^* \qquad (2.30)$$

where ε_{ij} is an error term which is assumed to be an independent random variable. It is well-known that under the usual assumptions of OLS regression, the estimators $\hat{\alpha}$, $\hat{\omega}$, $\hat{\gamma}$, and $\hat{\beta}$ are unbiased and consistent estimators of their respective population parameters.

Fotheringham and Webber [1980] used various branches of geographic theory to show that (2.30) is one equation of a set of simultaneous linear equations. For example, growth centre theory suggests that the mass of the small centres is a function of their interactions with the large centres. The greater the interaction a small centre has with a large centre, the more the small centre grows. One particular form of this simultaneous relationship given by Fotheringham and Webber is:

$$I_{ij}^* = \alpha^* + \omega\, m_i^* + \gamma\, m_j^* + \beta\, d_{ij}^* + \varepsilon_{ij}^* \qquad (2.31)$$

$$m_i^* = \rho + \phi \sum_{j=1}^{n} I_{ij}^* + \mu_i^* \qquad (2.32)$$

where μ_i^* is an error term.

If these equations do describe reality, then m^* is not independent of ε^* in (2.31) and hence biased and inconsistent parameter estimates are produced by OLS regression on (2.31). Thus, $\hat{\beta}$ would be a biased and inconsistent estimator of β due to the intrusion of a spatial structure effect. Spatial structure in this case is measured by the relationship between mass and interaction or, since interaction is a function of distance, between mass and distance. Fotheringham and Webber show how the inconsistency of $\hat{\beta}$ can be removed by calibrating the system given in (2.31) and (2.32) by two-stage least squares regression or by indirect least squares regression (the latter can be used since the system is exactly identified). The resulting equilibrium interaction equation is:

$$I_{ij}^* = a + b \sum_{j=1}^{n} d_{ij}^* + \gamma m_j^* + \beta d_{ij}^* + \mu_{ij}^* . \qquad (2.33)$$

Thus, $\hat{\beta}$ should be obtained from (2.33) and not from (2.31) if a simultaneous equation system exists. However, the value of $\hat{\beta}$ obtained from (2.33) still does not represent the total distance-decay effect on interaction, since distance terms occur twice in the model. This can be seen by expanding (2.33) in terms of a specific destination, k, to obtain:

$$I_{ik}^* = a + b \sum_{\substack{j=1 \\ j \neq k}}^{n} d_{ij}^* + \gamma m_k^* + (b + \beta) d_{ik}^* + \mu_{ik}^* . \qquad (2.34)$$

The total distance-decay effect is then measured by $(b + \beta)$.

Fotheringham and Webber have thus shown that there are two sources of error in estimating the distance-decay parameter from (2.31) rather than from (2.34). Firstly, the OLS estimator of β using (2.31) is biased and inconsistent. Secondly, (2.34) shows that the reduced form distance-decay parameter is $b + \beta$, not β. Their paper demonstrated theoretically that accurate estimates of β can only be obtained if other elements of the spatial system are modelled too: in particular, $\hat{\beta}$ depends upon spatial structure as well as upon interaction behaviour. In the example given above, growth centre theory is used to define a simultaneous equation system. There is, however, no single equation system which is correct, but different systems, utilising different branches of geographic theory, should be used in different problems.

However, although their theory is sound, three criticisms can be levelled at the Fotheringham and Webber approach. One is that the method appears to have limited application. If a simultaneous equation system is not present, then obviously the method is not applicable. There may be situations, for example, in which mass is not a function of distance, although the method can easily be generalised so that it depends upon mass being a function of general accessibility - a more frequently found relationship. This generalisation is given in Section 2.5 and removes much of the criticism of limited application.

A second criticism is that the relationship between mass and interaction is fairly arbitrary. Other functions of interaction may be equally plausible, *a priori*, but which do not give a linear equation in logarithms and would be much more difficult to calibrate.

The third criticism is the one that also applies to the work of Johnston and of Curry, Griffith and Sheppard. Fotheringham and Webber's work is again solely concerned with unconstrained interaction modelling and no explanation can be given for the effect of spatial structure that appears to be present in constrained modelling.

2.5 A Generalisation of the Fotheringham and Webber Method

Consider the stochastic version of the origin-specific unconstrained interaction model given in (1.33). This is written as:

$$I_{ij} = \alpha_i \, m_j^{\gamma_i} \, d_{ij}^{\beta_i} \, \varepsilon_{ij} \qquad (2.35)$$

where ε_{ij}^* is an error term having the usual properties assumed in OLS regression (see Fotheringham and Webber [1980]). Biased and inconsistent estimates of α_i, γ_i and β_i are obtained from the calibration of (2.35) by OLS regression when mass is a function of interaction. In the Fotheringham and Webber technique the mass of small centres is related to interactions with large centres. Thus, the discussion was confined to interactions between small and large centres and more accurate values of $\hat{\beta}$ could only be obtained for such interactions. However, if the relationship between mass and interaction given by Fotheringham and Webber is improved upon, a more general result can be obtained which can be applied to any origin-specific interaction. The distinction between small and large centres can then be eliminated.

The mass of a centre is in part a function of the volume of interaction between itself and all other centres. A centre having a large volume of interaction with surrounding centres will grow rapidly while a centre having little interaction with other centres will stagnate. However, the mass of a centre is not an equal function of all interactions.[1] Interactions with larger centres are more important to the growth of a particular centre than are interactions with smaller centres. The hypothesis is that the relationship between the mass of destinations and interaction is of the following form:

$$m_j = \delta \prod_{i=1}^{n} I_{ij}^{\lambda} m_i^{\phi} \nu_j \qquad (2.36)$$

[1] This is the assumption made by Fotheringham and Webber. Since they only discuss interactions between small centres and large centres though, and the large centres have approximately equal masses, their assumption is not an unlikely one. When all interactions are modelled, however, it is a poor assumption.

where n is the number of origins in the system and ν_j is an error term. This relationship holds for all centres in the system and each centre in the system is assumed to be both an origin and a destination. The total system can then be described, in natural logarithms, as:

$$I_{ij}^* = \alpha_i^* + \gamma_i m_j^* + \beta_i d_{ij}^* + \varepsilon_{ij}^* \qquad (2.37)$$

$$m_j^* = \delta^* + \lambda \sum_{i=1}^{n} I_{ij}^* + \phi \sum_{i=1}^{n} m_i^* + \nu_j^* \qquad (2.38)$$

Since both equations are exactly-identified, either two-stage least squares or indirect least squares regression can be used to obtain a consistent estimate of β. The reduced form equations of (2.37) and (2.38), from which $\hat{\beta}$ is derived, are, respectively:

$$I_{ij}^* = A_i + B_i \sum_{i=1}^{n} \beta_i d_{ij}^* + \beta_i d_{ij}^* + R_{ij} \qquad (2.39)$$

$$m_j^* = C + D \sum_{i=1}^{n} \beta_i d_{ij}^* + E_j \qquad (2.40)$$

where R_{ij} and E_j are error terms. The derivation of (2.39) and (2.40) is given in Appendix I. Estimating (2.39) by OLS regression produces a consistent estimate of β_i since none of the independent variables is correlated with the error term. However, the term $\beta_i d_{ij}$ occurs twice so that the interaction can be rewritten for a specific origin k as:

$$I_{kj}^* = A_k + B_k \sum_{\substack{i=1 \\ i\neq k}}^{n} \beta_i d_{ij}^* + (B_k \beta_k + \beta_k) d_{kj}^* + \varepsilon_{kj} \qquad (2.41)$$

The parameter to be estimated is then $(B_k \beta_k + \beta_k)$ and not β_k.

Thus, the Fotheringham and Webber technique can easily be generalised to estimate the distance-decay parameter specific to any set of origin interactions. However, the generalisation given here does not apply to the set of interactions for the complete spatial system. For example, if m_i is also a function of I_{ij}, then the simultaneous-equation system would be:

$$I_{ij}^* = \alpha^* + \gamma\, m_j^* + \rho\, m_i^* + \beta\, d_{ij}^* + \varepsilon_{ij}^* \qquad (2.42)$$

$$m_j^* = \delta^* + \lambda \sum_{i=1}^{n} I_{ij}^* + \phi \sum_{i=1}^{n} m_i^* + \nu_j^* \qquad (2.43)$$

$$m_i^* = \psi^* + \omega \sum_{j=1}^{n} I_{ij}^* + \tau \sum_{j=1}^{n} m_j^* + \chi_i^* \qquad (2.44)$$

Each equation of this system is underidentified and the parameters of the system cannot be calibrated by indirect or two-stage least squares regression.

2.6 Existing Methods to Account for Spatial Structure in Interaction Modelling

A sound theoretical method to obtain estimated distance-decay parameters which are free from any effect of spatial structure depends upon a thorough understanding of the relationship between the two. Since the previous sections have demonstrated that the latter is not yet available, it is highly unlikely that any existing method is complete. Four such methods are known. The first is the two-stage or indirect least squares estimation procedure outlined by Fotheringham and Webber, which is not totally satisfactory.

The second results from the work by Curry, Griffith and Sheppard and is given by Griffith and Jones [1980]. Using the same notation as in Section 2.3, Griffith and Jones propose that the unconstrained model which should be calibrated in order to obtain estimated distance-decay parameters free from spatial structure effects is not (1.10) but:

$$\tilde{I}_{ij} = \alpha \, \tilde{m}_i^{\,\omega} \, \tilde{m}_j^{\,\gamma} \, \tilde{d}_{ij}^{\,\beta} \qquad (2.45)$$

where "~" denotes a variable which has zero spatial autocorrelation. In matrix notation:

$$\tilde{M} = (I - \rho W) \; M \qquad (2.46)$$

$$\tilde{N} = (I - \gamma W) \; N \qquad (2.47)$$

$$\tilde{T} = (I - \rho W) \; T \; (I - \gamma W)^{Tr} \qquad (4.28)$$

where the typical elements of $\tilde{M}$, $\tilde{N}$ and $\tilde{T}$ are, respectively, $\tilde{m}_i$, $\tilde{m}_j$ and $\tilde{I}_{ij}$, while the typical elements of M, N and T are, respectively, m_i m_j and I_{ij}. W is a connectivity matrix and ρ and γ are spatial autocorrelation statistics. Tr represents the transpose of a matrix.

However, as Griffith and Jones note, (2.45) is very difficult to calibrate. There may be non-positive values in $\tilde{M}$, $\tilde{N}$ and $\tilde{T}$ which preclude the use of logarithmic regression and since the first partial derivatives with respect to the parameters involve logarithms, a non-linear regression technique cannot be employed for the same reason. Another, more serious disadvantage of this technique is that it is unlikely to remove all, or much, of the spatial structure effect from $\{\hat{\beta}_i\}$ since the use of spatial autocorrelation measures in this context has been shown to be very dubious (see Section 2.3).

The third method that has been proposed to account for the effect of spatial structure in interaction modelling is that given by Ewing [1974] and Cesario [1975]. Their method is specific to unconstrained spatial interaction models. Very simply, they state that biased parameter estimates will result from the calibration of (1.10) since it misspecifices reality: no account is taken of alternative destinations. Consider (1.10),

$$I_{ij} = \alpha\, m_i^{\omega}\, m_j^{\gamma}\, d_{ij}^{\beta} \quad . \tag{2.49}$$

The attractiveness of j for interaction with i is $m_j^{\gamma} d_{ij}^{\beta}$. As the mass of j increases, its attractiveness increases but as its distance from i increases, its attractiveness to i decreases. Ewing and Cesario proposed that this is not a true measure of attractiveness but that it should be standardised by dividing by the sum of attractiveness measures to all destinations. This produces a so-called "spatial choice" model,

$$I_{ij} = \alpha\, m_i^{\omega} \frac{m_j^{\gamma}\, d_{ij}^{\beta}}{\sum_{j=1}^{n} m_j^{\gamma}\, d_{ij}^{\beta}} \tag{2.50}$$

which, they argued, would control for competing destinations and would produce unbiased estimates of β.

However, it is obvious that (2.50) is simply a production-constrained interaction model in which $O_i = \alpha\, m_i^{\omega}$. Section 1.5 has already shown that distance-decay parameters estimated from this model appear to be as equally biased by spatial structure as those derived from (2.49). Also Ewing and Cesario's hypothesis that the denominator of (2.50) accounts for competing destinations is wrong. The inclusion of the denominator simply produces a relative accessibility measure for one destination with respect to all others. The model given in (2.50) does not necessarily differentiate between situations in which a destination is clustered with other destinations as opposed to being completely isolated. A destination could be isolated and be relatively accessible or it could be clustered and be relatively inaccessible. In either case, Ewing and Cesario's standardisation would not model what they propose. Their idea is a good one in that competing destinations are likely to be a factor in explaining interactions but their modelling of this wrong. The theme of competing destinations is pursued in later sections and is shown to be a prime reason for the spatial structure component in estimated distance-decay parameters.

The fourth method that has been proposed to eliminate the relationship between spatial structure and estimated distance-decay parameters is given by Gordon [1976]. In a similar manner to Ewing and Cesario, Gordon hypothesised that accessibility measures should be added to the unconstrained interaction model in order that unbiased estimates of β_i can be obtained. He proposed that this should be done by adding origin and destination constraints on the estimated interactions to produce the doubly-constrained spatial interaction model given in (1.21). Comparison of the calibration of the unconstrained and doubly-constrained models was undertaken by Gordon with the British freight flow data used by Chisholm and O'Sullivan [1973] and by Frost [1969]. The results showed that the variation in $\{\hat{\beta}_i\}$ was approximately halved when β_i was estimated from the doubly-constrained model as opposed to the unconstrained model.

However, there was no significant change in the correlation coefficient between $\hat{\beta}_i$ and a measure of the accessibility of i, which was high for both models. While the decrease in variation of $\{\hat{\beta}_i\}$ is encouraging, the strong relationship between accessibility and $\{\hat{\beta}_i\}$ suggests that there is some effect of spatial structure still not accounted for in a doubly-constrained model. Further investigation is needed into the nature of the balancing factors of constrained models and their relationship with $\hat{\beta}_i$. It would also be extremely useful if some method of controlling for spatial structure in an unconstrained model could be found which does not eliminate the useful property of being able to estimate such a model by regression techniques.

In constrained spatial interaction modelling, none of the existing theories relating spatial structure to distance-decay parameters is entirely satisfactory, while in constrained modelling such theories are non-existant. More investigation into the relationship between $\hat{\beta}_i$ and spatial structure is clearly needed and Chapter 3 develops a paradigm for such investigation.

CHAPTER THREE

A PARADIGM RELATING SPATIAL STRUCTURE TO ESTIMATED DISTANCE-DECAY PARAMETERS

3.1 Introduction

Chapters 1 and 2 introduced the topic of distance-decay parameters and outlined existing hypotheses relating the estimated values of distance-decay parameters to spatial structure. A major conclusion from the outline was that much more investigation into the nature of this relationship is needed. The purposes of this brief chapter are twofold. First, it indicates, and justifies, the directions subsequent research will take and second, it provides a paradigm for this research into which Chapters 4, 5 and 6 can be categorised. Thus, while it is very brief, this chapter can be considered as a cornerstone in this thesis since it connects existing theories on distance-decay parameters and spatial structure to new theories which are presented in following chapters.

3.2 Directions for Research in Subsequent Chapters

Chapter 1 described how the traditional interpretation of β is a purely behavioural one. Distance-decay parameters are assumed to describe interaction behaviour with respect to distance, and, as such, they are assumed to be measures of distance-disutility. Section 1.5, however, demonstrated that the spatial variation of $\hat{\beta}_i$ often has a regular pattern and that $\hat{\beta}_i$ and the accessibility of i are related.

This suggests that $\hat{\beta}_i$ may be related to spatial structure. Chapter 2 outlined attempts at demonstrating this relationship theoretically. None was totally satisfactory although some light was shed on important questions to be asked about the relationship between spatial structure and estimated distance-decay parameters.

The theories of Curry, Griffith and Sheppard and of Fotheringham and Webber are both essentially concerned with the effects of multicollinearity upon estimated parameter values in a regression model. However, only in the Fotheringham and Webber study is this made explicit and then their theory is primarily concerned with simultaneous equation systems. It would be interesting and useful to determine the exact relationship between multicollinearity and parameter estimation. What effect does multicollinearity have on parameter estimates? Is this the source of distance-decay parameter variations when these parameters are derived from an unconstrained spatial interaction model calibrated by OLS regression?

The work of Ewing, Cesario and Gordon raises the question of a possible relationship between the constraints of constrained spatial interaction models and $\hat{\beta}_i$ through the inclusion of balancing factors into the models. They only suggest that such a relationship may occur, however, and no theory is given for such a relationship. What is the exact relationship between $\hat{\beta}_i$ and the balancing factors of constrained spatial interaction models? Do the balancing factors eliminate the spatial variation of $\hat{\beta}_i$? The empirical evidence suggests that doubly-constrained interaction models may be better than singly-constrained models in reducing the spatial variation of $\hat{\beta}_i$. If this is so, why is it so? Alternatively, can the addition of constraints produce a relationship between $\hat{\beta}_i$ and spatial structure since the balancing factors are quasi-accessibility measures?[1]

The third area of serious questioning is that hinted at throughout Chapter 2. Every existing theory that has proposed a link between spatial structure and estimated distance-decay parameters is specific to an unconstrained interaction model. No theory is given which could explain such a relationship in unconstrained and constrained modelling, yet, Section 1.5 demonstrated that distance-decay parameters estimated from constrained models also appear to be a function of spatial structure. Is there some spatial structure effect that is common to both types of model, or more generally, to both types of calibration method? If there is, what is the relative importance of such a calibration-independent effect compared to calibration-specific effects such as multicollinearity in unconstrained models or $\hat{\beta}_i$ being a function of the balancing factors in constrained modelling?

Thus, if a relationship between $\hat{\beta}_i$ and spatial structure exists and is to be determined, three major questions need to be answered. One, in unconstrained interaction models, what is the relationship between $\hat{\beta}_i$ and the degree of multicollinearity between mass and distance? Two, in constrained interaction models, what is the relationship between $\hat{\beta}_i$ and the balancing factors included in the models? Three, in both types of model, is there a calibration-independent spatial structure effect and, if there is, how important is it in determining the spatial variation of $\hat{\beta}_i$? Answers to these questions correspond to the theoretical work given in Chapters 4, 5 and 6 respectively.

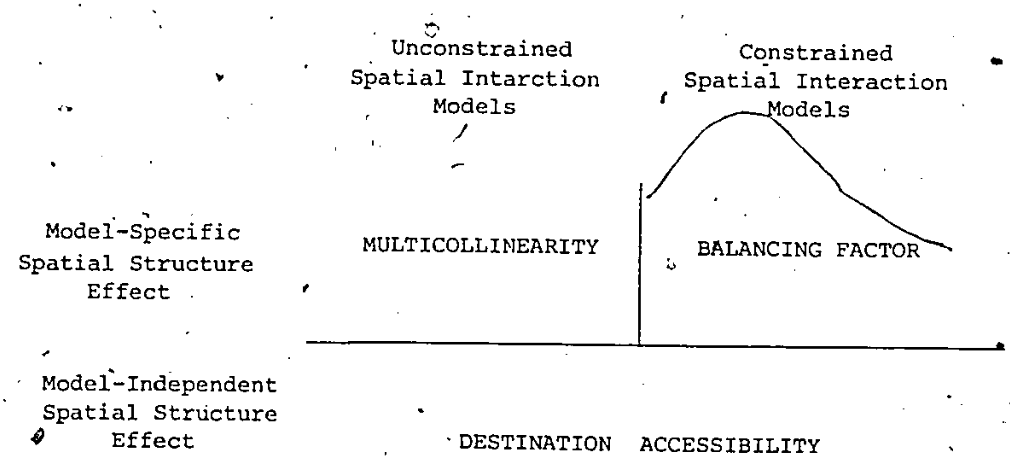
---

[1]For example, the right-hand side of (1.14) is the inverse of potential accessibility defined in Section 1.5.

3.3 The Paradigm

From the above reasoning, it is hypothesised that there are two effects of spatial structure present in estimated distance-decay parameters: a model-specific effect and a model-independent effect. The model-specific effect is subdivided into unconstrained and constrained model effects. In unconstrained models, the model-specific spatial structure effect is related to multicollinearity while in constrained models it is related to the presence of balancing factors which are functions of accessibility. The model-independent effect is hypothesised to be a function of the varying accessibility of centres within a spatial system. These hypothesised relationships between spatial structure and $\hat{\beta}_i$ are presented in Figure 3.1.

FIGURE 3.1: A Paradigm Relating Spatial Structure to $\hat{\beta}_i$.



A link between the model-independent spatial structure effect and the constrained model spatial structure effect is envisaged since the balancing factors are functions of accessibility.

The paradigm given in Figure 3.1 is presented at this stage, rather than as a conclusion, since it clarifies the organization of subsequent chapters. Chapter 4 is concerned with multicollinearity; Chapter 5 is concerned with balancing factor effects; and Chapter 6 is concerned with the effects of varying accessibility. It is useful to be aware of the connections between these chapters.

CHAPTER FOUR

MODEL-SPECIFIC SPATIAL STRUCTURE EFFECTS IN AN UNCONSTRAINED INTERACTION MODEL: MULTICOLLINEARITY AND $\hat{\beta}$

4.1 Introduction

Surprisingly little is known about the effects of multicollinearity on parameter estimates derived from OLS regression. Standard econometric texts such as Christ [1966] and Huang [1970] merely show that multicollinearity between independent variables has two main effects. It produces greater sample variance in the parameter estimators and so reduces the power of standard inferential statistical tests; and it confuses the interpretation of the exact parameter estimates derived. Little or nothing, however, is reported on the effects of varying degrees of multicollinearity. Multicollinearity is usually taken to be a severe problem or else it is ignored completely. Klein [1962, p. 101], for example, refers to multicollinearity being a severe problem ohly when the absolute value of the correlation coefficient between the two independent variables is greater than the square root of the coefficient of multiple determination in the regression. Huang [1970], on the other hand, suggests that the point at which multicollinearity becomes a problem in parameter estimation is more subjective and should be left to the individual investigator. Farrar and Glauber [1967] state that "harmful multicollinearity" is hard to identify but agree that Klein's rule may be appropriate. Alternatively, they suggest that multicollinearity becomes a problem when the correlation coefficient between the independent variables reaches 0.8 or 0.9.

An alternative to the use of the correlation coefficient between the two independent variables in assessing multicollinearity is to examine the determinant of the correlation matrix between independent variables (Haitovsky [1969]). If any of the independent variables is exactly correlated with another the correlation matrix becomes singular and its determinant will be zero. The determinant of a correlation matrix of orthogonal variables is 1. Haitovský presented a heuristic chi-square statistic to examine whether or not the determinant of a correlation matrix is significantly different from zero. If it is, multicollinearity is assumed not be a severe problem. However, although this is a different measure of multicollinearity, it still assumes a cut-off point to one side of which multicollinearity is a severe problem while to the other side it can be ignored.

All four studies assume that the effect on parameter estimation when the degree of multicollinearity is below a certain level is negligible. The effects of multicollinearity are assumed to be discontinuous although this discontinuity is not theoretically proven and since the degree of multicollinearity is continuous (ranging from 0.0 to 1.0 as measured by the correlation coefficient between the independent variables), the effects of multicollinearity are also likely to be continuous. Any departure from the orthogonality of independent variables produces multicollinearity and the severity of this multicollinearity will increase as the departure from orthogonality increases. There is no theoretical reason why multicollinearity should only be considered at a certain degree, for example when the correlation coefficient between two independent variables reaches 0.8. Is multicollinearity not likely to have some effects on parameter estimators when the correlation coefficient is 0.75?

A problem then in econometric and statistical literature is that there are no studies of the explicit effects of varying degrees of multicollinearity on parameter estimates. Such a study is presented here in terms of the unconstrained spatial interaction model.

4.2 Multicollinearity and the Unconstrained Spatial Interaction Model

Consider the stochastic version of (1.33) in natural logarithms:

$$I_{ij}^* = \alpha_i^* + \gamma_i m_j^* + \beta_i d_{ij}^* + \varepsilon_{ij}^* . \qquad (4.1)$$

The aim of this chapter is to demonstrate explicitly the pattern of estimated distance-decay parameters derived from the calibration of (4.1) when varying degrees of multicollinearity exist between m_j^* and d_{ij}^*. Although the analysis will be presented in terms of the origin-specific unconstrained model it can easily be extended to the general unconstrained model. The advantage of discussing the former model is that only one type of multicollinearity can affect the estimated distance-decay parameter, whereas in the latter there are two - m_j^* and d_{ij}^* can be highly correlated as can m_j^* and d_{ij}^*. Strong linear relationships between m_j^* and d_{ij}^* are plausible. Consider, for example, origin i being a growth centre. Population size is then likely to decline as distance from the growth centre increases. Alternatively, origin i could be a small inaccessible origin and as distance from this origin increases, the population size of centres increases as they become more accessible.

To study the effects of multicollinearity a simulation is undertaken in which the degree of multicollinearity between m_j^* and d_{ij}^* is gradually increased and β_i is estimated at each degree. The relationship between m_j* and d_{ij}* in part determines $\hat{\beta}_i$ and this can produce a bias in the estimate due to spatial structure. The relationships between d_{ij}* and I_{ij}* and between m_j*-and I_{ij}*can be constant and yet $\hat{\beta}_i$ will vary due to variations in the relationship between m_j* and d_{ij}* - this latter relationship being a measure of spatial structure. Such variation in $\hat{\beta}_i$ needs to be described accurately since it is the essence of the model - specific spatial structure effect in unconstrained interaction modelling.

Thus, the main assertion on which the following simulation is based is that the correlation between the logarithm of mass and the logarithm of distance in a spatial system may have no effect upon a person's perceived disutility of travel yet the strength of this correlation is a major determinant of $\hat{\beta}_i$. Hence, $\hat{\beta}_i$ can contain a measurement of spatial structure which is independent of travel behaviour. The extent to which spatial structure can determine $\hat{\beta}_i$ in this way is now shown.

4.3 The Simulation

Consider an area divided into zones, each of which contains centres of population. A zone centroid (the largest centre in the zone) is designated in each zone and interactions are modelled from the zone centroid to all of the other centres within the zone. All interactions are intrazonal. The zone centroid and the zone it is located in are denoted by an i subscript. A destination in any zone is denoted by a j subscript. In each zone, the location of the centres is identical so that $\{d_{ij}\}$ is identical for each zone. Also the set of interactions, $\{I_{ij}\}$, is identical for each zone, so that only $\{m_j\}$ varies between zones. Forty-one centres are assumed present in each zone: 40 ordinary centres and one zone centroid. The sets of interactions and distances used for all zones are given in Table 4.1.[1] The set of masses was varied 71 times so simulating 71 zones. For each variation of masses equation (4.1) was calibrated and the following statistics obtained.[2]

$\hat{\beta}_i$, the estimated distance-decay parameter for zone i,

$\hat{\gamma}_i$, the estimated parameter relating mass to interaction for zone i,

$R_i(m^*d^*)$, the correlation coefficient between $\{m^*\}$ and $\{d^*\}$ for zone i where m* is the mass of a destination j in natural logarithms and d* is the logarithmic distance between i and j. The j subscripts are omitted for conciseness. In all simulations $R_i(m^*d^*)$ was negative.

$R_i(m^*I^*)$, the correlation coefficient between $\{m^*\}$ and $\{I^*\}$ for zone i where I* is the logarithm of interaction from i to j. In all simulations this was positive.

$R_i(d^*I^*)$, the correlation coefficient for zone i between the logarithm of distances to the zone centroid and the logarithm of interaction from the zone centroid, is a constant for all zones since $\{d_{ij}\}$ and $\{I_{ij}\}$ are constant. The value of $R_i(d^*I^*)$ in these simulations is -0.7.

For each simulation the value of $\hat{\beta}_i$ can be graphed against $R_i(m^*d^*)$ and $R_i(m^*I^*)$. It is then solely variations in these two relationships that produce variations in $\hat{\beta}_i$. The relationship between distance and interaction remains constant. The interrelationship between $R_i(m^*d^*)$ and $R_i(m^*I^*)$ includes a measure of spatial structure and as this interrelationship varies (i.e. as $\{m_j\}$ varies), $\hat{\beta}_i$ varies. The relationship between $\hat{\beta}_i$ and $R_i(m^*d^*)$ and $R_i(m^*I^*)$ is not a simple linear one, however, as can be seen from Figure 4.1 which will now be described and then explained theoretically.

[1] There is no significance to the values of interaction and distance given for each zone. An attempt was made to produce a negative relationship between the two variables but otherwise the values are entirely random.

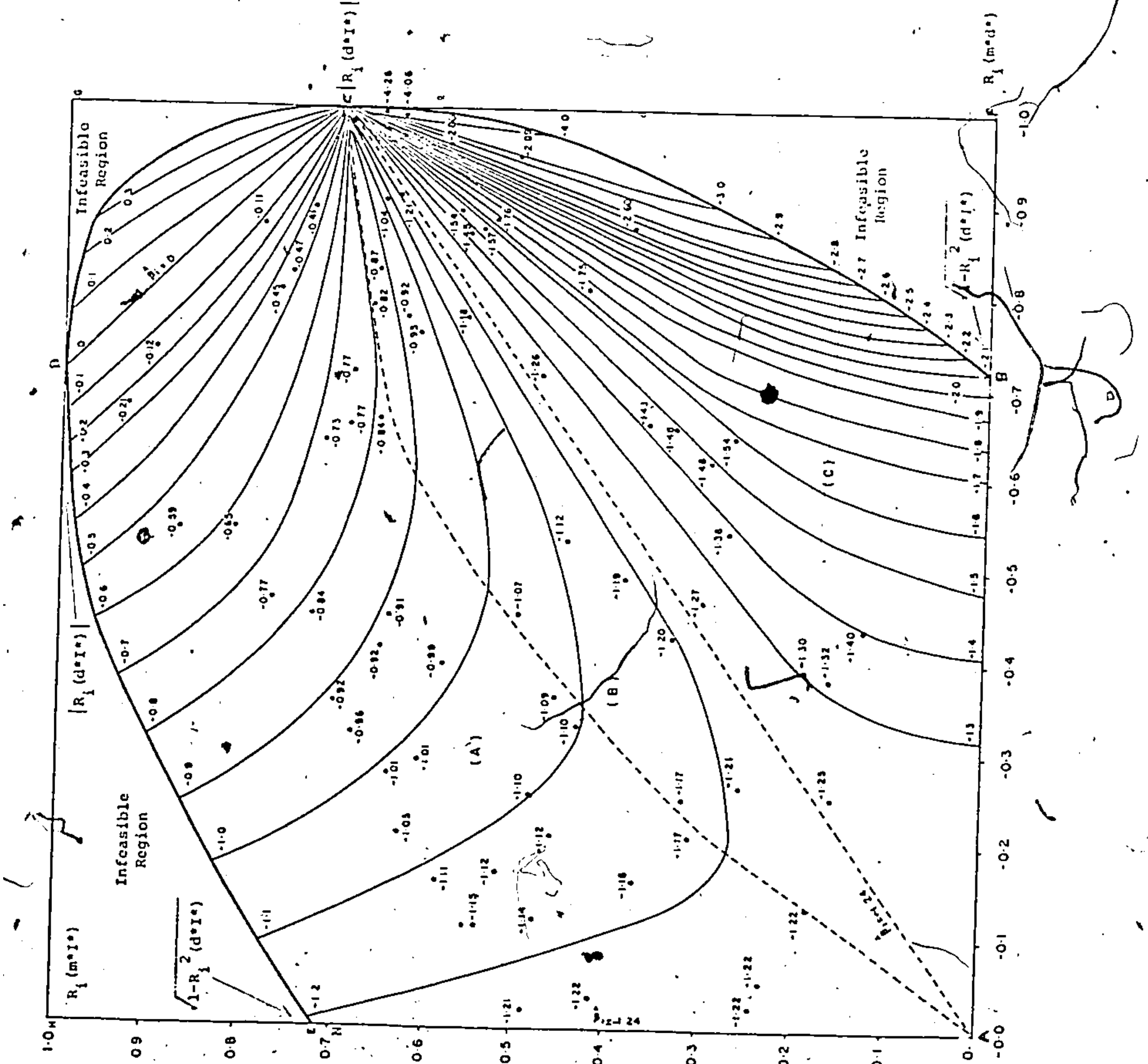
[2] It is important to note that for each zone, that is for each calibration of the model, $\{m_j^*\}$ and $\{\varepsilon_{ij}^*\}$ are independent.

TABLE 4.1

$\{I_{ij}\}$ AND $\{d_{ij}\}$ USED FOR ALL 71 SIMULATIONS

No. of Centre	I_{ij}	d_{ij}	No. of Centre	I_{ij}	d_{ij}
1	10	110	21	60	35
2	20	70	22	70	40
3	50	45	23	80	90
4	35	80	24	80	40
5	60	30	25	40	70
6	90	10	26	60	20
7	5	100	27	40	50
8	45	80	28	20	50
9	25	75	29	25	60
10	1	200	30	10	110
11	50	50	31	50	25
12	80	40	32	90	20
13	75	45	33	55	70
14	90	30	34	60	80
15	10	90	35	15	35
16	25	80	36	40	90
17	9	75	37	45	70
18	40	60	38	85	70
19	20	120	39	40	35
20	1	200	40	15	90

FIGURE 4.1 The Effect of Variations in $R_i(m^*d^*)$ and $R_i(m^*I^*)$ on $\hat{\beta}_i$



4.4 A Description of Figure 4.1

The horizontal axis of the graph given in Figure 4.1 is defined by $R_i(m^*d^*)$ which ranges from 0 to -1.0. The vertical axis of the graph is defined by $R_i(m^*I^*)$ which ranges from 0 to 1.0. The values inside the graph are the values of $\hat{\beta}_i$ derived from calibrating (4.1) with various $\{m_i\}$. Consequently, each $\hat{\beta}_i$ corresponds to a specific combination of $R_i(m^*d^*)$ and $R_i(m^*I^*)$ representing a zone. In all cases $\hat{\beta}_i$ was found to be negative since $R_i(d^*I^*)$ is negative, although it will be demonstrated theoretically that even under these circumstances, $\hat{\beta}_i$ can be positive. The values of $\hat{\beta}_i$ ranged from -4.26 to -0.01 indicating that even if the correlation between distance and interaction remains constant, $\hat{\beta}_i$ can show marked variation.

The solid lines within the graph represent combinations of $R_i(m^*d^*)$ and $R_i(m^*I^*)$ which produce equal values of $\hat{\beta}_i$. These contour lines were interpolated partly from theory and partly from the derived values of $\hat{\beta}_i$. They converge on the point where $R_i(m^*d^*) = -1.0$ and where $R_i(m^*I^*) = |R_i(d^*I^*)|$ which here is 0.7.

The broken lines within the graph delimit regions of homogeneity and their diagnostic properties are:

Region (A): The mass term enters the regression positively. For a given value of $R_i(m^*I^*)$, $\hat{\beta}_i$ increases as $|R_i(m^*d^*)|$ increases.

Region (B): The mass term enters the regression positively. For a given value of $R_i(m^*I^*)$, $\hat{\beta}_i$ decreases as $|R_i(m^*d^*)|$ increases.

Region (C): The mass term enters the regression negatively. For a given value of $R_i(m^*I^*)$, $\hat{\beta}_i$ decreases as $|R_i(m^*d^*)|$ increases.

It is obvious from Figure 4.1 that the relationship between $R_i(m^*d^*)$ and $\hat{\beta}_i$ is not always monotonic and is quite complex. The implication of this finding is that increasing multicollinearity does not have a predictable effect on the parameters of a regression model unless other relationships between variables are taken into account. For example, in Figure 4.1 if $R_i(m^*I^*)$ is 0.8, increasing multicollinearity produces increasing values of $\hat{\beta}_i$. If $R_i(m^*I^*)$ is 0.0, increasing multicollinearity produces decreasing values of $\hat{\beta}_i$. If $R_i(m^*I^*)$ is 0.5 increasing multicollinearity produces increasing values of $\hat{\beta}_i$ up to a certain degree of multicollinearity and then decreasing values of $\hat{\beta}_i$ as multicollinearity continues to increase beyond this certain degree.

There are three infeasible regions on the graph which represent combinations of $R_i(m^*I^*)$ and $R_i(m^*d^*)$ that are impossible given a fixed value of $R_i(d^*I^*)$. The infeasible regions are defined as lying outside the curve BE. No values of $\hat{\beta}_i$ are possible within these regions.

For combinations of $R_i(m^*d^*)$ and $R_i(m^*I^*)$ which lie within the rectangle AFCN, distance will enter a stepwise regression, in which interaction is the dependent variable, before mass since $|R_i(d^*I^*)| > R_i(m^*I^*)$. For combinations of $R_i(m^*d^*)$ and $R_i(m^*I^*)$ which lie within the rectangle NCGH, mass will enter a stepwise regression before distance since $R_i(m^*I^*) > |R_i(d^*I^*)|$.

The base value of $\hat{\beta}_i$ to which all other $\hat{\beta}_i$'s can be compared is the value of $\hat{\beta}_i$ obtained when the multicollinearity between m_j^* and d_{ij}^* is zero -- in this case, -1.24. The value is equivalent to the distance-decay parameter that would be obtained in a simple regression of distance on interaction. This base value occurs for all values of $R_i(m^*I^*)$ when $R_i(m^*d^*) = 0$ and it also occurs along the diagonal line AC which runs from the origin to the point at which $R_i(m^*d^*) = -1.0$ and $R_i(m^*I^*) = |R_i(d^*I^*)|$. Along this line mass does not enter the regression and $\hat{\gamma}_1 = 0$ even though $R_i(m^*I^*)$ may not be zero. This and other features of Figure 4.1 are now explained theoretically.

4.5 A Theoretical Explanation of Figure 4.1

The preceding section contains a simple description of the graph given in Figure 4.1. In this section various features of the graph are derived theoretically and the trends in $\hat{\beta}_i$ are explained in relation to variations in $R_i(m^*d^*)$ and $R_i(m^*I^*)$. Central to this theory is the derivation of mathematical expressions for $\hat{\beta}_i$ and $\hat{\gamma}_i$ from the normal equations of (4.1) - a technique outlined in most econometrics and statistics texts (*inter alia,* Christ [1966, pp. 380-383]). Since the derivation of such expressions is well-documented, they can simply be stated here as:

$$\hat{\beta}_i = \frac{\sum_{j=1}^{n}(m_j')^2 \cdot \sum_{j=1}^{n} d_{ij}' \cdot I_{ij}' - \sum_{j=1}^{n} m_j' \cdot I_{ij}' \cdot \sum_{j=1}^{n} m_j' \cdot d_{ij}'}{\sum_{j=1}^{n}(d_{ij}')^2 \cdot \sum_{j=1}^{n}(m_j')^2 - (\sum_{j=1}^{n} m_j' \cdot d_{ij}')^2} \quad (4.2)$$

and,

$$\hat{\gamma}_i = \frac{\sum_{j=1}^{n}(d_{ij}')^2 \cdot \sum_{j=1}^{n} m_j' \cdot I_{ij}' - \sum_{j=1}^{n} d_{ij}' \cdot I_{ij}' \cdot \sum_{j=1}^{n} m_j' \cdot d_{ij}'}{\sum_{j=1}^{n}(d_{ij}')^2 \cdot \sum_{j=1}^{n}(m_j')^2 - (\sum_{j=1}^{n} m_j' \cdot d_{ij}')^2}, \quad (4.3)$$

where n is the number of destinations available to the inhabitants of origin i,

$$m_j' = m_j^* - (\sum_{j=1}^{n} m_j^*)/n \quad , \tag{4.4}$$

$$d_{ij}' = d_{ij}^* - (\sum_{j=1}^{n} d_{ij}^*)/n \quad , \tag{4.5}$$

and

$$I_{ij}' = I_{ij}^* - (\sum_{j=1}^{n} I_{ij}^*)/n \quad . \tag{4.6}$$

Several features of the graph given in Figure 4.1 will now be stated as theorems and proved:

Theorem (i)

Given a fixed value of $R_i(d^*I^*) < 0$, there exist regions of the graph which have infeasible combinations of $R_i(m^*d^*)$ and $R_i(m^*I^*)$, and in which no $\hat{\beta}_i$ is possible. The boundary of these regions is given by the quadratic equation:

$$R_i^2(m^*I^*) + R_i^2(m^*d^*) - 2R_i(m^*I^*)\cdot R_i(d^*I^*)\cdot R_i(m^*d^*) + R_i^2(d^*I^*) - 1$$
$$= 0$$

Proof:

Following Merrill and Fox [1970, p. 402], the coefficient of determination of the regression given in (4.1) can be defined in the following manner:

$$R_i^2(I^* : m^*d^*) = \frac{R_i^2(m^*I^*) + R_i^2(d^*I^*) - 2R_i(m^*I^*)\cdot R_i(d^*I^*)\cdot R_i(m^*d^*)}{1 - R_i^2(m^*d^*)} \tag{4.7}$$

where $R_i^2(I^* : m^*d^*)$ is the proportion of variance of I_{ij}^* explained by m_j^* and d_{ij}^*, and $R_i^2(x,y)$ is used as the notation for $(R_i(x,y))^2$.

With $R_i(d^*I^*)$ given, the proof involves deriving the maximum possible values of $|R_i(m^*d^*)|$ given various values of $R_i(m^*I^*)$ or deriving the maximum possible values of $R_i(m^*I^*)$ given various values of $R_i(m^*d^*)$. Such maxima occur when $R_i^2(I^*:m^*d^*)$ is at a maximum, that is, when $R_i^2(I^*: m^*d^*) = 1$, i.e. when

$$R_i^2(m^*I^*) + R_i^2(m^*d^*) - 2R_i(m^*I^*) \cdot R_i(d^*I^*) \cdot R_i(m^*d^*) + R_i^2(d^*I^*) - 1 = 0 . \tag{4.8}$$

Q.E.D.

While it is clear that the maximum value of $R_i(m^*I^*)$ occurs when $R_i^2(I^*: m^*d^*)$ is at a maximum, it may not be obvious that $R_i^2(I^*: m^*d^*)$ and $|R_i(m^*d^*)|$ are positively related and that by increasing $|R_i(m^*d^*)|$ a greater proportion of the variance of $\{I_{ij}^*\}$ can be explained: the two coefficients seem unrelated. They are not, however, as the following example demonstrates. Suppose $R_i(m^*I^*) = 0.0$ and $R_i(d^*I^*) = -0.7$. From (4.8), the maximum value of $R_i^2(I^*: m^*d^*)$ occurs when $|R_i(m^*d^*)| = 0.71$. When $|R_i(m^*d^*)| < 0.71$, $R_i^2(I^*: m^*d^*) < 1.0$. The reason for this relationship is that the simple correlation between m^* and I^*, which is 0.0, becomes increasingly spurious as $|R_i(m^*d^*)|$ increases. The simple correlation of 0.0 is a resultant of two forces acting in opposite directions. Mass is related to interaction directly in one direction but it is related to interaction indirectly in the other direction. This indirect relationship occurs since mass is inversely related to distance which in turn is inversely related to interaction. As $|R_i(m^*d^*)|$ increases, the indirect relationship increases and for $R_i(m^*I^*)$ to remain at 0.0, the direct relationship must increase. Thus, as $|R_i(m^*d^*)|$ increases, $R_i^2(I^*: m^*d^*)$ can increase, and vice versa, because the simple correlation coefficient between m* and I* becomes increasingly spurious. In the example given here, the indirect relationship between m* and I* is positive (as distance increases, mass decreases and interaction decreases), so that the "hidden" direct relationship must be negative. This is evident in Figure 4.3 where $\hat{\gamma}_i < 0.0$ when $R_i(m^*I^*) = 0.0$ and $\hat{\gamma}_i$ becomes increasingly negative as $|R_i(m^*d^*)|$ increases.

Some interesting properties derived from (4.8) are that:

(a) When $R_i(m^*d^*) = 0$, max. $R_i(m^*I^*) = \sqrt{1 - R_i^2(d^*I^*)}$.

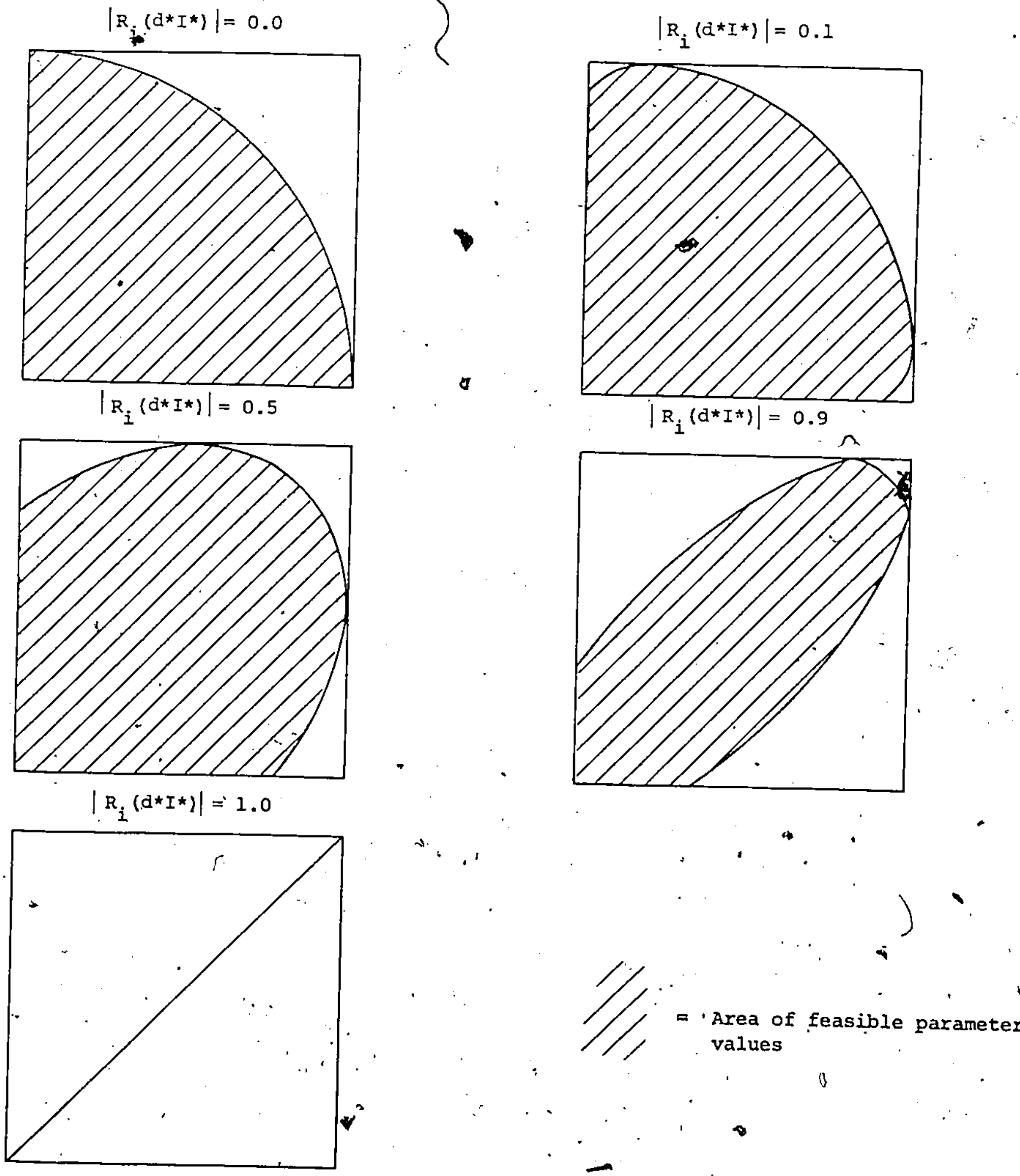
(b) When $R_i(m^*I^*) = 0$, max. $|R_i(m^*d^*)| = \sqrt{1 - R_i^2(d^*I^*)}$.

(c) When $|R_i(m^*d^*)| \to 1$, max. $R_i(m^*I^*) \to |R_i(d^*I^*)|$.

(d) When $R_i(m^*I^*) = 1$, max. $|R_i(m^*d^*)| = |R_i(d^*I^*)|$.

From the above four properties and from (4.8), it is obvious that as $R_i(d^*I^*)$ varies, the area of feasible $\hat{\beta}_i$ values will vary. As a result, Figure 4.1 can be considered as one "slice" out of a "mound" of feasible $\hat{\beta}_i$ values - the slice being taken at a particular value of $R_i(d^*I^*)$. Other "slices" will give differently-shaped areas of feasible values as shown by Figure 4.2. The graphs in Figure 4.2 were derived

Figure 4.2: Variations in the Area of Feasible Parameter Values with Variations in $|R_i(d^*I^*)|$.



by substituting values of $R_i(d^*I^*)$ into (4.8), and into the four properties listed above, and obtaining maximum values for $R_i(m^*d^*)$ and $R_i(m^*I^*)$. For example, if $R_i(d^*I^*) = 0.0$, (4.8) reduces to the form $x^2 + y^2 = 1$, which is the equation of a circle. If $|R_i(d^*I^*)| = 1.0$, (4.8) reduces to the form $x = y$, which is the equation of a straight line. Figure 4.2 indicates that as $|R_i(d^*I^*)|$ increases, the area of feasible parameter values decreases until at $|R_i(d^*I^*)| = 1.0$, the area degenerates into a straight line from the origin to the point at which $|R_i(m^*d^*)| = R_i(m^*I^*) = 1.0$.

A consequence of the varying areas of these "slices" is that since $\hat{\beta}_i$ varies with variations in $R_i(m^*d^*)$ and $R_i(d^*I^*)$, the variation of $\hat{\beta}_i$ will decrease as $|R_i(d^*I^*)|$ increases. Thus, the effects of multicollinearity on $\hat{\beta}_i$ will decrease as $|R_i(d^*I^*)|$ increases. When $|R_i(d^*I^*)| = 1.0$, there will be no effect on parameter estimates due to multicollinearity, and consequently no spatial structure bias in $\hat{\beta}_i$, unless $R_i(m^*d^*) = 1.0$, since the parameter estimates will be constant along the diagonal line of feasible values (see the proof of theorem (ii)). This diagonal line would be equivalent to the straight line AC in Figure 4.1. Thus, the effects of multicollinearity on parameter estimates cannot be described by the degree of multicollinearity alone. The relationships between all other variables in the regression - in this case measured by $R_i(m^*I^*)$ and $R_i(d^*I^*)$ - must also be considered. A correlation coefficient between m* and d* of 0.8, for example, may have very serious effects on parameter estimates if $|R_i(d^*I^*)|$ is low, whereas it will have no effect on parameter estimates if $|R_i(d^*I^*)| = 1.0$.

It has been shown how variations in $R_i(d^*I^*)$ produce variations in the effects of multicollinearity between m* and d*. In the following theorems and proofs, $R_i(d^*I^*)$ is considered fixed and the effects of variations in $R_i(m^*d^*)$ and $R_i(m^*I^*)$ on $\hat{\beta}_i$, which are described by Figure 4.1, are investigated theoretically.

Theorem (ii)

The straight line AC is defined by the conditions that $\hat{\gamma}_i = 0$ and $\hat{\beta}_i$ is a constant equal to its "base-value". The equation of this line is $R_i(m^*I^*) = R_i(m^*d^*)\cdot R_i(d^*I^*)$.

Proof:

Mass does not enter the regression when $\hat{\gamma}_i = 0$. From (4.3) this is when,

$$\sum_{j=1}^{n}(d_{ij}')^2\cdot\sum_{j=1}^{n}m_j'\cdot I_{ij}' - \sum_{j=1}^{n}d_{ij}'\cdot I_{ij}'\cdot\sum_{j=1}^{n}m_j'\cdot d_{ij}' = 0 . \qquad (4.9)$$

Rearranging (4.9) gives:

$$\frac{\sum_{j=1}^{n}m_j'\cdot I_{ij}'}{\sqrt{\sum_{j=1}^{n}(m_j')^2}\cdot\sqrt{\sum_{j=1}^{n}(I_{ij}')^2}} = \frac{\sum_{j=1}^{n}m_j'\cdot d_{ij}'}{\sqrt{\sum_{j=1}^{n}(d_{ij}')^2}\cdot\sqrt{\sum_{j=1}^{n}(m_j')^2}}\cdot\frac{\sum_{j=1}^{n}d_{ij}'\cdot I_{ij}'}{\sqrt{\sum_{j=1}^{n}(d_{ij}')^2}\cdot\sqrt{\sum_{j=1}^{n}(I_{ij}')^2}} \qquad (4.10)$$

which is equivalent to:

$$R_i(m^*I^*) = R_i(m^*d^*)\cdot R_i(d^*I^*) . \qquad (4.11)$$

Q.E.D.

The distance-decay parameter in (4.1) will be equal to its base-value whenever mass does not enter the regression since $\hat{\beta}_i$ is then simply determined by $R_i(d^*I^*)$ which is a constant.

In words, the relationship given in (4.11) indicates that mass does not enter the regression along the straight line AC because it has no direct effect on interaction. The relationship between mass and interaction indicated by the correlation coefficient is entirely due to mass being related to distance which in turn is related to interaction. Hence, when distance has entered the regression, the presence of a mass term explains no more of the variance of interaction and $\hat{\gamma}_i = 0$.

Theorem (iii):

The curved line AC is defined where $\frac{\delta \hat{\beta}_i}{\delta R_i(m^*d^*)} = 0$ *and it represents the boundary between regions (A) and (B). For combinations of* $R_i(m^*I^*)$ *and* $R_i(m^*d^*)$ *that lie above this line, as* $R_i(m^*d^*)$ *decreases,* $\hat{\beta}_i$ *increases. For combinations of* $R_i(m^*I^*)$ *and* $R_i(m^*d^*)$ *that lie below this line, as* $R_i(m^*d^*)$ *decreases,* $\hat{\beta}_i$ *decreases. The curved line AC is defined by the quadratic expression:*

$$R_i(m^*I^*) \cdot R_i^2(m^*d^*) - 2R_i(d^*I^*) \cdot R_i(m^*d^*) + R_i(m^*I^*) = 0.$$

Proof:

Dividing the numerator and denominator of equation (4.2) by

$$\sum_{j=1}^{n} (d_{ij}')^2 \cdot \sum_{j=1}^{n} (m_j')^2 \cdot \sqrt{\sum_{j=1}^{n} (I_{ij}')^2} \qquad \text{gives}$$

$$\hat{\beta}_i = \frac{R_i(d^*I^*) - R_i(m^*I^*) \cdot R_i(m^*d^*)}{[1 - R_i^2(m^*d^*)]} \cdot \frac{\sqrt{\sum_{j=1}^{n} (I_{ij}')^2}}{\sqrt{\sum_{j=1}^{n} (d_{ij}')^2}} \qquad (4.12)$$

Define $C \equiv \sqrt{\sum_{j=1}^{n} (I_{ij}')^2} \Big/ \sqrt{\sum_{j=1}^{n} (d_{ij}')^2}$ since $\{I_{ij}\}$ and $\{d_{ij}\}$

are constant, and then the derivative of $\hat{\beta}_i$ w.r.t. $R_i(m^*d^*)$ is:

$$\frac{\delta \hat{\beta}_i}{\delta R_i(m^*d^*)} = \frac{2R_i(d^*I^*) \cdot R_i(m^*d^*) - R_i(m^*I^*) \cdot [1 + R_i^2(m^*d^*)]}{[1 - R_i^2(m^*d^*)]^2} \cdot C \tag{4.13}$$

For any given value of $R_i(m^*I^*)$ the curved line AC represents the point at which the derivative $\delta\hat{\beta}_i/\delta R_i(m^*d^*) = 0$. Thus, the curved line AC is defined when:

$$R_i(m^*I^*) = \frac{2R_i(d^*I^*) \cdot R_i(m^*d^*)}{1 + R_i^2(m^*d^*)} \tag{4.14}$$

which is equivalent to the quadratic equation in $R_i(m^*d^*)$:

$$R_i(m^*I^*) \cdot R_i^2(m^*d^*) - 2R_i(d^*I^*) \cdot R_i(m^*d^*) + R_i(m^*I^*) = 0 .$$

Q.E.D.

Equation (4.12) indicates that as $|R_i(m^*d^*)| \rightarrow 1, |\hat{\beta}_i| \rightarrow \infty$. This explains the rapid decrease in $\hat{\beta}_i$ when $|R_i(m^*d^*)| > 0.9$ in Figure 4.1.

Theorem (iv):

Along the line AE where $R_i(m^*d^*) = 0$, $\hat{\beta}_i$ does not vary and is equal to its "base-value".

Proof:

When $R_i(m^*d^*) = 0$, $\sum_{j=1}^{n} m_j' . d_{ij}' = 0$ and (4.2) can be rewritten as:

$$\hat{\beta}_i = \frac{\sum_{j=1}^{n} (m_j')^2 \cdot \sum_{j=1}^{n} d_{ij}' \cdot I_{ij}'}{\sum_{j=1}^{n} (d_{ij}')^2 \cdot \sum_{j=1}^{n} (m_j')^2} \tag{4.15}$$

Dividing through by $\sum_{j=1}^{n} (m_j')^2$ gives:

$$\hat{\beta}_i = \frac{\sum_{j=1}^{n} d_{ij}' \cdot I_{ij}'}{\sum_{j=1}^{n} (d_{ij}')^2} \tag{4.16}$$

which is a constant since $\{I_{ij}\}$ and $\{d_{ij}\}$ are constant. The value of $\hat{\beta}_i$ given in (4.16) is identical to the value of $\hat{\beta}_i$ derived for any combination of $R_i(m^*I^*)$ and $R_i(m^*d^*)$ along the straight line AC. This is proved as follows. The straight line AC is characterised by the property that:

$$R_i(m^*I^*) = R_i(m^*d^*) \cdot R_i(d^*I^*) ,$$

or equivalently, where

$$\frac{\sum_{j=1}^{n} m_j' \cdot I_{ij}'}{\sqrt{\sum_{j=1}^{n} (m_j')^2 \cdot \sum_{j=1}^{n} (I_{ij}')^2}} = \frac{\sum_{j=1}^{n} m_j' \cdot d_{ij}'}{\sqrt{\sum_{j=1}^{n} (m_j')^2 \cdot \sum_{j=1}^{n} (d_{ij}')^2}} \cdot \frac{\sum_{j=1}^{n} d_{ij}' \cdot I_{ij}'}{\sqrt{\sum_{j=1}^{n} (d_{ij}')^2 \cdot \sum_{j=1}^{n} (I_{ij}')^2}} \tag{4.17}$$

Equation (4.17) can be simplified to:

$$\sum_{j=1}^{n} m_j' \cdot I_{ij}' = \frac{\sum_{j=1}^{n} m_j' \cdot d_{ij}' \cdot \sum_{j=1}^{n} d_{ij}' \cdot I_{ij}'}{\sum_{j=1}^{n} (d_{ij}')^2} \tag{4.18}$$

Substituting (4.18) into (4.2) and rearranging gives:

$$\hat{\beta}_i = \frac{\sum_{j=1}^{n} d_{ij}'\cdot I_{ij}'}{\sum_{j=1}^{n} (d_{ij}')^2}$$

Since this value is the same as that given in (4.16) the value of $\hat{\beta}_i$ is the same along AE as it is along the straight line AC.

Q.E.D.

Theorem (v):

Along the line AB where $R_i(m^*d^*)$ varies and $R_i(m^*I^*) = 0$, $\hat{\beta}_i$ varies.

Proof:

When $R_i(m^*I^*) = 0$, $\sum_{j=1}^{n} m_j'\cdot I_{ij}' = 0$ and (4.2) can be rewritten as:

$$\hat{\beta}_i = \frac{\sum_{j=1}^{n}(m_j')^2 \cdot \sum_{j=1}^{n} d_{ij}'\cdot I_{ij}'}{\sum_{j=1}^{n} (d_{ij}')^2 \cdot \sum_{j=1}^{n} (m_j')^2 - (\sum_{j=1}^{n} m_j'\cdot d_{ij}')^2} \tag{4.19}$$

Since, by definition, the denominator of (4.19) varies as $R_i(m^*d^*)$ varies, $\hat{\beta}_i$ will vary as $R_i(m^*d^*)$ varies. Again, as $|R_i(m^*d^*)| \longrightarrow 1$, the denominator of (4.19) $\longrightarrow 0$ and $|\hat{\beta}_i| \longrightarrow \infty$.

Q.E.D.

Theorem (vi):

Along the straight line DC, distance does not enter the regression and $\hat{\beta}_i = 0$. This line is defined where the equality $R_i(d^*I^*) = R_i(m^*d^*).R_i(m^*I^*)$ is true.

Proof:

From (4.2), when $\hat{\beta}_i = 0$,

$$\sum_{j=1}^{n} (m_j')^2 \cdot \sum_{j=1}^{n} d_{ij}' \cdot I_{ij}' - \sum_{j=1}^{n} m_j' \cdot I_{ij}' \cdot \sum_{j=1}^{n} m_j' \cdot d_{ij}' = 0 . \tag{4.20}$$

Dividing both sides of (4.20) by $\sum_{j=1}^{n} (m_j')^2 \cdot \sqrt{\sum_{j=1}^{n} (d_{ij}')^2} \cdot \sqrt{\sum_{j=1}^{n} (I_{ij}')^2}$

and rearranging gives:

$$\frac{\sum_{j=1}^{n} d_{ij}' \cdot I_{ij}'}{\sqrt{\sum_{j=1}^{n} (d_{ij}')^2} \cdot \sqrt{\sum_{j=1}^{n} (I_{ij}')^2}} = \frac{\sum_{j=1}^{n} m_j' \cdot I_{ij}'}{\sqrt{\sum_{j=1}^{n} (m_j')^2} \cdot \sqrt{\sum_{j=1}^{n} (I_{ij}')^2}} \cdot \frac{\sum_{j=1}^{n} m_j' \cdot d_{ij}'}{\sqrt{\sum_{j=1}^{n} (m_j')^2} \cdot \sqrt{\sum_{j=1}^{n} (d_{ij}')^2}} \tag{4.21}$$

which is equivalent to:

$$R_i(d^*I^*) = R_i(m^*I^*) \cdot R_i(m^*d^*) . \tag{4.22}$$

Q.E.D.

In words, the relationship in (4.22) indicates that distance does not enter the regression along the straight line DC because it has no direct effect on interaction. The relationship between distance and interaction is entirely due to distance being related to mass which in turn is related to interaction. Hence, when mass has already entered the regression, the presence of a distance term explains no more of the variance of interaction and $\hat{\beta}_i = 0$.

Theorem (vii):

When $R_i(d^*I^*) < 0$, $\hat{\beta}_i$ can be $\gtrless 0$. In Figure 4.1 this occurs in the feasible region which lies between the straight line DC and the curved line DC. In the rest of the feasible region $\hat{\beta}_i < 0$.

Proof:

As can be seen from equation (4.2) and the proof of theorem (vi), $\hat{\beta}_i$ will be positive when,

$$R_i(d^*I^*) > R_i(m^*I^*) \cdot R_i(m^*d^*) \,.$$

Since, $R_i(d^*I^*) < 0$ and $R_i(m^*d^*) < 0$, $\hat{\beta}_i$ will be positive, when,

$$R_i(m^*I^*) \cdot |R_i(m^*d^*)| > |R_i(d^*I^*)| \,. \tag{4.23}$$

Squaring both sides of the inequality given in (4.23) and rearranging gives the condition that $\hat{\beta}_i > 0$ when,

$$R_i^{\,2}(m^*d^*) > \frac{R_i^{\,2}(d^*I^*)}{R_i^{\,2}(m^*I^*)} \,. \tag{4.24}$$

From (4.8) the boundary of the feasible region occurs when:

$$R_i^{\,2}(m^*d^*) = 1 + 2R_i(d^*I^*) \cdot R_i(m^*d^*) \cdot R_i(m^*I^*) - R_i^{\,2}(m^*I^*) - R_i^{\,2}(d^*I^*) \,. \tag{4.25}$$

Values of $R_i^{\,2}(m^*d^*)$ above the value of the expression given in

(4.25) are infeasible. This is easily seen by the fact that the maximum derived value of $R_i^2(m^*d^*)$ is 1.0 when $R_i(m^*I^*) = |R_i(d^*I^*)|$. Hence, $\hat{\beta}_i > 0$, iff;

$$\frac{R_i^2(d^*I^*)}{R_i^2(m^*I^*)} \leq R_i^2(m^*d^*) < 1 + 2R_i(d^*I^*) \cdot R_i(m^*d^*) \cdot R_i(m^*I^*) - R_i^2(m^*I^*) - R_i^2(d^*I^*) \quad (4.26)$$

If there is a combination of values for $R_i(m^*d^*)$, $R_i(d^*I^*)$ and $R_i(m^*I^*)$ which satisfies (4.26) then this proves that there will be a region on the graph where $\hat{\beta}_i > 0$ even though $R_i(d^*I^*) < 0$. In fact there are many combinations of the three correlation coefficients which satisfy (4.26). One such combination is when $R_i(m^*d^*) = -0.9$, $R_i(d^*I^*) = -0.7$ and $R_i(m^*I^*) = 0.9$. Hence, by simply alter ing $\{m_j\}$ and keeping $\{I_{ij}\}$ and $\{d_{ij}\}$ constant $\hat{\beta}_i$ can vary from extremely large negative values through to positive values.

Q.E.D.

Theorem (viii):

For a fixed value of $R_i(m^*d^*)$, as $R_i(m^*I^*)$ increases, $\hat{\beta}_i$ increases.

Proof:

From (4.12) the derivative of $\hat{\beta}_i$ w.r.t. $R_i(m^*I^*)$ is:

$$\frac{\delta\hat{\beta}_i}{\delta R_i(m^*I^*)} = \frac{-R_i(m^*d^*)}{[1 - R_i^2(m^*d^*)] \cdot \sqrt{\sum_{j=1}^{n} (I_{ij}')^2} \cdot \sqrt{\sum_{j=1}^{n} (d_{ij}')^2}}$$

$$= \frac{-R_i(m^*d^*)}{1 - R_i^{\,2}(m^*d^*)} \cdot C \geq 0 , \qquad (4.27)$$

$$\text{where } C = [\sqrt{\sum_{j=1}^{n} (\bar{I}_{ij}')^2} \cdot \sqrt{\sum_{j=1}^{n} (d_{ij}')^2}]^{-1} .$$

Hence, as $R_i(m^*I^*)$ increases, $\hat{\beta}_i$ increases for any value of $R_i(m^*d^*) > 0$. This is simply a statement of the fact that as $R_i(m^*I^*)$ increases, the relationship between distance and interaction, as measured by the simple correlation coefficient between the two variables, becomes increasingly spurious unless mass and distance are uncorrelated. As a result, $|\hat{\beta}_i|$ decreases.

Q.E.D.

Theorem (ix):

There exists a feasible region on the graph where mass enters the regression positively and for a given value of $R_i(m^*I^*)$, as $R_i(m^*d^*)$ decreases, $\hat{\beta}_i$ increases. This is region (A).

Proof:

From equation (4.11) and the proof of theorem (ii), mass enters the regression positively whenever,

$$R_i(m^*I^*) > R_i(m^*d^*) \cdot R_i(d^*I^*) . \qquad (4.28)$$

From equation (4.14) and the proof of theorem (iii), $\hat{\beta}_i$ increases as $R_i(m^*d^*)$ decreases whenever,

$$R_i(m^*I^*) \quad > \quad \frac{2R_i(d^*I^*) \cdot R_i(m^*d^*)}{1 + R_i^{\,2}(m^*d^*)} \quad . \tag{4.29}$$

Since $2/[1 + R_i^{\,2}(m^*d^*)] \geq 1$, the inequality in (4.28) is true whenever the inequality in (4.29) is true. From equation (4.8) and the proof of theorem (i), the upper boundary of region (A) is given by the equation

$$R_i^{\,2}(m^*I^*) + R_i^{\,2}(m^*d^*) - 2R_i(m^*I^*) \cdot R_i(d^*I^*) \cdot R_i(m^*d^*) + R_i^{\,2}(d^*I^*) - 1 = 0 \quad . \tag{4.30}$$

Thus region (A) exists whenever,

$$\frac{2R_i(d^*I^*) \cdot R_i(m^*d^*)}{1 + R_i^{\,2}(m^*d^*)} \quad < \quad R_i(m^*I^*) \quad \leq \quad R_i^{\,+}(m^*I^*) \tag{4.31}$$

where $R_i^{\,+}(m^*I^*)$ is a value of $R_i(m^*I^*)$ satisfying (4.30).

Q.E.D.

An explanation for the trend in $\hat{\beta}_i$ in region (A) is as follows. The model calibrated is:

$$\hat{I}_{ij} = \hat{\alpha}_i \, m_j^{\hat{\gamma}_i} \, d_{ij}^{\hat{\beta}_i} \qquad \text{(hypothesis: } \hat{\gamma}_i > 0,\ \hat{\beta}_i < 0). \tag{4.32}$$

If $R_i(m^*d^*) < 0$, then the "distance effect" of the mass term in (4.32) can be represented explicitly as:

$$\hat{I}_{ij} = \hat{\alpha}_i \left(\frac{1}{d_{ij}}\right)^{\hat{\gamma}_i} d_{ij}^{\hat{\beta}_i} \qquad \text{(hypothesis: } \hat{\gamma}_i > 0,\ \hat{\beta}_i < 0) \tag{4.33}$$

which is equivalent to

$$\hat{I}_{ij} = \hat{\alpha}_i \, d_{ij}^{\hat{\gamma}_i} \, d_{ij}^{\hat{\beta}_i} \quad \text{(hypothesis: } \hat{\gamma}_i < 0, \hat{\beta}_i < 0) . \tag{4.34}$$

As $|R_i(m*d*)|$ increases, the effect of distance upon interaction is increasingly measured by $\hat{\gamma}_i$ and since $\hat{\gamma}_i$ and $\hat{\beta}_i$ are complementary, $\hat{\beta}_i$ increases.

Theorem (x):

There exists a feasible region on the graph where mass enters the regression positively and for a given value of $R_i(m*I*)$, as $R_i(m*d*)$ decreases, $\hat{\beta}_i$ decreases. This is region (B).

Proof:

From equation (4.11) and the proof of theorem (ii), mass enters the regression positively whenever,

$$R_i(m*I*) > R_i(m*d*) \cdot R_i(d*I*) . \tag{4.35}$$

From equation (4.14) and the proof of theorem (iii), $\hat{\beta}_i$ decreases as $R_i(m*d*)$ decreases whenever,

$$R_i(m*I*) < \frac{2R_i(d*I*) \cdot R_i(m*d*)}{1 + R_i^2(m*d*)} . \tag{4.36}$$

Thus region (B) exists whenever,

$$R_i(m*d*) \cdot R_i(d*I*) < R_i(m*I*) < \frac{2R_i(d*I*) \cdot R_i(m*d*)}{1 + R_i^2(m*d*)} . \tag{4.37}$$

Q.E.D.

An explanation of the trend in $\hat{\beta}_i$ in region (B) is as follows. Since $R_i(m*d*) < 0$, the "distance effect" of the mass term in the interaction model can be represented explicitly as:

$$\hat{I}_{ij} = \hat{\alpha}_i \left(\frac{1}{d_{ij}}\right)^{\hat{\gamma}_i} d_{ij}^{-\hat{\beta}_i} \quad \text{(hypothesis: } \hat{\gamma}_i > 0,\ \hat{\beta}_i < 0) \quad (4.38).$$

This is equivalent to,

$$\hat{I}_{ij} = \hat{\alpha}_i\, d_{ij}^{\hat{\gamma}_i}\, d_{ij}^{\hat{\beta}_i} \quad \text{(hypothesis: } \hat{\gamma}_i < 0,\ \hat{\beta}_i < 0). \quad (4.39)$$

As $|R_i(m*d*)|$ increases within region (B), $\hat{\gamma}_i$ decreases until the line AC is reached when $\hat{\gamma}_i = 0.0$ (see the proof of theorem (ii)). Thus, as $\hat{\gamma}_i$ decreases, more of the pure distance effect on interaction is measured by $\hat{\beta}_i$ which decreases.

Theorem (xi):

There exists a feasible region on the graph where mass enters the regression negatively even though $R_i(m*I*)$ is positive. As $R_i(m*d*)$ decreases, $\hat{\beta}_i$ decreases. This is region (C).

Proof:

From equation (4.11) and the proof of theorem (ii), mass enters the regression negatively whenever,

$$R_i(m*I*) < R_i(m*d*) \cdot R_i(d*I*) \quad (4.40)$$

From equation (4.8) and the proof of theorem (i), the lower boundary of region (C) is given by the equation:

$$R_i^2(m^*I^*) + R_i^2(m^*d^*) - 2R_i(m^*I^*) \cdot R_i(d^*I^*) \cdot R_i(m^*d^*) + R_i^2(d^*I^*) - 1 = 0 \quad . \tag{4.41}$$

Thus region (C) exists whenever,

$$R_i^+(m^*I^*) \leq R_i(m^*I^*) < R_i(m^*d^*) \cdot R_i(d^*I^*) \quad , \tag{4.42}$$

where $R_i^+(m^*I^*)$ is a value of $R_i(m^*I^*)$ satisfying (4.41).

Q.E.D.

An explanation of the trend in $\hat{\beta}_i$ in region (C) is as follows. Since $R_i(m^*d^*) < 0$ and $\hat{\gamma}_i < 0$, the interaction equation in terms of distance can be represented as:

$$\hat{I}_{ij} = \hat{\alpha}_i \, d_{ij}^{\hat{\gamma}_i} \, d_{ij}^{\hat{\beta}_i} \qquad \text{(hypothesis: } \hat{\gamma}_i > 0,\ \hat{\beta}_i < 0\text{)}. \tag{4.43}$$

As $|R_i(m^*d^*)|$ increases, the "pseudo-distance" term becomes increasingly inversely related to the real distance term and $\hat{\gamma}_i$ increases. Since $\hat{\gamma}_i$ and $\hat{\beta}_i$ have opposite signs, as $\hat{\gamma}_i$ increases, $\hat{\beta}_i$ becomes increasingly negative to compensate for an increasingly strong positive distance effect in the regression.

4.6 The Graph of $\hat{\gamma}_i$ with Variations in Multicollinearity

This section represents a sidestep from the general discussion on distance-decay parameters and spatial structure but it is included here since it is relevant to the topic of multicollinearity and parameter estimates. The parameter $\hat{\gamma}_i$ is a measure of the elasticity of interaction with respect to mass, and multicollinearity between m* and d* will affect the estimates of this parameter when these estimates are derived from OLS regression. Figure 4.3 represents the variation of $\hat{\gamma}_i$ under varying degrees of multicollinearity and under varying values of $R_i(m^*I^*)$. All estimates of γ_i are derived from the same simulation described in Section 4.3.

The pattern of $\hat{\gamma}_i$ described by Figure 4.3 differs from the pattern of $\hat{\beta}_i$ described by Figure 4.1 in several respects. The reason for the differences is that in the simulation $\{d_{ij}\}$ and $\{I_{ij}\}$ are constant so that the main determinant of $\hat{\beta}_i$, $R_i(d^*I^*)$, is a constant. Since $\{m_j\}$ varies in each simulation, the main determinant of $\hat{\gamma}_i$, $R_i(m^*I^*)$, is not a constant. This produces variation in $\hat{\gamma}_i$ along both axes of Figure 4.3 whereas $\hat{\beta}_i$ only varies along the horizontal axis of Figure 4.1. Another difference between the two graphs is the positioning of the line dividing regions of homogeneity with respect to parameter variations. In Figure 4.1 this is the curved line AC which in terms of $|R_i(m^*d^*)|$ and $R_i(m^*I^*)$ co-ordinates runs from (0,0) to $(1, |R_i(d^*I^*)|)$. In Figure 4.3 the respective line is the curved line YC which runs from point Y to the point where $|R_i(m^*d^*)| = 1.0$ and $R_i(m^*I^*) = |R_i(d^*I^*)|$.

The main differences between Figures 4.3 and 4.1 are now explained theoretically.

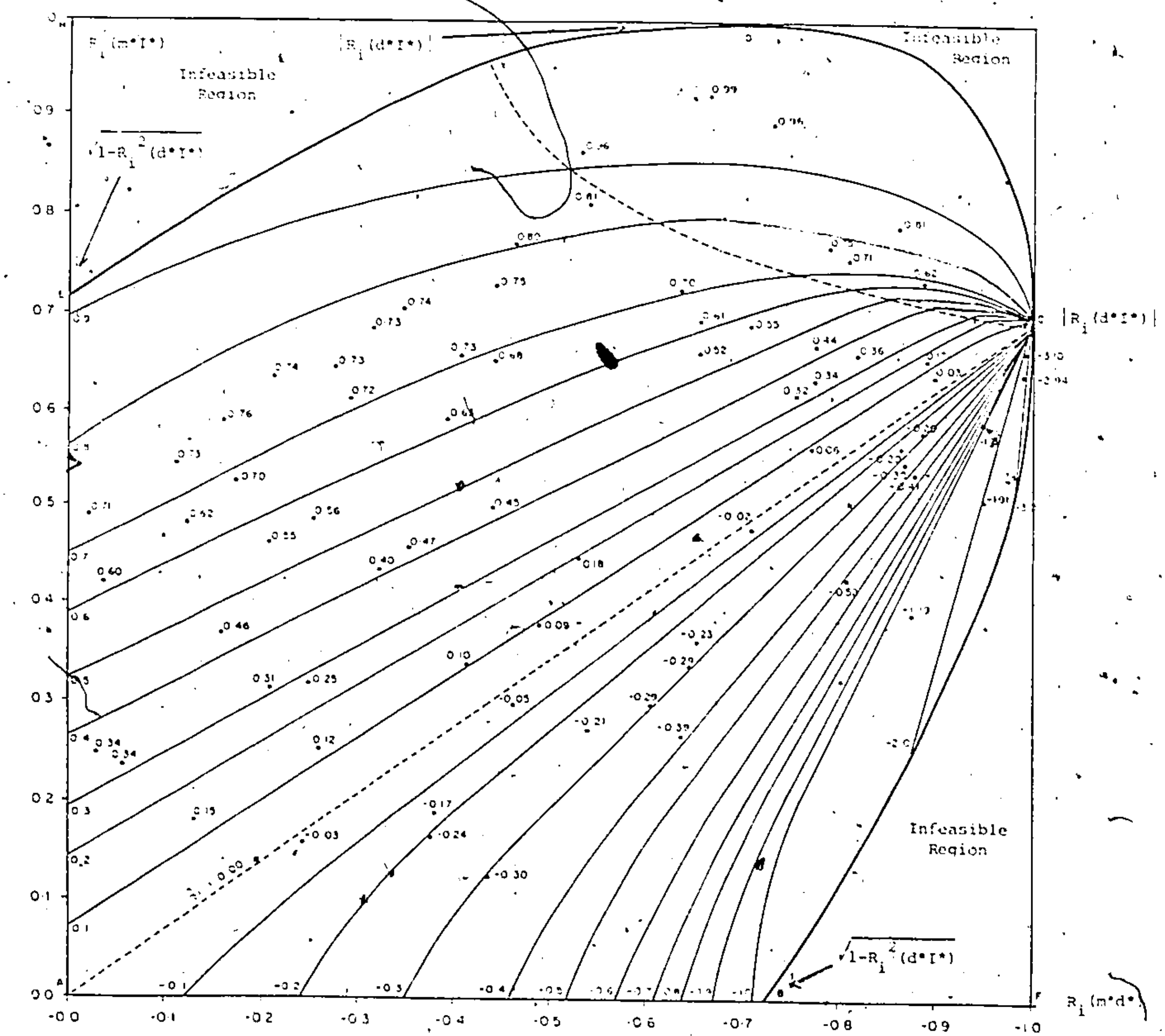
Theorem (xii):

Along the line AE where $R_i(m^*d^*) = 0$, $\hat{\gamma}_i$ varies.

Proof:

When $R_i(m^*d^*) = 0$, $\sum_{j=1}^{n} m_j{}'' d_{ij}{}' = 0$ and (4.3) can be rewritten as:

Figure 4.3: The Effect of Variations in $R_i(m^*d^*)$ and $R_i(m^*I^*)$ on $\hat{Y}_i$



$$\hat{\gamma}_i = \frac{\sum_{j=1}^{n} (d_{ij}')^2 \cdot \sum_{j=1}^{n} m_j' \cdot I_{ij}'}{\sum_{j=1}^{n} (d_{ij}')^2 \cdot \sum_{j=1}^{n} (m_j')^2} \tag{4.44}$$

Dividing through by $\sum_{j=1}^{n} (d_{ij}')^2 \cdot \sqrt{\sum_{j=1}^{n} (m_j')^2} \cdot \sqrt{\sum_{j=1}^{n} (I_{ij}')^2}$ gives:

$$\hat{\gamma}_i = R_i(m^*I^*) \cdot \frac{\sqrt{\sum_{j=1}^{n} (I_{ij}')^2}}{\sqrt{\sum_{j=1}^{n} (m_j')^2}} \tag{4.45}$$

Since $\{I_{ij}\}$ is constant, then as $R_i(m^*I^*)$ increases, $\hat{\gamma}_i$ increases.[1]

Q.E.D.

Theorem (xiii):

Along the line AD where $R_i(m^*I^*) = 0$, $\hat{\gamma}_i$ varies.

Proof:

When $R_i(m^*I^*) = 0$, $\sum_{j=1}^{n} m_i' \cdot I_i' = 0$ and (4.3) can be rewritten as:

---

[1]The assumption has been made here that the term $\sqrt{\sum_{j=1}^{n} (m_j')^2}$ does not decrease as $R_i(m^*I^*)$ increases; a situation which could result in no increase in $\hat{\gamma}_i$. While there was no actual constraint on the variance of the mass terms, there was an attempt to keep this variance constant. However, since $\{m_j\}$ varies at each simulation, it is likely that $\sum_{j=1}^{n} (m_j')^2$ will also vary at each simulation. This would produce much more variation in the graph of $\hat{\gamma}_i$ than is shown in Figure 4.3.

$$\hat{\gamma}_i = \frac{-\sum_{j=1}^{n} d_{ij}' \cdot I_{ij}' \cdot \sum_{j=1}^{n} m_j' \cdot d_{ij}'}{\sum_{j=1}^{n} (d_{ij}')^2 \cdot \sum_{j=1}^{n} (m_j')^2 - (\sum_{j=1}^{n} m_j' \cdot d_{ij}')^2} \quad (4.46)$$

As $|R_i(m^*d^*)| \longrightarrow 1$, the denominator of (4.46) $\longrightarrow 0$ while the numerator increases. Thus, as $|R_i(m^*d^*)|$ increases, $|\hat{\gamma}_i|$ becomes large and as $R_i(m^*d^*) \longrightarrow -1$, $\hat{\gamma}_i \longrightarrow -\infty$.

Q.E.D.

Theorem (xiv):

The line YC is defined by the condition that $\frac{\partial \hat{\gamma}_i}{\partial R_i(m^*d^*)} = 0$ *and the equation of this line is* $R_i(m^*I^*) = R_i(d^*I^*) \cdot [1+R_i^2(m^*d^*)]/2R_i(m^*d^*)$. *For combinations of* $R_i(m^*I^*)$ *and* $R_i(m^*d^*)$ *that lie above this line, as* $|R_i(m^*d^*)|$ *increases,* $\hat{\gamma}_i$ *increases. For combinations of* $R_i(m^*I^*)$ *and* $R_i(m^*d^*)$ *that lie below this line, as* $|R_i(m^*d^*)|$ *increases,* $\hat{\gamma}_i$ *decreases.*

Proof:

Dividing the numerator and denominator of (4.3) by $\sum_{j=1}^{n} (d_{ij}')^2 \cdot \sum_{j=1}^{n} (m_j')^2 \cdot \sqrt{\sum_{j=1}^{n} (I_{ij}')^2}$ gives

$$\hat{\gamma}_i = \frac{R_i(m^*I^*) - R_i(d^*I^*) \cdot R_i(m^*d^*)}{1 - R_i^2(m^*d^*)} \cdot \frac{\sqrt{\sum_{j=1}^{n} (I_{ij}')^2}}{\sqrt{\sum_{j=1}^{n} (m_j')^2}} \quad (4.47)$$

Define $C \equiv \frac{\sqrt{\sum_{j=1}^{n} (I_{ij}')^2}}{\sqrt{\sum_{j=1}^{n} (m_j')^2}}$ since $\sqrt{\sum_{j=1}^{n} (I_{ij}')^2}$ is non-zero constant

and $\sqrt{\sum_{j=1}^{n} (m_j')^2}$ is assumed to be constant,[1] and then the derivative of $\hat{\gamma}_i$ w.r.t. $R_i(m^*d^*)$ is:

$$\frac{\delta\hat{\gamma}_i}{\delta R_i(m^*d^*)} = \frac{[1 - R_i^2(m^*d^*)][- R_i(d^*I^*)] - [- 2R_i(m^*d^*)][R_i(m^*I^*) - R_i(d^*I^*) \cdot R_i(m^*d^*)]}{[1 - R_i^2(m^*d^*)]^2} \cdot C \tag{4.48}$$

For any given value of $R_i(m^*I^*)$ the line YC represents the point at which the derivative $\delta\hat{\gamma}_i/\delta R_i(m^*d^*) = 0$. Thus, the line YC is defined when:

$$R_i(m^*I^*) = \frac{R_i(d^*I^*)[1 + R_i^2(m^*d^*)]}{2R_i(m^*d^*)} \tag{4.49}$$

Q.E.D.

Equation (4.48) indicates that as $|R_i(m^*d^*)| \longrightarrow 1$, $\delta\hat{\gamma}_i/\delta R_i(m^*d^*) \longrightarrow -\infty$ when $R_i(m^*I^*) < |R_i(d^*I^*)|$. This explains the rapid decrease of $\hat{\gamma}_i$ in this region of Figure 4.3.

Point Y on Figure 4.3 is determined by the intersection of the line YC and AE. Thus at Y the two conditions hold that:

---

[1] See preceding footnote.

$$R_i^{\,2}(m^*I^*) + R_i^{\,2}(m^*d^*) - 2R_i(m^*I^*)\cdot R_i(d^*I^*)\cdot R_i(m^*d^*) + R_i^{\,2}(d^*I^*) - 1 = 0 \tag{4.50}$$

and,

$$R_i(m^*I^*) = \frac{R_i(d^*I^*)[1 + R_i^{\,2}(m^*d^*)]}{2R_i(m^*d^*)} . \tag{4.51}$$

To determine the coordinates of Y, substitute (4.51) into (4.50) and rearrange to give:

$$\frac{R_i^{\,2}(d^*I^*)}{4R_i^{\,2}(m^*d^*)} + R_i^{\,2}(m^*d^*)[1 + \frac{3R_i^{\,2}(d^*I^*)}{4}] = 1 - \frac{R_i^{\,2}(d^*I^*)}{2} \tag{4.52}$$

in which there is only one unknown, $R_i(m^*d^*)$. In this situation, when $R_i(d^*I^*) = -0.7$, $R_i(m^*d^*) = -0.44$. Substituting these values into (4.50) gives $R_i(m^*I^*) = 0.95$. Thus, the coordinates of point Y are (- 0.44, 0.95). The minimum value of (4.51) is 0.7 when $|R_i(m^*d^*)| = 1.0$ and this determines an end point of the line being at C.

Theorem (xv):

When $\hat{\gamma}_i = 0$, $\hat{\beta}_i$ remains constant (the straight line AC). However, when $\hat{\beta}_i = 0$, $\hat{\gamma}_i$ varies.

Proof:

The normal equations for $\hat{\beta}_i$ and $\hat{\gamma}_i$ can be stated as follows:[1]

$$\hat{\beta}_i = \frac{\sum_{j=1}^{n} d_{ij}' \cdot I_{ij}' - \hat{\gamma}_i \sum_{j=1}^{n} m_j' \cdot d_{ij}'}{\sum_{j=1}^{n} (d_{ij}')^2} \tag{4.53}$$

---

[1] See, for example, Beals [1972, p. 269].

$$\hat{\gamma}_i = \frac{\sum_{j=1}^{n} d_{ij}' \cdot I_{ij}' - \hat{\beta}_i \sum_{j=1}^{n} (d_{ij}')^2}{\sum_{j=1}^{n} m_j' \cdot d_{ij}'} \quad (4.54)$$

From (4.53) when $\hat{\gamma}_i = 0$,

$$\hat{\beta}_i = \frac{\sum_{j=1}^{n} d_{ij}' \cdot I_{ij}'}{\sum_{j=1}^{n} (d_{ij}')^2} \quad (4.55)$$

which is the value of $\hat{\beta}_i$ along the straight line AC in Figure 4.1. This is a constant since $\{I_{ij}\}$ and $\{d_{ij}\}$ are constant.

From (4.54), when $\hat{\beta}_i = 0$,

$$\hat{\gamma} = \frac{\sum_{j=1}^{n} d_{ij}' \cdot I_{ij}'}{\sum_{j=1}^{n} m_j' \cdot d_{ij}'} \quad (4.56)$$

which is the value of $\hat{\gamma}_i$ along the straight line DC in Figure 4.1. This is not a constant since $\left|\sum_{j=1}^{n} m_j' \cdot d_{ij}'\right|$ increases as $|R_i(m^*d^*)|$ increases.

Q.E.D.

The implication of this result is that in this simulation, given fixed interactions and distances, if the mass term does not enter the regression, $\hat{\beta}_i$ is a fixed value. However, if the distance term does not enter the regression, $\hat{\gamma}_i$ can vary over a range of values.

4.7 Implications of the Results

(1) The Multicollinearity Problem

Figure 4.1 supports an earlier assertion on the variance of parameter estimates. As $|R_i(m*d*)| \longrightarrow 1.0$, any slight variation in one of the correlation coefficients produces an extreme variation in $\hat{\beta}_i$. As $|R_i(m*d*)| \longrightarrow 1.0$, the variance of $\hat{\beta} \longrightarrow \infty$. When $|R_i(m*d*)|$ is low there is little variance in $\hat{\beta}_i$. However, a cut-off point at which the effect of multicollinearity becomes too severe is not obvious. Figure 4.1 demonstrates the earlier hypothesis that the effects of multicollinearity are continuous. It has also been shown that the effects of multicollinearity cannot be assessed simply by a measure of the relationship between the two relevant independent variables. In the case of the unconstrained interaction model the effect of multicollinearity on parameter estimates is not only a function of $R_i(m*d*)$ but also of $R_i(m*I*)$ and $R_i(d*I*)$. From Figure 4.1 the effects of multicollinearity appear to be more severe when $R_i(m*I*)$ is low. From the proof of theorem (i) the effects of multicollinearity are more severe when $|R_i(d*I*)|$ is low. Thus, the effects of multicollinearity when $|R_i(m*d*)| = 0.7$, $R_i(m*I*) = 0.2$ and $|R_i(d*I*)| = 0.2$ are likely to be more severe and more of a problem in parameter estimation than when $|R_i(m*d*)| = 0.9$, $R_i(m*I*) = 0.8$ and $|R_i(d*I*)| = 0.8$. Simply comparing the degree of multicollinearity would lead to the opposite conclusion.

Another implication of these results is that the effect of multicollinearity on parameter estimation is unpredictable. There is rarely a monotonic relationship between $R_i(m*d*)$ and $\hat{\beta}_i$. Only when $R_i(m*I*) = 0.0$ and when $R_i(m*I*) > |R_i(d*I*)|$ is this so. This again indicates how the effects of multicollinearity are dependent upon other relationships.

(ii) Parameter Estimation

The pattern of estimated distance-decay values described by Figure 4.1 has been theoretically derived. A model-specific spatial structure bias in $\hat{\beta}_i$ is given explicitly - $\hat{\beta}_i$ can vary due to variations in $R_i(m^*d^*)$ even though such variations may not have any effect on people's perceived disutility of travel. The mechanism of this confusion is demonstrated in the proofs of theorems (ix), (x) and (xi). When multicollinearity is present, the relationships between each independent variable and the dependent variable are obscured. In the case of the unconstrained interaction model when m_j^* is a function of d_{ij}^*, $\hat{\gamma}$ is, in part, a measure of the relationship between d_{ij}^* and I_{ij}^*. Similarly, $\hat{\beta}_i$ is, in part, a measure of the relationship between m_j^* and I_{ij}^* and variations in masses affect not only $\hat{\gamma}_i$ but also $\hat{\beta}_i$. The obfuscation is greatest when $|R_i(m^*d^*)|$ approaches 1.0, that is, at extreme degrees of multicollinearity. This is the spatial structure effect discussed by Curry [1972, p. 132] when he states of the true friction of distance and spatial structure:

> "It is clear that where there is a genuine friction of distance, the two quite separate entities are being hopelessly confounded. It is only in the unlikely event of zero autocorrelation that the distance exponent may be read directly as a friction term."

The latter point is also borne out in Figure 4.1. The portion of the vertical axis between A and E represents situations in which there is zero autocorrelation between mass terms or where $R_i(m^*d^*) = 0.0$. Here

$\hat{\beta}_i$ is invariant and equal to its "base value". Variations in the set of masses along this line produce no variation in $\hat{\beta}_i$.

The confusion in the interpretation of $\hat{\beta}_i$ described above can have spatial implications. Given variations in the masses of origins or in the distances to destinations, $R_i(m^*d^*)$ will vary between origins. There could be a systematic variation from one extreme at the most central origin to the other extreme at the most peripheral origin. If the underlying interaction behaviour was constant over space, and if $R_i(m^*d^*)$ varied systematically, a systematic spatial variation in $\hat{\beta}_i$ would result. However, the large variation in $\hat{\beta}_i$ described in earlier empirical studies could only result from a large systematic variation in $R_i(m^*d^*)$. There is no empirical evidence to suggest that such variation in $R_i(m^*d^*)$ is an empirical regularity. Chapter 7 presents evidence that a large variation in $\hat{\beta}_i$ results even when the variation in $R_i(m^*d^*)$ is small. Another spatial structure effect accounts for the variation in $\hat{\beta}_i$. For individual origins, however, where multicollinearity in the data set is very high, the estimates of β_i should be treated with great caution since they will contain the type of spatial structure bias outlined in this chapter.

(iii) The Aggregation Problem

As Openshaw [1977] reports, the estimated distance-decay parameter in spatial interaction modelling is determined in part by the zoning system used. For example, consider New York City as a zone centroid for two zones - the first zone is New York State and the second is the North-East U.S.A.. Calibrating (4.1) using interaction data for both zones is likely to produce different $\hat{\beta}_i$ values for New York City. This is the essence of the aggregation or zoning system problem in spatial interaction modelling. The origin-specific distance-decay parameter will vary as the zoning system varies. As Openshaw notes [1977, pp. 169-170]:

> "The effects of the zoning system on the correlation coefficient is well known but the sensitivity of many other statistical techniques and that of spatial interaction models has not yet been demonstrated."

This study demonstrates the link between correlation coefficients and the distance-decay parameter of an unconstrained spatial interaction model. It also demonstrates the sensitivity of the distance-decay parameter to changes in the zoning system. In an unconstrained interaction model the reason for a specific β_i having different estimated values at different levels of data aggregation is that the correlation coefficients $R_i(m^*d^*)$ and $R_i(m^*I^*)$ vary. Different levels of data aggregation or different zoning systems are likely to have different values of $R_i(m^*I^*)$ and $R_i(m^*d^*)$ for each origin. These varying values of the correlation coefficients produce variations in $\hat{\beta}_i$ in accordance with the trends given in Figure 4.1. If an alternative zoning-system was used which did not alter the correlation coefficients, then $\hat{\beta}_i$ would be identical for both zoning systems.

4.8 Summary

This chapter has demonstrated that a model-specific spatial structure effect can be included in the estimation of β_i in an unconstrained spatial interaction model. The presence of multicollinearity confuses the interpretation of $\hat{\beta}_i$ and $\hat{\gamma}_i$ such that $\hat{\beta}_i$ can include a measure of the mass-interaction relationship and $\hat{\gamma}_i$ can include a measure of the distance-interaction relationship.

Earlier work concerned with identifying this model-specific spatial structure ~~bias~~ has concentrated on the effects of spatial autocorrelation. This is shown to be too vague. Spatial autocorrelation statistics in this respect are only general substitutes for the more exact determinants of $\hat{\beta}_i$ - the correlation coefficients between mass and distance, mass and interaction and between distance and interaction. Theoretically, $\hat{\beta}_i$ is a function of correlation coefficients and not of a spatial autocorrelation statistic. Since $\hat{\beta}_i$ can be biased by multicollinearity, it may be possible to remove this bias by conventional methods of eliminating multicollinearity (see Huang [1970, pp. 154-158]). Alternatively, the technique outlined by Fotheringham and Webber [1980] may be useful in this respect since it redefines m_j* in terms of a sum of logarithmic distances.

While the approach described in this chapter to identify a model-specific spatial structure bias in an unconstrained interaction model is more useful than the use of spatial autocorrelation statistics, the actual degree of this type of spatial structure bias is in doubt. If multicollinearity is extreme then $\hat{\beta}_i$ is likely to be strongly biased and misleading. However, there is no evidence to suggest that a large systematic variation in multicollinearity is an empirical regularity to the extent that it might explain the variance of parameter values described in Chapter 1.

Thus, the evidence presented in this analysis supports both the work of Curry, Griffith and Sheppard [1975] and of Cliff, Martin and Ord [1974, 1975, 1976]. Spatial structure does appear to determine $\hat{\beta}$ to some extent due to multicollinearity - a point consistent with the views of Curry et al. but this model-specific spatial structure effect is unlikely to be the cause of the regular spatial pattern of $\hat{\beta}_i$ values commonly found in empirical studies - a point consistent with the views of Cliff et al. Since multicollinearity is unlikely to be the main spatial structure effect included in $\hat{\beta}_i$, other types of spatial structure bias in interaction modelling are now investigated.

CHAPTER FIVE

MODEL-SPECIFIC SPATIAL STRUCTURE EFFECTS IN CONSTRAINED INTERACTION MODELS: BALANCING FACTORS AND $\hat{\beta}$

5.1 Introduction

Section 3.2 posited a relationship between the balancing factors of constrained interaction models and the estimated distance-decay parameter. From (1.35), (1.37) and (1.38) it is evident that the balancing factors are quasi-accessibility measures and as such are measures of spatial structure. The question investigated here is: "what is the relationship between balancing factors and estimated distance-decay parameters?". If a relationship exists, this would be evidence of a link between spatial structure and $\hat{\beta}$ in constrained interaction modelling. The question is answered by the use of simple calculus and an understanding of the estimation of β.

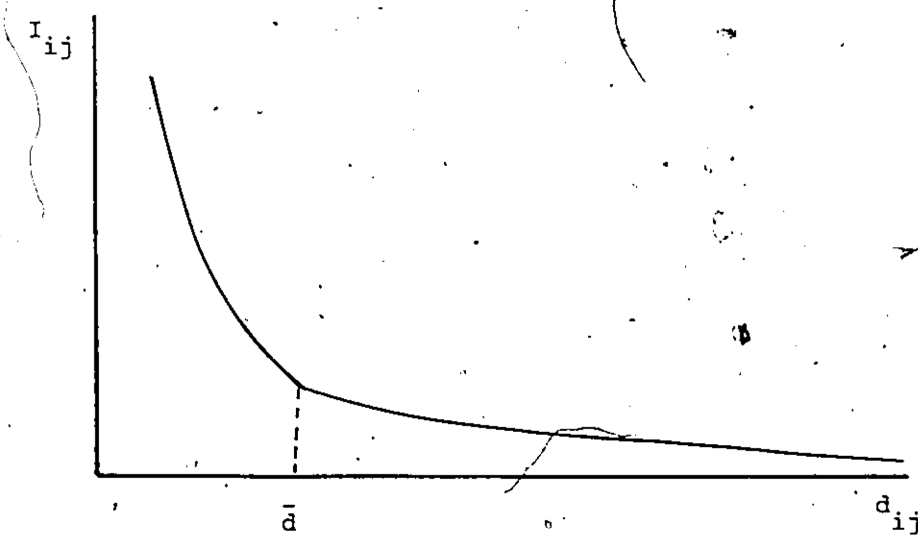
Consider Figure 5.1a in which actual interaction, I_{ij}, is graphed against distance, and Figure 5.1b in which predicted interaction, $\hat{I}_{ij}$, is graphed against distance. The calibration of a constrained interaction model is undertaken with the objective of maximising the similarity between the two curves shown in Figures 5.1a and 5.1b usually utilising the constraint that $\hat{\bar{d}}$, the predicted mean distance travelled, equals $\bar{d}$, the true mean distance travelled.[1] The only degree of freedom

[1] Other constraints such as on predicted total outflows, predicted total inflows and total predicted interaction may be used in conjunction with the mean distance constraint but they are of no concern here.

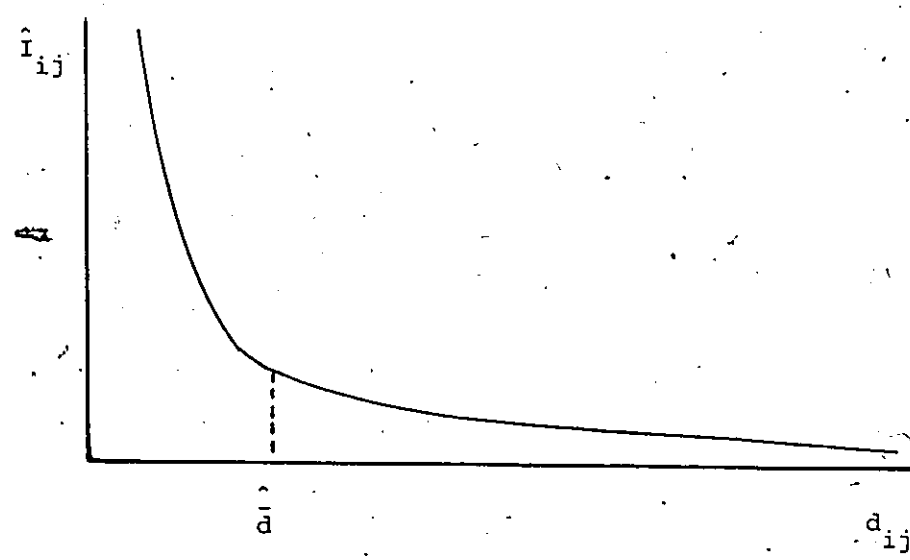
with which to achieve this maximisation is $\hat{\beta}$, the estimated distance-decay parameter. This estimate is then a result of achieving the closest possible resemblance between the actual interaction-distance relationship and the predicted interaction-distance relationship.

FIGURE 5.1: Relationships between Actual and Predicted Interactions and Distance

(a)



(b)



Assume that in a spatial system consisting of accessible and inaccessible centres, interaction behaviour with respect to distance is the same for all origins. With the constraints on flow totals, β_i is estimated so that $\hat{\bar{d}}_i = \bar{d}_i$ and the similarity between the actual and predicted interaction-distance relationship is maximised for each origin. If β_i is interpreted as being a purely behavioural parameter, the expectation would be that $\hat{\beta}_i$ is a constant for all origins since the derivative of actual interaction with respect to distance is the same for all origins. The subsequent analysis demonstrates that this expectation is not realised unless all origins are equally accessible. The accessibility of origin i determines, in part, the derivative of predicted interaction with respect to distance and $\hat{\beta}_i$ adjusts to compensate for this effect so that the predicted derivatives are constant regardless of accessibility. If the predicted derivatives were not constant, then the predicted interaction-distance relationship would not be equal to the actual interaction-distance relationship for each origin. Hence, the estimated distance-decay parameter is related to the accessibility of i which is a measure of spatial structure. $\hat{\beta}_i$ will not be constant for all origins even if actual interaction behaviour with respect to distance is constant, unless each origin is equally accessible. The exact nature of this relationship is described below.

5.2 A Singly-Constrained Spatial Interaction Model

The origin-specific, singly-constrained interaction model given in (1.34) can be written, when calibrated, as:

$$\hat{I}_{ik} = \frac{O_i \; D_k \; d_{ik}^{\hat{\beta}_i}}{\sum_{j=1}^{n} D_j \; d_{ij}^{\hat{\beta}_i}} \tag{5.1}$$

where subscript k denotes a particular destination. From (5.1) the partial derivative of predicted interaction with respect to distance to the particular destination, k, is given by:[1]

$$\frac{\partial \hat{I}_{ik}}{\partial d_{ik}} = \frac{\hat{\beta}_i \; O_i \; D_k \; d_{ik}^{\hat{\beta}_i - 1} \left(\sum_{j=1}^{n} D_j \; d_{ij}^{\hat{\beta}_i} - D_k \; d_{ik}^{\hat{\beta}_i} \right)}{\left(\sum_{j=1}^{n} D_j \; d_{ij}^{\hat{\beta}_i} \right)^2} \tag{5.2}$$

or alternatively,

$$\frac{\partial \hat{I}_{ik}}{\partial d_{ik}} = \frac{\hat{\beta}_i \; O_i \; D_k \; d_{ik}^{\hat{\beta}_i - 1} \left(\sum_{\substack{j=1 \\ j \neq k}}^{n} D_j \; d_{ij}^{\hat{\beta}_i} \right)}{\left(\sum_{j=1}^{n} D_j \; d_{ij}^{\hat{\beta}_i} \right)^2} , \tag{5.3}$$

which is only negative when $\hat{\beta}_i$ is negative - the standard expectation. Hence, if $\hat{\beta}_i < 0$, as the distance between i and k increases, the predicted interaction between i and k decreases. However, (5.3) indicates that the rate at which predicted interaction declines as distance increases does not depend only upon $\hat{\beta}_i$. Let A_i represent the accessibility of origin i and define,

$$A_i = \sum_{j=1}^{n} D_j \; d_{ij}^{\hat{\beta}_i} . \tag{5.4}$$

---

[1] See Appendix II.

Such an accessibility measure has been discussed by Hansen [1959] and used, *inter alia*, by Fotheringham [1979]. Substituting (5.4) into (5.2) and using the definition of $\hat{I}_{ij}$ given in (5.1), (5.2) can be rewritten as:

$$\frac{\partial \hat{I}_{ik}}{\partial d_{ik}} = \frac{\hat{\beta}_i \hat{I}_{ik} d_{ik}^{-1}(A_i - D_k d_{ik}^{\hat{\beta}_i})}{A_i} \qquad (5.5)$$

and then it is seen that the rate at which predicted interaction between i and j declines as distance increases is a function of the accessibility of origin i. The relationship between the derivative and accessibility is complex and is discussed fully in Section 5.5.

From (5.3), the second derivative of predicted interaction with respect to distance is given by:[1]

$$\frac{\partial^2 \hat{I}_{ik}}{\partial d_{ik}^2} = -\hat{\beta}_i \hat{I}_{ik} d_{ik}^{-2}\left(1 - \frac{D_k d_{ik}^{\hat{\beta}_i}}{A_i}\right)\left(1 - \hat{\beta}_i + 2\hat{\beta}_i \frac{D_k d_{ik}^{\hat{\beta}_i}}{A_i}\right) \qquad (5.6)$$

which, if $\hat{\beta}_i < 0$, is positive when[2],

$$(\hat{\beta}_i + 1) D_k d_{ik}^{\hat{\beta}_i} > (\hat{\beta}_i - 1) \sum_{\substack{j=1 \\ j \neq k}}^{n} D_j d_{ij}^{\hat{\beta}_i} .$$

The inequality is true except under extreme conditions.[3] Hence, when

---

[1]See Appendix III

[2]See Appendix IV

[3]For example, the inequality may not hold when d_{ik} is very small, $\hat{\beta}_i < -1$, n is small, D_k is large and D_j is small $\forall j \neq k$.

$\hat{\beta}_i < 0$ and the inequality is true, predicted interaction decreases at a decreasing rate as distance increases - the standard expectation. Assuming that actual interaction from each origin decreases at a decreasing rate as distance increases, for the predicted interaction-distance relationship to be equal to the actual relationship for each origin, $\hat{\beta}_i < 0 \; \forall \; i$. However, if the actual interaction-distance relationship is the same for all origins, $\hat{\beta}_i$ cannot be equal for all origins since, from (5.5), the accessibility of an origin determines, in part, the predicted interaction-distance relationship, and origin accessibility varies. If the predicted interaction-distance relationship is to be the same for all origins, $\hat{\beta}_i$ has to vary to compensate for the variation in origin accessibility. The nature of this relationship is investigated in Section 5.5 where a more representative derivative is obtained - one that does not depend upon the destination chosen as does the derivative in (5.5). Use of the simple derivative given in (5.5), however, can first give useful insights into interaction-distance elasticities.

5.3 A Note on Interaction Elasticities

Let $e(\hat{I}_{ik})$ denote the elasticity of predicted interaction with respect to distance to one particular destination, k, where $e(\hat{I}_{ik})$ is defined as:

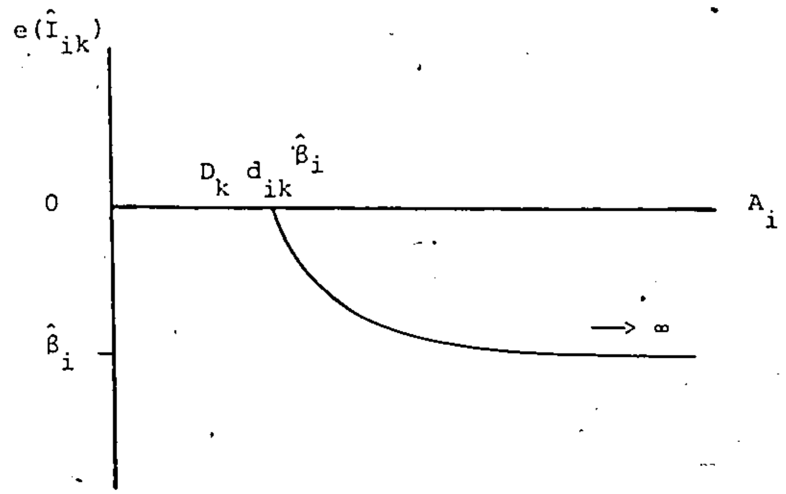
$$e(\hat{I}_{ik}) \equiv \frac{d_{ik}}{\hat{I}_{ik}} \cdot \frac{\partial \hat{I}_{ik}}{\partial d_{ik}} \quad (5.7)$$

Substituting (5.5) into (5.7) and rearranging gives,

$$e(\hat{I}_{ik}) = \hat{\beta}_i (1 - \frac{D_k \, d_{ik}^{\hat{\beta}_i}}{A_i}) \quad . \qquad (5.8)$$

The elasticity of predicted interaction with respect to distance is a function of the distance over which interaction takes place and also of the accessibility of i. Assume that d_{ik} is a constant for all i and then the relationship between $e(\hat{I}_{ik})$ and A_i is given in Figure 5.2. As $A_i \longrightarrow \infty$, $e(\hat{I}_{ik}) \longrightarrow \hat{\beta}_i$. The elasticity can never be greater than zero since in (5.8), $A_i \geq D_k \, d_{ik}^{\hat{\beta}_i}$. At the minimum value of A_i, $D_k \, d_{ik}^{\hat{\beta}_i}$, $e(\hat{I}_{ik}) = 0$; the interpretation then is that there is only one possible destination and as distance to that destination increases, there can be no change in interaction patterns to alternative destinations. Given an outflow from origin i, all of that outflow must terminate at destination k. Hence, $\hat{I}_{ik}$ is a constant with respect to d_{ik} and $e(\hat{I}_{ik}) = 0$.

FIGURE 5.2: Interaction-Distance Elasticities and Accessibility



Equation (5.8) indicates that $e(\hat{I}_{ik})$ decreases as the accessibility of i increases. Consider the elasticity of true interactions, $e(I_{ik})$, being constant for all origins: that is, interaction behaviour is constant for all origins. Then $e(\hat{I}_{ik})$ can only be constant for all origins if $\hat{\beta}_i$ varies since A_i is assumed to vary between origins. From (5.8), $e(\hat{I}_{ik})$ decreases as A_i increases and as $\hat{\beta}_i$ decreases. To ensure that $e(\hat{I}_k)$ is constant for all i, $\hat{\beta}_i$ must increase as A_i increases. The resulting pattern of origin-specific distance-decay parameters will be one in which accessible origins have low negative values and inaccessible origins have high negative values. This is the pattern found in the empirical studies discussed in Section 1.5 and the result is reinforced below.

5.4 Deriving "Average" Derivatives

The preliminary analysis in the preceding sections indicates the methodology used in determining the relationship between $\hat{\beta}_i$ and spatial structure which is specific to constrained interaction modelling. However, the derivative of predicted interaction with respect to distance which has been used so far is with respect to one particular destination, k. The value of the derivative will vary as the destination k varies since in (5.2) the terms $D_k\, d_{ik}^{\hat{\beta}_i - 1}$, $\sum_{\substack{j=1 \\ j \neq k}}^{n} D_j\, d_{ij}^{\beta_i}$ and $\sum_{j=1}^{n} D_j\, d_{ij}^{\beta_i}$ are not constant for all k. A bias would result if the derivative given in (5.2) were used in a comparison between origins when k is closer to one origin than another. To accurately compare derivatives between origins in a spatial system, an "average" derivative is required which is taken with respect to all destinations. At least two formulations appear reasonable for an "average" derivative where the average derivative is denoted by $\left(\frac{\partial \hat{I}_{ik}}{\partial d_{ik}}\right)^{Av}$:

$$\left(\frac{\partial \hat{I}_{ik}}{\partial d_{ik}}\right)^{Av} = \frac{\sum_{k=1}^{n} \frac{\partial \hat{I}_{ik}}{\partial d_{ik}}}{n} \tag{5.9}$$

and,

$$\left(\frac{\partial \hat{I}_{ik}}{\partial d_{ik}}\right)^{Av} = \frac{\sum_{k=1}^{n} \frac{\partial \hat{I}_{ik}}{\partial d_{ik}} \cdot D_k}{\sum_{k=1}^{n} D_k} \tag{5.10}$$

The average derivative given in (5.10) was chosen for all subsequent analyses since the derivative with respect to any particular destination is weighted by the value of the attractiveness of that destination. At equal distances from an origin, a destination with an attractiveness of 100 units should be a more important component of the average derivative than a destination with an attractiveness of 10 units. This is modelled in (5.10) but not in (5.9). In terms of (5.3), (5.10) can be rewritten as:

$$\left(\frac{\partial \hat{I}_{ik}}{\partial d_{ik}}\right)^{Av} = \frac{\hat{\beta}_i \, O_i}{\sum_{k=1}^{n} D_k} \cdot \sum_{k=1}^{n} \frac{D_k^{\,2} \, d_{ik}^{\,\hat{\beta}_i - 1} \left(\sum_{\substack{j=1 \\ j \neq k}}^{n} D_j \, d_{ij}^{\,\hat{\beta}_i}\right)}{\left(\sum_{j=1}^{n} D_j \, d_{ij}^{\,\hat{\beta}_i}\right)^2} \tag{5.11}$$

5.5 Accessible and Inaccessible Origins

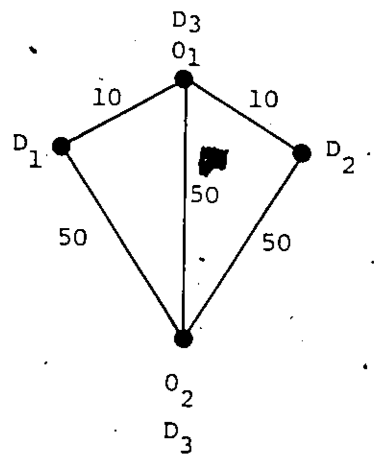
In a closed[1] spatial system consider an origin that is accessible ($\sum_{j=1}^{n} D_j d_{ij}^{\hat{\beta}_i}$ is large) and an origin that is inaccessible ($\sum_{j=1}^{n} D_j d_{ij}^{\hat{\beta}_i}$ is small). Assume that for each origin the number of interactions leaving is equal to the number of interactions arriving and assume that this value is the same for the accessible origin and the inaccessible origin. Both origins are then confronted with the same destination opportunities but at different distances. Figure 5.3a illustrates such a spatial system and Figures 5.3b and 5.3c illustrate the destination opportunities and distance relationships faced by an accessible origin and an inaccessible origin respectively, in the system.

For the spatial system given in Figure 5.3 a comparison is made of the average derivatives $\left(\frac{\partial \hat{I}_{1k}}{\partial d_{1k}}\right)^{Av}$ and $\left(\frac{\partial \hat{I}_{2k}}{\partial d_{2k}}\right)^{Av}$. To make an accurate comparison, these values must be compared at the same value of d_{ij} since it has been shown that $\frac{\partial \hat{I}_{ij}}{\partial d_{ij}}$ varies with d_{ij}. The methodology used to relate $\hat{\beta}_i$ to spatial structure is similar to that given in Section 5.3. Assume that actual interaction behaviour with respect to distance, *ceteris paribus*, is the same for all origins. Then, if the traditional interpretation of $\hat{\beta}_i$ is true, $\hat{\beta}_i$ should be equal for all origins since it is held to be solely a measure of interaction behaviour. A corollary of this is that if $\hat{\beta}_i$ is assumed to be a constant for all origins, the average derivative of predicted interactions with respect to distance

[1]All interactions are within the system.

FIGURE 5.3: Spatial Systems

(a) The Spatial Structure

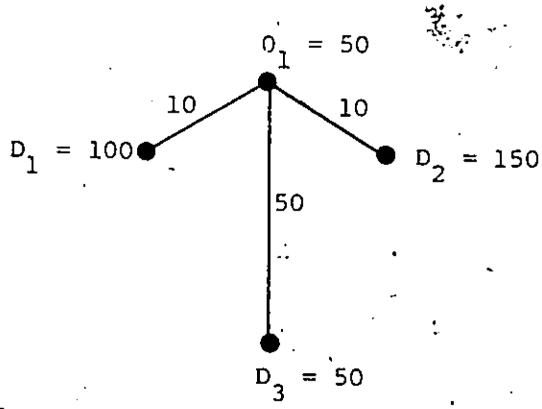


O_1 is an accessible origin

O_2 is an inaccessible origin

<u>50</u> is the distance between centres

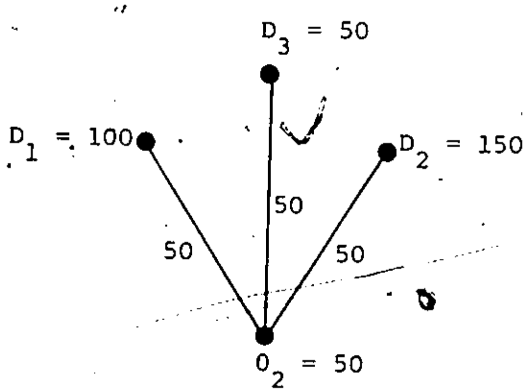
(b) Destination Opportunities Faced by the Accessible Origin O_1



$O_1 = 50$ is the number of interactions originating from O_1.

$D_3 = 50$ is the number of interactions terminating at D_3.

(c) Destination Opportunities Faced by the Inacessible Origin O_2



should be a constant for all origins. Thus, for the spatial system in Figure 5.3, a comparison is made of the average derivative for the accessible and inaccessible origin under the assumptions of constant real interaction behaviour and constant $\hat{\beta}_i$ values. If the average derivative of predicted interaction with respect to distance is not equal for each origin, this would indicate that spatial structure is a component of $\hat{\beta}_i$ since $\hat{\beta}_i$ will vary in such a way that it ensures predicted interaction behaviour is as similar to actual interaction behaviour as possible. The nature of this variation, if present, and its relationship with spatial structure is investigated.

Rearranging the numerator of (5.11) gives,

$$\left(\frac{\partial \hat{I}_{ik}}{\partial d_{ik}}\right)^{Av} = \frac{\hat{\beta}_i O_i}{\sum_{k=1}^{n} D_k} \cdot \sum_{k=1}^{n} \frac{D_k^{\,2} d_{ik}^{\,\hat{\beta}_i - 1} \cdot \sum_{j=1}^{n} D_j d_{ij}^{\,\hat{\beta}_i} - D_k^{\,3} d_{ik}^{\,2\hat{\beta}_i - 1}}{\left(\sum_{j=1}^{n} D_j d_{ij}^{\,\hat{\beta}_i}\right)^2} \tag{5.12}$$

Define,

$$A_i(k) \equiv \sum_{j=1}^{n} D_j d_{ij}^{\,\hat{\beta}_i} \equiv \sum_{\substack{j=1 \\ j\neq k}}^{n} D_j d_{ij}^{\,\hat{\beta}_i} + D_k d_{ik}^{\,\hat{\beta}_i} \tag{5.13}$$

$$C_i(k) \equiv D_k^{\,3} d_{ik}^{\,2\hat{\beta}_i - 1} \tag{5.14}$$

$$B_i(k) \equiv D_k^{\,2} d_{ik}^{\,\hat{\beta}_i - 1} \tag{5.15}$$

where the notation (k) denotes a variable whose value depends upon destination k. Using the above definitions, (5.12) can be rewritten as:

$$\left(\frac{\partial\,\hat{I}_{ik}}{\partial\,d_{ik}}\right)^{Av} = \frac{\hat{\beta}_i\,O_i}{\sum_{k=1}^{n} D_k} \cdot \sum_{k=1}^{n} \frac{B_i(k)\cdot A_i(k) - C_i(k)}{A_i(k)^2} \qquad (5.16)$$

From (5.16) the relationship between $\left(\frac{\partial\,\hat{I}_{ik}}{\partial\,d_{ik}}\right)^{Av}$ and $A_i(k)$ is described by,

$$\frac{\partial\left(\frac{\partial\,\hat{I}_{ik}}{\partial\,d_{ik}}\right)^{Av}}{\partial\,A_i(k)} = \frac{\hat{\beta}_i\,O_i}{\sum_{k=1}^{n} D_k} \cdot \sum_{k=1}^{n} \frac{2C_i(k) - A_i(k)\cdot B_i(k)}{A_i(k)^3} \qquad (5.17)$$

where A_i in the LHS of (5.17) is assumed to vary between origins and not simply as d_{ik} varies. Assuming $\hat{\beta}_i < 0$, the cross partial derivative given in (5.17) is negative, and hence the derivative of predicted interaction with respect to distance becomes increasingly negative as the accessibility of the origin increases, when:

$$2C_i(k) > A_i(k) \cdot B_i(k) \quad \forall\, k\,, \qquad (5.18)$$

that is, when,

$$2D_k^{\,3}\, d_{ik}^{\,2\hat{\beta}_i - 1} > \sum_{j=1}^{n} D_j\, d_{ij}^{\,\hat{\beta}_i} \cdot D_k^{\,2}\, d_{ik}^{\,\hat{\beta}_i - 1} \quad \forall\, k\,, \qquad (5.19)$$

or, more simply, when,

$$D_k\, d_{ik}^{\,\hat{\beta}_i} > \sum_{\substack{j=1\\ j\neq k}}^{n} D_j\, d_{ij}^{\,\hat{\beta}_i} \quad \forall\, k\,. \qquad (5.20)$$

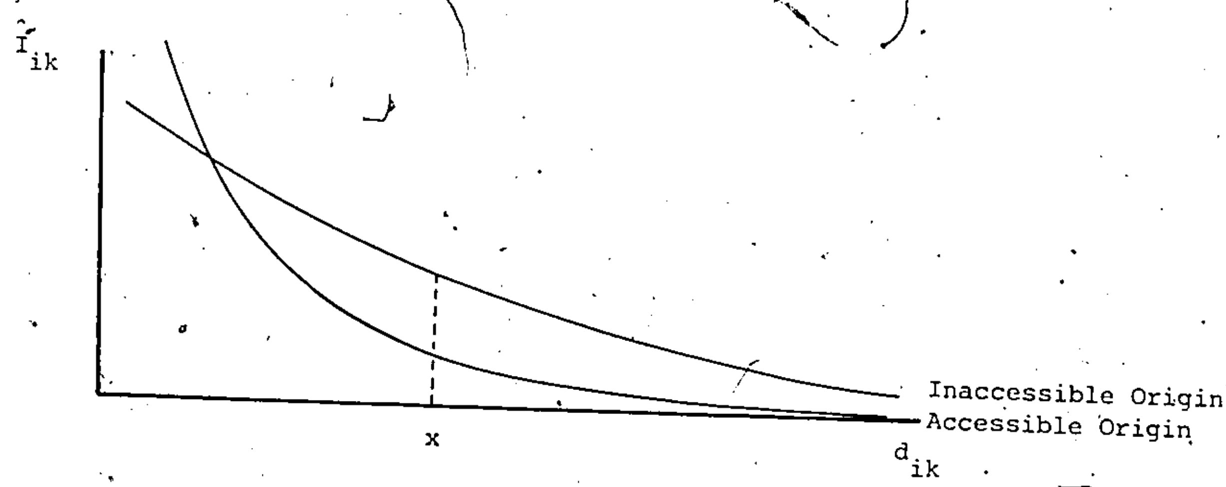
Alternatively, if for any subset of the k's, $A_i(k) \cdot B_i(k) > 2C_i(k)$,

$$\frac{\left(\frac{\partial \hat{I}_{ik}}{\partial d_{ik}}\right)^{Av}}{A_i(k)} < 0 \text{ iff:}$$

$$\sum_{p\varepsilon P} \frac{2C_i(k) - A_i(k)\cdot B_i(k)}{A_i(k)^3} > \left| \sum_{q\varepsilon Q} \frac{2C_i(k) - A_i(k)\cdot B_i(k)}{A_i(k)^3} \right|, \tag{5.21}$$

where $p\varepsilon P$ is the subset of k for which condition (5.18) is true and $q\varepsilon Q$ is the subset of k for which condition (5.18) is false. Condition (5.18) is obviously a stronger condition than (5.21). Both conditions, however, are only true when d_{ik} is small. From (5.20), for example, if D_j is a constant for all j, condition (5.18) is false whenever $d_{ik} \geq d_{ij}$ where $j \neq k$. As d_{ik} increases, the derivative given in (5.17) becomes less negative and at a certain value of d_{ik} it will become positive. Thus, when d_{ik} is small, as the accessibility of an origin increases, $\left(\frac{\partial \hat{I}_{ik}}{\partial d_{ik}}\right)^{Av}$ becomes more negative and vice versa. When d_{ik} is large, as the accessibility of an origin increases, $\left(\frac{\partial \hat{I}_{ik}}{\partial d_{ik}}\right)^{Av}$ becomes less negative and vice versa. These relationships are graphed in Figure 5.4.

FIGURE 5.4: Theoretical Relationships between Predicted Interaction and Distance for an Accessible and an Inaccessible Origin.

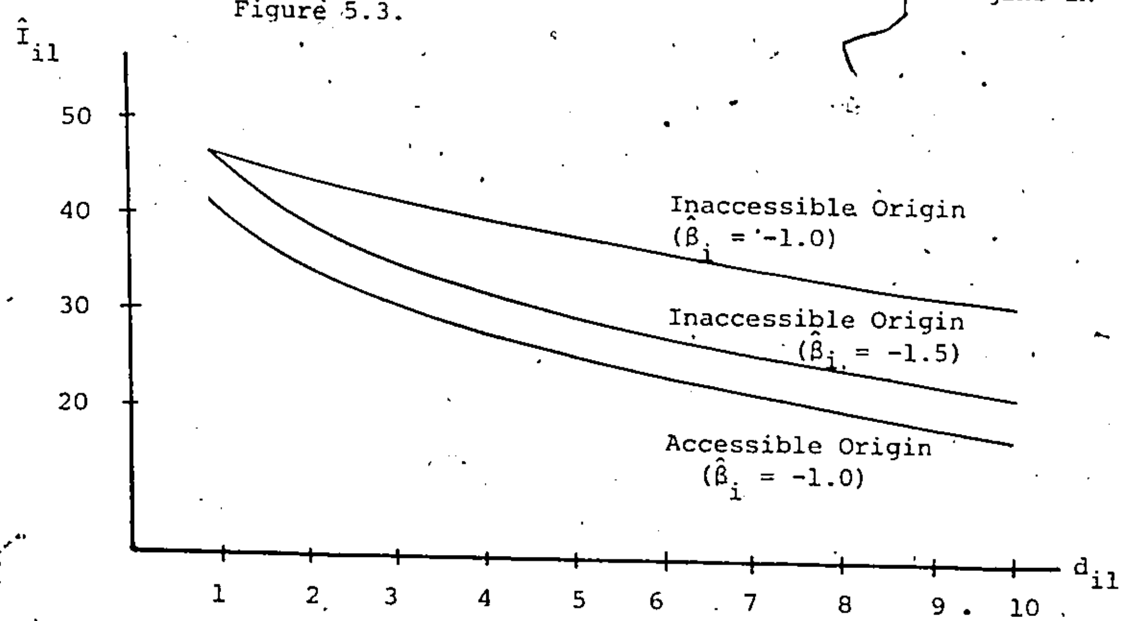


When d_{ik} is between o and x, $\dfrac{\partial \left(\dfrac{\partial \hat{I}_{ik}}{\partial d_{ik}}\right)^{Av}}{\partial A_i(k)} < 0$ and $\left(\dfrac{\partial \hat{I}_{ik}}{\partial d_{ik}}\right)^{Av}$ is more negative for the accessible origin. When $d_{ik} > x$, $\dfrac{\left(\dfrac{\partial \hat{I}_{ik}}{\partial d_{ik}}\right)^{Av}}{\partial A_i(k)} > 0$ and $\left(\dfrac{\partial \hat{I}_{ik}}{\partial d_{ik}}\right)^{Av}$ is less negative for the more accessible origin. The implication of this is that if true interaction-distance relationships are constant over space and $\hat{\beta}_i$ is assumed constant for all i, then the predicted interaction-distance relationships would vary with the accessibility of i according to Figure 5.4. More accessible origins would have steeper predicted interaction-distance relationships than inaccessible origins. To obtain a constant set of predicted interaction-distance relationships, $\hat{\beta}_i$ would have to vary between origins. Specifically, from (5.18) and Figure 5.4, the estimated distance-decay parameter would have to be more negative for inaccessible origins and less negative for accessible origins. This is the frequently-described empirical finding discussed in Section 1.5.

A numerical example of the theoretical findings is obtained by applying the interaction model given in (5.1) to the two origins given in Figure 5.3, one of which is defined as accessible and the other inaccessible. Arbitrarily setting the distance-decay parameter at -1.0 for both origins and using the destination, D_1, a set of predicted interactions was calculated over varying distances of D_1 from both origins. These are graphed in Figure 5.5. The rate of decrease of predicted interactions with respect to increased distance is much steeper for the accessible origin over short distances than for the inaccessible origin. As distance increases, however, the rates tend to converge and their relative steepnesses would eventually reverse. If the actual interaction-distance relationships were constant for both origins, the estimated distance-decay parameter for the inaccessible origin would have to be more negative than the parameter for the accessible origin. Figure 5.5 shows that as a rough approximation $\hat{\beta}_i = -1.5$ for the inaccessible origin gives much the same slope as $\hat{\beta}_i = -1.0$ for the accessible origin. Thus, if actual interaction behaviour with respect to distance is constant over space, $\hat{\beta}_i$ must vary to model such behaviour and this variation is systematic.

Ceteris paribus, $\hat{\beta}_i$ is more negative for inaccessible origins and less negative for accessible origins in order to counteract the effects of spatial structure which result from the balancing factors being accessibility measures. This result is explored more fully below.

FIGURE 5.5: Numerical Relationships between Predicted Interaction and Distance for the Accessible and Inaccessible Origins in Figure 5.3.



For the spatial systems given in Figures 5.3b and 5.3c, the average derivative given in (5.17) was calculated at various distances and the results are given in Table 5.1. Predicted interaction declines more rapidly from the accessible origin than from the inaccessible origin as distance increases up to a distance of between 10 and 20 units. Beyond this distance, predicted interaction from the inaccessible origin declines more rapidly as distance increases - a further demonstration of what was theoretically derived in Figure 5.4.

TABLE 5.1

AVERAGE DERIVATIVES OF PREDICTED INTERACTION WITH RESPECT TO DISTANCE FOR THE ACCESSIBLE AND INACCESSIBLE ORIGINS IN FIGURE 5.3: $O_1 = O_2 = 50$.

d_{ik}	$\left(\frac{\partial \hat{I}_{1k}}{\partial d_{1k}}\right)^{Av}$	$\left(\frac{\partial \hat{I}_{2k}}{\partial d_{2k}}\right)^{Av}$	
1	-5.43	-1.79	$\left(\frac{\partial \hat{I}_{1k}}{\partial d_{1k}}\right)^{Av} < \left(\frac{\partial \hat{I}_{2k}}{\partial d_{2k}}\right)^{Av}$
10	-1.13	-0.90	
20	-0.49	-0.56	$\left(\frac{\partial \hat{I}_{2k}}{\partial d_{2k}}\right)^{Av} < \left(\frac{\partial \hat{I}_{1k}}{\partial d_{1k}}\right)^{Av}$
50	-0.13	-0.22	

1 = Accessible Origin

2 = Inaccessible Origin

So far, the assumption has been made that O_i, the total outflow from origin i, is a constant for the accessible and inaccessible origins. This may be so, for example, if zones are chosen of equal population size. However, it is unlikely in reality that O_i would be equal for accessible and inaccessible origins - consider cities as origins, for example. It is more likely in this case that the total outflow from an accessible origin is greater than that from an inaccessible origin and it is of interest to investigate the effects of varying O_i on the derivative of predicted interactions with respect to distance. This information is difficult to obtain from an equation such as (5.17) since varying O_i alters a D_j which in turn alters many of the terms in that expression. It is easier to consider a numerical example such as that given in Table 5.1.

In Figure 5.3b let $O_1 = 100$ and consequently in Figure 5.3c, $D_3 = 100$. The average derivatives of O_1 and O_2 are then given in Table 5.2. The results indicate that the disparity between the average derivatives of predicted interaction with respect to distance for accessible and inaccessible origins increases over short distances. The rate at which predicted interaction from the accessible origin declines as distance increases is much greater than when the total outflows were equal for both origins. The reason for this is that in (5.17) O_1 has increased from 50 to 100 and there has been a corresponding increase in the absolute value of the derivative. There is no change in the accessibility of origin 1 since when $i = j$, D_j is assumed not to contribute to the accessibility of i. The rate at which predicted interaction from the inaccessible origin declines as distance increases is only slightly greater than when $O_1 = 50$. The reason for this change is that the inaccessible origin has become slightly more accessible since a destination, D_3, has doubled its size.

One other scenario can be envisaged: that in which the population of, or outflow from, an inaccessible origin is greater than the population of, or outflow from, an accessible origin. In Figure 5.3c let $O_2 = 100$ and consequently in Figure 5.3b, $D_3 = 100$. The average derivatives for O_1 and O_2 are then given in Table 5.3.

TABLE 5.2

AVERAGE DERIVATIVES OF PREDICTED INTERACTION
WITH RESPECT TO DISTANCE FOR THE ACCESSIBLE
AND INACCESSIBLE ORIGINS IN FIGURE 5.3: $O_1 = 100$, $O_2 = 50$

d_{ik}	$\left(\frac{\partial \hat{I}_{ik}}{\partial d_{1k}}\right)^{Av}$	$\left(\frac{\partial \hat{I}_{2k}}{\partial d_{2k}}\right)^{Av}$	
1	-11.89	-1.84	
10	- 2.37	-0.99	$\left(\frac{\partial \hat{I}_{1k}}{\partial d_{1k}}\right)^{Av} < \left(\frac{\partial \hat{I}_{2k}}{\partial d_{2k}}\right)^{Av}$
20	- 0.91	-0.60	
50	- 0.20	-0.22	$\left(\frac{\partial \hat{I}_{2k}}{\partial d_{2k}}\right)^{Av} < \left(\frac{\partial \hat{I}_{1k}}{\partial d_{1k}}\right)^{Av}$

1 = Accessible Origin

2 = Inaccessible Origin

TABLE 5.3

AVERAGE DERIVATIVES OF PREDICTED INTERACTION
WITH RESPECT TO DISTANCE FOR THE ACCESSIBLE
AND INACCESSIBLE ORIGINS IN FIGURE 5.3: $O_1 = 50$, $O_2 = 100$.

d_{ik}	$\left(\frac{\partial \hat{I}_{1k}}{\partial d_{1k}}\right)^{Av}$	$\left(\frac{\partial \hat{I}_{2k}}{\partial d_{2k}}\right)^{Av}$	
1	-5.53	-3.58	$\left(\frac{\partial \hat{I}_{1k}}{\partial d_{1k}}\right)^{Av} < \left(\frac{\partial \hat{I}_{2k}}{\partial d_{2k}}\right)^{Av}$
10	-1.15	-1.80	
20	-0.48	-1.12	$\left(\frac{\partial \hat{I}_{2k}}{\partial d_{2k}}\right)^{Av} < \left(\frac{\partial \hat{I}_{1k}}{\partial d_{1k}}\right)^{Av}$
50	-0.12	-0:44	

1 = Accessible Origin 2 = Inaccessible Origin

The results in Table 5.3 reinforce the previous findings on the relationship between population, accessibility and $\hat{\beta}_i$. The disparity between the average derivative of predicted interaction with respect to distance for accessible and inaccessible origins is reduced over short distances and, following the above discussion, the variance of predicted distance-decay parameters will be reduced.

Thus, if actual interaction-distance relationships are constant over space but accessibility varies with origin location, the predicted distance-decay parameter will be more negative for inaccessible origins in order to model reality accurately. If the more accessible origins have greater populations than the inaccessible origins, the disparity in the predicted distance-decay parameters between origins is amplified. If the more inaccessible origins have greater populations than the accessible origins, the disparity in the predicted distance-decay parameters is diminished. Hence, it is not only differences in accessibility that affect $\left(\frac{\partial \hat{I}_{ik}}{\partial d_{ik}}\right)^{Av}$, and consequently $\hat{\beta}_i$, but it is also differences in population size.

5.6 A Doubly-Constrained Spatial Interaction Model

The preceding theory indicates that in a singly-constrained interaction model $\hat{\beta}_i$ is a function of spatial structure. The balancing factor, Z_i, is larger for inaccessible and smaller for accessible origins since its denominator is an accessibility term. Consequently, as one particular distance increases, the relative decrease in Z_i, and hence $\hat{I}_{ij}$, is less when i is inaccessible than when i is accessible, *ceteris* *paribus*. If the relationship between I_{ij} and d_{ij} when all other variables are constant is invariant over space, $\hat{\beta}_i$ has to be more negative for inaccessible origins than for accessible origins in order to model reality.

The same process can occur in a doubly-constrained interaction model such as the calibrated version of (1.36) which is written below for convenience:

$$\hat{I}_{ij} = Z_i O_i B_j D_j d_{ij}^{\hat{\beta}_i} \tag{5.22}$$

where,

$$Z_i = [\sum_{j=1}^{n} B_j D_j d_{ij}^{\hat{\beta}_i}]^{-1} \tag{5.23}$$

and,

$$B_j = [\sum_{i=1}^{m} Z_i O_i d_{ij}^{\hat{\beta}_i}]^{-1} \tag{5.24}$$

when Z_i is large for inaccessible origins and small for accessible origins. However, since Z_i and B_j are estimated iteratively, it is very difficult to interpret the derivatives of (5.22) and the methodology outlined for the analysis of spatial structure on $\hat{\beta}_i$ derived from singly-constrained models cannot be followed. Instead, the following discussion is basically an analogy to the singly-constrained case and it relies partly on reasoned conjecture and partly on existing empirical findings.

From (5.23) Z_i can be considered as a measure of the inaccessibility of an origin to the set of destinations in the spatial system while from (5.24) B_j can be considered as a measure of the inaccessibility of a destination to the set of origins. Thomas [1977], for example, indicates that in the calibration of a doubly-constrained model using journey-to-work data on Merseyside, England, the Z_i's and B_j's were generally larger for zones that were inaccessible to destinations and origins respectively. This is also true in the empirical work presented in Chapter Seven. The spatial pattern of Z_i is then similar to that described for the singly-constrained model and the resulting spatial structure effect on $\hat{\beta}_i$ is similar.

However, the variance of Z_i is likely to be lower for the doubly-constrained model than for the singly-constrained model and the spatial structure effect may be diminished. The Z_i term in the doubly-constrained model can be considered a function of two parts: $D_j\, d_{ij}^{\hat{\beta}_i}$ and B_j. For an accessible origin, the latter is large when the former is small since j would be an inaccessible destination. Similarly, B_j would be small when $D_j\, d_{ij}^{\hat{\beta}_i}$ is large since j would be an accessible destination. This would reduce the relative size of the denominator of Z_i and increase the relative magnitude of Z_i when i is accessible. For an inaccessible origin, B_j is large when $D_j\, d_{ij}^{\hat{\beta}_i}$ is large (j is inaccessible) and vice versa. This would increase the relative size of the denominator of Z_i and decrease the relative magnitude of Z_i when i is inaccessible. Thus, the overall effect of adding B_j to the formulation of Z_i is to reduce the variation of Z_i. As a result of this reduced variation, following the discussion in the preceding section and assuming actual interaction behaviour to be constant, the spatial variation of $\hat{\beta}_i$ estimated from a doubly-constrained spatial interaction model would be less than that of $\hat{\beta}_i$ estimated from a singly-constrained model. The estimated distance-decay parameter will still include a measure of spatial structure when derived from a doubly-constrained model but the relative effect of this determinant on $\hat{\beta}_i$ will depend upon the spatial variation of B_j. The greater the variation of B_j (i.e. the greater the relative magnitude of B_j for inaccessible destinations and the less the relative magnitude of B_j for accessible destinations), the less will be the spatial structure effect included in $\hat{\beta}_i$. A corollary of this is that if the spatial pattern of the B_j's is reversed from that assumed above and B_j is large when j is accessible and small when j is inaccessible, the variance of the Z_i's would be greater, and the spatial structure effect on $\hat{\beta}_i$ would be greater, than in the singly-constrained case.

The inclusion of the second balancing factor, B_j, in the doubly-constrained model has no effect on the relationship between $\hat{\beta}_i$ and O_i described for the singly-constrained model in Section 5.5. As O_i increases, $\hat{\beta}_i$ becomes less negative, *ceteris paribus*. This explains the apparent anomaly in the empirical findings of Stillwell [1978] described in Section 1.5. The estimated distance-decay parameter for Scotland, an inaccessible origin, is close to zero while the parameter estimates for other inaccessible origins such as those in Wales are very highly negative. Wales, however, with a population less than that of Scotland is divided into six regions in Stillwell's study whereas Scotland is treated as one region. Consequently, for each origin in Wales, O_i is relatively small and $\hat{\beta}_i$ is highly negative. For Scotland, O_i is large and $\hat{\beta}_i$ is therefore much less negative, *ceteris paribus*. If Scotland were disaggregated into smaller regions in Stillwell's study, the $\hat{\beta}_i$ values for the individual regions would be much more negative. This has interesting implications for the aggregation problem in interaction modelling as outlined by, *inter alia*, Openshaw [1977].

Hence, with reference to the estimated distance-decay parameter, a doubly-constrained spatial interaction model is likely to have the same spatial structure bias as a singly-constrained model. Even if true interaction-distance relationships are constant over space, the estimated distance-decay parameter will be more negative for inaccessible origins than for accessible origins. Whether the spatial variation of $\hat{\beta}_i$ is greater or less than when $\hat{\beta}_i$ is derived from a singly-constrained model depends on the variance and distribution of the B_j terms. If B_j is large for inaccessible destinations and small for accessible destinations, the spatial variation of, and spatial structure bias in, $\hat{\beta}_i$ is reduced and vice versa.

5.7 Intraurban Interactions

Much of the preceding theory has been couched in terms of inter-urban interaction. The theory also holds for intraurban interaction and the translation of the theory produces some interesting insights in the light of a recent empirical study.

Consider a fixed quantity of employment opportunities in a city and journeys-to-work are predicted between zones of equal population. Generally, the more concentrated are employment opportunities towards the centre of the city, the greater will be the average accessibility of an origin zone. As a city grows and employment opportunities become increasingly located towards the periphery of the city, and even outside it, the average accessibility to such destinations will decline. Thus, as city size increases, the average accessibility to employment opportunities will decline. Smith and Hutchinson [1978] lend support to this argument. Using journey-to-work data for 30 Canadian cities, they report that the mean trip length from a zone increases as the accessibility of that zone decreases, and that the average mean trip length[1] increases with city size. If accessibility is measured by average mean trip length, city size explains over 60% of the variance of accessibility.[2]

If true interaction behaviour with respect to distance is constant for all city sizes, then, relating to the preceding theory, larger cities should have more negative $\hat{\beta}_i$ values. As city size increases, average accessibility decreases, and $\hat{\beta}_i$ will become more negative. Evidence that this pattern occurs is also given by Smith and Hutchinson. The relationship between the natural logarithm of $\hat{\beta}_i$ ($\hat{\beta}_i^*$) and the natural logarithm of population size (P_i^*) in their study is:

$$\hat{\beta}_i^* = 1.42 - \underset{(7.04)}{0.237}\, P_i^* \tag{5.26}$$

where P_i^* explains 64% of the variance of $\hat{\beta}_i^*$ and the bracketed figure is the t-statistic for the population coefficient. This relationship explains the apparently anomalous finding in Smith and Hutchinson's results that $\hat{\beta}_i$ becomes more negative as mean trip length increases. Mean trip length and accessibility are closely related and as

[1] The mean trip length from each zone is averaged for all zones in a city.

[2] My calculation.

accessibility increases, $\hat{\beta}_i$ becomes more negative. The relationship between mean trip length and $\hat{\beta}_i$ may not be a result of varying interaction behaviour but of varying spatial structure. When differences in accessibility are controlled for, for example when $\hat{\beta}_i$ is regressed on population and mean trip length, $\hat{\beta}_i$ becomes less negative as mean trip length increases.

5.8 Conclusions

In singly- and doubly-constrained spatial interaction models, the estimated distance-decay parameter is a function of spatial structure whether the interaction is interurban or intraurban. The spatial structure effect identified here is specific to constrained models since $\hat{\beta}_i$ is shown to be a function of the balancing factor Z_i which is a function of accessibility. As origin accessibility increases, $\hat{\beta}_i$ becomes less negative and vice versa, producing a spatial pattern consistent with the empirical findings described in Section 1.5. The relationship between $\hat{\beta}_i$ and the accessibility of i would probably be even more apparent in empirical work if total outflows were equal for all origins since it has been shown that $\hat{\beta}_i$ is an increasing function of O_i. The variance of $\hat{\beta}_i$ could thus be attributed to interaction behaviour being constant over space but origin accessibility and population size varying. Consequently, using constrained interaction models, it may not always be necessary to explain variations in $\hat{\beta}_i$ as something "abnormal". Variations in $\hat{\beta}_i$ are the normal result of constant interaction behaviour and varying accessibility over space. In reality, however, the variance of $\hat{\beta}_i$ is likely to be a result of varying spatial structure and varying interaction behaviour. Empirical work in Chapter Seven attempts to identify the relative strengths of these two effects.

CHAPTER SIX

MODEL-INDEPENDENT SPATIAL STRUCTURE EFFECTS:
COMPETING DESTINATIONS AND β

6.1 Introduction

Chapters Four and Five discussed theoretically how the estimated distance-decay parameter can be a function of spatial structure in unconstrained and constrained models respectively. Both spatial structure effects, however, were specific to a particular type of interaction model, whereas in Chapter Three it was postulated that there is an unknown spatial structure effect which is independent of the interaction model calibrated. This postulate was supported by the similarity between the spatial pattern of estimated, origin-specific, distance-decay parameters derived from unconstrained and constrained interaction models.

If such a model-independent spatial structure effect exists, it would indicate that the set of principles underlying present interaction models is either wrong or incomplete, since this set is the only common denominator of unconstrained and constrained interaction models. Consequently, the principles underlying present interaction models and their implications for interaction modelling are now examined. The set of these principles is indeed incomplete and the addition of another principle yields a new type of interaction model. This new type of model is tested empirically in Chapter Seven. Since the discussion in this chapter is concerned with basic principles, in some places it necessarily relies more on reasoned argument than on mathematical justification.

6.2 The Underlying Principles of Present Spatial Interaction Models

Three principles can be identified which characterise the composition and structure of present spatial interaction models.[1]

(i) The number of interactions from an origin is an increasing function of the emissiveness of that origin.

(ii) The number of interactions to a destination is an increasing function of the attractiveness of that origin.

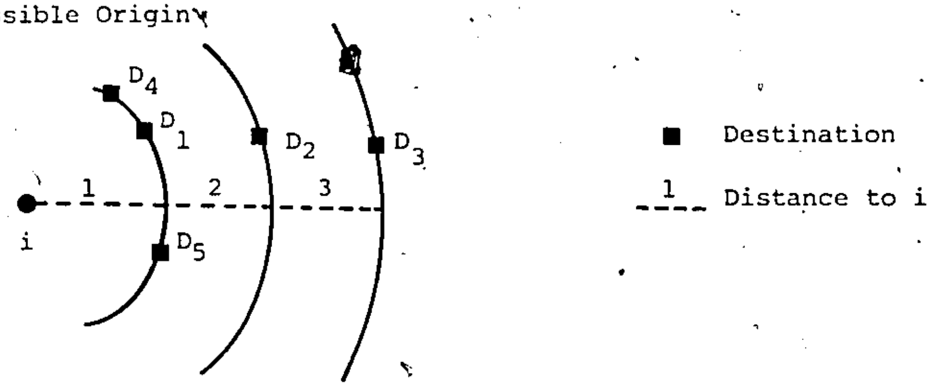
(iii) The number of interactions between any two points is a decreasing function of the distance separating these two points.

These three principles combine to produce the interaction models given in Section 1.3. Consider the operation of these models in the two different spatial systems shown in Figure 6.1. In both Figures 6.1a and 6.1b, centre i is the origin to be investigated and the attractiveness of each destination faced by origin i is assumed to be a constant. The destinations are at distances of one, two and three units from i and origin i is more accessible in Figure 6.1a than in Figure 6.1b. For each interaction model it is originally assumed that $\hat{\beta}_i$ is not a function of spatial structure and then this assumption is proved false.

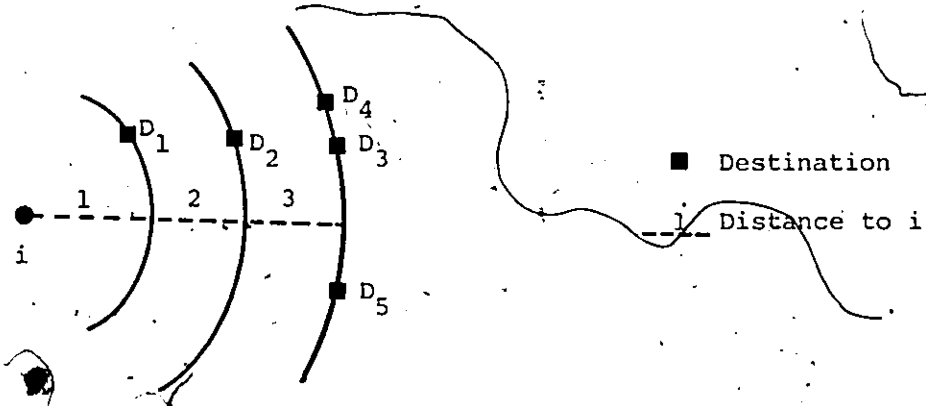
[1] A repetition of some of the ideas given in Section 1.3 is necessary here.

FIGURE 6.1: Scenarios of Spatial Structure

(a) An Accessible Origin



(b) An Inaccessible Origin



Predictions from an Unconstrained Interaction Model

In the spatial systems given in Figure 6.1, the emissiveness of i and the attractiveness of j are constant. Thus, as the distance between i and j increases, the volume of predicted interaction between i and j decreases if $\hat{\beta}_i$ is negative and this relationship is independent of the spatial structure of destinations. This is proved as follows.

The calibrated version of the unconstrained interaction model given in (1.33) is:

$$\hat{I}_{ij} = \hat{\alpha}_i \, m_j^{\hat{\gamma}_i} \, d_{ij}^{\hat{\beta}_i} \qquad (6.1)$$

The ratio of predicted interactions to two destinations, j and k, is:

$$\frac{\hat{I}_{ij}}{\hat{I}_{ik}} = \frac{\hat{\alpha}_i \, m_j^{\hat{\gamma}_i} \, d_{ij}^{\hat{\beta}_i}}{\hat{\alpha}_i \, m_k^{\hat{\gamma}_i} \, d_{ik}^{\hat{\beta}_i}} \qquad (6.2)$$

Since $m_j = m_k$, the ratio simplifies to:

$$\frac{\hat{I}_{ij}}{\hat{I}_{ik}} = \left(\frac{d_{ij}}{d_{ik}} \right)^{\hat{\beta}_i} , \qquad (6.3)$$

whence, if $d_{ij} < d_{ik}$, and if $\hat{\beta}_i < 0$, $\hat{I}_{ij} > \hat{I}_{ik}$. The result given in (6.3) is independent of spatial structure if $\hat{\beta}_i$ is independent of spatial structure. As an example, consider the spatial systems in Figures 6.1a and 6.1b. The ratio of the predicted volume of interaction terminating at D_1 to the amount terminating at D_2 is $(\frac{1}{2})^{\hat{\beta}_i}$ and the ratio of the predicted volume of interaction terminating at D_2 to the amount terminating at D_3 is $(\frac{2}{3})^{\hat{\beta}_i}$. If the true values of these ratios vary as the spatial structure of destinations varies, $\hat{\beta}_i$ would have to vary in order to model reality accurately.

Predictions from a Production-Constrained Interaction Model

The same results apply to the predictions from a production-constrained interaction model as those obtained for an unconstrained model. The ratios of predicted interactions are constant regardless of the spatial structure of destinations if $\hat{\beta}_i$ is independent of spatial structure.[1] This is proved as follows.

The calibrated version of the production-constrained model given in (1.34) is:

$$\hat{I}_{ij} = Z_i \, O_i \, D_j \, d_{ij}^{\hat{\beta}_i} \quad . \tag{6.4}$$

The ratio of predicted interactions to two destinations, j and k, is:

$$\frac{\hat{I}_{ij}}{\hat{I}_{ik}} = \frac{Z_i \, O_i \, D_j \, d_{ij}^{\hat{\beta}_i}}{Z_i \, O_i \, D_k \, d_{ik}^{\hat{\beta}_i}} \tag{6.5}$$

Since $D_j = D_k$, the ratio simplifies to:

$$\frac{\hat{I}_{ij}}{\hat{I}_{ik}} = \left(\frac{d_{ij}}{d_{ik}}\right)^{\hat{\beta}_i} , \tag{6.6}$$

which is equal to the result given in (6.3) and which is independent of the spatial structure of destinations.

The only difference between the predictions from an unconstrained and a production-constrained interaction model is that if the unconstrained model is calibrated and then new destinations are added to the system, the existing predicted interaction pattern remains unchanged. In the

---

[1] The apparent inconsistency between this statement and the results of the analysis in Chapter Five is discussed in Appendix V. It is shown that there is no inconsistency.

production-constrained model, the value of the balancing factor, Z_i, would decrease as more destinations were added to the system and the volume of interaction to each existing destination would decrease. The ratios of predicted interactions terminating at various destinations would remain the same in both models, however.

Predictions from a Doubly-Constrained Interaction Model

The above results for the unconstrained and production-constrained interaction models do not hold for the doubly-constrained model. The ratio of predicted interactions to two destinations, j and k, whose inflows are equal, is:

$$\frac{\hat{I}_{ij}}{\hat{I}_{ik}} = \frac{B_j}{B_k} \left(\frac{d_{ij}}{d_{ik}} \right)^{\hat{\beta}_i} \quad . \tag{6.7}$$

This is proved as follows.

The calibrated version of the doubly-constrained interaction model given in (1.36) is:

$$\hat{I}_{ij} = Z_i \, O_i \, B_j \, D_j \, d_{ij}^{\hat{\beta}_i} \quad . \tag{6.8}$$

The ratio of predicted interactions to j and k is then:

$$\frac{\hat{I}_{ij}}{\hat{I}_{ik}} = \frac{Z_i \, O_i \, B_j \, D_j \, d_{ij}^{\hat{\beta}_i}}{Z_i \, O_i \, B_k \, D_k \, d_{ik}^{\hat{\beta}_i}} \quad . \tag{6.9}$$

Since $D_j = D_k$, the ratio simplifies to:

$$\frac{\hat{I}_{ij}}{\hat{I}_{ik}} = \frac{B_j}{B_k} \left(\frac{d_{ij}}{d_{ik}}\right)^{\hat{\beta}_i} . \qquad (6.10)$$

B_j has already been interpreted as a measure of the inaccessibility of destination j to all other centres in the system and hence B_j and B_k will vary as the accessibility of j and k varies. Thus, as the spatial structure of destinations varies, the ratio of predicted interactions to j and k varies.[1] The nature of this variation can be described by reference to Figure 6.1. If B_j measures the inaccessibility of a destination to all origins and every centre is assumed to be both a destination and an origin, the ratio $\frac{B_1}{B_2}$ will be lower for the spatial system in Figure 6.1a than in 6.1b. The accessibility of D_1, with respect to the accessibility of D_2, is greater in Figure 6.1a than in 6.1b. From (6.7) this produces a lower ratio, $\frac{\hat{I}_{i1}}{\hat{I}_{i2}}$, for the spatial system given in Figure 6.1a than in 6.1b. Similarly, the ratio $\frac{\hat{I}_{i2}}{\hat{I}_{i3}}$ will be lower for the spatial system given in Figure 6.1a than in 6.1b. Thus, as origin i becomes more accessible, the predicted interaction-distance relationship from that origin becomes less steep and consequently $\hat{\beta}_i$ becomes less negative. These results also hold for an attraction-constrained interaction model since the balancing factor B_j is present in this model.

[1] The assumption is made that there is more than one origin interacting with each destination since if there is only one origin, all ratios would be equal to one. D_j, the total inflow into j, is assumed to be constant for all j. If there is only one origin, $\hat{I}_{ij}$ must be constant for all j and hence all ratios would be equal to one.

Thus, given the assumptions made earlier, the unconstrained and production-constrained spatial interaction models would predict that the interaction-distance relationships are constant regardless of the spatial structure of destinations faced by an origin. The prediction from doubly-constrained and attraction-constrained interaction models, however, would be that as origins become more accessible, interaction-distance relationships become less steep. From the present understanding of interaction decision-making behaviour (as summarised in principles (i) - (iii) above), there is no reason to suspect that the latter prediction represents reality. However, the following section describes a theory of interaction decision-making in which true interaction-decision relationships do become less steep as the accessibility of an origin increases. If this theory is true in reality, present spatial interaction models, especially unconstrained and production-constrained models, are mis-specified since the behaviour described by the theory is ignored in their formulations. This mis-specification produces estimates of $\hat{\beta}_i$ which are biased by spatial structure and relieving the mis-specification produces estimates of β_i which are solely measures of interaction behaviour. The separation between the spatial structure determinant and the behavioural determinant of β_i is achieved.

6.3 The Theory of Competing Destinations

Consider individuals in an origin who have already made a decision to interact but whose destinations are unknown. Present interaction models are based on the assumption that the destinations which individuals choose to interact with are a result of a single decision-making process. The net utility to be derived from interacting with any particular destination (in terms of the attractiveness of the destination and its distance from the origin) is compared to that derived for all other destinations and interactions are predicted on the basis of this comparison. If a destination is relatively unattractive and peripheral to the origin, for example, relatively few individuals will be predicted to interact with that destination.

However, many types of interaction can be considered a result of a two-stage decision-making process. The first stage is that individuals choose a broad region with which to interact. The second stage is that individuals then choose a specific destination with which to interact from the set of destinations contained within the broad region. Consider, for example, the decision to migrate in search of employment. An individual in a region of high unemployment will be aware that there are other regions which hold better prospects of employment. Once the individual has decided to move, his first locational decision is to choose one of these broad regions to which to move. The second decision would then be to choose a specific city within the broad region chosen. For example, an unemployed individual in Detroit who has made a decision to move may be aware that better employment opportunities exist elsewhere in the North-East, or in the Mid-West, South-East, etc. The individual must first choose a broad region in which to concentrate his search for work and then choose a specific destination within that region. Another example of the two-stage decision-making process is in choosing a vacation destination. Assume that an individual maximises the distance from his origin subject to some budget constraint. This results in an annulus of destinations being considered. For example, an individual in

Indianapolis may decide to go to New England, Florida, California or the Pacific North-West. His first locational decision is to choose one of these broad regions; the second locational decision is to choose a specific destination within the chosen region.

The above two examples of the two-stage decision-making process refer to interurban interaction. The theory also applies to intraurban interaction. Consider the primary type of intraurban interaction - the journey-to-work. An individual's workplace is fixed; his decision which determines his journey-to-work is to locate his home. The individual is likely to have predetermined criteria which his journey-to-work and housing must meet: for example, he may wish to be within twenty minutes' walking distance of his place of employment or he may prefer to substitute travel convenience for more land by living in suburban areas or outside of the city. In either case, the first decision made is that of choosing a particular ring of locations from his workplace. The second decision made is choosing a specific location within the ring of possible locations.

The implication of interactions being a result of a two-stage decision-making process, and not a one-stage process, is that if such decision-making behaviour is constant for all origins, constant proportions of the total outflow for each origin will terminate in given annulii around each origin. For example, in the journey-to-work situation x% of total outflow from any origin would terminate at destinations up to two miles from the origin; y% would terminate at destinations between two and five miles from the origin; z% would terminate at destinations beyond five miles. Thus, under the assumption of constant interaction behaviour, it is these proportions that remain constant.

As the spatial structure of destinations varies, however, the volume of interaction terminating at individual destinations varies. For example, out of the x% of the total outflow that terminates at destinations up to two miles from the origin, the actual number of interactions terminating at specific destinations within this annulus will be greater, the fewer number of destinations there are in the annulus, *ceteris paribus*. An example is given here to clarify these points before their implications to interaction modelling and distance-decay parameters are discussed below.

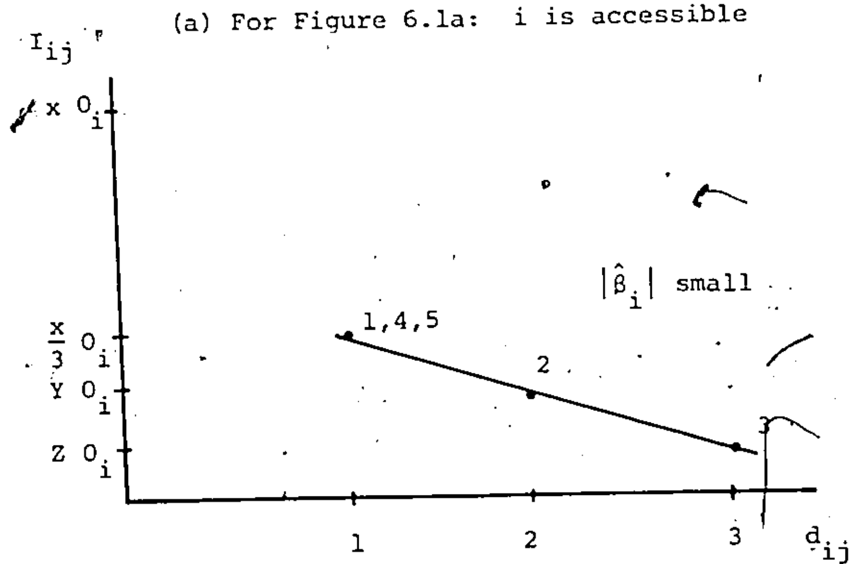
In Figure 6.1, assume that a constant proportion of the total outflow from origin i terminates at destinations less than, or equal to, a distance one unit away.[1] Also assume that constant proportions of the total outflow from i terminate at destinations greater than one unit but less than or equal to two units from i and at destinations greater than two units but less than three units from i. Let these proportions be x, y and z, respectively and if O_i is the total outflow from origin i, the number of interactions over these distances are xO_i, yO_i and zO_i, respectively. In Figure 6.1a, the number of interactions to specific destinations are then; $I_{i1} = \frac{x}{3} O_i$, $I_{i2} = y\ O_i$, $I_{i3} = z\ O_i$, $I_{i4} = \frac{x}{3} O_i$, $I_{i5} = \frac{x}{3} O_i$. In Figure 6.1b, the interactions are; $I_{i1} = x\ O_i$, $I_{i2} = y\ O_i$, $I_{i3} = \frac{z}{3} O_i$, $I_{i4} = \frac{z}{3} O_i$, $I_{i5} = \frac{z}{3} O_i$. The relationships between interaction and distance for the spatial systems given in Figures 6.1a and 6.1b are shown in Figure 6.2, where a principle is demonstrated which is

---

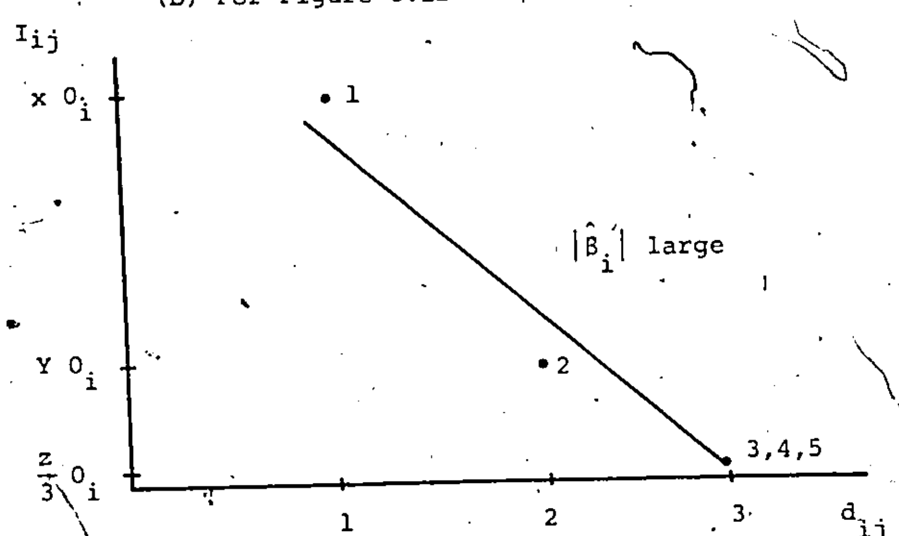
[1] Since it has already been assumed that all destinations lie at exactly one, two, or three units from i, the "less than" is irrelevant. It is included, however, since it aids in understanding a more general example.

FIGURE 6.2: Interaction-Distance Relationships for the Spatial Systems given in Figure 6.1 Assuming Two-Stage Decision-Making Behaviour.

(a) For Figure 6.1a: i is accessible



(b) For Figure 6.1b: i is inaccessible



ignored in present spatial interaction modelling: *the more accessible is a destination to all other destinations in a spatial system, the less attractive is that destination for interaction from any origin, ceteris paribus*. In terms of competition, there is a fixed volume of interaction terminating at a certain distance range from an origin. The more destinations there are within this distance range, the greater is the competition to attract this interaction and the lower is the volume of interaction terminating at any one destination. The fewer destinations there are within this distance range, the less is the competition and the greater is the volume of interaction terminating at any one destination.

Such behaviour has the following implication for distance-decay parameters estimated from present interaction models. If an origin is accessible to destinations, there will be many destinations in close proximity to the origin and each of these destinations will be accessible with respect to all other destinations. Consequently, the volume of interaction terminating at one specific destination in close proximity to the origin will be relatively small, *ceteris paribus*. In other words, there is a constant volume of interaction that terminates within a given radius from the origin and the more destinations that are within that radius, the lower will be the interaction to each specific destination. If an origin is inaccessible to destinations, there will be few destinations in close proximity and the interaction terminating at each specific destination will be relatively large, *ceteris paribus*. Conversely, there will be few destinations very distant from an accessible origin (the inaccessible centres) and interaction volumes to these destinations will be relatively large. There will be more destinations which are very distant from an inaccessible origin (the accessible centres) and interaction volumes to specific destinations at these distances will be relatively small. Thus, the implication of this interaction decision-making behaviour is that if the accessibility of destinations is not accounted for in interaction models, the more accessible is an origin, the less steep will be the resulting interaction-distance relationship and the less negative will be the estimated distance-decay parameter for that origin. The reason why several empirical studies have obtained positive estimates of distance-decay parameters can be seen from Figures 6.1a and 6.2a. If $\frac{x\ O_i}{m} < \frac{z\ O_i}{n}$, where m and n are the number of destinations at one and three units from i respectively, $\hat{\beta}_i$ will be positive. Since the inequality will only occur when m is large compared to n, origin i would have to be very accessible for $\hat{\beta}_i$ to be positive. The positive $\hat{\beta}_i$ values reported in existing empirical studies are always for the most accessible origins. This is also seen in the empirical findings reported in Chapter Seven.

6.4 The Mis-specification of Present Spatial Interaction Models

The diagnostic feature of interaction behaviour resulting from a two-stage decision-making process is that the more accessible is a destination to all other destinations, the less attractive it is for interaction from any origin, *ceteris paribus*. The implication of this relationship for estimated distance-decay parameters is that if competing destinations are not accounted for explicitly in a spatial interaction model, the true interaction-distance relationship is not measured correctly. The relationship between interaction and distance, of which $\hat{\beta}_i$ is a measure, must be measured holding constant all other variables which determine interaction. When differences in the accessibility of destinations to other destinations are not accounted for, mis-specification of the interaction model and of the interaction-distance relationship occurs. Estimating distance-decay parameters when differences in competition are not accounted for is similar to measuring interaction-distance relationships when differences in attractiveness are not accounted for - nonsensical results can occur such as the positive $\hat{\beta}_i$ values reported in several empirical studies.

The analysis in Section 6.2, however, indicates that the mis-specification described here is not equal for all interaction models. The degree of mis-specification in a particular model is indicated by the flexibility of the predicted interactions to changes in spatial structure. As an origin becomes more accessible, fewer interactions from that origin will terminate at each of the destinations in close proximity to the origin since the number of such destinations increases. Thus, an accurately-specified interaction model will reflect the fact that the ratio between interactions terminating at a nearby destination and interactions terminating at a peripheral destination will decrease as the origin becomes more accessible. Section 6.2 indicates that the unconstrained and production-constrained interaction models are totally inflexible in this way to changes in spatial structure. In the terminology of Section 6.2, $\frac{\hat{I}_{ij}}{\hat{I}_{ik}}$ is a constant regardless of spatial structure. The mis-specification in each of these models is relatively major. Section 6.2 indicates that doubly-constrained and attraction-constrained interaction models are more flexible to changes in spatial structure:

$\frac{\hat{I}_{ij}}{\hat{I}_{ik}} = \frac{B_j}{B_k} \left(\frac{d_{ij}}{d_{ik}} \right)^{\hat{\beta}_i}$ which, if destination j is more accessible than destination k, will decrease as the origin becomes more accessible. Thus, the mis-specification in these models is likely to be minor although this is dependent upon the spatial pattern of the B_j term. If B_j is small when j is accessible, the mis-specification will be minor: if B_j is small when j is inaccessible, the mis-specification will be major.

Similar conclusions are reached by Hua and Porell [1979, p. 107] although their conclusions are based on an intuitive guess of what interaction patterns should look like. The results given here are based on a theory of interaction decision-making behaviour. Hua and Porell state that the unconstrained interaction model is mis-specificed since it does not take into account ~~the relevant~~ location of other origins and destinations but the constrained models are probably correctly specified. Their result is shown to be only partially correct.

6.5 ~~A New~~ Set of Interaction Models: Competing Destinations Models

Alleviating the mis-specification of spatial interaction models described above and removing the spatial structure bias in $\hat{\beta}_i$ is a simple matter. The relationship between the accessibility of a destination and the volume of interaction terminating at that destination needs to be made explicit. Since this is an inverse relationship, the models given in (1.33), (1.34) and (1.36) are rewritten as:

(i) An unconstrained, origin-specific competing destinations model:

$$I_{ij} = \alpha_i \, m_j^{\gamma_i} \, d_{ij}^{\beta_i} \, A_{ij}^{\delta_i} \tag{6.11}$$

where A_{ij} is the accessibility of destination j to all other destinations available to origin i and is defined as:

$$A_{ij} = \sum_{\substack{k=1 \\ (k \neq i,\ k \neq j)}}^{n-2} m_k \, d_{jk}^{\sigma_i} \qquad (6.12)$$

The parameter σ_i measures the importance of distance in determining the perception of accessibility. Its relationship with β_i is discussed in Chapter Seven.

(ii) A production-constrained, origin-specific competing destinations model:

$$I_{ij} = Z_i \, O_i \, D_j \, A_{ij}^{-1} \, d_{ij}^{\beta_i} \qquad (6.13)$$

where,

$$Z_i = \Big[\sum_{j=1}^{n} D_j \, A_{ij}^{-1} \, d_{ij}^{\beta_i}\Big]^{-1} \qquad (6.14)$$

and,

$$A_{ij} = \sum_{\substack{k=1 \\ (k \neq i,\ k \neq j)}}^{n-2} D_k \, d_{jk}^{\sigma_i} \qquad (6.15)$$

The constraint given in (6.14) ensures that the total predicted outflow from each destination is equal to its true value.

(iii) A doubly-constrained, origin-specific competing destinations model:

$$I_{ij} = Z_i \, O_i \, B_j \, D_j \, A_{ij}^{-1} \, d_{ij}^{\beta_i} \qquad (6.16)$$

where,

$$Z_i = [\sum_{j=1}^{n} B_j D_j A_{ij}^{-1} d_{ij}^{\beta_i}]^{-1} \quad (6.17)$$

$$B_j = [\sum_{i=1}^{m} Z_i O_i A_{ij}^{-1} d_{ij}^{\beta_i}]^{-1} \quad (6.18)$$

and,

$$A_{ij} = \sum_{k=1}^{n-2} D_k d_{jk}^{\sigma_i} \quad (6.19)$$

$$(k \neq i, k \neq j)$$

The constraint given in (6.17) ensures that the total predicted outflow from each destination is equal to its true value. The constraint in (6.18) ensures that the total predicted inflow into each destination equals its true value. The variable D_k in (6.19) can be defined as in (6.16) or it can simply be measured by population size. For completeness, the equivalent attraction-constrained model is also given although this model is not discussed further.

(iv) An attraction-constrained, destination-specific competing destinations model:

$$I_{ij} = O_i B_j D_j A_{ij}^{-1} d_{ij}^{\beta_j} \quad (6.20)$$

where,

$$B_j = [\sum_{i=1}^{m} O_i A_{ij}^{-1} d_{ij}^{\beta_i}]^{-1} \quad (6.21)$$

and,

$$A_{ij} = \sum_{k=1}^{n-2} D_k d_{jk}^{\sigma_i} \quad (6.22)$$

$$(k \neq i, k \neq j)$$

In (6.11) the expected sign of $\hat{\delta}_i$ is negative - as the accessibility of a destination increases, the volume of interaction terminating at that destination decreases, *ceteris paribus*. The term A_{ij}, as defined in (6.12), (6.15), (6.19) and (6.22), measures the accessibility of destination j to all other destinations k that are available to i as perceived by individuals at origin i. It is a measure of how destination j competes with all other destinations for interactions originating at i. It is important to note that the accessibility of j is measured with respect to all other possible destinations - not only to the particular destinations used in an analysis. The spatial structure effect described here is the result of not taking the relationship between one destination and all others into account. It is not an effect which could be removed by a measure of the accessibility of j to all other destinations in the analysis unless the sample of centres used in the analysis is a good representation of the population set.

The interrelationship between the accessibility equations given in (6.12), (6.15), (6.19) and (6.22) and each interaction equation has varying degrees of complexity. The two can be considered as completely independent models. If the sample set of centres between which interaction is measured is not a good representation of the population set of centres, and if σ_i is estimated independently for all i (for example, assuming $\sigma_i = -1.0$ for all i), then A_{ij} can be derived independently of the interaction equation. However, given the same poor sample of centres, σ_i and β_i can be estimated iteratively. An estimate of σ_i is made for all i; A_{ij} is calculated for all i and for all j; and the interaction equation calibrated. The estimates of the β_i's are then used as estimates of the $\hat{\sigma}_i$'s and the A_{ij}'s recalculated etc.. The process continues until convergence of $\hat{\sigma}_i$ and $\hat{\beta}_i$ is reached or until the researcher is satisfied that accurate estimates of σ_i and β_i have been achieved.[1] A further degree of interrelationship between the equation for A_{ij} and the interaction equation can be achieved if the sample set of centres is a good representation of the population set. The set of destinations used in the definition of accessibility can be taken from the set of destinations between which interactions are modelled and the two equations become one model. The degree of interrelationship between the two equations chosen by a researcher must first depend upon the sample of centres between which interaction is measured and then on a decision as to the value of slightly better estimates of β_i at a cost of substantially more computing time. These different degrees of interrelationship are investigated empirically in Chapter Seven.

If each centre in a spatial system is both an origin and a destination (such is usually the case in interurban interaction patterns, for example), an approximation to the definition of A_{ij} given in (6.12), (6.15), (6.19) and (6.22) is:

$$A_{ij} \simeq A_j = \sum_{\substack{i=1 \\ i \neq j}}^{m} m_i \, d_{ij}^{\sigma} \quad , \tag{6.23}$$

where σ_i is assumed to equal σ for all i and then A_j is equal for all i.

---

[1] In the empirical analysis to be discussed subsequently, it was found that there were no substantial differences in the calibration of the interaction model when σ_i was approximated by -1.0 for all i and when σ_i and β_i were estimated iteratively. The former process, however, is much less time-consuming.

This has the advantage that only n such terms need be calculated instead of m x n. The approximation is only a good one, however, when σ is constant for all i and the number of alternative centres is large since,

$$A_{ij} = A_j - m_i d_{ij}^{\sigma} \tag{6.24}$$

and A_j has to be large relative to $m_i\ d_{ij}^{\sigma}$ in order to ignore $m_i\ d_{ij}^{\sigma}$. In the unconstrained and production-constrained competing destinations models given in (6.11) and (6.13), A_{ij} is simply substituted by A_j and the resulting models calibrated. The substitution of A_{ij} by A_j in the doubly-constrained and attraction-constrained competing destinations models is more interesting, however. By making the approximation in (6.23), the doubly-constrained competing destinations model given in (6.20) is structurally identical to the original doubly-constrained interaction model given in (1.36). This is proved as follows.

Using the substitution given in (6.23), the doubly-constrained competing destinations model is written as:

$$I_{ij} = Z_i\ O_i\ B_j\ D_j\ A_j^{-1}\ d_{ij}^{\beta_i} \tag{6.25}$$

where,

$$Z_i = [\sum_{j=1}^{n} B_j\ D_j\ A_j^{-1}\ d_{ij}^{\beta_i}]^{-1} \tag{6.26}$$

and,

$$B_j = A_j [\sum_{i=1}^{m} Z_i\ O_i\ d_{ij}^{\beta_i}]^{-1} \tag{6.27}$$

Define,

$$B_j' \equiv [\sum_{i=1}^{m} Z_i\ O_i\ d_{ij}^{\beta_i}]^{-1} \tag{6.28}$$

and then (6.27) can be written as:

$$B_j = A_j B'_j \tag{6.29}$$

Substituting (6.29) into (6.25) gives:

$$I_{ij} = Z_i O_i B'_j D_j d_{ij}^{\beta_i} \tag{6.30}$$

where,

$$Z_i = [\sum_{j=1}^{n} B'_j D_j d_{ij}^{\beta_i}]^{-1} \tag{6.31}$$

and B'_j is defined in (6.28). The model given in (6.30) is identical to that original doubly constrained interaction model given in (1.36). The same result applies to the attraction-constrained destination-specifi[illegible] interaction model. Using the substitution given in (6.23), [illegible] constrained model given in (6.20) is written as:

$$I_{ij} = O_i B_j D_j A_j^{-1} d_{ij}^{\beta_j} \tag{6.32}$$

where,

$$B_j = A_j [\sum_{i=1}^{m} O_i d_{ij}^{\beta_j}]^{-1} \tag{6.33}$$

Define:

$$B'_j \equiv [\sum_{i=1}^{m} O_i d_{ij}^{\beta_j}]^{-1} \tag{6.34}$$

Substituting (6.34) into (6.33) gives:

$$B_j = A_j \cdot B'_j \tag{6.35}$$

and (6.32) is rewritten as:

$$I_{ij} = O_i \, B_j' \, D_j \, d_{ij}^{\beta_j} \tag{6.36}$$

which is the original attraction-constrained model. Thus, if A_j as defined in (6.23) is a good approximation of A_{ij} and each centre is both an origin and a destination, the original doubly-constrained and attraction-constrained models are not mis-specified, and the competing destinations models are identical to the original interaction models. In the case of the attraction-constrained model, the operation of the model can be considered as allocating interactions from many origins to one destination and as such, the concept of competing destinations is controlled for in the original model formulation. The same reasoning can also be applied to the original doubly-constrained interaction model. In both cases, it is the inclusion of the B_j term which accounts for differences in competing destinations and removes the problem of mis-specification. This was discussed in Section 6.4.

If each centre in a spatial system is not both an origin and a destination, however, doubly-constrained and attraction-constrained gravity models are not identical to the equivalent competing destinations models. The competing destinations variable, A_{ij}, measures the accessibility of one particular destination to all other destinations. The balancing factor, B_j, measures the accessibility of one particular destination to all origins in the system. The two terms are only similar when each centre is both an origin and a destination. Such is usually the case in inter-urban interactions but is not always so in intra-urban interactions where origins and destinations may be quite separate.

Examples are shopping plazas and residential locations in shopping trip analysis and basic industry sites and residential areas in journey-to-work analysis. In each case, centres are either an origin or a destination but not both. The mis-specification of doubly and attraction-constrained gravity models can easily be seen. In shopping trip analysis, for example, B_j would measure the accessibility of a shopping plaza to all residential areas whereas A_{ij} would measure the accessibility of a particular shopping plaza to all other shopping plazas. Thus, it is only for interurban interactions and intraurban interactions where each centre or zone is both an origin and a destination, that the doubly-constrained and attraction-constrained gravity models are not mis-specified. In situations where each centre is not both an origin and a destination, the models are mis-specified.

Finally, in unconstrained and singly-constrained competing destinations models, other formulations for A_{ij} could be used than that given in (6.12) and (6.15). Sheppard [1979b] and Pirie [1979] discuss some of these alternatives. However, the formulation given here has been well-tested (*inter alia*, Fotheringham [1979]) and conforms well with our intuitive notion of accessibility. The arbitrariness of the definition is no greater than that of the definition of the attractiveness, emissiveness and distance variables already included in the interaction models.

6.6 The Derivation of the Constrained Competing Destinations Models by Entropy - Maximising Techniques

The constrained competing destinations models given in (6.13) and (6.16) can be derived theoretically by entropy-maximising techniques in a similar manner to that used by Wilson [1970b] to derive the gravity models given in (1.34) and (1.36)[1]. In Wilson's formalism, the entropy function maximised is Shannon's [1948] measure of entropy (H) which is defined as:

$$H = -\sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} \ln p_{ij} \qquad (6.37)$$

where p_{ij} is the probability of an individual travelling between i and j. To obtain interaction models of the form given in (1.34) and (1.36), (6.37) is maximised subject to the constraints:

$$\sum_{j=1}^{n} p_{ij} = 1 \qquad (6.38)$$

$$\sum_{i=1}^{m} p_{ij} = 1 \qquad (6.39)$$

and,

$$\sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} d_{ij} = D \qquad (6.40)$$

where D is the expected distance (or cost) travelled in the system. Maximising (6.37) subject to (6.38), (6.39) and (6.40) produces a doubly-constrained gravity model with the same structure as that given in (1.36). If the constraint given in (6.39) is omitted, a production-constrained

[1] Henceforth, the models given in (6.11), (6.13) and (6.16) will be termed competing destinations models while the equivalent models given in (1.33), (1.34) and (1.36) will be termed gravity models.

gravity model such as that given in (1.34) results. If the constraint given in (6.38) is omitted, an attraction-constrained gravity model results.

An alternative and more general measure of entropy to that given in (6.37) is Kullback's Information Gain (KIG) [1959] which is defined as:

$$KIG = -\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij} \ln \left(\frac{p_{ij}}{q_{ij}} \right) \qquad (6.41)$$

where q_{ij} is the prior probability of an individual travelling between i and j given no knowledge of the constraints operating on the interaction system. The probability p_{ij} is then defined as the posterior probability of an individual travelling between i and j. Maximising Kullback's Information Gain maximises the gain of information as a result of knowing $\{p_{ij}\}$ instead of $\{q_{ij}\}$. It is well known that if no prior knowledge of p_{ij} is known apart from that contained in the constraints, $q_{ij} = \frac{1}{mn}$ for all i and for all j and maximising (6.41) is equivalent to maximising (6.37). Shannon's entropy measure is a special case of Kullback's Information Gain where prior probabilities are equal. Thus, in maximising (6.37), Wilson has assumed equal prior probabilities in the absence of the knowledge contained in the constraints (6.38)-(6.40). The earlier discussion has shown this to be a false assumption: the more accessible is one destination to all other possible destinations in a spatial system, the smaller is the probability that an individual in any origin will terminate his interaction in that destination, *ceteris paribus*. Consequently, q_{ij} is defined as:

$$q_{ij} = \frac{A_{ij}^{-1}}{\sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij}^{-1}} , \qquad (6.42)$$

and the new entropy measure to be maximised is then:

$$KIG = -\sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} \ln (p_{ij} A_{ij} \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij}^{-1}) \qquad (6.43)$$

Since $\sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij}^{-1}$ is a constant for all i and for all j, (6.43) is maximised when,

$$KIG = -\sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} \ln (p_{ij} A_{ij}) \qquad (6.44)$$

is maximsed. Maximising (6.44) subject to the constraints given in (6.38) - (6.40) produces a doubly-constrained competing destinations model such as that given in (6.16). Maximising (6.44) subject to the constraints given in (6.38) and (6.40) produces a production-constrained competing detinations model such as that given in (6.13). Wilson's formalism can be seen as maximising entropy subject to a partial amount of known information on the interaction system: a mis-specified model results. The above formalism maximises an entropy function subject to the full amount of known information and the mis-specification is relieved.

6.7 Parameter Estimates from Competing Destinatins Models

Gravity models have been shown to be mis-specificed by not accounting for competing destinations. As a result, the parameter estimates derived from the calibration of such models are biased: in particular, estimates of the distance-decay parameter are biased in such a way that accessible origins have less negative parameter estimates than inaccessible origins even if interaction behaviour is constant over space. The reason for this spatial pattern is given in the explanation of Figure 6.2. The competing destinations models, however, are not mis-specified in this way and more accurate parameter estimates, free from the effect of spatial structure described in this chapter, will result from the calibration of these models. An empirical test on both sets of models, analysing parameter estimates and model performance, is given in Chapter Seven. However, it is useful here to give a brief example showing why parameter estimates derived from the competing destinations models are more accurate.

The distance-decay parameter is a measure of the relationship between interaction and distance when all other variables influencing interaction are controlled for. Figure 6.2 describes this relationship for the unconstrained and production-constrained gravity models applied to the spatial systems of Figure 6.1 (it is assumed in Figure 6.1 that each centre is both an origin and a destination and that the mis-specification in doubly and attraction-constrained gravity models is minor). The interaction-distance relationship is simply measured by the graph of I_{ij} and d_{ij} since the only other variable influencing interaction, the attractiveness of a destination, is a constant. When the equivalent competing destinations models given in (6.11) and (6.13) are applied to the spatial systems of Figure 6.1 however, the accessibility of a destination must be controlled for and the interaction-distance relationship is measured by the graph of $I_{ij} \cdot A_{ij}$ and d_{ij}. This has the following effect. In Figure 6.1a, destinations 1, 4 and 5 are accessible to all other destinations while destination 3 is inaccessible. Thus, $I_{i1} \cdot A_{i1}$, $I_{i4} \cdot A_{i4}$ and $I_{i5} \cdot A_{i5}$ will be larger relative to $I_{i3} \cdot A_{i3}$ than I_{i1}, I_{i4} and I_{i5} were to I_{i3}. This has the effect of increasing the slope of the line given in Figure 6.2a and increasing $|\hat{\beta}_i|$. In Figure 6.1b, destination 1 is inaccessible to other destinations while destinations 3, 4 and 5 are accessible to other destinations. Thus, $I_{i1} \cdot A_{i1}$ will be smaller relative to $I_{i3} \cdot A_{i3}$, $I_{i4} \cdot A_{i4}$ and $I_{i5} \cdot A_{i5}$ than I_{i1} was to I_{i3}, I_{i4} and I_{i5}. Hence, the slope of the line given in Figure 6.2b will decrease and $|\hat{\beta}_i|$ will decrease. Consequently, when β_i is estimated from the competing destinations models, the variance and the spatial pattern of $\{\hat{\beta}_i\}$ is reduced. If the particular spatial structure effect described here is the only such effect and interaction behaviour is constant over space, $\hat{\beta}_i$ derived from competing destinations models will be constant over space.

The bias in the estimated distance-decay paramter can be seen explicity in the unconstrained model. The unconstrained gravity model given in (1.33) is written in logarithmic form as:

$$I_{ij}^* = \alpha^* + \gamma_i m_j^* + \beta_i d_{ij}^* + \varepsilon_{ij} , \qquad (6.45)$$

and in similar form the competing destinations model of (6.11) is:

$$I_{ij}^* = \tilde{\alpha}^* + \tilde{\gamma}_i m_j^* + \tilde{\beta}_i d_{ij}^* + \delta_i A_{ij}^* + \mu_{ij} , \qquad (6.46)$$

where ε_{ij} and μ_{ij} are error terms assumed to have the usual least-squares properties and "~" denotes a parameter value in (6.46) which differentiates it from the equivalent value in (6.45). A_{ij}^*, the logarithm of the accessibility of destination j as perceived by residents of origin i, can be represented as a linear function of the logarithmic distance between i and j, for each origin:

$$A_{ij}^* = \phi_i + \Pi_i d_{ij}^* + \nu_{ij} \quad , \tag{6.47}$$

where ϕ_i and Π_i are parameters to be estimated and ν_{ij} is an error term assumed to have the usual least-squares properties. Substituting (6.47) into (6.46) gives:

$$I_{ij}^* = \tilde{\alpha}^* + \tilde{\gamma}_i m_j^* + \tilde{\beta}_i d_{ij}^* + \delta_i(\phi_i + \Pi_i d_{ij}^* + \nu_{ij}) + \varepsilon_{ij} \quad , \tag{6.48}$$

which, on rearranging, becomes:

$$I_{ij}^* = \tilde{\alpha}^* + \delta_i \phi_i + \tilde{\gamma}_i m_j^* + (\tilde{\beta}_i + \delta_i \Pi_i) d_{ij}^* + \delta_i \nu_{ij} + \varepsilon_{ij} \quad , \tag{6.49}$$

and is equivalent to (6.45). Thus,

$$\beta_i = \tilde{\beta}_i + \delta_i \Pi_i \tag{6.50}$$

or, when calibrated,

$$\hat{\beta}_i = \hat{\tilde{\beta}}_i + \hat{\delta}_i \hat{\Pi}_i \quad . \tag{6.51}$$

Since $\hat{\tilde{\beta}}$ is an unbiased estimator of β_i, the bias in $\hat{\beta}_i$ is given by the term $\hat{\delta}_i \hat{\Pi}_i$ which can be interpreted as an indirect elasticity of interaction with respect to distance. $\hat{\delta}_i$ is the estimated elasticity of interaction with respect to the accessibility of a destination and $\hat{\Pi}_i$

is the estimated elasticity of the accessibility of a destination with respect to distance. The estimated distance-decay parameter from (6.45) is biased since it includes this indirect effect. The competing destinations model in (6.46) makes this effect explicit in the model and the estimated distance-decay parameter is unbiased. If $\hat{\delta}_i = 0$ and/or $\hat{\Pi}_i = 0$, there is no bias in $\hat{\beta}_i$. The former indicates that there is no log-linear relationship between the accessibility of a destination and interaction to that destination: the latter indicates that there is no log linear relationship between the distance to a destination and its accessibility. If $\hat{\delta}_i = 0$, the unconstrained model given in (1.33) is correctly specified. If $\hat{\Pi}_i = 0$, however, and $\hat{\delta}_i \neq 0$, $\hat{\beta}_i = \hat{\bar{\beta}}$ but the competing destinations model of (6.11) would give more accurate predictions of interaction than the gravity model of (1.33).

From (6.51) it is easy to guess at the direction of the bias in $\hat{\beta}_i$. Assume that $\hat{\delta}_i < 0$ for all i and it is constant over space, or does not vary systemmatically (there is no reason to suspect otherwise). Also assume that $\hat{\bar{\beta}}_i$ is constant over space. Then, for accessible origins, $\hat{\Pi}_i$ is negative (as the distance between the origin and destination increases, the destination becomes less accessible), and $\hat{\delta}_i \; \hat{\Pi}_i > 0$ so that $\hat{\beta}_i > \hat{\bar{\beta}}_i$. Conversely, for inaccessible origins, $\hat{\Pi}_i > 0$, $\hat{\delta}_i \; \hat{\Pi}_i < 0$ and $\hat{\beta}_i < \hat{\bar{\beta}}_i$. Since $\hat{\bar{\beta}}_i < 0$, $\hat{\beta}_i$ will be more negative for inaccessible origins than for accessible origins. If $\hat{\delta}_i \; \hat{\Pi}_i > |\hat{\bar{\beta}}_i|$, $\hat{\beta}_i$ will be positive and since $\hat{\Pi}_i$ would have to be large and positive for this condition to occur, origin i must be very accessible. This was mentioned previously - only the most accessible origins could have $\hat{\beta}_i$ values that are so upwardly biased as to be positive.

6.8 A Comparison of the Competing Destinations Models with other Models of "Alternative Opportunities".

Ewing [1974] and Cesario, [1975] proposed a so-called "spatial choice" model as the solution to the spatial structure effect on $\hat{\beta}_i$ derived from an unconstrained interaction model and this solution was described and discarded in Section 2.6. The similarity between the "spatial choice" model and the unconstrained competing destinations model given in (6.11) invites comparison. The two models are rewritten below for convenience.

Ewing and Cesario's "spatial choice" model:

$$I_{ij} = \alpha_i \frac{m_j^{\gamma_i} d_{ij}^{\beta_i}}{\sum_{j=1}^{n} m_j^{\gamma_i} d_{ij}^{\beta_i}} \quad \text{(hypothesis: } \gamma_i > 0,\ \beta_i < 0). \tag{6.52}$$

The unconstrained competing destinations model:

$$I_{ij} = \alpha_i \frac{m_j^{\gamma_i} d_{ij}^{\beta_i}}{\left(\sum_{\substack{k=1 \\ (k \neq i,\ k \neq j)}}^{n-2} m_k d_{jk}^{\sigma_i}\right)^{\delta_i}} \quad \text{(hypothesis: } \gamma_i > 0,\ \beta_i < 0,\ \delta_i > 0\text{)}. \tag{6.53}$$

Obviously, the differences between the two models lie in their respective denominators. Besides the competing destinations model having an extra degree of freedom in δ_i, there are two main differences. The denominator of the spatial choice model measures the analysis-specific accessibility of origin i. If the model is made origin-specific, then this variable is a constant. The denominator of the competing destinations model measures the accessibility of destination j to all other possible destinations available to origin i, whether these destinations are included in the interaction analysis or not. The latter measure removes the mis-specification in the unconstrained gravity model: the former simply constrains the total predicted outflow from each origin to be equal to its true value. The mis-specification of the original gravity model is not relieved. Consequently, the model proposed by Ewing and Cesario as a solution to the spatial structure problem in unconstrained modelling will not remove any of the spatial structure effect described here.

Gordon's [1976] suggestion to remove spatial structure effects from $\{\hat{\beta}_i\}$, also described in Section 2.6, is a more viable one. He suggested adding origin and destination constraints to an unconstrained gravity model: in effect suggesting the calibration of doubly-constrained gravity models instead of unconstrained gravity models. Since the doubly-constrained gravity model has been shown to be properly specified when σ_i is a constant for all i and when each centre is both an origin and a destination, Gordon's suggestion is a reasonable one. However, in accepting the replacement of an unconstrained gravity model by a doubly-constrained gravity model, the useful properties of the unconstrained model are lost. Also, the doubly-constrained gravity model is mis-specified when centres are not both origins and destinations. A better solution to the spatial structure bias in unconstrained gravity modelling is the unconstrained competing destinations model given in (6.11).

Britton [1971] suggested the addition of variables to the unconstrained gravity model which were very similar to origin and destination constraints. However, Britton suggested that these new variables be calculated prior to the calibration of the model and simply be added as extra regression terms. Since the destination constraint can be considered a measure of the accessibility of a particular destination to all others when each centre is both an origin and a destination, Britton's model is similar to the unconstrained competing destinations model given in (6.11). The model can be criticised, however, since the origin accessibility variable is irrelevant and the measurement of destination accessibility is incorrect when each centre is not both an origin and a destination.

Long [1970] proposed an "alternative opportunity model" to account for competing destinations. His basic model is an unconstrained gravity model with the addition of a variable measuring the mean distance from an origin to alternative destinations. Empirical testing of the model indicated that interaction to a particular destination increases as the mean distance from the origin to all other destinations increases. However, this is not a measure of the competing destinations effect described here although the two may be similar in some situations. Long's "alternative opportunities" variable does not necessarily distinguish between destinations which are isolated and destinations which are in close proximity to one another. Consider for example, all destinations lying on a circle of given radius from an origin but the destinations are unevenly spaced. The mean distance to alternative opportunities will be equal for each particular destination but the accessibility of each destination to all other destinations will not be. There are some situations in which Long's "alternative opportunity" variable and the competing destinations variable would have completely opposite effects on predicted interactions. If a destination is very inaccessible to an origin and to all other destinations, and the mean distance to alternative destinations from the origin is small, Long's model would predict lower interactions to that destination than would the original gravity model. The competing destinations model, on the other hand, would predict greater interaction to that destination than would the original gravity model.

6.9 A Comparison of the Competing Destinations Models with the Intervening Opportunities Model

The intervening opportunities model was first formulated by Stouffer [1940] and has been developed and tested by many authors, such as Ullman [1956], Thomazinis [1962] and Porter [1964][1]. The model can be written in many forms which makes comparison difficult, but there are two basic ones. In the first, interactions are expressed solely as a function of intervening opportunities (Stouffer's original concept) while in the second, interactions are expressed as a function of intervening opportunities and distance (Ullman's contribution).[2] The relationship between interaction and intervening opportunities is expected to be negative: the greater are the opportunities for interactions to terminate between i and j, the lower will be the volume of interaction terminating at j, *ceteris paribus*. The first formulation of the intervening opportunities model is dismissed here as being a mis-specified interaction

[1] Thomazinis termed his model a competing opportunities model although it is an intervening opportunities model in which all destinations lie in arbitrarily-defined annulii around an origin. Destinations lying in annulii between the origin and a specific destination, j, are defined as competing opportunities and are inversely related to the volume of interaction between the origin and destination j. Destinations lying in annulii beyond that in which destination j lies, however, are not considered as competing opportunities. A discussion of Thomazinis' model is given by Finney [1972].

[2] The models also contain variables measuring the emissiveness of origins and the attractiveness of destinations but these are irrelevant to this discussion.

model since it does not include a distance variable. The second formulation, however, is very similar to that of the competing destinations models and it is interesting to consider why this formulation has never gained universal acceptance. There are two main reasons for this: one is that the theoretical derivation of the intervening opportunities model is weak,[1] and the second, which is related, is that the measurement of intervening opportunities is very subjective. How is an intervening opportunity defined and where do intervening opportunities begin and end? Porter's diagram of intervening opportunities reproduced by Taaffe and Gauthier [1973] is a typical example of the confusion involved in measuring such an abstract concept. The failure to obtain accurate calibrations of Tomazinis' competing opportunities model in both the Penn-Jersey Transport Study [1964] and in a study of trip destinations in Washington D.C. by Heanue and Pyers [1966], was also due to the difficulty in measuring the concept of intervening opportunity (see Finney [1972]).

The competing destinations models, on the other hand, suffer from neither of these problems. The theoretical derivation of the models is sound and results from a correct understanding of interaction decision-making behaviour. In fact, the theoretical derivation of these models is sounder than that of the respective gravity models which have been shown to be mis-specified. As a result, the measurement of competing

---

[1] It is interesting to note that Wilson's [1970b] entropy-maximising derivation of the intervening opportunities model (Stouffer's original form) is very suspect and in Wilson's words, depends upon "a rather strange cost-constraint" (p.155). The entropy-maximising derivation of the competing destinations models given here is free from such suspicion. This relates to the point made above that Stouffer's original formulation is mis-specified.

destinations is no more subjective than the measurement of the other variables in the model. An exact formulation for the measurement of competing destinations is derived and the subjectivity in measuring intervening opportunities, that is, where do such opportunities begin and end, is eliminated. Competing destinations and intervening opportunities are obviously similar concepts: the advantage of the former is that it has a theoretical derivation and a more rigorous definition. The concept of intervening opportunities can be considered as an earlier, less accurate attempt to measure the relationship between competing destinations and interaction described here. A quotation from Greenwood and Sweetland [1972, pp. 668-669] on this point is interesting:

> "... while distance itself does not adequately reflect the opportunity costs of migration because it accounts for 'too much', neither is Stouffer's measure of intervening opportunities adequate in itself, for it accounts for 'too little'. No acceptable method has yet been suggested to disentangle the effect associated with opportunity costs from the other effects generally associated with distance."

The theory that interaction patterns are a result of two types of decision-making behaviour, and the consequent relationship between competing destinations and interaction, produces a method for which this latter statement is invalid.

6.10 Concluding Comment

Two types of interaction decision-making behaviour have been described above: one-stage and two-stage decision-making. Observed interaction patterns are a result of both types of decision-making behaviour. However, the present set of interaction models assumes that interaction patterns result solely from the one-stage decision-making process. The effect of the two-stage decision-making process is to make destinations that are accessible to all other destinations, less attractive for interaction, *ceteris paribus*. Since this relationship is not modelled by the present set of gravity models, they are mis-specified. In the case of the doubly-constrained and attraction-constrained gravity models, however, this mis-specification is likely to be slight when each centre in the analysis is both an origin and a destination. Relief of the mis-specification is achieved by producing a new set of interaction models which account for competing destinations and calibrating these competing destinations models will produce estimates of β_i that are unaffected by the spatial structure effect described in this analysis. The sets of such parameter estimates derived from competing destinations models will have less variation and exhibit less spatial pattern than those derived from present gravity models. Both sets of models are tested empirically in Chapter Seven.

CHAPTER SEVEN

EMPIRICAL EVIDENCE OF MODEL INDEPENDENT SPATIAL STRUCTURE EFFECTS: A COMPARISON OF THE GRAVITY AND COMPETING DESTINATIONS MODELS

7.1 Introduction

The results of calibrating the competing destinations models given in (6.11), (6.13) and (6.16) are compared to the results of calibrating the gravity models given in (1.33), (1.34) and (1.36). The main purpose of this comparison is to demonstrate that origin-specific, distance-decay parameters estimated from present gravity models are biased by spatial structure and that this bias can be removed by calibrating the equivalent competing destinations models. In Chapter Six it was predicted theoretically that the removal of the spatial structure bias results in more negative parameter estimates for accessible origins and less negative parameter estimates for inaccessible origins. Hence, the spatial pattern and spatial variation of parameter estimates are prime topics of investigation. Also of interest are the parameter estimates themselves since those derived from the competing destinations models can be interpreted as purely behavioural measures of the distance-interaction relationship from each origin, while those derived from the gravity models are primarily indices of accessibility. Of secondary importance is a comparison of an origin-specific goodness-of-fit statistic for both sets of models to identify any regular spatial pattern in the relationship between the accessibility of a destination and interaction to that destination.

7.2 Data Measurement

The data used to calibrate both sets of interaction models are 1970 airline passenger interaction data published by the United States Civil Aeronautics Board [1971]. These data are a ten percent sample of all airline passenger journeys on domestic routes within the United States during 1970. The sample was taken continuously throughout 1970: each airline ticket sold with a number ending in zero was selected and the origin and final destination on the ticket were noted. Round-trip tickets were counted as two separate trips. The volume of interaction and airline (great circle) distances are given between every city with at least one commercial airport in the United States.

The data were chosen primarily for their accuracy and comprehensiveness. Many of the interactions in the 10% sample are of the order of 10^3 while some are as high as 10^5. These sample figures are used throughout the analysis. The sampling error statistic given by the Civil Aeronautics Board indicates that the maximum percentage error of the sample is less than 10% at the 95% confidence level when the sample size is over 400[1]. The sampling statistic and further details of the data are given on pages v - xii of the above reference.

---

[1] For example, when the sample size is 500, the estimated population size is 5,000 and 95 times out of 100 the true population figure will be in the range 5,000±416 where 416 is 8.3% of 5,000.

A slight drawback of the interaction data, however, for estimating origin-specific parameters is that the interaction given is that between two centres i and j and it is the sum of the interaction from i to j and the interaction from j to i. For any average centre, 50% of the interactions result from decision-making behaviour by the inhabitants of that centre, and 50% result from decision-making behaviour by the inhabitants of all other centres. Consequently, the origin-specific parameters estimated for such data primarily reflect characteristics of each origin but also reflect, to a very minor degree, characteristics of each of the other centres in the system. However, since this latter effect will be fairly constant for each origin, especially when the number of centres in the sample is large, the origin-specific parameters resulting from the data can accurately be considered as a reflection of interaction behaviour from each origin.

One hundred cities were selected as a basis for the analysis and interaction and distance matrices of dimension 100 x 100 were obtained from the published data set. The cities chosen were those in the 100 largest Standard Metropolitan Statistical Areas (SMSA's) in terms of their 1970 populations. The names of these SMSA's are given in Table 7.1. The size of the sample was determined by two considerations: it had to be small enough to be manageable on the computing system and it had to be large enough to produce a representative sample of origin-specific parameters.[1]

[1]Calibrating an origin-specific doubly-constrained interaction model with a 100 x 100 interaction matrix requires approximately 20 minutes of central memory time on a CDC Cyber, 170/730.

TABLE 7.1

SMSA'S USED IN THE INTERACTION MATRIX[1]

*1. Akron-Canton, Ohio.
2. Albany-Schenectady-Troy, N.Y.
*3. Albuquerque, N.M.
4. Allentown-Bethlehem-Easton, Pa.-NJ.
5. Atlanta, Ga.
6. Augusta, Ga.-S.C.
*7. Austin, Tex.
*8. Bakersfield, Calif.
*9. Baltimore, Md.
10. Baton Rouge, La.
*11. Beaumont-Port Arthur-Orange, Tex.
12. Binghampton, N.Y. -Pa.
13. Birmingham, Ala.
*14. Boston, Mass.
15. Bridgeport, Conn.
16. Buffalo, N.Y.
17. Charleston, S.C.
18. Charleston, W.Va.
19. Charlotte-Gastonia, N.C.
20. Chattanooga, Tenn.-Ga.
21. Chicago, Ill.
22. Cincinnati, Ohio-Ky.-Ind.
23. Cleveland, Ohio.
*24. Colorado Springs, Colo.
25. Columbia, S.C.
*26. Columbus, Ohio.
*27. Corpus Christi, Tex.
*28. Dallas-Fort Worth, Tex.
*29. Dayton, Ohio.
*30. Denver-Boulder, Colo.
31. Des Monies, Iowa.
32. Detroit, Mich.
33. Duluth-Superior, Minn.-Wis.
*34. El Paso, Tex.
35. Erie, Pa.
36. Evansville, Ind.-Ky.
*37. Flint, Mich.
38. Fort Wayne, Ind.
*39. Fresno, Calif.
40. Grand Rapids, Mich.
41. Greensboro-Winston Salem-High Point, N.C.
42. Greenville-Spartanburg, S.C.
43. Harrisburg-York, Pa.
44. Hartford, Conn.
*45. Houston, Tex.
46. Indianapolis, Inc.
47. Jackson, Miss.
48. Jacksonville, Fla.
49. Kansas City, Mo.-Kans.
50. Knoxville, Tenn.
*51. Las Vegas, Nev.
52. Lexington, Ky.
53. Little Rock-North Little Rock, Ark.
*54. Los Angeles-Long Beach, Calif.
55. Louisville, Ky.-Ind.
56. Madison, Wis.
57. Memphis, Tenn.-Ark.-Miss.
58. Miami, Fla.
59. Milwaukee, Wis.
60. Minneapolis-St. Paul, Minn.-Wis.
61. Mobile, Ala.
62. Moline-Davenport-Rock Is., Iowa-Ill.
63. Nashville-Davidson, Tenn.
64. New Orleans, La.
65. New York-Newark, N.Y.-N.Y.
66. Norfolk, Va.
67. Oklahoma City, Okla.
68. Omaha, Nebr.-Iowa.
69. Orlando, Fla.
70. Pensacola, Fla.
71. Philadelphia, Pa.-N.J.
*72. Phoenix, Ariz.
73. Pittsburgh, Pa.
*74. Portland, Oreg.-Wash.
75. Providence-Warwick-Pawtucket, R.I.-Mass.
76. Raleigh-Durham, N.C.
77. Richmond, Va.
78. Rochester, N.Y.
79. Rockford, Ill.
*80. Sacramento, Calif.
81. St. Louis, Mo.-Ill.
*82. Salinas-Seaside-Monterey, Calif.
*83. Salt Lake City, Utah.
*84. San Antonio, Tex.
*85. San Diego, Calif.
*86. San Francisco, Calif.
*87. Seattle-Everett, Wash.
88. Shreveport, La.
89. South Bend, Ind.
*90. Spokane, Wash.
91. Syracuse, N.Y.
92. Tampa-St. Petersburg, Fla.
93. Toledo, Ohio-Mich.
*94. Tucson, Ariz.
95. Tulsa, Okla.
96. Utica-Rome, N.Y.
97. Washington, D.C.-Md.-Va.
98. West Palm Beach-Boca Raton, Fla.
99. Wichita, Kans.
100. Youngstown-Warren, Ohio.

[1]Refer to Figure 7.2 for the location of each SMSA.
*Indicates an SMSA also included in a 30 x 30 interaction matrix analysed later.

The attractiveness and emissiveness of each centre were measured in two ways depending upon the type of model being calibrated. In the doubly-constrained models they were measured by total inflows into a destination and total outflows out of an origin, respectively. In the unconstrained models, the attractiveness and emissiveness of a centre were both measured by its 1970 population size. In the production-constrained models, total outflows from an origin were used as a measure of the emissiveness of that origin but since either of the above attractiveness measures is commonly used, two versions of both the production-constrained gravity model and the production-constrained competing destinations model were calibrated. In one version, the attractiveness of a destination is measured by total inflows into that destination, while in the other, attractiveness is measured by population size. It is interesting to note whether one definition of attractiveness is superior to the other in terms of parameter estimation. Both versions of the production-constrained models are now formally defined.

7.3 Two Versions of the Production-Constrained Gravity and Competing Destinations Models

As mentioned above, both unconstrained models and both doubly-constrained models to be compared are defined unambiguously. However, there are two versions of the production-constrained models given in (1.34) and (6.13) depending upon the definition of the attractiveness of a centre. Both versions of each model will be calibrated and compared. The two versions of the production-constrained gravity model are:

Version I

$$\hat{I}_{ij} = Z_i \; O_i \; D_j \; d_{ij}^{\hat{\beta}_i} \tag{7.1}$$

where,

$$Z_i = [\sum_{j=1}^{n} D_j \; d_{ij}^{\hat{\beta}_i}]^{-1} \tag{7.2}$$

and,

$$D_j = \sum_{i=1}^{m} I_{ij} \tag{7.3}$$

I_{ij} is the true interaction between i and j and $\hat{I}_{ij}$ is a prediction of I_{ij}.

Version II

$$\hat{I}_{ij} = Z_i \; O_i \; D_j \; d_{ij}^{\hat{\beta}_i} \tag{7.4}$$

where,

$$Z_i = [\sum_{j=1}^{n} D_j \; d_{ij}^{\hat{\beta}_i}]^{-1} \tag{7.5}$$

and

$$D_j = m_j \tag{7.6}$$

where m_j is the population size of j.

The two versions of the production-constrained competing destinations model are:

Version I

$$\hat{I}_{ij} = Z_i \; O_i \; D_j \; A_{ij}^{-1} \; d_{ij}^{\hat{\beta}_i} \tag{7.7}$$

where,

$$Z_i = [\sum_{j=1}^{n} D_j A_{ij}^{-1} d_{ij}^{\hat{\beta}_i}]^{-1} \tag{7.8}$$

and,

$$D_j = \sum_{i=1}^{m} I_{ij} \tag{7.9}$$

Version II

$$\hat{I}_{ij} = Z_i O_i D_j A_{ij}^{-1} d_{ij}^{\hat{\beta}_i} \tag{7.10}$$

where,

$$Z_i = [\sum_{j=1}^{n} D_j A_{ij}^{-1} d_{ij}^{\hat{\beta}_i}]^{-1} \tag{7.11}$$

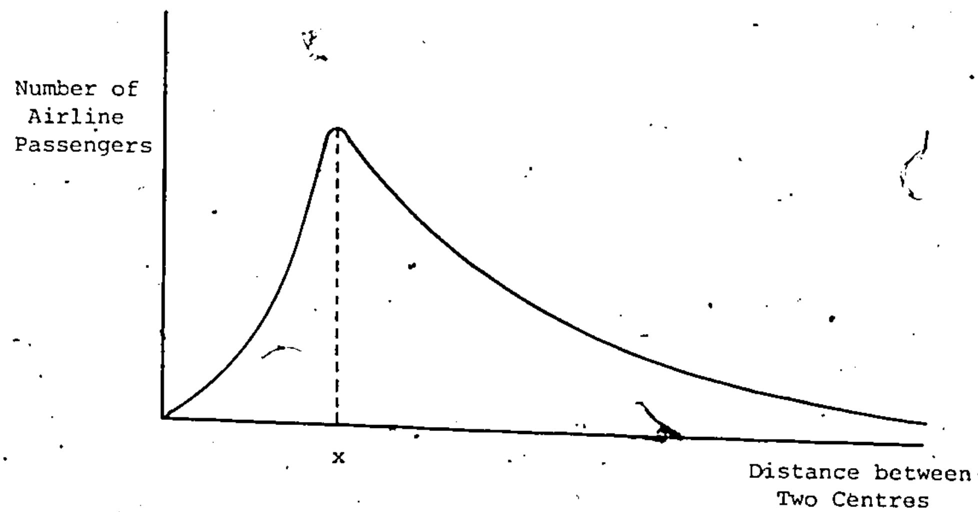
and,

$$D_j = m_j \tag{7.12}$$

7.4 The Problem of Short-Distance Interactions

Distance-decay parameters estimated from airline interaction data contain a potential spatial structure bias which is peculiar to this type of data. Consider two centres between which interaction is taking place. Few people would travel by commercial aircraft if the two centres are separated by only a very short distance. Up to some distance, interaction by air will increase as the distance between the two centres increases. Beyond this distance, interaction by air will decrease as distance between the centres increases due to a friction of distance effect. This situation is described in Figure 7.1.

FIGURE 7.1: The Relationship between the Number of Airline Passengers and Distance Travelled.



In estimating distance-decay parameters, where a measure of the friction of distance is obtained, the interest is only in interactions which occur over distances greater than x in Figure 7.1. If all interactions are used in estimating β, then the resulting estimate will be biased since it would also include a measure of the rate at which interaction increases as distance increases up to distance x. This would have the effect of reducing the absolute value of $\hat{\beta}$. It is evident that such a bias is a spatial structure bias when origin-specific distance-decay parameters are estimated. The absolute value of $\hat{\beta}_i$ will be less for an origin having destinations at distances less than x than the equivalent value for an origin not having destinations at distances less than x, *ceteris paribus*. Different values of $\hat{\beta}_i$ would not necessarily reflect varying perceived travel disutilities between centres but merely that some centres have destinations in close proximity while others have not. This spatial structure effect needs to be removed from the data before the set of distance-decay parameters is estimated since it is peculiar to this type of data.[1] The spatial structure bias that is general to the calibration of gravity models with any interaction data needs to be identified as a separate entity from the bias described above.

To remove the bias in $\hat{\beta}_i$ described by the pattern of interaction in Figure 7.1, all interactions which take place over distances less than x are eliminated from the model calibration. Hence, the variable n in all model formulations is replaced by n_i, the number of destinations greater than distance x from origin i. In so doing, estimates of $\hat{\beta}_i$ are obtained which purport to measure solely the decline of interaction as distance increases. In order to determine the distance x for the interaction data used in this study, all interactions over distances less than 200 miles were divided into classes of distances of 20 - mile intervals. An unconstrained competing destinations model was calibrated on the data in each class and the class with the shortest distance interactions and having a negative $\hat{\beta}$ was the 160 - 179 mile class. When this and the 140 - 159 class were disaggregated to 10 - mile interval classes, the 160 -

[1] Actually, all interaction data when graphed against distance will have a similar shape to that given in Figure 7.1 since the volume of interaction over zero distance is zero. However, in most cases, the distance x is so small that interactions over distances less than x can be ignored. In cases such as airline interaction or shipping interaction, x is large enough that this effect cannot be ignored.

169 mile class produced the first negative value of $\hat{\beta}$. Further disaggregation was impossible due to sample size. Thus, interactions which took place over distances of less than 160 miles were excluded from all subsequent model calibrations. It is interesting to note that this distance is similar to that given by Iklé [1954] who suggested that the substitution of ground transportation by air transportation begins at distances between 150 and 200 miles.

7.5 The Problem of Zero Interactions

It is assumed that no real flows between centres are zero (otherwise the models would be mis-specified) and that zero interaction values result from the sampling procedure used. Zero interaction values cause problems when natural logarithms of those values need to be taken. Such problems occur when the log-linear form of the unconstrained gravity model is calibrated by OLS regression. In logarithms, zero interaction values would be $-\infty$ and the regression would be indeterminable. Thus, although the problem of zero interactions is only of minor consequence in the data set used here (out of 9900 interactions, only 150 are zero), the problem has to be solved where such interactions occur. There are three possible solutions: the first is to omit all zero interactions from the analysis: the second is to reduce the size of the interaction matrix by removing all origins and destinations which have zero interactions; and the third is to approximate zero interactions by some reasonable value which is large enough so that its logarithm does not unduly bias the calibration results and yet which measures the fact that some interactions are very low. The merits of these three solutions are now discussed and the reasons are given for choosing the third solution.

By omitting all occurrences where $I_{ij} = 0$ from the analysis, the problem of taking the logarithm of zero is solved. However, the parameter estimates obtained from such a method will be biased. For example, if d_{1j} is small when $I_{1j} = 0$ but d_{2j} is large when $I_{2j} = 0$, these different relationships should be measured by different distance-decay parameters but they will not be if all $I_{ij} = 0$ are omitted. The resulting $\hat{\beta}_i$ values will thus be biased by these omissions.

The second solution will remove all zero interactions as in the first solution but it will do so in an unbiased manner. However, by omitting certain origins entirely from the analysis, the explanatory power of the model is reduced and information may be required regarding some of the origins for which there are zero interactions. It is conceivable that for some interaction matrix every row of the matrix contains at least one zero element and hence no model calibration by OLS regression could be undertaken using that matrix.

The third method was chosen to solve the problem of zero interactions and all zero interactions were approximated by setting $I_{ij} = 1$. This seems a reasonable approximation when the order of magnitude of some of the interactions is 10^5. The approximation is equivalent to constraining the logarithm of interaction to be non-negative. No significant bias is introduced into the calibration procedure as it would be if I_{ij} were set to a smaller value and yet interactions which are in fact zero still play an important role in determining $\hat{\beta}_i$. To demonstrate the unbiasedness of distance-decay parameter estimates when zero interactions are approximated by unit interactions, the unconstrained gravity model in

(1.33) was calibrated for each row of the 100 x 100 interaction matrix described above. The origin-specific distance-decay parameters exhibited a marked relationship to origin accessibility and this relationship was the same for origins having no zero interactions as for origins having one or more zero interactions. The linear relationship between $\hat{\beta}_i$ and A_i (the accessibility of origin i) for the 42 origins having no zero interactions is:

$$\hat{\beta}_i = -2.31 + .00000706\ A_i \qquad R^2 = .76 \qquad (7.13)$$
$$(.23) \quad (.00000062)$$

The figures in brackets are the standard errors of the parameter estimates. The linear relationship between $\hat{\beta}_i$ and A_i for the 58 origins having one or more zero interactions is:

$$\hat{\beta}_i = -2.57 + .00000733\ A_i \qquad R^2 = .74 \qquad (7.14)$$
$$(.21) \quad (.00000058)$$

When both sets of origins are combined, the relationship is:

$$\hat{\beta}_i = -2.44 + .00000712\ A_i \qquad R^2 = .73 \qquad (7.15)$$
$$(.15) \quad (.00000044)$$

The relationship between $\{\hat{\beta}_i\}$ and $\{A_i\}$ is virtually identical whether $\{\hat{\beta}_i\}$ is derived from origins having no zero interactions and origins having one or more zero interactions which are approximated by unit interactions. Thus, this method appears to solve the problem of zero interactions since the resulting distance-decay parameter estimates are not biased by the approximation. However, out of the 150 zero interactions which occur in the matrix, two origins, Rockford and Bridgeport, contain 53 and 18 zero interactions, respectively, and the parameter estimates for each of these origins will be treated with suspicion. The parameter estimates for all other origins, however, will be minimally biased by any zero interactions.

7.6 <u>Notes on the Calibration of Gravity and Competing Destinations Models</u>

There exists an extensive literature on the calibration of interaction models (*inter alia*, Batty and Mackie [1972], Stetzer [1976] and Openshaw [1976]), and this section merely outlines some characteristics of the calibration procedures employed here. The calibration of the unconstrained interaction models given in (1.33) and (6.11) was undertaken by a packaged OLS regression computer program. The data for each origin were assumed to be free from measurement error which is reasonable given that the data have the degree of accuracy described above. The calibration of the constrained interaction models was undertaken by maximum likelihood estimation. The computer programs for the calibration of (6.16) and (7.10) are given in Appendices VI and VII respectively. It is a simple matter to alter these programs to calibrate the models given in (7.1), (7.4), (7.7) and (1.36). The production-constrained models were calibrated by a Newton-Raphson interative procedure while the doubly-constrained models were calibrated using a first-order iterative procedure: both techniques are outlined in Batty and Mackie [1972]. For the doubly-constrained interaction models, the first-order iteration procedure is probably a faster calibrating procedure than the Newton-Raphson technique when the number of origins is large.[1]

[1]Batty and Mackie indicated that for a small matrix (30 x 30), the Newton-Raphson procedure was more efficient than the first-order iteration procedure when a single parameter was estimated. No tests investigated, however, the effects of matrix size or multiple parameter estimation on calibration efficiency.

The latter procedure consists of calculating and inverting an m x n matrix of derivatives (where m is the number of origins and n is the number of destinations) which becomes very time consuming as m and n become large.

In the set of competing destinations models, the accessibility of a destination as perceived by residents of origin i was calculated with respect to the remaining 98 centres in the analysis (the origin and destination are excluded from the calculation). This sample of alternative destinations was assumed to be a good representation of the population set of such centres. The attractiveness of a centre, a component of the accessibility terms, was measured in each model by the 1970 population size of the SMSA. The parameter, σ_i, was assumed to equal -1.0 for all i, although in one analysis σ_i was estimated iteratively with β_i but no advantage was found to this (see Section 7.10).

As a crude measure of how well each calibrated model replicated the existing interaction data, an adjusted R_i^2 statistic was used (see *inter alia*, Huang [1970, p. 81]). This statistic is defined by:

$$\bar{R}_i^2 = 1 - \frac{\sum_{j=1}^{n_i} (I_{ij} - \hat{I}_{ij})^2}{\sum_{j=1}^{n_i} (I_{ij} - \bar{I}_{ij})^2} \cdot \frac{n_i - 1}{n_i - k_i - 1} \qquad (7.16)$$

where,

I_{ij} is the measured interaction between i and j,

$\hat{I}_{ij}$ is the model prediction of I_{ij},

$$\bar{I}_{ij} = \frac{\sum_{j=1}^{n_i} I_{ij}}{n_i}$$

n_i is the number of destinations greater than, or equal to, 160 miles from origin i,

and,

k_i is the number of parameters estimated for origin i in each model calibration.

Since n_i is always very large compared to k_i for the 100 x 100 interaction matrix, the adjustment, $(n_i - 1)/(n_i - k_i - 1)$ could safely be ignored. It is included, however, since it proves useful in a later analysis of a 30 x 30 interaction matrix. Despite the potential failings of the $\bar{R}_i^2$ statistic, as outlined by, *inter alia*, Wilson [1976], its use can be defended on several grounds. The interpretation of the $\bar{R}^2$ statistic is immediately understood, ranging from -1.0 to +1.0 and it has been used in many other studies so that direct comparisons of model fits can be made. It is not influenced by the magnitude of the data as are some other goodness-of-fit measures such as phi and chi-squared statistics. In a practical sense, an advantage of using $\bar{R}^2$ in this study is that it is given as output in the packaged regression programs used to calibrate the unconstrained models. Finally, as Hutchinson and Smith [1979] note, no goodness-of-fit statistic is without its disadvantages: here, those of $\bar{R}^2$ are outweighed by its advantages.

7.7 Calibration Results

Origin-specific parameter estimates are given and compared for both sets of interaction models described in Chapter Six. Since the emphasis is on the spatial pattern and spatial variation of these estimates and there are 100 estimates for each model calibration, the results are mapped in a series of figures. The key to these figures is given in Table 7.1 and Figure 7.2: the former gives the name and number of each SMSA, while the latter gives its location in the United States. Quantitative evidence is presented on the relationship between the sets of parameter estimates and the set of origin accessibilities, as a measure of spatial pattern. However, in order to obtain a visual impression of the varying degrees of spatial pattern exhibited by the sets of parameter estimates, the accessibility of each origin to all other centres in the analysis is given in Figure 7.3. Lines joining origins of approximately equal accessibility are drawn to clarify the spatial pattern of these data.[1] The pattern is one in which origins in the North-East are very accessible and accessibility generally decreases as distance from the North-East increases. The pattern is interrupted only around California where there is a slight increase in accessibility. It is interesting to compare the spatial pattern of each set of parameter estimates with the spatial pattern of origin accessibilities.

The results of calibrating individual models are now presented and the results from each gravity model are compared to those from each equivalent competing destinations model. For clarity, the results presented in this section concentrate on the spatial variation, spatial pattern and the interpretation of the parameter estimates. This concentration is justified given the *a priori* theoretical expectation that

[1]In this, and in each of the subsequent figures where contour lines are drawn, there exist exceptions to the general trend these lines represent.

FIGURE 7.2: Location of the 100 SMSA's listed in Table 7.1

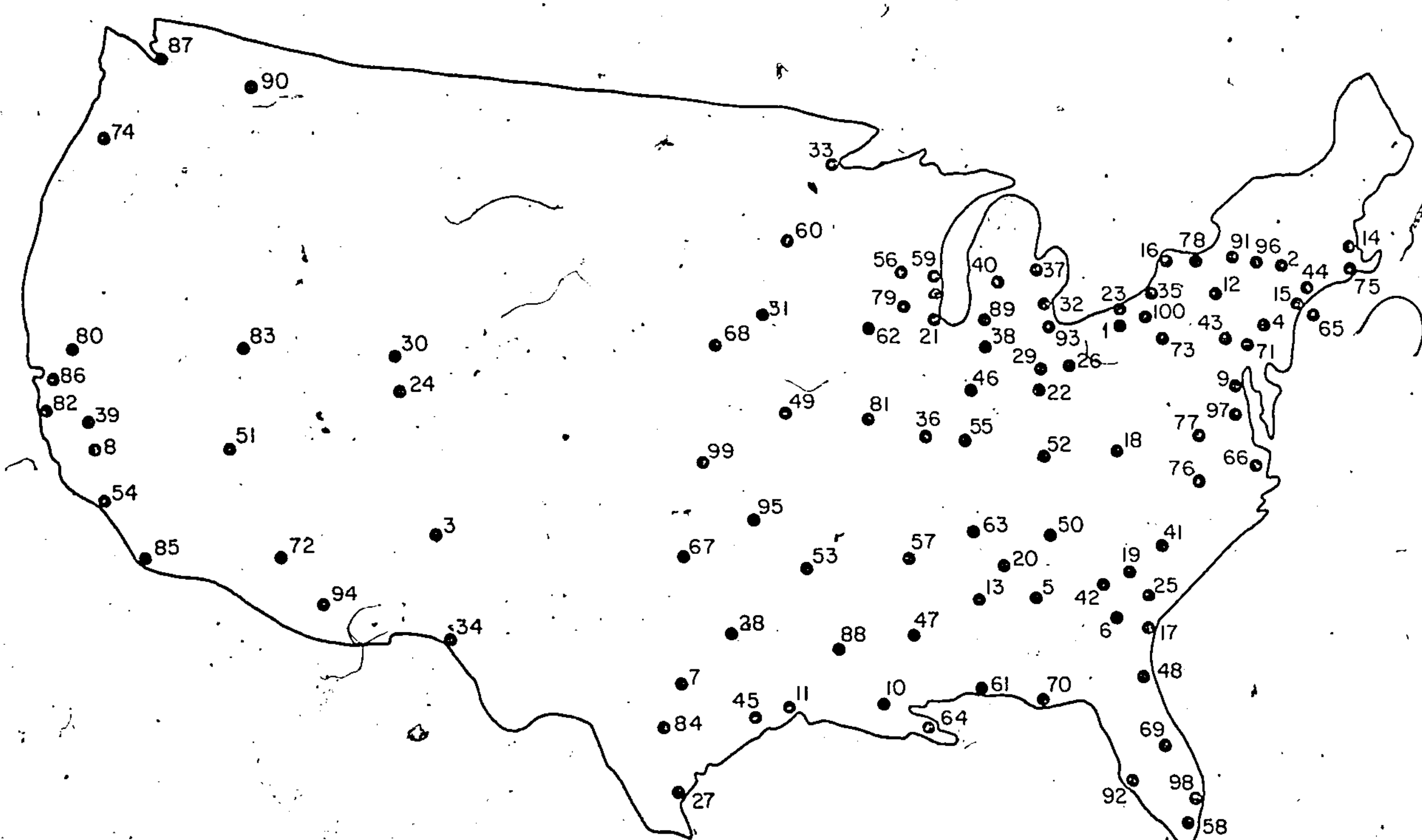
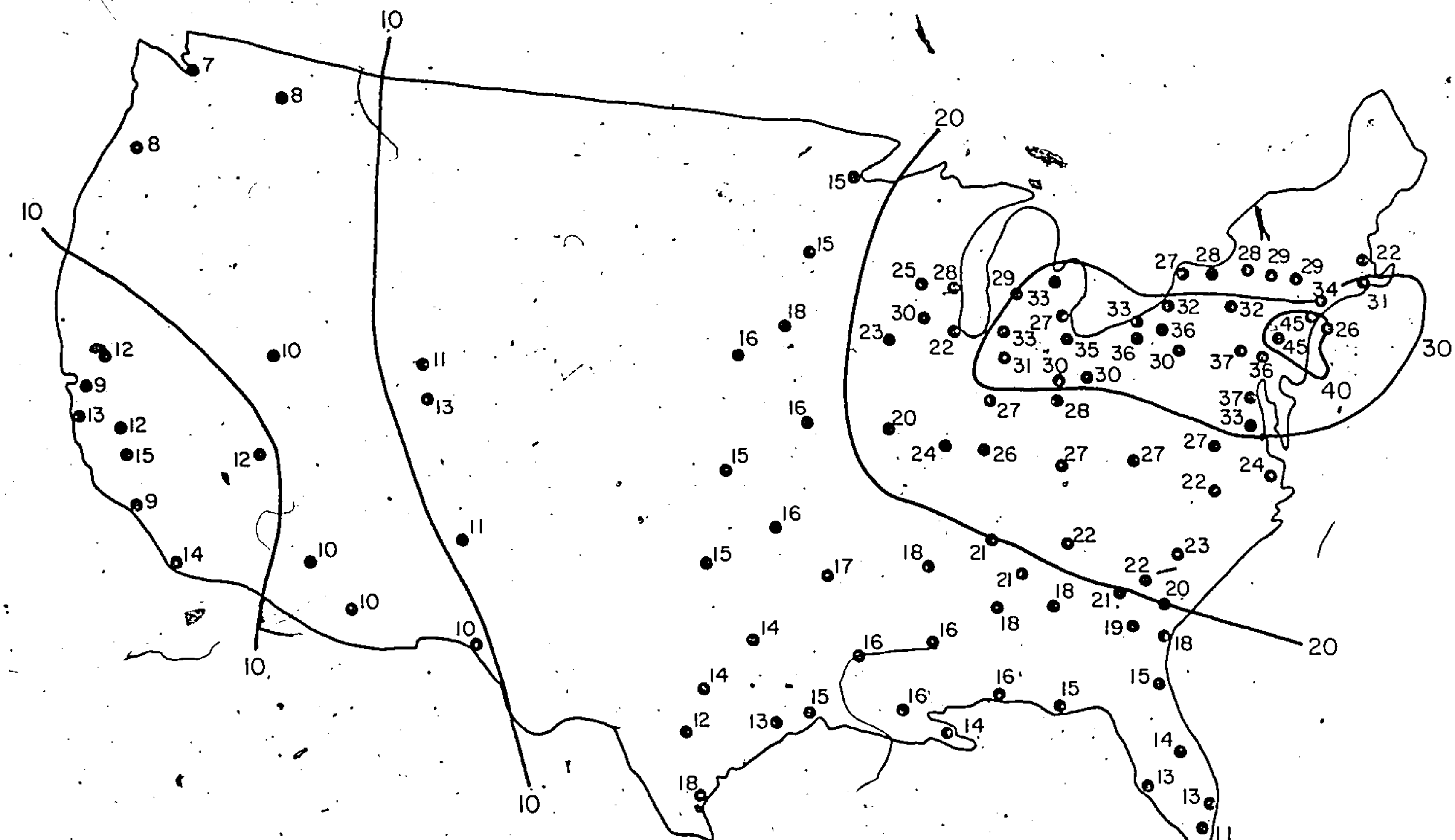


FIGURE 7.3: The Spatial Variation of Accessibility[1] (figures in 10,000's)



[1] $A_i = \sum_{j=1}^{n} m_j \, d_{ij}^{-1}$

the gravity models produce parameter estimates which are biased by spatial structure and the variation in such estimates is due to variation in spatial structure. Many other issues are raised in the comparison of the calibration results and some of these are discussed in later sections.

(i) Unconstrained Interaction Models

The distance-decay parameter estimates from the unconstrained gravity model in (1.33) are mapped in Figure 7.4. There is a marked spatial pattern in these estimates which is very similar to that of origin accessibility in Figure 7.3. Origins in the North-East have very low negative, or even positive, distance-decay parameter estimates and the estimates become more negative moving westwards and southwards, rising only slightly around California. The linear relationship between $\hat{\beta}_i$ and A_i,[1] as measured by a Pearson correlation coefficient (R), is .85 and regressing $\hat{\beta}_i$ on A_i and A_i^2 gives a coefficient of determination (R^2) of .84. The strength of this relationship can be seen in Figure 7.5 where $\hat{\beta}_i$ is graphed against the non-linear function of A_i - as A_i increases, $\hat{\beta}_i$ increases. The spatial pattern of $\hat{\beta}_i$ described in Figure 7.4 is the same as that described in previous empirical studies (see Section 1.5).

The variation in the distance-decay parameter estimates is also similar to that described in earlier empirical studies. In Figure 7.4, $\hat{\beta}_i$ ranges from a high of +0.5 for Flint, Michigan to a low of -2.6 for

---

[1] A_i, origin accessibility, is defined by: $A_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} m_j / d_{ij}$

FIGURE 7.4: The Spatial Variation of $\hat{\beta}_i$ Derived from (1.33): $I_{ij} = \alpha_i \, m_j^{\gamma_i} \, d_{ij}^{\beta_i}$

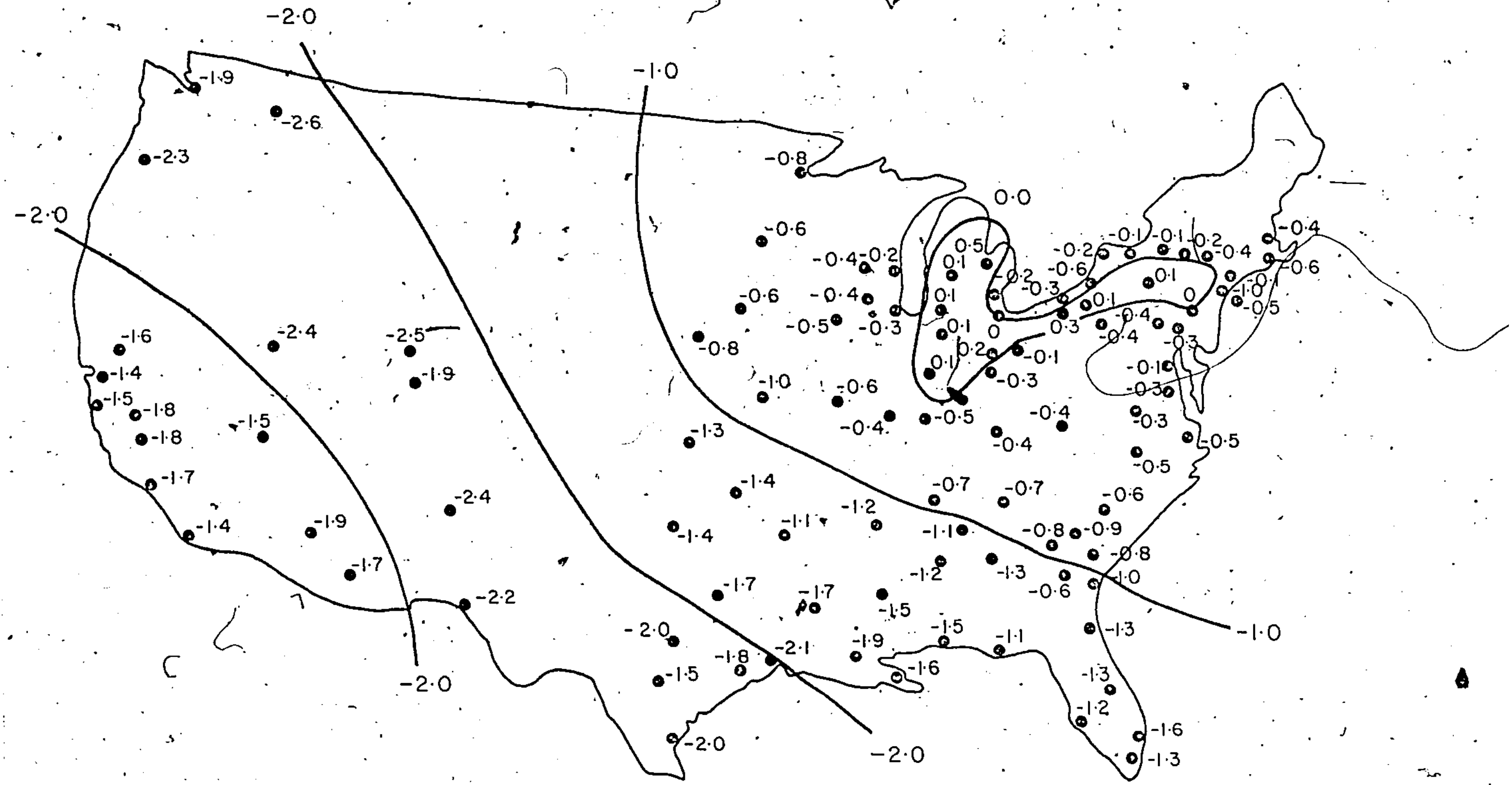
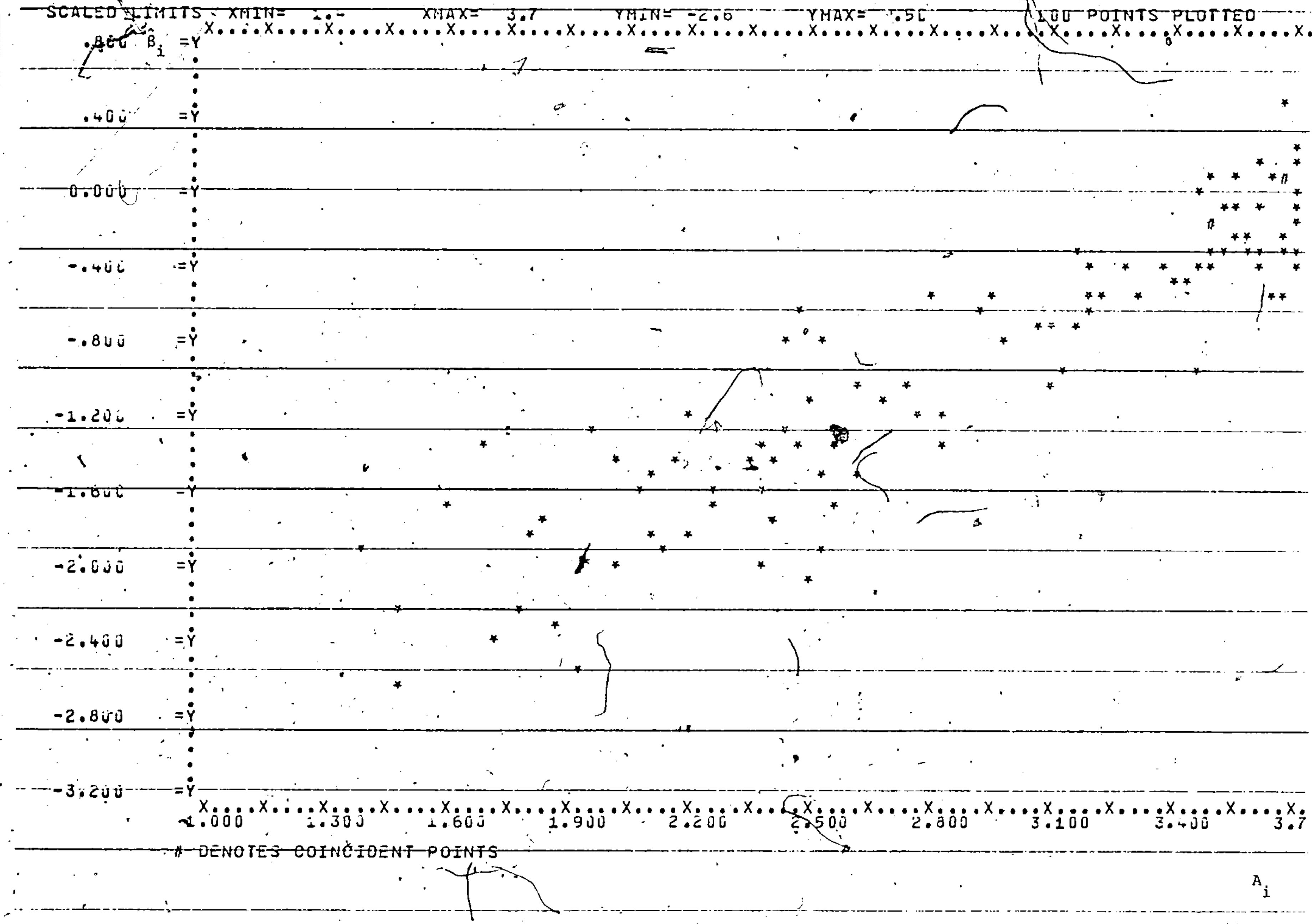


FIGURE 7.5: The Relationship between $\hat{\beta}_i$ from (1.33) and a Non-Linear Function of A_i



Spokane, Washington. Following the traditional interpretation of $\hat{\beta}_i$, this range represents a tremendous variation in interaction behaviour throughout the United States. Nine centres out of 100 have positive $\hat{\beta}_i$ values, while there are nine centres with estimated distance-decay parameters equal to or below -2.0. The former seemingly indicate that as distance from an origin increases, interaction increases, *ceteris paribus*, while the latter seemingly indicate that interaction from an origin decreases very rapidly as distance from that origin increases, *ceteris paribus*. The coefficient of variation for the set of estimates is 81.0.

The mean parameter estimate in Figure 7.4 is -0.9 and cities with less negative parameter estimates than this can be considered as "jet-setters" while cities with more negative parameter estimates can be considered as "parochial". It is counter-intuitive that, on the basis of the evidence given in Figure 7.4, cities such as Albany, Syracuse and Utica would be considered as "jet-setters" while cities such as Los Angeles, San Francisco and Las Vegas would be considered as "parochial".

The results from the unconstrained competing destinations model indicate how misleading are the parameter estimates from (1.33). The estimates from the calibration of (6.11) are given in Figure 7.6 and these can be compared to the equivalent values given in Figure 7.4. The spatial pattern and spatial variation in the parameter estimates have been greatly reduced: the correlation coefficient between $\{A_i\}$ and $\{\hat{\beta}_i\}$ is only 0.4 and a linear function of A_i explains less than a quarter of the variance of $\{\hat{\beta}_i\}$ it explained for the unconstrained gravity model estimates. There is still a band of less-negative parameter estimates in the North-East and Mid-West which, although more negative than the equivalent values in

FIGURE 7.6: The Spatial Variation of $\hat{\beta}_i$ Derived from (6.11): $I_{ij} = \alpha_i \, m_j^{\gamma_i} \, d_{ij}^{\beta_i} \, A_{ij}^{\delta_i}$,

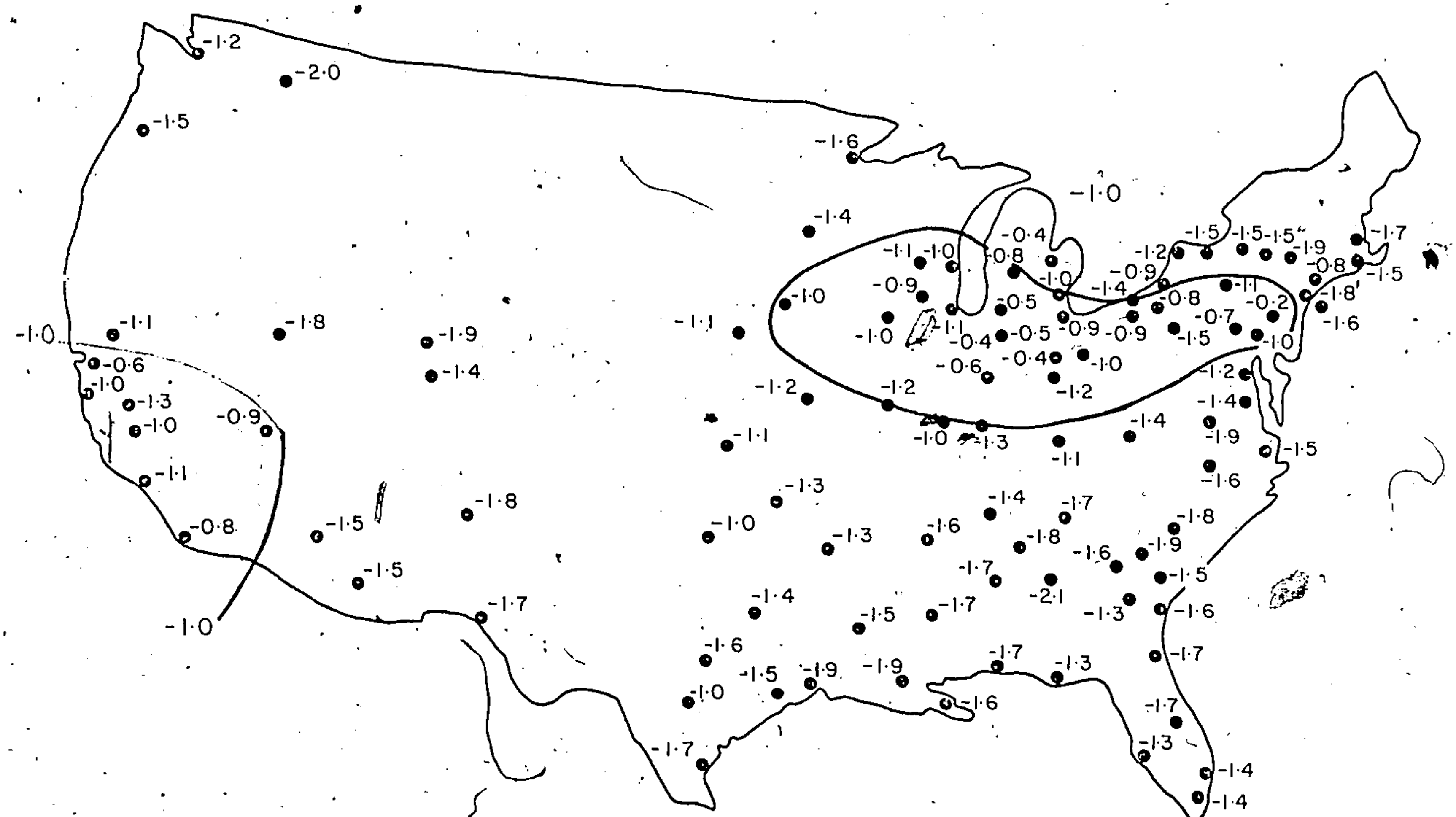


Figure 7.4, perhaps indicate that not all the spatial structure bias in $\hat{\beta}_i$ has been removed. There are no positive distance-decay parameter estimates in Figure 7.6 and the variation of $\{\hat{\beta}_i\}$ has been significantly reduced: the coefficient of variation of $\{\hat{\beta}_i\}$ is 30.5, compared to 81.0 for the gravity model estimates.[1]

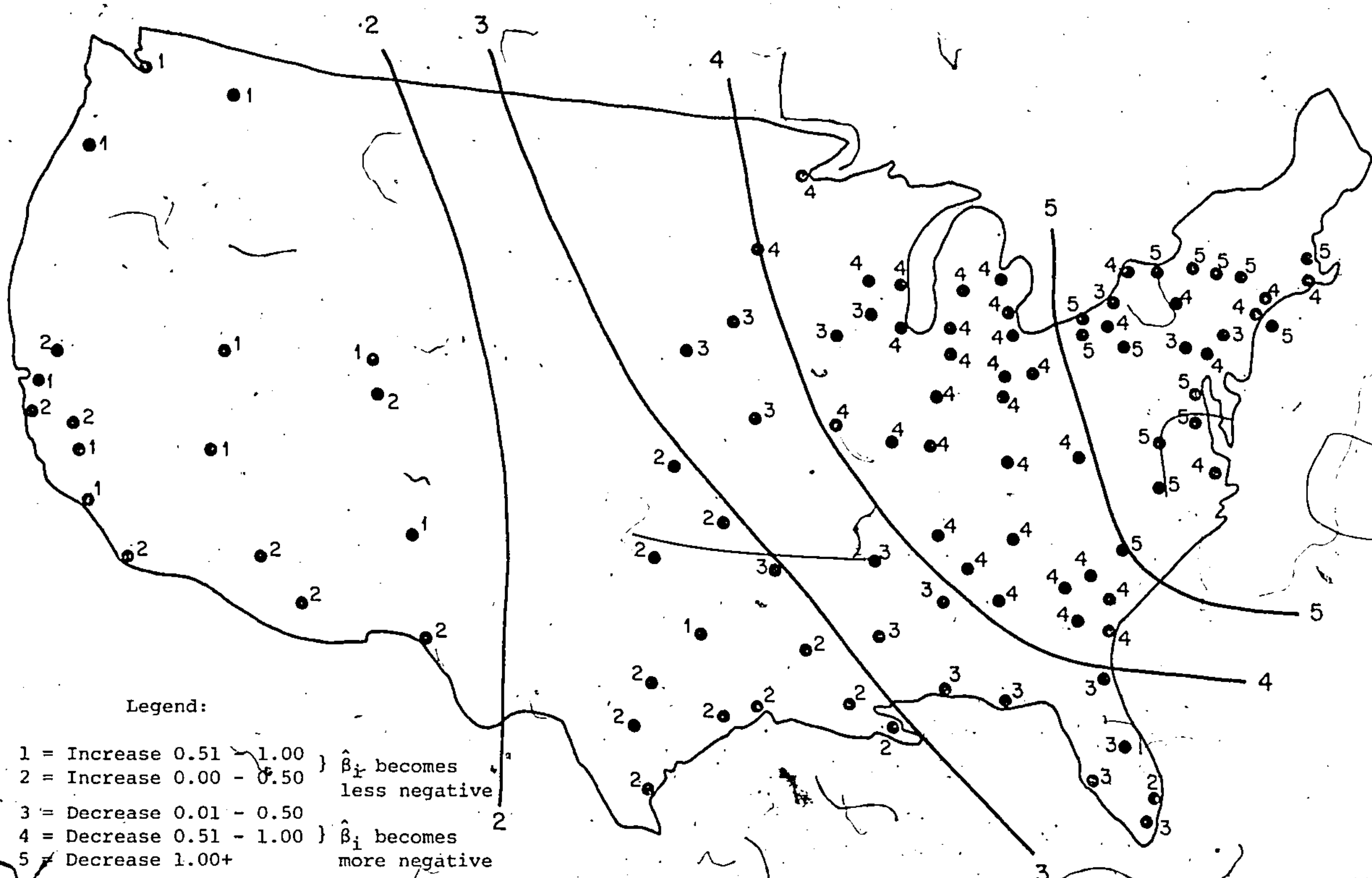
Since the unconstrained competing destinations model is correctly specified, the estimated distance-decay parameters in Figure 7.6 can be interpreted as measures of the perception of distance as a barrier to interaction by the inhabitants of each origin. This perception is fairly constant throughout the United States although some exceptions can be identified. The mean value of $\hat{\beta}_i$ is -1.3 indicating the "jet-setter" cities now include the Californian cities and Las Vegas while the "parochial" cities include Albany, Syracuse and Utica: a complete reversal of the inferences from the unconstrained gravity model. Some small cities in the Mid-West still appear to be "jet-setters", however, and it may be that not all the spatial structure effect has been removed by the competing destinations model. This possibility is discussed in Chapter Eight.

The change in $\hat{\beta}_i$ when A_{ij} is added to the unconstrained gravity model has a marked spatial pattern which is shown in Figure 7.7. These values indicate the direction and magnitude of the error in estimating distance-decay parameters from (1.33) instead of from (6.11). The parameter estimates for the cities in the North-East and Mid-West all become much more negative when estimated from (6.11) while those for

---

[1] A t-statistic given in Table 7.2 indicates that this is a significant reduction in variance.

FIGURE 7.7: Changes in $\hat{\beta}_i$ when A_{ij} is Added to the Unconstrained Gravity Model



the Western cities all become much less negative. The spatial structure bias in parameter estimates is greatest for origins of extreme accessibility: the most and least accessible origins in the system. Parameter estimates for origins of medium accessibility are only slightly biased. A similar result is obtained for changes in the goodness-of-fit statistic and it is explained in Section 7.8.

In the calibration of the unconstrained competing destinations model, the estimated elasticity of interaction with respect to the perceived accessibility of a destination, $\hat{\delta}_i$, is significantly negative at the 95% confidence level for 93 out of the 100 origins.[1] This reinforces the hypothesis that interaction patterns are in part a result of a two-stage decision-making process and that as a destination becomes more accessible, it becomes less attractive for interaction, *ceteris paribus*. There is only a small variation in $\{\hat{\delta}_i\}$ (the coefficient of variation is 37.9%) and this variation is only weakly related to variation in origin accessibility (R^2 = .16).

(ii) Production-Constrained Models

In Section 7.3 two sets of production-constrained models were defined, the difference between the sets being the definition of D_j. The calibration results proved to be similar for each set and in this section the results from only one set will be described: the set being the models given in (7.7) and (7.10) where $D_j = m_j$. The differences in

[1] For six origins, $\hat{\delta}_i$ is insignificantly negative and for one origin it is insignificantly positive.

calibration results between the two sets are discussed in Section 7.9.

The parameter estimates obtained from calibrating the production-constrained gravity model given in (7.7) are mapped in Figure 7.8. The results exhibit the same spatial pattern as exhibited by the parameter estimates derived from the unconstrained gravity model shown in Figure 7.4. Origins in the North-East and Mid-West have very small negative parameter estimates while those in the South and West have large negative estimates.[1] The correlation coefficient between A_i and $\hat{\beta}_i$ is 0.64.

The spatial variation of parameter estimates in Figure 7.8 is again large, the coefficient of variation being 52.2. While no positive estimates are derived from the calibration of the singly-constrained gravity model, 27 origins have parameter estimates whose absolute values are less than, or equal to, 0.5. Again, the behavioural interpretation given to these estimates is suspect. With a mean $\hat{\beta}_i$ of -0.9, most Californian cities and Las Vegas are more parochial than average and most of the Mid-Western cities have parameter estimates very close to zero (indicating an almost complete disregard for distance by the inhabitants of these cities).

Figure 7.9 represents the set of parameter estimates resulting from the calibration of the production-constrained competing-destinations model given in (7.10). The relationship between $\hat{\beta}_i$ and origin accessibility is eliminated (the correlation coefficient between the two is -.01) and the parameter estimates are remarkably constant over space.

---

[1] There are two parameter estimates which are obvious exceptions to this generalisation. Rockford, Illinois and Bridgeport, Connecticut, both very accessible origins, have parameter estimates of -2.2 and -2.1 respectively. As mentioned earlier, the interaction from these origins is zero in many cases. This is probably due to the airports in both origins being overshadowed by the far larger airports in nearby Chicago and New York City, respectively, and the parameter estimates reflect this situation.

FIGURE 7.8: The Spatial Variation of $\hat{\beta}_i$ Derived from (7.7): $I_{ij} = Z_i \; O_i \; D_j \; d_{ij}^{\;\beta_i} \quad (D_j = m_j)$

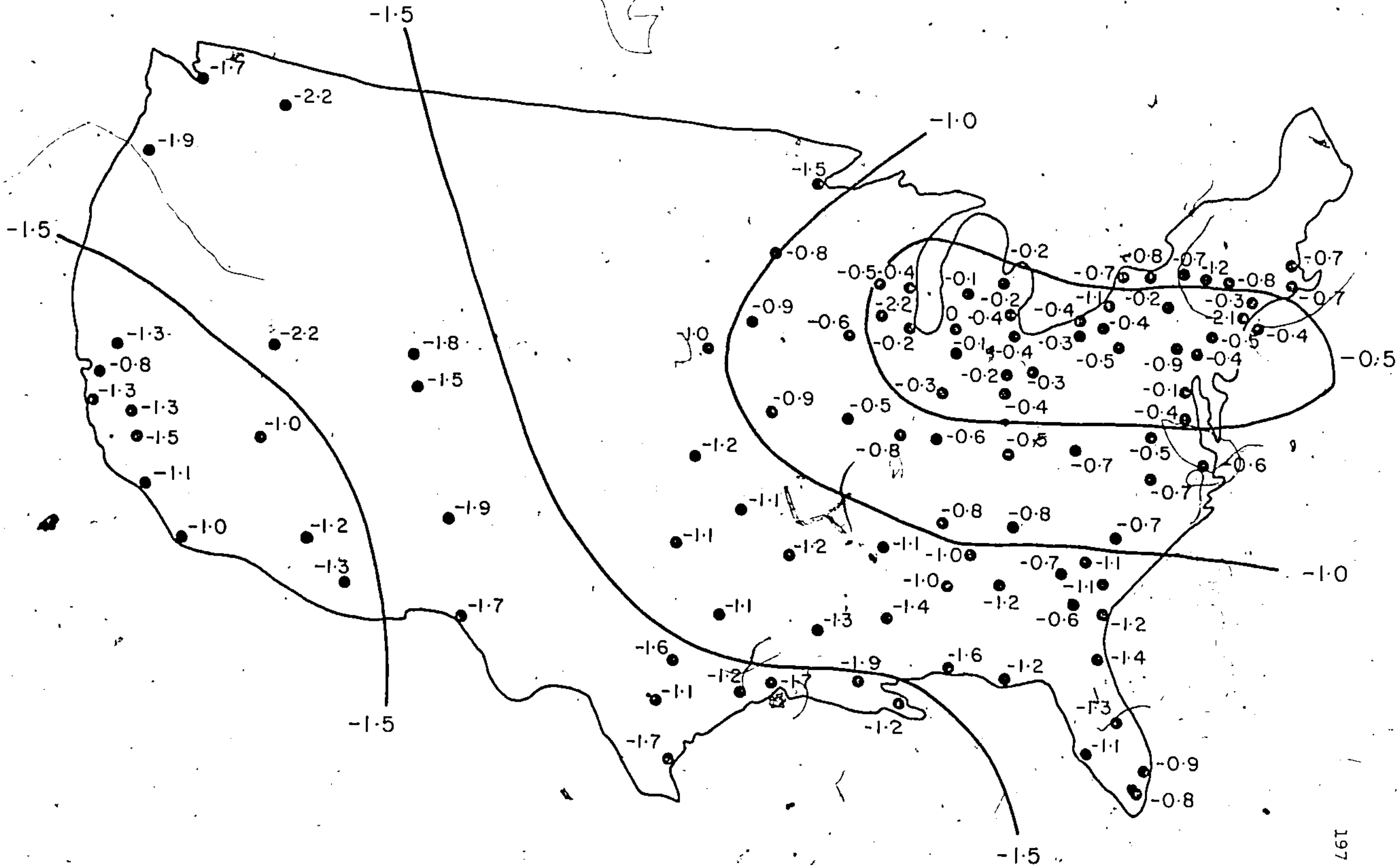
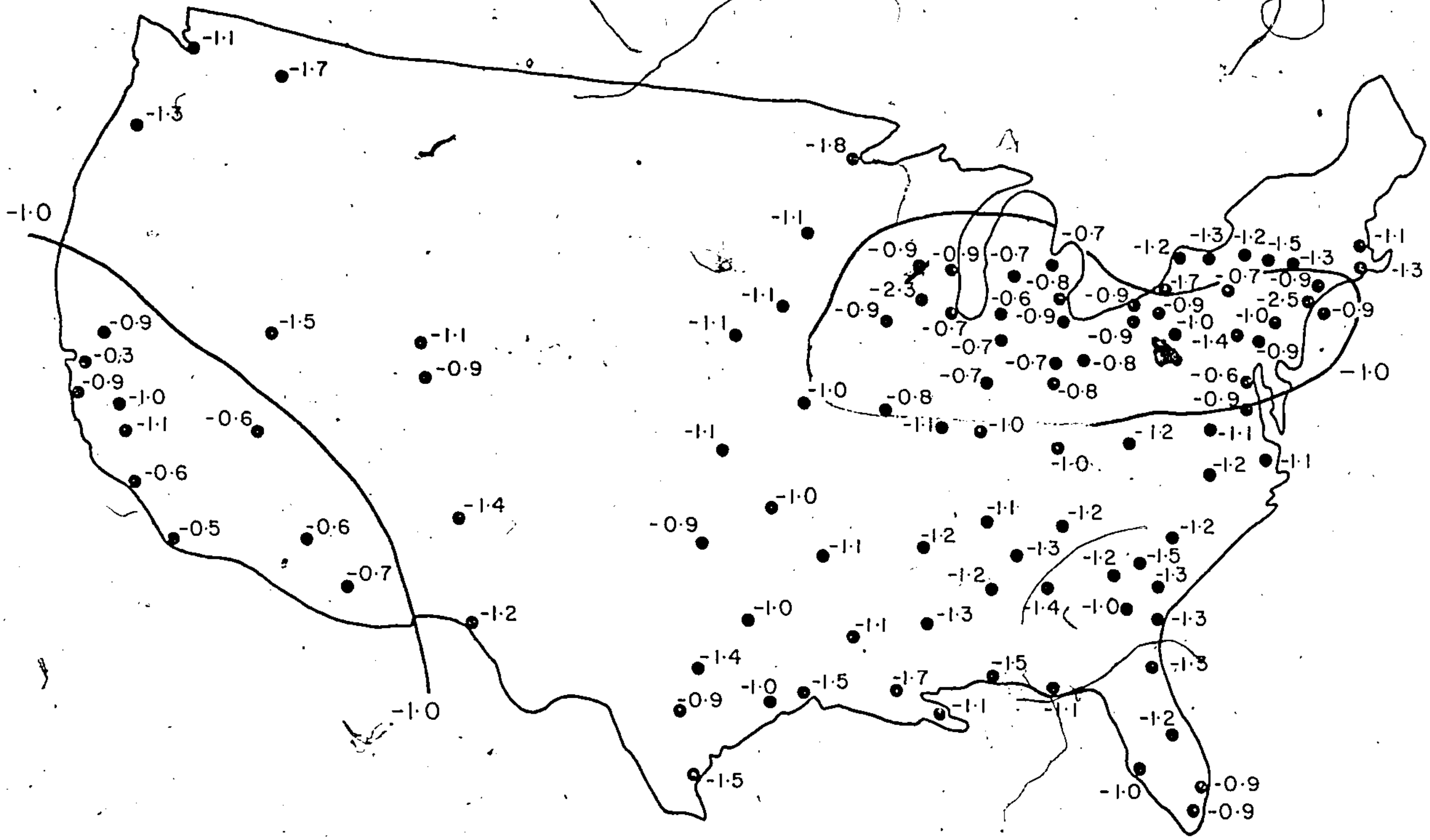


FIGURE 7.9: The Spatial Variation of $\hat{\beta}_i$ Derived from (7.10): $I_{ij} = Z_i\ O_i\ D_j\ A_{ij}^{-1}\ d_{ij}^{\beta_i}$ $(D_j = m_j)$

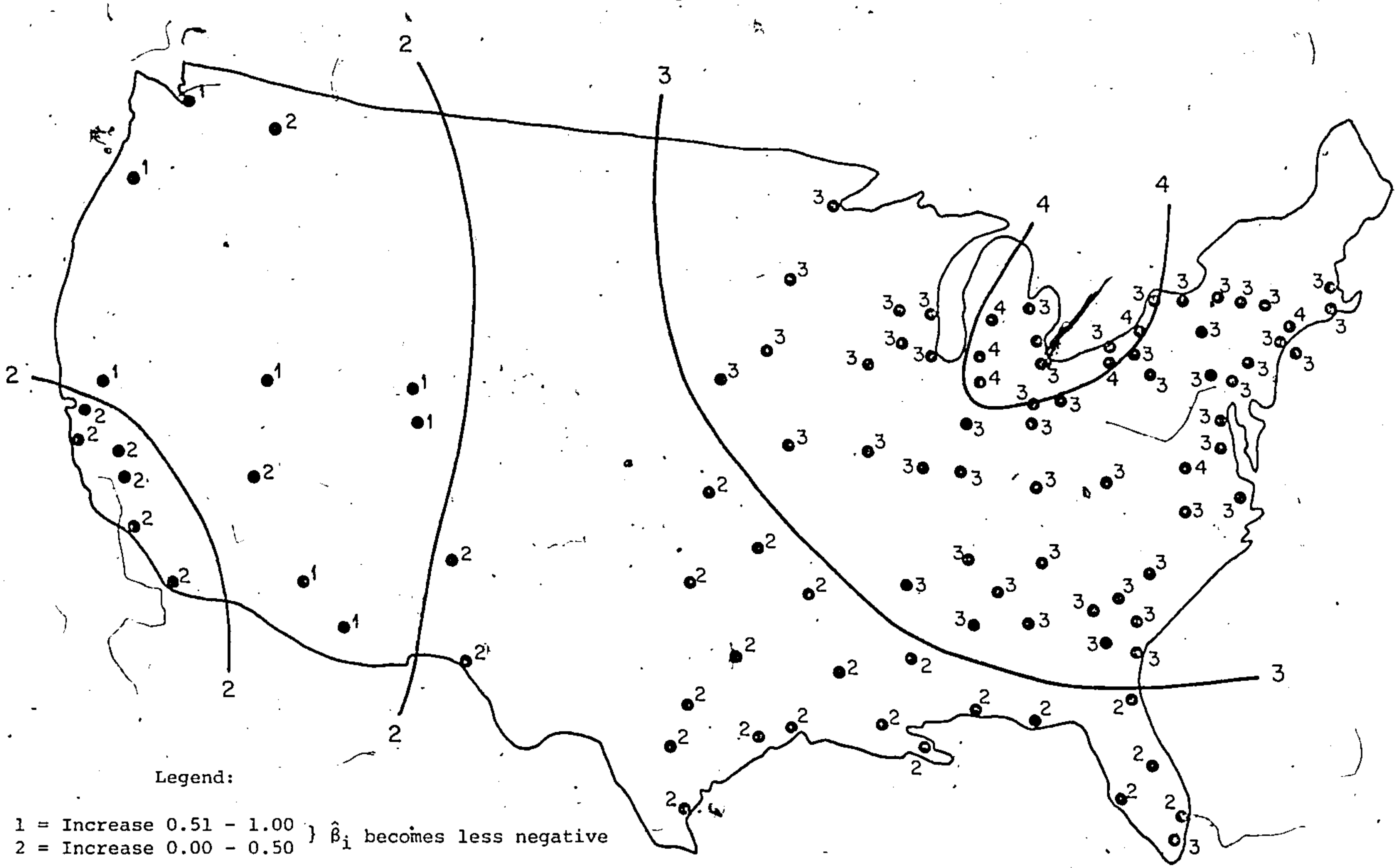


There are no estimates with absolute values less than 0.5 and there are only five estimates with absolute values greater than 1.5. The coefficient of variation for the set of parameter estimates is 30.2 which is a significant decrease from 52.2, the value obtained in the calibration of the production-constrained gravity model.

The interpretation of the parameter estimates given in Figure 7.9 as measures of interaction behaviour is also more reasonable. With the mean $\hat{\beta}_i$ equal to -1.1, the "jet-setter" cities can be identified as San Francisco, Las Vegas, Los Angeles, San Diego and Phoenix while the "parochial" cities can be identified as Spokane, Salt Lake City, Corpus Christi, Baton Rouge and Duluth. The large negative parameter estimates for Rockford and Bridgeport are again ignored as being influenced by an inordinate number of zero interactions.

The changes in $\{\hat{\beta}_i\}$ when the production-constrained competing destinations model is calibrated, as opposed to the production-constrained gravity model, are mapped in Figure 7.10. The pattern is similar to that for the unconstrained models: origins of extreme accessibility exhibit the greatest changes in $\hat{\beta}_i$, indicating that the spatial structure bias in the original gravity model is greatest for origins of extremely high or low accessibility. Origins of high accessibility have parameter estimates that are biased upwards by spatial structure while origins of low accessibility have parameter estimates that are biased downwards by spatial structure. These biases are removed when the competing destinations model is calibrated.

FIGURE 7.10: Changes in $\hat{\beta}_i$ when A_{ij} is added to the Singly-Constrained Gravity Model ($D_j = m_j$).



Legend:

1 = Increase 0.51 - 1.00 } $\hat{\beta}_i$ becomes less negative
2 = Increase 0.00 - 0.50

3 = Decrease 0.01 - 0.50 } $\hat{\beta}_i$ becomes more negative
4 = Decrease 0.51 - 1.00

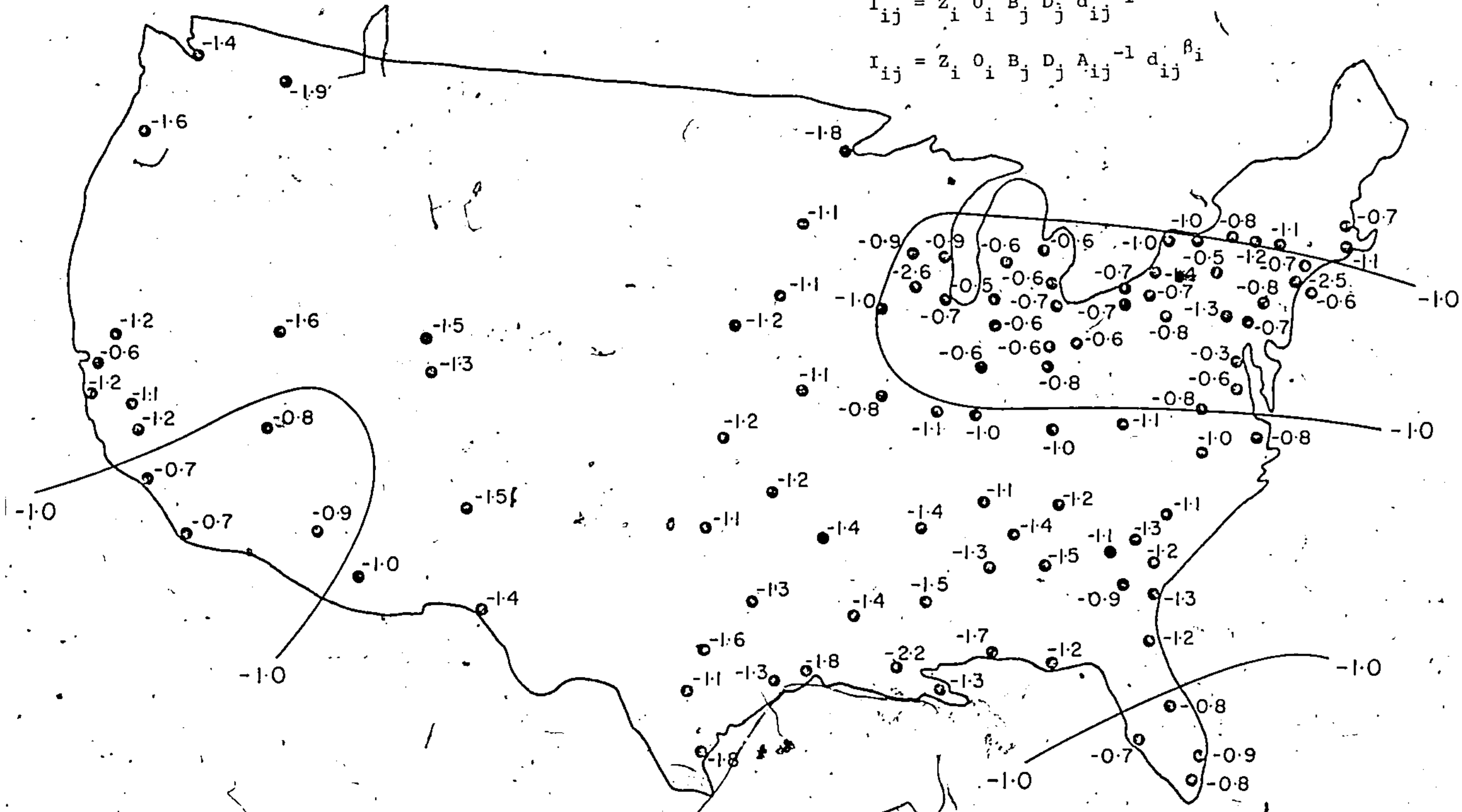
(iii) Doubly-Constrained Models

In each model calibration it is assumed that in the formulation of A_{ij}, $\sigma_i = -1.0$ for all i. With this assumption and with each centre being both an origin and a destination, A_j as defined in (6.23) is a good approximation of A_{ij}. The relationship between A_j and A_{ij} is given in (6.24). Hence, with reference to the discussion in Section 6.5, the doubly-constrained competing destinations model defined in (6.16) is virtually identical to the doubly-constrained gravity model defined in (1.36) and the results from the two models can be considered simultaneously. Consequently, only one set of parameter estimates will be discussed; these estimates can be derived from either model and are assumed to be free from the spatial structure effect which results from model mis-specification. The set of parameter estimates is mapped in Figure 7.11 and is very similar to that given for the unconstrained and singly-constrained competing destinations models in Figures 7.6 and 7.9. There is very little spatial pattern evident in the estimates (the correlation coefficient between $\{\hat{\beta}_i\}$ and $\{A_i\}$ is .34) and there is very little variation (the coefficient of variation of $\{\hat{\beta}_i\}$ is 37.5). With a mean $\hat{\beta}_i$ of -1.1, the estimates conform to expectations: the "jet-setter" cities include San Francisco, Los Angeles, Miami, Washington and New York and the "parochial" cities include Spokane, Salt Lake City, Baton Rouge, Mobile and Duluth. The band of slightly less negative parameter estimates still exists in the North-East and Mid-West, again perhaps an indication that not all the spatial structure effect is removed. However, the spatial pattern, spatial variation and interpretation of the parameter estimates all indicate that for the interaction data used here,

FIGURE 7.11: The Spatial Variation of $\hat{\beta}_i$ Derived from (1.36) and (6.16):

$$I_{ij} = Z_i \; O_i \; B_j \; D_j \; d_{ij}^{\beta_i}$$

$$I_{ij} = Z_i \; O_i \; B_j \; D_j \; A_{ij}^{-1} \; d_{ij}^{\beta_i}$$

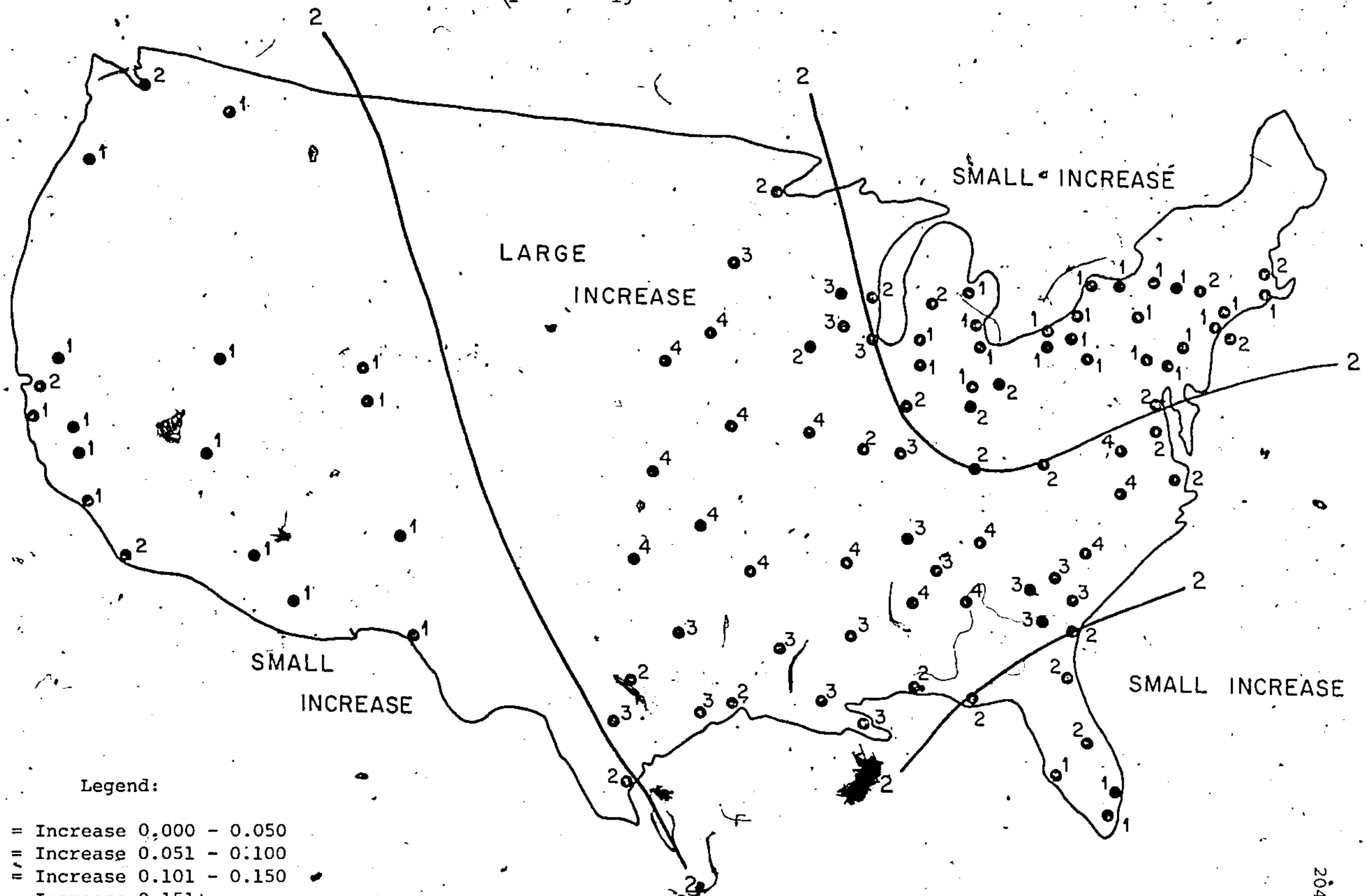


the doubly-constrained gravity model and the doubly-constrained competing destinations model are equivalent and that the original gravity model is not seriously mis-specified.

7.8 Spatial Variations in Model Performance

The goodness-of-fit statistic used to test the performance of the calibrated models is $\bar{R}_i^2$ which is defined in (7.16). It is of interest to compare the change in model performance between each gravity model and the equivalent competing destinations model. Since $\bar{R}_i^2$ is an origin-specific statistic, the change in $\bar{R}_i^2$ ($\Delta\bar{R}_i^2$) has a spatial pattern which can be mapped. Figure 7.12 represents this pattern for the unconstrained models defined in (1.33) and (6.11) and Figure 7.13 represents the pattern for the singly-constrained models defined in (7.7) and (7.10). In the unconstrained case, the smallest increases in $\bar{R}_i^2$ when A_{ij} is added to the gravity model occur for origins of extremely high or low accessibility, while the largest improvements in model performance occur for origins of medium accessibility. In the singly-constrained case, the pattern is very similar: origins in the most accessible region of the country around Ohio and Indiana and origins in the least accessible origins such as the West and Florida, show large decreases in $\bar{R}_i^2$ when accessibility is added to the model, while origins of medium accessibility in the middle of the country show large increases in $\bar{R}_i^2$. This is the reverse of the pattern described in Figures 7.7 and 7.10 of changes in $\hat{\beta}_i$ when A_{ij} is added to the unconstrained and singly-constrained gravity models. Both patterns can be explained by the spatial relationship

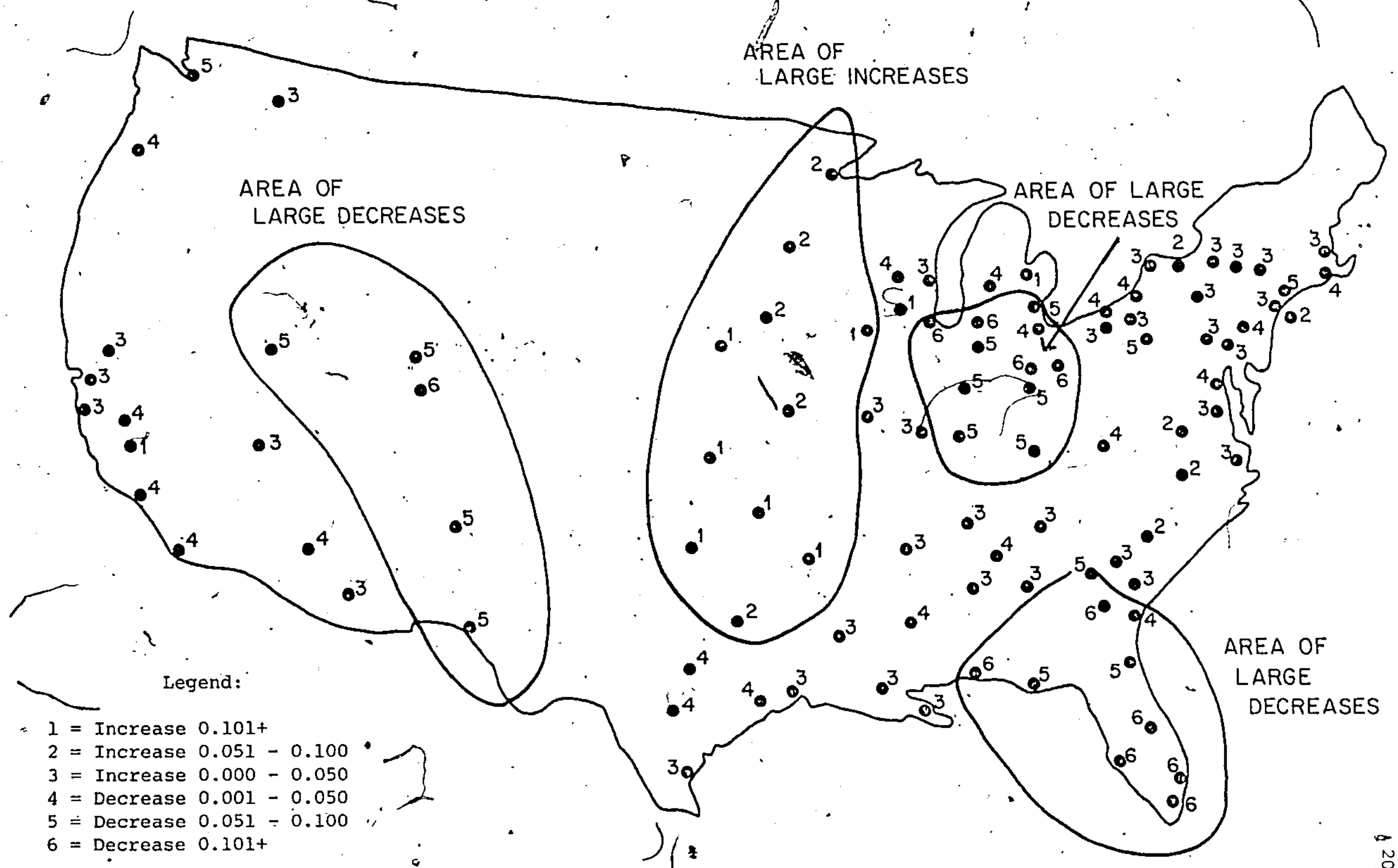
FIGURE 7.12: Increases in $\bar{R}_i^2$ when A_{ij} is Added to the Unconstrained Gravity Model



Legend:

1 = Increase 0.000 - 0.050
2 = Increase 0.051 - 0.100
3 = Increase 0.101 - 0.150
4 = Increase 0.151+

FIGURE 7.13: Changes in $\bar{R}_i^2$ when A_{ij} is Added to the Singly-Constrained Gravity Model ($D_j = m_j$)



Legend:

1 = Increase 0.101+
2 = Increase 0.051 - 0.100
3 = Increase 0.000 - 0.050
4 = Decrease 0.001 - 0.050
5 = Decrease 0.051 - 0.100
6 = Decrease 0.101+

between d_{ij} and A_{ij} (see Appendix VIII).[1] For the accessible origins in the North-East and Mid-West, A_{ij} and d_{ij} are highly negatively correlated: as distance from origin i increases, the accessibility of the destination decreases. Hence, the addition of A_{ij} to the original gravity models adds little or nothing to the explanatory power of the model[2] but it does correct the spurious positive or low negative correlation between d_{ij} and I_{ij}. Consequently, $\Delta\bar{R}_i^2$ is either low or negative while the change in $\hat{\beta}_i$ is large and negative.

For the inaccessible origins in the West and Florida, A_{ij} and d_{ij} are highly positively correlated so the addition of A_{ij} to the original gravity models again adds little or nothing to the explanatory power of the models but it does correct the spurious highly negative correlation between d_{ij} and I_{ij}. Consequently, $\Delta\bar{R}_i^2$ is either very slightly positive or negative while the change in $\hat{\beta}_i$ is large and positive.

For origins of medium accessibility, A_{ij} and d_{ij} are weakly correlated: in one direction as d_{ij} increases, A_{ij} increases, while in

---

[1] Appendix VIII gives the degree of multicollinearity between d_{ij}* and A_{ij}*, the logarithms of d_{ij} and A_{ij}, respectively, for each origin since this was given as output in the calibration of the unconstrained competing destinations models. The discussion is given in terms of the relationship between d_{ij} and A_{ij} in order to be more general since the logarithms of these variables are not part of the calibration of the singly-constrained interaction models. It is assumed that the degree of multicollinearity between d_{ij}* and A_{ij}* reflects the degree of multicollinearity between d_{ij} and A_{ij}.

[2] The explanatory power of the model is here taken to mean how well it replicates the existing interaction data for each origin. Since the competing destinations models are correctly specified, they will give better predictions of interactions in other spatial systems than would the gravity models even though the explanatory power of the two model types is equal on the existing data.

the other direction as d_{ij} increases, A_{ij} decreases. Hence, the addition of A_{ij} to the original models adds significantly to their explanatory power but does not alter the relationship between d_{ij} and I_{ij} which, since A_{ij} and d_{ij} are only weakly correlated, is not a spurious one. Consequently, $\Delta\bar{R}_i^{\,2}$ is large and positive while the change in β_i is minimal. The pattern of the changes in $\hat{\beta}_i$ discussed above was predicted theoretically in Section 6.7.

It is interesting to note that the spatial variations in $\bar{R}_i^{\,2}$ given in Figures 7.12 and 7.13 are reflected in the mean $\bar{R}_i^{\,2}$ for each model given in Table 7.2. For the unconstrained case, the competing destinations model has a much higher mean $\bar{R}_i^{\,2}$ value than the gravity model while for the two production-constrained cases, the values are very similar. The situation probably results from the fact that in the unconstrained competing destinations model, if A_{ij} is a poor explanatory variable of I_{ij}, $\hat{\delta}_i$ can be zero or near zero while in the production-constrained competing destinations models given in (7.7) and (7.10), no such freedom exists and $\hat{\delta}_i$ is set at -1.0. With hindsight, the constrained models should perhaps include an unknown parameter on the accessibility variable which can be estimated in the calibration of the models.

The goodness-of-fit statistics described here are somewhat misleading since they only reflect how well the calibrated models fit the data used to calibrate them. A much better goodness-of-fit test would be to see how well the calibrated models fit other data. In this respect, the competing destinations models should be superior to the original gravity models since the parameter estimates of the former vary less over space. It has long been a problem that existing spatial interaction models replicate the data used to calibrate them well but fail to replicate other data equally well (see, for example, Hyman and Wilson [1969], Taaffe and Gauthier [1973, p. 98], and Southworth [1980]). Such a problem is due to the large variation in parameter estimates which results from the mis-specification of the models. Since the competing destinations models are free from this mis-specification, they should perform much better in spatial systems other than the one in which they are calibrated.

7.9 A Comparison of the Calibration Results for the Unconstrained, Singly-Constrained and Doubly-Constrained Models

A set of estimated distance-decay parameters has been derived from the calibration of each of seven previously-defined interaction models (the doubly-constrained gravity and competing destinations models are treated as one model). Certain characteristics of each parameter set and the model from which it is derived are given in Table 7.2. The obvious comparison in the table is pairwise across the rows so that the superiority of the competing destinations models over the gravity models can easily be seen.

In the first column, the correlation coefficient between $\{\hat{\beta}_i\}$ and $\{MTL_i\}$, the set of origin mean trip lengths, is given. As mentioned in Section 1.5, the expected direction of this relationship is positive but it is often reported as being negative. From Table 7.2 it is clear that the negative relationships reported are a result of mis-specified gravity models being calibrated. When the competing destinations models in (6.11), (7.7), (7.10) and (6.16) are calibrated, the relationship is always positive,

TABLE 7.2

CHARACTERISTICS OF SEVEN $\{\hat{\beta}_i\}$'s: 100 x 100 MATRIX

Model Equation		Correlation Coefficients between $\{\hat{\beta}_i\}$ and: $\{MTL_i\}$	$\{A_i\}$	$\{m_i\}$	$\bar{\hat{\beta}}_i$	$\hat{\beta}$	$\bar{\bar{R}}_i^2$ $(\hat{\beta}_i)$	$\bar{\bar{R}}_i^2$ $(\hat{\beta})$	# of $\hat{\beta}_i$'s > -0.5 or < -2.0	$\hat{v}$	Z statistic for difference in $\hat{v}$
UC	(1.33)	-.49	.85	.05*	-0.92	n.c.	.62	n.c.	40	81.0	8.25
	(6.11)	.07*	.40	.02*	-1.31	n.c.	.71	n.c.	4	30.5	
SC $D_j = \sum_{i=1}^{m} I_{ij}$	(7.1)	.01*	.49	.28	-0.98	-0.74	.88	.81	17	43.7	3.2[illegible]
	(7.7)	.60	-.25	.26	-1.13	-0.92	.84	.79	5	31.[illegible]	
SC $D_j = m_j$	(7.4)	-.13*	.64	.24	-0.93	-0.71	.81	.73	25	52.2	5.16
	(7.10)	.47	-.01	.28	-1.07	-0.88	.80	.76	3	30.2	
DC	(1.36) (6.16)	.15*	.34	.32	-1.10	-0.85	.89	.83	4	37.5	

*Not significantly different from zero at the 99% confidence level.

n.c. Not calculated.

albeit insignificantly so in the case of (6.11) and (6.16).[1]

In the second column, the correlation coefficient between $\{\hat{\beta}_i\}$ and $\{A_i\}$, the set of origin accessibilities, is a measure of the spatial pattern exhibited by the parameter estimates. In Section 1.5, it was hypothesised that the strong spatial pattern generally exhibited by parameter estimates was an indication of $\hat{\beta}_i$ being a function of spatial structure. The hypothesis is supported by the evidence in Table 7.2: for the mis-specified gravity models, (1.33), (7.1) and (7.4), there is a strong positive relationship between $\{\hat{\beta}_i\}$ and $\{A_i\}$ which is significantly reduced when the equivalent competing destinations models are calibrated. The proportion of the variance of $\{\hat{\beta}_i\}$ explained by $\{A_i\}$ declines from .72 to .16 for the unconstrained models; from .25 to .06 for the singly-constrained models in which D_j is measured by $\sum_{i=1}^{m} I_{ij}$; and from .41 to .00 for the singly-constrained models in which D_j is measured by m_j. The proportion of the variance of $\{\hat{\beta}_i\}$ explained by $\{A_i\}$ for the doubly-constrained interaction models is only .11 which suggests that the models are correctly specified and the resulting parameter estimates are not strongly biased by spatial structure.

[1] In Table 7.2, the probability of a sample correlation coefficient being non-zero when the population correlation coefficient is zero is determined by a t-test where t is defined as:

$$t = \frac{r - \rho}{\sqrt{(1-r^2)/(n-2)}}$$

with n-2 degrees of freedom. The sample coefficient is denoted by r and ρ denotes the population coefficient which is assumed to be zero. A two-tailed test is used in all cases.

The third column represents the correlation ceofficient between $\{\hat{\beta}_i\}$ and $\{m_i\}$, the set of origin populations. Several authors, *inter alia* Olsson [1976], have hypothesised that the larger the population of an origin, the less negative will be the distance-decay parameter for that origin. An *a priori* hypothesis was that this relationship would be more evident when the effects of spatial structure are removed from the parameter estimates. However, there is very little difference in the coefficients between the relevant models. The results indicate that for all the constrained models, whether gravity or competing destinations models, there is a slight positive relationship between $\{\hat{\beta}_i\}$ and $\{A_i\}$, reinforcing Olsson's hypothesis, while for both unconstrained models, there is no relationship, refuting Olsson's hypothesis. It is not clear why this difference exists between the unconstrained and constrained model results.

The fourth column gives the mean $\hat{\beta}_i$ value for each model and the fifth column gives the single $\hat{\beta}$ value obtained if the models are calibrated for the whole system rather than being origin-specific. In all cases, a comparison of the mis-specified gravity model with the correctly-specified competing destinations model indicates that the overall spatial structure bias is a positive one. This is probably a result of there being many accessible origins in the system, the estimated distance-decay parameters of which are biased upwards by spatial structure. In the case of a 30 x 30 interaction matrix described later where most of the origins are inaccessible, the overall spatial structure bias is a negative one. The single parameter estimates for the whole system indicate that it is not only origin-specific parameters that are biased by spatial structure when obtained from the calibration of gravity models: the bias in the single parameter is also very large. For systems that have many accessible origins, the bias in the single distance-decay parameter estimate will be upwards: for systems that have many inaccessible origins, the bias will be downwards. In each case, the competing destinations models will give unbiased estimates.

The sixth and seventh columns are goodness-of-fit statistics for the models calibrated for each origin and for the models calibrated for the whole system, respectively. The former produces origin-specific parameter estimates while the latter produces a single estimate. In each case the goodness-of-fit statistic reported is the mean value of $\bar{R}_i^2$ since the emphasis is on how well the models predict interaction from each origin. As expected, the origin-specific models replicate existing outflows from each origin better than the general models. A comparison of the goodness-of-fit statistic for the equivalent unconstrained gravity and competing destinations models indicates that the latter replicates the existing data much more accurately than the former while the singly-constrained models are roughly equal in this respect. As mentioned earlier, this probably results from the relationship between I_{ij} and A_{ij} being reflected in an estimated parameter, $\hat{\delta}_i$, in the unconstrained case, whereas no such degree of freedom exists in the constrained cases. Also, how well the calibrated models replicate existing data is not the best test of a model's utility : the model needs to be able to predict interactions equally well in other spatial systems. Since the parameter estimates of the competing destinations models are less variable over space than those from the gravity models, the competing destinations models should be superior in this respect.

The last three columns refer to the variance of the parameters estimated from each model. It was hypothesised in Section 1.5 that the variance of parameter estimates obtained from gravity models is inflated because the estimates are functions of spatial structure and that when this relationship is eliminated, the variance of the parameter estimates will be reduced. The results in Table 7.2 substantiate this hypothesis. The coefficient of variation for each set of parameter estimates ($\hat{v}$) derived from a competing destinations model is significantly less than that for the estimates derived from the equivalent gravity model.[1] Another measure of the variation in parameter estimates is given by the number of $\hat{\beta}_i$ values which are greater than -0.5 or less than -2.0 in each model calibration, which again indicates that the sets of parameter estimates derived from the competing destinations models have fewer extreme values than those derived from the equivalent gravity models.

To demonstrate further the similarities and dissimilarities between the sets of estimated distance-decay parameters, a matrix of correlation coefficients between the parameter sets is given in Table 7.3.

---

[1] The Z statistic for testing the significance of the difference between two coefficients of variation was derived as follows. As Gregory [1963, p. 143] notes, the standard error of the coefficient of variation is $\hat{v}/\sqrt{2n}$, where n is the sample size, and the standard error of the difference between two sample coefficients is $\sqrt{(\hat{v}_1^2/2n_1) + (\hat{v}_2^2/2n_2)}$. Thus, a Z statistic defined as:

$$Z = \frac{\hat{v}_1 - \hat{v}_2}{\sqrt{(\hat{v}_1^2/2n_1) + (\hat{v}_2^2/2n_2)}}$$

is normally distributed and can be used to define the probability limits of the sample v's being derived from the same population v's. A two-tailed test was used in all cases.

TABLE 7.3

CORRELATION COEFFICIENTS BETWEEN SEVEN $\{\hat{\beta}_i\}$'s: 100 x 100 MATRIX

$\{\hat{\beta}_i\}$ DERIVED FROM MODEL*		UC (1.33)	UC (6.11)	SC $(D_j = \sum_{i=1}^{m} I_{ij})$ (7.1)	SC $(D_j = \sum_{i=1}^{m} I_{ij})$ (7.7)	SC $(D_j = m_j)$ (7.4)	SC $(D_j = m_j)$ (7.10)	DC (1.36) (6.16)
UC	(1.33)	-	.53	.73	.02*	.84	.25	.58
	(6.11)	.53	-	.52	.41	.54	.54	.51
SC $(D_j = \sum_{i=1}^{m} I_{ij})$	(7.1)	.73	.52	-	.63	.97	.77	.94
	(7.7)	.02*	.41	.63	-	.46	.94	.77
SC $(D_j = m_j)$	(7.4)	.84	.54	.97	.46	-	.67	.87
	(7.10)	.25	.54	.77	.94	.67	-	.86
DC	(1.36) (6.16)	.58	.51	.94	.77	.87	.86	-

*Not significantly different from zero at the 99% confidence level.

The sets of parameter estimates derived from the competing destinations models are similar to each other while being dissimilar to the sets of estimates derived from the gravity models. This is to be expected since the competing destinations models produce estimates of $\hat{\beta}_i$ which are comparatively free from spatial structure effects while the parameter estimates from the gravity models are determined to a large extent by spatial structure. However, it is worrying that some correlation coefficients, such as those between the unconstrained competing destinations model and the doubly-constrained model, are not larger. It appears that the estimated distance-decay parameter obtained for an origin is in part a function of the model used to estimate it and a problem then arises in determining which model gives the most accurate estimate of the true distance-decay parameter. Such a problem may result because model-specific spatial structure effects are greater for one model than another.

Finally, to demonstrate that the calibration results given for the 100 x 100 interaction matrix are not peculiar to these data, a 30 x 30 interaction matrix was constructed from the original matrix and each model was calibrated using these data. The thirty centres were chosen so that interactions would be modelled primarily between inaccessible centres (the 100 x 100 matrix is dominated by accessible centres). In this way, it is shown that the relevant accessibility measure used in the competing destinations model is still a measure of the accessibility of a destination to all other possible destinations and not simply to those used in the analysis. In the 100 x 100 interaction matrix, origins in the North-East have a high accessibility to all possible centres in the system (i.e., the U.S. city system) and to all other centres in the analysis. Origins in the West have a low accessibility to all possible centres in the system and to all other centres in the analysis. Thus, the effect of overall accessibility is difficult to distinguish from the effect of analysis-specific accessibility. The 30 centres chosen for the second analysis include 24 centres in the Western U.S.A. and 6 centres in the North-East. These centres are denoted by an asterisk in Table 7.1. The North-Eastern origins have low analysis-specific accessibilities but high general accessibilities. The Western origins have high analysis-specific accessibilities but low general accessibilities. The results of the gravity model calibrations on the data indicate that the $\hat{\beta}_i$ values have the same spatial pattern as described for the 100 x 100 interaction matrix: large negative $\hat{\beta}_i$ values are associated with Western origins and small negative or positive $\hat{\beta}_i$ values are associated with origins in the N.E. This indicates that in gravity models, $\hat{\beta}_i$ is determined by the general accessibility of an origin to all destinations, whether these destinations are in the analysis or not.

The summary results of the model calibrations are given in Tables 7.4 and 7.5. They are not discussed in detail since they simply reinforce the conclusions from Tables 7.2 and 7.3. However, a few differences in the two sets of tables are worth noting. The positive relationship between $\{MTL_i\}$ and $\{\hat{\beta}_i\}$ decreases when the competing destinations models are calibrated, as opposed to increasing in the 100 x 100 example. This is because for the 30 x 30 matrix, the relationship between $\{MTL_i\}$ and $\{\hat{\beta}_i\}$ is spurious due to a strong positive relationship between $\{MTL_i\}$ and $\{A_i\}$. When the relationship between $\{\hat{\beta}_i\}$ and $\{A_i\}$ is reduced in the competing destinations models, the simple relationship between $\{\hat{\beta}_i\}$ and $\{MTL_i\}$ is also reduced.

TABLE 7.4

CHARACTERISTICS OF SEVEN $\{\hat{\beta}_i\}$'s: 30 x 30 MATRIX

Model	Equation	Correlation Coefficients between $\{\hat{\beta}_i\}$ and: $\{MTL_i\}$	$\{A_i\}$	$\{m_i\}$	$\bar{\hat{\beta}}_i$	$\hat{\beta}$	$\bar{\bar{R}}_i^2$ $(\hat{\beta}_i)$	$\bar{\bar{R}}_i^2$ $(\hat{\beta})$	# of $\hat{\beta}_i$'s > -0.5 or < -2.0	$\hat{v}$	t-statistic for differences in $\hat{v}$[1]
UC	(1.33)	.67	.87	.17*	-1.33	n.c.	.71	n.c.	14	70.7	3.03
	(6.11)	.51	.56	.22*	-1.18	n.c.	.76	n.c.	8	51.9	
SC $D_j = \sum_{i=1}^{m} I_{ij}$	(7.1)	.54	.38*	.35*	-0.97	-0.82	.92	.83	4	41.3	0.50*
	(7.7)	-.02*	-.29*	.29*	-1.03	-0.82	.90	.84	4	43.4	
SC $D_j = m_j$	(7.4)	.67	.83	.24*	-1.10	-0.99	.87	.74	8	60.6	3.63
	(7.10)	.45	.53	.29*	-1.08	-0.91	.87	.81	4	41.7	
DC	(1.36) (6.16)	.71	.72	.35*	-1.07	-0.97	.94	.88	4	38.4	

*Not significantly different from zero at the 99% confidence level.

n.c. Not calculated

[1] A t-statistic is used since the sample size is small

TABLE 7.5

CORRELATION COEFFICIENTS BETWEEN SEVEN $\{\hat{\beta}_i\}$'s: 30 x 30 MATRIX

$\{\hat{\beta}_i\}$ DERIVED FROM MODEL:		UC (1.33)	UC (6.11)	SC ($D_j = \sum_{i=1}^{m} I_{ij}$) (7.1)	SC ($D_j = \sum_{i=1}^{m} I_{ij}$) (7.7)	SC ($D_j = m_j$) (7.4)	SC ($D_j = m_j$) (7.10)	DC (1.36) (6.16)
UC	(1.33)	–	.82	.52	-.13*	.96	.77	.83
	(6.11)	.82	–	.61	.22*	.81	.82	.77
SC ($D_j = \sum_{i=1}^{m} I_{ij}$)	(7.1)	.52	.61	–	.72	.65	.80	.87
	(7.7)	-.13*	.22*	.72	–	.03*	.44	.36*
SC ($D_j = m_j$)	(7.4)	.96	.81	.65	.03*	–	.87	.90
	(7.10)	.77	.82	.80	.44	.87	–	.88
DC	(1.36) (6.16)	.83	.77	.87	.36*	.90	.88	–

*Not significantly different from zero at the 99% confidence level.

Olsson's hypothesis that $\hat{\beta}_i$ is less negative for larger origins is not supported in any model calibration using the 30 x 30 interaction matrix. The correlation coefficient between $\{m_i\}$ and $\{\hat{\beta}_i\}$ is insignificant in every case.

As mentioned earlier, the bias in the parameter estimates estimated from the gravity models is reversed from the 100 x 100 matrix and is a downwards one since most of the origins are inaccessible. When the competing destinations models are calibrated, the mean $\hat{\beta}_i$ values and the single $\hat{\beta}_i$ values generally become less negative.

In Table 7.5, three correlation coefficients are unexpectedly low: those between $\{\hat{\beta}_i\}$ from (7.7) and the $\{\hat{\beta}_i\}$'s from (6.11), (7.10) and (6.16). It appears that the parameter estimates from the competing destinations model given in (7.7) bear little resemblance to those derived from the other competing destinations models. A partial explanation for this may be that the attractiveness variables, $\sum_{i=1}^{m} I_{ij}$ and m_j, are inversely related to each other for interactions in the 30 x 30 matrix. If destination j is located in the West, $\sum_{i=1}^{m} I_{ij}$ is probably higher than average but m_j is probably lower than average. If j is located in the North-East, $\sum_{i=1}^{m} I_{ij}$ is probably lower than average but m_j is probably higher than average. This inverse relationship between the attractiveness variables only occurs because the 30 x 30 interaction matrix is an unusual one in which very inaccessible centres predominate, but it does indicate the importance of correctly measuring the attractiveness variable.

7.10 The Relationship between $\{\hat{\sigma}_i\}$ and $\{\hat{\beta}_i\}$

The definition of the accessibility of a destination to all other destinations, as perceived by residents of origin i, is:

$$A_{ij} = \sum_{\substack{k=1 \\ (k\neq i,\ k\neq j)}}^{n-2} m_k d_{jk}^{\sigma_i} \quad , \tag{7.17}$$

where the notation is the same as that in Section 6.5. The preceding section demonstrated that the addition of this variable to gravity models improves the performance of these models. However, a problem in the calculation of A_{ij} is the determination of the parameter σ_i, which is a measure of how important distance is in determining the perception of accessibility by residents of origin i. Since it would be difficult to obtain accurate data on residents' perceptions of accessibility, σ_i has to be estimated. One estimation method is to assume that people's perception of distance as a deterrent to interaction is equal to their perception of distance in determining accessibility, that is, $\hat{\sigma}_i = \hat{\beta}_i$. Although both parameters are measures of the perception of distance, there is no theoretical justification for the two to be equal. A second approach is to assume that σ_i is equal for all origins and to estimate this single value from prior knowledge of the relationship between $\{A_{ij}\}$ and $\hat{\sigma}_i$. Both approaches were used to determine sets of accessibility measures and using each $\{A_{ij}\}$, the production-constrained competing destinations model given in (7.10) was calibrated. To indicate the relationship between $\{\hat{\sigma}_i\}$ and $\{\hat{\beta}_i\}$, a goodness-of-fit statistic $(\bar{\bar{R}}_i^2)$, a measure of the variance of the parameter estimates $(\hat{v})$, and the mean parameter estimate $(\bar{\hat{\beta}}_i)$ for each calibration are given in Table 7.6.

TABLE 7.6

THE EFFECT OF VARYING $\{\hat{\sigma}_i\}$ ON CALIBRATION RESULTS FOR (7.10)

$\hat{\sigma}_i$	$\bar{\bar{R}}_i^2$	$\hat{v}$	$\hat{\bar{\beta}}_i$
0.00 $\forall$ i[1]	.81	52.2	-0.93
-0.26 $\forall$ i	.82	46.9	-0.99
-0.59 $\forall$ i	.82	37.5	-1.03
-0.86 $\forall$ i	.81	33.6	-1.05
-1.00 $\forall$ i	.80	30.2	-1.07
-1.15 $\forall$ i	.78	31.6	-1.10
-1.33 $\forall$ i	.75	32.5	-1.12
-1.71 $\forall$ i	.68	36.6	-1.16
$\hat{\beta}_i$ $\forall$ i	.79	38.0	-1.09

[1] When $\hat{\sigma}_i = 0.00 \; \forall \; i$, the model is very similar to the equivalent gravity model given in (7.4) and the results given are for this latter model.

The results indicate that the best estimates of σ_i (in terms of maximising $\bar{\bar{R}}_i^2$ and minimising $\hat{v}$) are obtained by setting $\sigma_i = -1.0$ for all i. The goodness-of-fit of the model is reasonably high and the variance of the parameter estimates is at a minimum. When $\hat{\sigma}_i$ is set at a less negative value, the goodness-of-fit increases slightly but the variance of the parameter estimates also increases, making the model less useful for prediction. When $\hat{\sigma}_i$ is set at a more negative value, the goodness-of-fit decreases and the variance of the parameter estimates increases slightly. Allowing $\hat{\sigma}_i$ and $\hat{\beta}_i$ to be estimated iteratively does not appear to be a suitable method of obtaining accurate estimates of $\bar{\beta}_i$. Only two iterations were carried through, however, due to the time involved in the procedure.

The relationship between $\{\hat{\sigma}_i\}$ and $\{A_{ij}\}$ is demonstrated by Fotheringham [1979]. The more negative is $\hat{\sigma}_i$, the more peaked is the distribution of A_{ij} around large destinations. When $\hat{\sigma}_i$ is zero, the distribution of A_{ij} is virtually uniform. Thus, the relationship between I_{ij} and A_{ij} is best described when A_{ij} has a "moderately-peaked" distribution, that is when $\hat{\sigma}_i = -1.0$ for all i. This value of $\hat{\sigma}_i$ is very similar to the mean $\hat{\beta}_i$ values derived from the model calibrations although more empirical work is needed to determine whether this is a general result or one which is specific to this data set. What is known, is that the addition of an accessibility variable as defined in (7.17) relieves the mis-specification of gravity models. Thus, it is recommended that the accessibility variable be calibrated with $\hat{\sigma}_i = -1.0$ for all i and although this definition is somewhat subjective, it is no more subjective than that of the other variables present in the models.

7.11 Concluding Comments

The results of the model calibrations show conclusively that the competing destinations models given in (6.11), (7.7) and (7.10) are superior interaction models to the gravity models given in (1.33), (7.1) and (7.4). The latter models are mis-specified since they do not explicitly indicate that when a destination becomes more accessible to other destinations, it becomes less attractive for interaction, *ceteris paribus*. It was also shown that under certain conditions the doubly-constrained gravity model in (1.36) is correctly specified and equivalent to the doubly-constrained competing destinations model in (6.16). These conditions are that each centre is both an origin and a destination; the number of centres between which interaction is measured is large; and that σ_i is constant for all origins.

The mis-specification of gravity models causes the estimates of distance-decay parameters obtained from these models to be biased by spatial structure, so that they cannot be interpreted as measures of interaction behaviour. Accessible origins have parameter estimates that are biased upwards by spatial structure and inaccessible origins have parameter estimates biased downwards by spatial structure. Consequently, the set of origin-specific distance-decay parameters resulting from a gravity model calibration exhibits a strong spatial pattern and a large variance, both of which are spurious. When a correctly-specified set of interaction models, the set of competing destinations models, is calibrated, the spatial structure bias in parameter estimates is reduced and the estimates can be considered solely functions of interaction behaviour. As a result, the spatial pattern and variance of the origin-specific parameter estimates are negligible. Thus, the estimated distance-decay parameters from gravity models are behaviourally meaningless and are primarily indices of accessibility. The estimates from competing destinations models on the other hand can be interpreted as purely behavioural measures of the perception of distance as an impediment to interaction.

CHAPTER EIGHT

SUMMARY, IMPLICATIONS AND CONCLUSIONS

8.1 Summary

Three ways have been identified by which some measure of spatial structure can be included in the estimates of distance-decay parameters when these estimates are derived from the calibration of gravity models. In unconstrained gravity models, multicollinearity between the mass and distance variables results in estimates of distance-decay parameters which are not solely a measure of the relationship between distance and interaction but are also a measure of the relationship between mass and interaction. In singly and doubly-constrained gravity models, distance-decay parameter estimates are a function of the models' balancing factors which are in turn functions of accessibility. These two spatial structure effects are model-specific: the third spatial structure effect identified is a general one which applies to all gravity models. Unconstrained, singly-constrained and doubly-constrained gravity models are mis-specified since they do not measure the relationship between interaction and destination accessibility: as a destination becomes more accessible to all other destinations, it becomes less attractive for interaction, *ceteris paribus*. This mis-specification results in distance-decay parameter estimates that are strongly biased by spatial structure. Accessible origins have parameter estimates that are biased upwards by spatial structure and inaccessible origins have parameter estimates that are biased downwards by spatial structure. A new set of interaction models (competing destinations models) is derived in Chapter Six and a comparison of the calibration results from gravity and competing destinations models demonstrates that the latter models are correctly specified and have parameter estimates which are minimally biased by spatial structure.

The empirical analysis in Chapter Seven suggests that model mis-specification is the main reason why distance-decay parameters estimated from gravity models are biased by spatial structure. There is evidence, however, that the two model-specific spatial structure effects may also be important in some situations. The patterns of origin-specific parameters derived from the calibration of the competing destinations models indicate that accessible origins in the Mid-West and North-East have slightly less negative parameter estimates than other origins. These patterns may be a reflection of actual interaction behaviour with respect to distance, or they may be a result of model-specific spatial structure effects: severe multicollinearity between the distance and accessibility terms in the unconstrained competing destinations model and $\hat{\beta}_i$ being a function of the origin balancing factor in constrained models. Other evidence that these model-specific spatial structure effects are important is that the sets of parameter estimates from each of the competing destinations models are not always highly correlated with each other. It may be that model-specific spatial structure effects are more severe and parameter estimates more biased by spatial structure in one model calibration than in another.

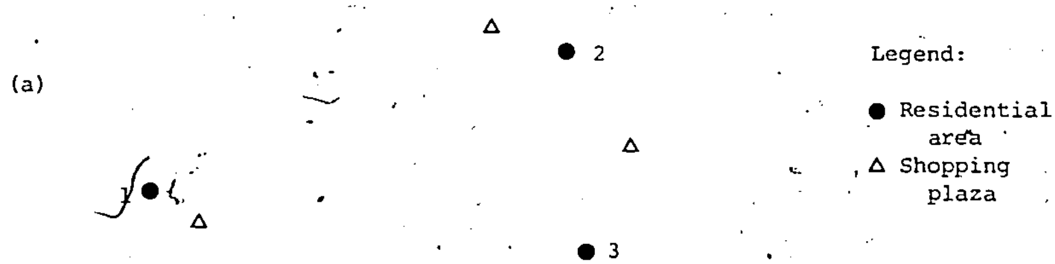
8.2 Implications

Estimates of distance-decay parameters obtained from present interaction models are primarily indices of accessibility and are behaviourally meaningless. This fact has several implications. The use of such parameters, for example, as descriptive statistics of interaction behaviour is misleading. If origin-specific distance-decay parameters are obtained by the calibration of a gravity model, little information is conveyed by these parameters that is not already known. The parameter estimates simply reflect which origins are accessible and which are inaccessible. The spatial variation of parameter estimates is a result of varying spatial structure and it has little or nothing to do with varying interaction behaviour. Empirical studies which discuss the spatial variation of parameter estimates in terms of varying interaction behaviour (see Section 1.5) are misleading. Except when spatial structure is constant, information can only be gained on the perception of distance as a deterrent to interaction when estimates of distance-decay parameters are obtained from the calibration of competing destinations models. Obviously, when spatial structure is constant, variations in gravity model estimates can be attributed to variations in behaviour. For example, when parameter estimates from different population cohorts are compared over the same spatial structure (Stillwell [1977, 1978]), or when parameter estimates for different modes of interaction are compared over the same spatial structure (Alcaly [1967]), variation in the estimates is a result of different perceptions of distance as a deterrent to interaction.

Studies which have compared estimates of distance-decay parameters through time are also suspect. Since spatial structure is likely to vary over time, distance-decay parameter estimates will vary even if the perception of distance as a deterrent to interaction remains constant. For example, a general finding from investigations of the temporal stability of distance-decay parameters is that $\hat{\beta}_i$ becomes less negative over time (*inter alia*, Hägerstrand [1957]). Since these estimates are derived from gravity models, it is impossible to determine whether the increase in $\hat{\beta}_i$ over time is due to increased accessibility or due to varying perceptions of distance.

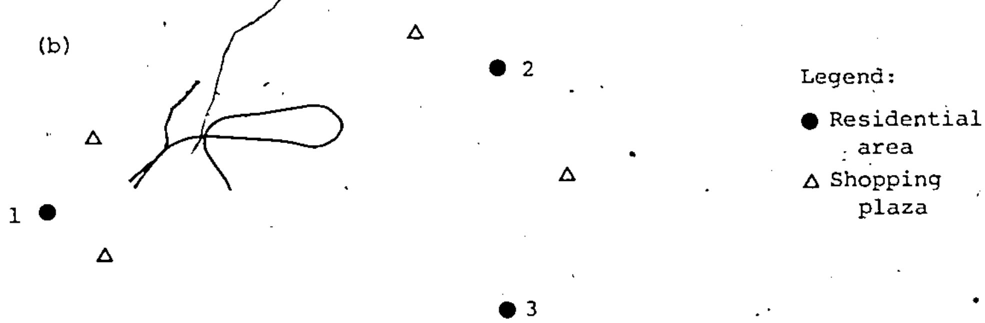
A major implication of the findings in this thesis concerns the prediction of interactions. Two types of prediction can be considered. One occurs when a spatial interaction model is calibrated on a spatial system where interaction data are known and then it is used to predict interactions in spatial systems where such data are not available: the change in spatial structure is a macro one. The second occurs when a spatial interaction model is calibrated in a spatial system and then is used to predict interactions in the same system given a change in the spatial structure of the system: the change in spatial structure is a micro one. The parameters of the calibrated model are assumed to remain constant in both cases but this thesis has shown that distance-decay parameters derived from gravity models will not be constant when there is a change in spatial structure. Consequently, there will be misleading predictions of interaction resulting from the use of gravity models or similar mis-specified interaction models. Southworth [1980] indicates the magnitude of the errors in prediction which result from mistakenly assuming parameter estimates to be constant when there is variation in spatial structure.[1] The errors in assuming constant distance-decay parameters when there are macro changes in spatial structure are evident in the results of the gravity model calibrations in Chapter Seven (for example, if Seattle's parameter estimate was assumed to equal Boston's). The errors in assuming constant distance-decay parameters when there are micro changes in spatial structure are also quite evident. Consider shopping trips in the spatial systems given in Figures 8.1a and 8.1b. In each spatial system, the perception of distance as a deterrent to interaction and the attractiveness of destinations are assumed to be constant over space.

FIGURE 8.1: Two Urban Spatial Systems



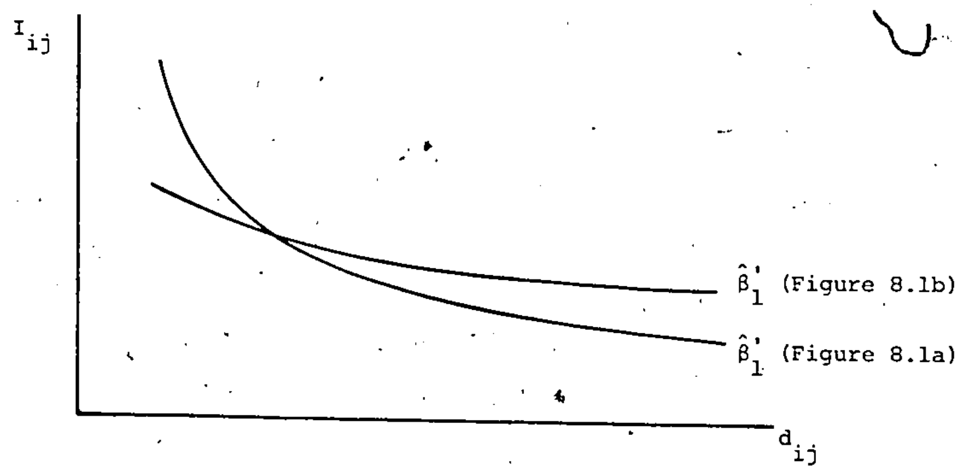
[1]Southworth's results concern the temporal instability of distance-decay parameter estimates. When 1962 estimates are used to forecast 1971 interactions, the errors in prediction are approximately two to three times larger than when 1971 estimates are used.

FIGURE 8.1: Two Urban Spatial Systems



In Figure 8.1a, the accessibilities of each residential area to the three shopping plazas (A_i) are ranked $A_2 > A_1 > A_3$. Following the analysis in Chapter Six, if an origin-specific gravity model were calibrated in this system, the distance-decay parameter estimates would be ranked $\hat{\beta}_2 > \hat{\beta}_1 > \hat{\beta}_3$. Suppose a new shopping plaza is located in close proximity to origin 1 so that the new origin accessibilities ($A_1{}'$) are ranked $A_1{}' > A_2{}' > A_3{}'$. The spatial structure of this system is given in Figure 8.1b. If a gravity model were calibrated for this system, the ranking of the distance-decay parameter estimates ($\beta_i{}'$) would be $\hat{\beta}_1{}' > \hat{\beta}_2{}' > \hat{\beta}_3{}'$ and thus predicting interactions using the original parameters $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ would give misleading results. Consider, for example, interactions from origin 1. The relationship between interaction and distance for the spatial systems given in Figures 8.1a and 8.1b is shown in Figure 8.2

FIGURE 8.2: Distance-Interaction Relationships for Origin 1 in Figure 8.1.



The distance-interaction relationship indicated by the gradient $\hat{\beta}_1$ is the prediction of the "true" distance-interaction relationship described by $\hat{\beta}_1'$. It is evident that if $\hat{\beta}_1$ is assumed to be a constant even though there is a change in spatial structure, misleading predictions result. Short distance interactions would be over-predicted while long distance interactions would be under-predicted. The two-stage decision-making behaviour described in Chapter Six indicates that no matter how many shopping plazas are located in close proximity to Origin 1, there will always be some residents of that origin who will shop in distance plazas. The remaining proportion of the residents who shop in close proximity will be divided between the increasing number of such opportunities and the number of interactions terminating at each specific destination will decrease.

A competing destinations model would reflect changes in spatial structure and their effects on interaction patterns through changes in the model's variables rather than through changes in the model's parameters. The parameters would remain constant since the behavioural relationship between interaction and distance remains constant - interaction patterns change solely because of changes in spatial structure. Consequently, the poor predictive performance of conventional spatial interaction models can be attributed to the spurious spatial variation of parameter estimates resulting from model mis-specification and not because of, as Hyman and Gleave [1976] conclude, the existence of intrinsically different spatial behaviour.

It is not only simple gravity models that will produce misleading predictions when $\hat{\beta}$ is assumed to be invariant to changes in spatial structure. Any multi-equation model which incorporates a gravity model will produce suspect results, and especially if the gravity model is used in an iterative procedure. An example of such a mis-use of gravity models is the Lowry model (Lowry [1964]) which forms the basis for most urban land use modelling (see, *inter alia*, Batty [1972]). In the Lowry model, $\hat{\beta}$ is predetermined and assumed to be constant. An attraction-constrained gravity model is used to predict the distribution of residential population within an urban system and a production-constrained gravity model is used to predict the distribution of retail employment. Both models are used iteratively until the convergence of each distribution is achieved. Consequently, the errors in assuming $\hat{\beta}$ to be invariant to the changes in spatial structure which are an integral part of the model, are compounded and this produces potentially large errors in prediction.

A final implication of these results concerns other types of spatial interaction modelling such as Rushton's revealed space preference (Rushton [1969]) and logit modelling (Domencich and McFadden [1975]). Any spatial interaction model which does not measure or account for the relationship between the accessibility of a destination and interaction to that destination will be mis-specified and will give suspect results. In the case of Rushton's work where all possible destinations are ranked in order to preference using the criteria of size and distance to the destination, two errors are made. The first is that individuals do not consider and rank all destinations. General locations, subjectively defined by the individual, are compared, and once a general location is chosen, specific destinations within that general area are considered and ranked in order of preference. The second, which is related, is that an important criterion in determining the attractiveness of a destination for interaction is the accessibility of the destination to all other destinations but this is not included in Rushton's analysis.

Similary, logit models which are used to analyse interaction patterns and which do not explicitly measure the relationship between interaction and destination accessibility are mis-specified. The relationship between distance and interaction is not measured accurately in such models since the distance-decay parameter estimate is biased by spatial structure in the same manner as gravity model parameter estimates.

8.3 Conclusions

Sayer [1976, p. 227] states, on the basis of Curry's [1972] earlier analysis, that:

> "parameters in gravity models are nothing more than abstract spatial statistics of very limited substantive meaning."

This statement is verified in several ways in this thesis and a new set of interaction models is derived whose parameter estimates do not suffer such criticism. The interpretation of distance-decay parameters is shown to be suspect when the parameters are estimated from any interaction model other than a competing destinations model. Competing destinations models are correctly specified interaction models since they take into account the relationship between the accessibility of a destination and the volume of interaction terminating at that destination. Distance-decay parameter estimates derived from competing destinations models are virtually invariant to changes in spatial structure, which is a property that makes the models very useful for prediction. Gravity models, on the other hand, are mis-specified interaction models since they ignore the relationship between destination accessibility and interaction. This mis-specification results in parameter estimates being a function of spatial structure and renders the models less useful for prediction. The calibration of any spatial interaction model which does not explicitly include a distance variable and a competing destinations variable, produces parameter estimates which are behaviourally meaningless and predictively useless.

APPENDIX I

THE DERIVATION OF EQUATIONS (2.39) and (2.40)

The simultaneous equation system given in (2.37) and (2.38) is:

$$I_{ij}^* = \alpha_i^* + \gamma_i m_j^* + \beta_i d_{ij}^* + \varepsilon_{ij}^* \quad (I.1)$$

$$m_j^* = \delta^* + \lambda \sum_{i=1}^{n} I_{ij}^* + \phi \sum_{i=1}^{n} m_i^* + \nu_j^* \quad . \quad (I.2)$$

Substituting (I.1) into (I.2) and expanding gives:

$$m_j^* = \delta^* + \lambda \sum_{i=1}^{n} \alpha_i^* + \lambda m_j \sum_{i=1}^{n} \gamma_i + \lambda \sum_{i=1}^{n} \beta_i d_{ij}^* + \lambda \sum_{i=1}^{n} \varepsilon_{ij}^*$$

$$+ \phi \sum_{i=1}^{n} m_i^* + \nu_j^* \quad . \quad (I.3)$$

Define,

$$\Theta \equiv \delta^* + \lambda \sum_{i=1}^{n} \alpha_i^* + \phi \sum_{i=1}^{n} m_i^* \quad , \quad (I.4)$$

and substitute into (I.3) and rearrange:

$$m_j^* (1 - \lambda \sum_{i=1}^{n} \gamma_i) = \Theta + \lambda \sum_{i=1}^{n} \beta_i d_{ij}^* + \lambda \sum_{i=1}^{n} \varepsilon_{ij}^* + \nu_j^* \quad , \quad (I.5)$$

Define,

$$\omega \equiv 1 - \lambda \sum_{i=1}^{n} \gamma_i \quad (I.6)$$

and substitute into (I.5) and rearrange:

$$m_j^* = \frac{\Theta}{\omega} + \frac{\lambda}{\omega} \sum_{i=1}^{n} \beta_i d_{ij}^* + \frac{\lambda}{\omega} \sum_{i=1} \varepsilon_{ij}^* + \frac{\nu_j^*}{\omega} \quad , \quad (I.7)$$

which can be rewritten as:

$$m_j^* = C + D \sum_{i=1}^{n} \beta_i \, d_{ij}^* + E_j \qquad \text{(I.8)}$$

which is (2.40). Substituting (I.8) into (I.1) gives (2.39).

APPENDIX II

THE DERIVATION OF EQUATION (5.2)

Equation (5.1) is written as:

$$\hat{I}_{ik} = \frac{O_i \, D_k \, d_{ik}^{\hat{\beta}_i}}{\sum_{j=1}^{n} D_j \, d_{ij}^{\hat{\beta}_i}} \qquad \text{(II.1)}$$

Define,

$$u \equiv O_i \, D_k \, d_{ik}^{\hat{\beta}_i} \qquad \text{(II.2)}$$

and then,

$$\frac{\partial u}{\partial d_{ik}} = \hat{\beta}_i \, O_i \, D_k \, d_{ik}^{\hat{\beta}_i - 1} \quad . \qquad \text{(II.3)}$$

Define,

$$v \equiv \sum_{j=1}^{n} D_j \, d_{ij}^{\hat{\beta}_i} \qquad \text{(II.4)}$$

and then,

$$\frac{\partial v}{\partial d_{ik}} = \hat{\beta}_i \, D_k \, d_{ik}^{\hat{\beta}_i - 1} \quad . \qquad \text{(II.5)}$$

Hence, by the quotient rule of derivatives,

$$\frac{\partial \hat{I}_{ik}}{\partial d_{ik}} = \frac{\sum_{j=1}^{n} D_j \, d_{ij}^{\hat{\beta}_i} \cdot \hat{\beta}_i \, O_i \, D_k \, d_{ik}^{\hat{\beta}_i - 1} - O_i \, D_k \, d_{ik}^{\hat{\beta}_i} \cdot \hat{\beta}_i \, D_k \, d_{ik}^{\hat{\beta}_i - 1}}{\left(\sum_{j=1}^{n} D_j \, d_{ij}^{\hat{\beta}_i}\right)^2} \, , \qquad \text{(II.6)}$$

which on rearranging the numerator gives (5.2)

APPENDIX III

THE DERIVATION OF EQUATION (5.6)

Equation (5.3) is written as:

$$\frac{\partial \hat{I}_{ik}}{\partial d_{ik}} = \frac{\hat{\beta}_i\, O_i\, D_k\, d_{ik}^{\hat{\beta}_i - 1} \left(\sum_{\substack{j=1 \\ j \neq k}}^{n} D_j\, d_{ij}^{\hat{\beta}_i} \right)}{\left(\sum_{j=1}^{n} D_j\, d_{ij}^{\hat{\beta}_i} \right)^2} \qquad \text{(III.1)}$$

Define,

$$C_i \equiv \hat{\beta}_i\, O_i\, D_k \sum_{\substack{j=1 \\ j \neq k}}^{n} D_j\, d_{ij}^{\hat{\beta}_i} , \qquad \text{(III.2)}$$

so that,

$$\frac{\partial \hat{I}_{ik}}{\partial d_{ik}} = \frac{C_i\, d_{ik}^{\hat{\beta}_i - 1}}{\left(\sum_{j=1}^{n} D_j\, d_{ij}^{\hat{\beta}_i} \right)^2} . \qquad \text{(III.3)}$$

Define,

$$u = C_i\, d_{ik}^{\hat{\beta}_i - 1} \qquad \text{(III.4)}$$

and then,

$$\frac{\partial u}{\partial d_{ik}} = (\hat{\beta}_i - 1)\, C_i\, d_{ik}^{\hat{\beta}_i - 2} . \qquad \text{(III.5)}$$

Define,

$$v = \left(\sum_{j=1}^{n} D_j\, d_{ij}^{\hat{\beta}_i} \right)^2 \qquad \text{(III.6)}$$

and then,

$$\frac{\partial v}{\partial d_{ik}} = 2\hat{\beta}_i\, D_k\, d_{ik}^{\hat{\beta}_i - 1} . \sum_{j=1}^{n} D_j\, d_{ij}^{\hat{\beta}_i} . \qquad \text{(III.7)}$$

Hence, by the quotient rule of derivatives,

$$\frac{\partial^2 \hat{I}_{ik}}{\partial d_{ik}^{\;2}} = \frac{(\sum_{j=1}^{n} D_j\ d_{ij}^{\hat{\beta}_i})^2 . (\hat{\beta}_i - 1) C_i\ d_{ik}^{\hat{\beta}_i - 2} - C_i\ d_{ik}^{\hat{\beta}_i - 1} . 2\hat{\beta}_i D_k\ d_{ik}^{\hat{\beta}_i - 1} . \sum_{j=1}^{n} D_j d_{ij}^{\hat{\beta}_i}}{(\sum_{j=1}^{n} D_j\ d_{ij}^{\hat{\beta}_i})^4} \tag{III.8}$$

$$= \frac{C_i [(\sum_{j=1}^{n} D_j\ d_{ij}^{\hat{\beta}_i}) . (\hat{\beta}_i - 1) d_{ik}^{\hat{\beta}_i - 2} - 2\hat{\beta}_i\ D_k\ d_{ik}^{2\hat{\beta}_i - 2}]}{(\sum_{j=1}^{n} D_j\ d_{ij}^{\hat{\beta}_i})^3} . \tag{III.9}$$

Now,

$$A_i \equiv \sum_{j=1}^{n} D_j\ d_{ij}^{\hat{\beta}_i} , \tag{III.10}$$

$$\hat{I}_{ik} \equiv \frac{O_i\ D_k\ d_{ik}^{\hat{\beta}_i}}{\sum_{j=1}^{n} D_j\ d_{ij}^{\hat{\beta}_i}} , \tag{III.11}$$

and,

$$C_i \equiv \hat{\beta}_i\ O_i\ D_k \sum_{\substack{j=1 \\ j \neq k}}^{n} D_j\ d_{ij}^{\hat{\beta}_i} \tag{III.12}$$

so that (III.9) can be rewritten as:

$$\frac{\partial^2 \hat{I}_{ik}}{\partial d_{ik}^{\;2}} = \frac{\hat{\beta}_i\ d_{ik}^{-2}\ \hat{I}_{ik} \sum_{\substack{j=1 \\ j \neq k}}^{n} D_j\ d_{ij}^{\hat{\beta}_i} [A_i (\hat{\beta}_i - 1) - 2\hat{\beta}_i\ D_k\ d_{ik}^{\hat{\beta}_i}]}{A_i^{\;2}} , \tag{III.13}$$

which is equivalent to,

$$\frac{\partial^2 \hat{I}_{ik}}{\partial d_{ik}^{\;2}} = \frac{\hat{\beta}_i \hat{I}_{ik}\ d_{ik}^{-2} (A_i - D_k\ d_{ik}^{\hat{\beta}_i}) (A_i\ \hat{\beta}_i - A_i - 2\hat{\beta}_i\ D_k\ d_{ik}^{\hat{\beta}_i})}{A_i^{\;2}} \tag{III.14}$$

and which, on dividing through by A_i^2 and rearranging, gives (5.6).

APPENDIX IV

THE NECESSARY AND SUFFICIENT CONDITIONS FOR THE SECOND DERIVATIVE OF PREDICTED INTERACTIONS WITH RESPECT TO DISTANCE BEING POSITIVE IN A PRODUCTION-CONSTRAINED INTERACTION MODEL

From (5.6) the relevant derivative is:

$$\frac{\partial^2 \hat{I}_{ik}}{\partial d_{ik}^2} = -\hat{\beta}_i \hat{I}_{ik} d_{ik}^{-2} \left(1 - \frac{D_k d_{ik}^{\hat{\beta}_i}}{A_i}\right)\left(1 - \hat{\beta}_i + 2\hat{\beta}_i \frac{D_k d_{ik}^{\hat{\beta}_i}}{A_i}\right) . \quad \text{(IV.1)}$$

Since, $\hat{I}_{ik} > 0$, $d_{ik}^{-2} > 0$, $-\hat{\beta}_i > 0$, $1 - \frac{D_k d_{ik}^{\hat{\beta}_i}}{A_i} > 0$, $\frac{\partial^2 \hat{I}_{ik}}{\partial d_{ik}^2} > 0$ iff,

$$\left(1 - \hat{\beta}_i + 2\hat{\beta}_i \frac{D_k d_{ik}^{\hat{\beta}_i}}{A_i}\right) > 0 , \quad \text{(IV.2)}$$

or equivalently,

$$1 - \hat{\beta}_i > -2\hat{\beta}_i \frac{D_i d_{ik}^{\hat{\beta}_i}}{A_i} . \quad \text{(IV.3)}$$

Define,

$$A_i \equiv \sum_{j=1}^{n} D_j d_{ij}^{\hat{\beta}_i} \quad \text{(IV.4)}$$

and then the inequality can be rewritten as:

$$(1 - \hat{\beta}_i)\left(\sum_{\substack{j=1\\j\neq k}}^{n} D_j\, d_{ij}^{\hat{\beta}_i} + D_k\, d_{ik}^{\hat{\beta}_i}\right) > -\, 2\hat{\beta}_i\, D_k\, d_{ik}^{\hat{\beta}_i} \qquad \text{(IV.5)}$$

or

$$(1 - \hat{\beta}_i) \sum_{\substack{j=1\\j\neq k}}^{n} D_j\, d_{ij}^{\hat{\beta}_i} > (-\, 1 - \hat{\beta}_i)\, D_k\, d_{ik}^{\hat{\beta}_i} \qquad \text{(IV.6)}$$

On transforming this becomes

$$(\hat{\beta}_i + 1) D_k\, d_{ik}^{\hat{\beta}_i} > (\hat{\beta}_i - 1) \sum_{\substack{j=1\\j\neq k}}^{n} D_j\, d_{ij}^{\hat{\beta}_i} \qquad \text{(IV.7)}$$

which is the condition given in Section 5.2.

Alternatively, if $(\hat{\beta}_i + 1)\, D_k\, d_{ik}^{\hat{\beta}_i} < (\hat{\beta}_i - 1) \sum_{\substack{j=1\\j\neq k}}^{n} D_j\, d_{ij}^{\hat{\beta}_i}$,

$\hat{\beta}_i$ must be greater than zero if $\dfrac{\partial^2 I_{ik}}{\partial d_{ik}^{\,2}} > 0$.

APPENDIX V

A DISCUSSION ON THE APPARENT INCONSISTENCY BETWEEN THE PREDICTION FROM A PRODUCTION-CONSTRAINED GRAVITY MODEL IN CHAPTER SIX AND THE RESULTS OF THE ANALYSIS IN CHAPTER FIVE

In Chapter Five it was shown that if $\hat{\beta}_i$ were a constant for all origins, the derivative of predicted interactions with respect to distance, in a production-constrained gravity model, would vary betweeen origins due to differences in accessibility. In Chapter Six, it was shown for the same model, that if $\hat{\beta}_i$ were a constant for all origins, the ratio of interaction with one destination to interaction with another would be constant regardless of the spatial structure of destinations. The two results appear contradictory although they are in fact complementary. The relationship between the two results can be seen by reference to an example.

In Figure 6.1, following the analysis of Chapter Five, let D_1 approximate an "average" destination and then the ratio of predicted interaction to D_1 when $d_{i1} = 1$, to predicted interaction to D_1 when $d_{i1} = 2$ would increase as i became more accessible. Following the analysis in Chapter Six, the ratio of predicted interaction terminating at a destination one unit from i, to that terminating at a destination two units away, would remain constant regardless of the accessibility of the origin. The two results are not incompatible. In the analysis of Chapter Five, when $d_{i1} = 2$, there is no longer a destination at one unit from i. In Chapter Six, when the volume of interaction terminating at a destination two units from i is measured, there is a destination one unit from i. The decrease in interaction to this destination would be greater in the spatial system given in Figure 6.1a where i is accessible since two other destinations are now closer to i than D_1. In Figure 6.1b, there would be a smaller decrease in the interaction to D_1 since it would still be, with D_2, the closest destination to i. Thus, the rate at which I_{i1} decreases as d_{i1} increases, is greater when i is accessible. This is the conclusion from the analysis in Chapter Five also.

```
*
*                    CALIBRATION OF AN ORIGIN-SPECIFIC,DOUBLY-
*                         CONSTRAINED SPATIAL INTERACTION MODEL
*
*                              A. STEWART FOTHERINGHAM
*
*INPUTS....INTERACTION MATRIX(INTFUL)
*..........DISTANCE MATRIX(RDIJ)
*..........ORIGIN-SPECIFIC MEAN TRIP LENGTHS(MTL)
*..........ORIGIN AND DESTINATION FLOW TOTALS(TOTALS)
*
*OUTPUTS....A(I),B(I),BETA(I),ORIGIN-SPECIFIC R**2'S
*
*THIS PROGRAM IS SET TO READ O(I)=SUMJTIJ AND
*                                   D(J)=SUMITIJ
*
*THE CALIBRATION PROCEDURE IS 1ST ORDER ITERATION. SEE BATTY
*+ MACKIE(E&P,1972). THIS IS THOUGHT TO BE A FASTER PROCEDURE
*THAN NEWTON-RAPHSON WHEN THERE IS A LARGE NUMBER OF ZONES.
*THE TIME TAKEN TO COMPUTE AN INVERSE MATRIX OF DERIVATIVES
*FOR THE N-R METHOD INCREASES RAPIDLY AS THE NUMBER OF
*ZONES INCREASES
*
*
        PROGRAM MAXLE(OUTPUT,MTL,TAPE6=MTL,TOTALS,TAPE7=TOTALS
       +,RDIJ,TAPE8=RDIJ,INTFUL,TAPE11=INTFUL,ACC,TAPE12=ACC)
*
*WORKING WITH 100*100 MATRIX
*
        DIMENSION PARA(100),SPRED(100),PRED(100),P(100),DERIV(100)
       +,IO(100),ID(100),SOBS(100),ERROR(100)
        DIMENSION INT(100),RINT(100),STAT(100)
        DIMENSION IACC(100)
        COMMON DIS(100,100),A(100),T(100),UU(100),WW(100),B(100)
        COMMON ACC(100)
*
*N IS THE NUMBER OF ORIGINS
*S IS A SENSITIVITY PARAMETER WHICH DETERMINES HOW CLOSE
*TO THE ACTUAL MTL WE GET
*LIMIT IS THE NUMBER OF ITERATIONS OF THE CALIBRATION PROCEDURE
*ASOBS IS THE AVERAGE MTL IN THE WHOLE SYSTEM. IT IS NOT
*NEEDED HERE BUT IS USED IN CALCULATING AN AVERAGE BETA
*FOR THE SYSTEM
```

```
*
      N=100
      S=.001
      LIMIT=50
      READ(6,1)ASOBS
    1  FORMAT(F12.5)
*
*SOBS ARE THE ORIGIN-SPECIFIC MTL'S
*
      READ(6,2)(SOBS(I),I=1,N)
    2  FORMAT(5F12.5)
*
* SET ALL THE B(J)'S TO 1.0 ON FIRST ITERATION
* STAT(I) IS THE DIFFERENCE BETWEEN ACTUAL AND PREDICTED MTL'S
* PARA(I) IS BETA(I)
*
      DO 10 I=1,N
      B(I)=1.0
      STAT(I)=1.0
      PARA(I)=1.0
*
* DIS(I,J) IS THE DISTANCE MATRIX
*
      READ(8,4)(DIS(I,J),J=1,N)
   10  CONTINUE
    4  FORMAT(16F5.0)
*
* IO(I) IS THE ROW FLOW TOTAL OF I
* ID(J) IS THE COLUMN FLOW TOTAL OF J
*
      READ(7,5)(IO(I),I=1,N)
      READ(7,5)(ID(J),J=1,N)
      DO 17 J=1,N
      READ(12,34) IACC(J)
       ACC(J)=FLOAT(IACC(J))
      ACC(J)=ACC(J)/10000.0
   34  FORMAT(I10)
*
* UU(J) IS NOW D(J)
* WW(I) IS NOW O(I)
*
      UU(J)=FLOAT(ID(J))
      WW(J)=FLOAT(IO(J))
      UU(J)=UU(J)/ACC(J)
   17  CONTINUE
    5  FORMAT(5I10)
*
* KK IS A COUNTER OF THE CALIBRATION ITERATIONS
*
      KK=0
  999  CONTINUE
      KK=KK+1
*
```

```
* THE SUBROUTINE MODEL CALCULATES A(I) AND B(J) AND THE
* PREDICTED MTL(I)(SPRED) GIVEN THE CURRENT INFORMATION
*
      CALL MODEL(N,S,STAT,PARA,SPRED)
*
* FIRST ORDER ITERATION TO DERIVE NEW BETA(I)'S
*
      DO 40 I=1,N
      STAT(I)=SOBS(I)-SPRED(I)
      PARA(I)=PARA(I)*(SPRED(I)/SOBS(I))
   40 CONTINUE
      IF(KK.LT.LIMIT)GO TO 999
      WRITE 880
  880 FORMAT(///3X,7H ORIGIN,8X,4HA(I),11X,4HB(J),8X,7HBETA(I),
     +8X,3HRSQ//)
*
* DERIVE THE ORIGIN-SPECIFIC R**2'S
      DO 700 I=1,N
      READ(11,702)(INT(J),J=1,N)
  702 FORMAT(10I8)
      RSUM=0.0
      RSUMI=0.0
      RMEAN=0.0
      TMEAN=0.0
      TSUMI=0.0
      KKK=0
      DO 703 J=1,N
      IF(DIS(I,J).LT.160.0) GO TO 703
      T(J)=A(I)*B(J)*WW(I)*UU(J)*DIS(I,J)**(-PARA(I))
      KKK=KKK+1
      RINT(J)=FLOAT(INT(J))
      RMEAN=RMEAN+RINT(J)
      TMEAN=TMEAN+T(J)
  703 CONTINUE
      TMEAN=TMEAN/KKK
      RMEAN=RMEAN/KKK
      DO 704 J=1,N
      IF(DIS(I,J).LT.160.0) GO TO 704
      RSUM=RSUM+(T(J)-TMEAN)*(RINT(J)-RMEAN)
      RSUMI=RSUMI+(RINT(J)-RMEAN)**2
      TSUMI=TSUMI+(T(J)-TMEAN)**2
  704 CONTINUE
      RSQ=(RSUM/(RSUMI*TSUMI)**.5)**2
      WRITE 890,I,A(I),B(I),PARA(I),RSQ
  890 FORMAT(6X,I3,3X,F14.10,3X,F14.2,3X,F9.5,4X,F5.3)
  700 CONTINUE
      STOP
      END
```

```
      SUBROUTINE MODEL(N,S,STAT,PARA,PRED)
      DIMENSION PARA(100),SPRED(100),PRED(100),P(100)
      DIMENSION STAT(100)
      COMMON DIS(100,100),A(100),T(100),UU(100),WW(100)
      COMMON B(100)
      COMMON ACC(100)
      SS=5.0
      AAA=0.0
  555 DO 717 I=1,N
      IF(ABS(STAT(I)).LT.S) GO TO 717
      B(I)=1.0
      A(I)=0.0
      DO 712 J=1,N
      IF(DIS(I,J).LT.160.0) GO TO712
      A(I)=A(I)+B(J)*UU(J)*DIS(I,J)**(-PARA(I))
  712 CONTINUE
      A(I)=1.0/A(I)
  717 CONTINUE
      DO 711 I=1,N
      IF(ABS(STAT(I)).LT.S) GO TO 711
      B(I)=0.0
      DO 713 J=1,N
      IF(DIS(I,J).LT.160.0) GO TO 713
      B(I)=B(I)+A(J)*WW(J)*DIS(I,J)**(-PARA(J))
  713 CONTINUE
      B(I)=ACC(I)/B(I)
  711 CONTINUE
      AAA=AAA+1.0
      IF(AAA.LT.SS) GO TO 555
      DO 1000 I=1,N
      IF(ABS(STAT(I)).LT.S) GO TO1000
      SUM=0.0
      SUMI=0.0
      DO 1030 J=1,N
      IF(DIS(I,J).LT.160.0)GO TO 1030
      T(J)=A(I)*B(J)*WW(I)*UU(J)*DIS(I,J)**(-PARA(I))
      DIS(I,J)=ALOG(DIS(I,J))
      SUM=SUM+T(J)
      SUMI=SUMI+T(J)*DIS(I,J)
      DIS(I,J)=EXP(DIS(I,J))
 1030 CONTINUE
      PRED(I)=SUMI/SUM
 1000 CONTINUE
      RETURN
      END
```

APPENDIX VII

```
*
*             CALIBRATION OF A SINGLY CONSTRAINED
*              COMPETING DESTINATIONS MODEL
*
*                 A.S.FOTHERINGHAM
*
* FOR FULL COMMENTS ON THIS PROGRAM REFER TO ACDUB OR NEWDUB
* PROGRAMS FOR CALIBRATING DOUBLY-CONSTRAINED MODELS. THE
*COMMENTS GIVEN HERE ARE          SPECIFIC TO THIS PROGRAM.
*
*
      PROGRAM MAXLE(OUTPUT,MTL,TAPE6=MTL,TOTALS,TAPE7=TOTALS
     +,RDIJ,TAPE8=RDIJ,POP,TAPE9=POP,ACC,TAPE10=ACC,INTFUL,
     +TAPE11=INTFUL)
      DIMENSION PARA(100),SPRED(100),PRED(100),P(100),DERIV(100)
     +,IO(100),ID(100),SOBS(100),ERROR(100)
      DIMENSION STAT(100)
      DIMENSION INT(100),RINT(100)
      DIMENSION IACC(100),ACC(100)
       COMMON DIS(100,100),A(100),T(100),UU(100),WW(100)
      N=100
      S=.001
* S IS THE CONVERGENCE LIMIT FOR THE PARAMETER ESTIMATES(
*I.E. DIFFERENCE BETWEEN NEW AND PREVIOUS ESTIMATE).
      D=160.0
* D IS THE ESTIMATED CUT-OFF DISTANCE BELOW WHICH AIRLINE
* INTERACTIONS ARE SUBSTITUTED BY OTHER FORMS OF INTERACTION.
      LIMIT=5
* LIMIT IS THE MAXIMUM NUMBER OF ITERATIONS OF THE NEWTON-
* RAPHSON CALIBRATION PROCEDURE-SEE BATTY AND MACKIE(E&P1972)
      READ(6,1)ASOBS
    1  FORMAT(F12.5)
      READ(6,2)(SOBS(I),I=1,N)
    2  FORMAT(5F12.5)
      DO 10 I=1,N
      STAT(I)=1.0
      PARA(I)=1.0
      READ(8,4)(DIS(I,J),J=1,N)
   10  CONTINUE
    4  FORMAT(16F5.0)
      READ(7,5)(IO(I),I=1,N)
      READ(9,6)(ID(J),J=1,N)
      READ(10,7)(IACC(J),J=1,N)
      DO 17 J=1,N
      UU(J)=FLOAT(ID(J))
```

```
      ACC(J)=FLOAT(IACC(J))
      UU(J)=(1/ACC(J))*UU(J)
      WW(J)=FLOAT(IO(J))
   17 CONTINUE
    5 FORMAT(5I10)
    6 FORMAT(8I10)
    7 FORMAT(I10)
      AINC=1.0/EXP(20.)
      KK=0
  999 CONTINUE
      KK=KK+1
      CALL MODEL(N,S,D,STAT,PARA,SPRED)
      DO 40 I=1,N
      STAT(I)=SOBS(I)-SPRED(I)
   40 CONTINUE
      DO 70 L=1,N
      P(L)=PARA(L)
      P(L)=P(L)+AINC
   70 CONTINUE
      CALL MODEL(N,S,D,STAT,P,PRED)
      DO 50 I=1,N
      IF(ABS(STAT(I)).LT.S) GO TO 50
      DERIV(I)=(SPRED(I)-PRED(I))/AINC
      DERIV(I)=1.0/DERIV(I)
   50 CONTINUE
      DO 80 I=1,N
      IF(ABS(STAT(I)).LT.S) GO TO 80
      ERROR(I)=0.0
      ERROR(I)=ERROR(I)+DERIV(I)*STAT(I)
      PARA(I)=PARA(I)-ERROR(I)
   80 CONTINUE
      IF(KK.LT.LIMIT)GO TO 999
      WRITE 880
  880 FORMAT(///5X,7H ORIGIN,10X,4HA(I),13X,7HBETA(I),10X,3HRSQ//)
      DO 700 I=1,N
      READ(11,702)(INT(J),J=1,N)
  702 FORMAT(10I8)
      RSUM=0.0
      RSUMI=0.0
      RMEAN=0.0
      TMEAN=0.0
      TSUMI=0.0
      KKK=0
      DO 703 J=1,N
      IF(DIS(I,J).LT.D) GO TO 703
      T(J)=A(I)*WW(I)*UU(J)*DIS(I,J)**(-PARA(I))
      KKK=KKK+1
      RINT(J)=FLOAT(INT(J))
```

```
      RINT(J)=FLOAT(INT(J))
      RMEAN=RMEAN+RINT(J)
      TMEAN=TMEAN+T(J)
  703 CONTINUE
      TMEAN=TMEAN/KKK
      RMEAN=RMEAN/KKK
      DO 704 J=1,N
      IF(DIS(I,J).LT.D) GO TO 704
      RSUM=RSUM+(T(J)-TMEAN)*(RINT(J)-RMEAN)
      RSUMI=RSUMI+(RINT(J)-RMEAN)**2
      TSUMI=TSUMI+(T(J)-TMEAN)**2
  704 CONTINUE
      RSQ=(RSUM/(RSUMI*TSUMI)**.5)**2
      WRITE 890,I,A(I),PARA(I),RSQ
  890 FORMAT(8X,I3,5X,F15.4,6X,F10.6,5X,F5.3)
  700 CONTINUE
      STOP
      END
      SUBROUTINE MODEL(N,S,D,STAT,PARA,PRED)
      DIMENSION PARA(100),SPRED(100),PRED(100),P(100)
      DIMENSION STAT(100)
      COMMON DIS(100,100),A(100),T(100),UU(100),WW(100)
      DO 1000 I=1,N
      IF(ABS(STAT(I)).LT.S) GO TO 1000
      A(I)=0.0
      DO 1010 J=1,100
      IF(DIS(I,J).LT.D) GO TO 1010
      A(I)=A(I)+UU(J)*DIS(I,J)**(-PARA(I))
 1010 CONTINUE
      A(I)=1.0/A(I)
      SUM=0.0
      SUMI=0.0
      DO 1030 J=1,100
      IF(DIS(I,J).LT.D) GO TO 1030
      T(J)=A(I)*WW(I)*UU(J)*DIS(I,J)**(-PARA(I))
      DIS(I,J)=ALOG(DIS(I,J))
      SUM=SUM+T(J)
      SUMI=SUMI+T(J)*DIS(I,J)
      DIS(I,J)=EXP(DIS(I,J))
 1030 CONTINUE
      PRED(I)=SUMI/SUM
 1000 CONTINUE
      RETURN
      END
```

APPENDIX VIII

THE DEGREE OF MULTICOLLINEARITY BETWEEN A_{ij}* AND d_{ij}*
FOR EACH ORIGIN IN THE 100 x 100 AIRLINE INTERACTION MATRIX

#[1]	$R_i(A_{ij}^*, d_{ij}^*)$	#	$R_i(A_{ij}^*, d_{ij}^*)$	#	$R_i(A_{ij}^*, d_{ij}^*)$	#	$R_i(A_{ij}^*, d_{ij}^*)$
1	-.90	26	-.85	51	.69	76	-.76
2	-.91	27	.33	52	-.73	77	-.86
3	.67	28	.25	53	-.09	78	-.94
4	-.89	29	-.82	54	.68	79	-.63
5	-.49	30	.55	55	-.71	80	.66
6	-.56	31	-.33	56	-.62	81	-.46
7	.32	32	-.86	57	-.28	82	.66
8	.66	33	-.58	58	-.31	83	.71
9	-.89	34	.67	59	-.69	84	.36
10	.01	35	-.94	60	-.46	85	.68
11	.18	36	-.60	61	-.17	86	.66
12	-.92	37	-.86	62	[illegible]	87	.61
13	-.36	38	-.79	[illegible]	[illegible]	88	.13
14	-.88	39	.68	[illegible]	[illegible]	89	-.75
15	-.89	40	-.80	65	[illegible]	[illegible]	.63
16	-.94	41	-.76	66	[illegible]	91	-.93
17	-.56	42	-.66	67	[illegible]23	92	-.25
18	-.84	43	-.91	68	-.15	93	-.84
19	-.70	44	-.89	69	-.39	94	.71
20	-.55	45	.23	70	-.17	95	.07
21	-.69	46	-.72	71	-.88	96	-.91
22	-.78	47	-.11	72	.71	97	-.89
23	-.90	48	-.37	73	-.92	98	-.37
24	.57	49	-.14	74	.64	99	-.12
25	-.61	50	-.64	75	-.89	100	-.93

[1]Refer to Table 7.1 for name of origin.

BIBLIOGRAPHY

Alcaly, R.E., 1967. "Aggregation and Gravity Models: Some Empirical Evidence", *Journal of Regional Science*, 7, 61-73.

Batty, M., 1972. "Recent Developments in Land Use Modelling: A Review of British Research", *Urban Studies*, 9, 155-177.

Batty, M. and S. Mackie, 1972. "The Calibration of Gravity, Entropy and Related Models of Spatial Interaction", *Environment and Planning*, 4, 205-233.

Beals, R.E., 1972. *Statistics for Economists*, Rand McNally: Chicago.

Beaumont, J., 1978. "Spatial Interaction Models and the Location-Allocation Problem", *WP240*, School of Geography, University of Leeds.

Black, W.R., 1971. "The Utility of the Gravity Model and Estimates of its Parameters in Commodity Flow Studies", *Proceedings of the Association of American Geographers*, 3, 28-32.

Black, W.R., 1972. "Inter-Regional Commodity Flows: Some Experiments with the Gravity Model", *Journal of Regional Science*, 12, 107-118.

Black, W.R. and R. Larson, 1972. "A Comparative Evaluation of Alternate Friction Factors in the Gravity Model", *The Professional Geographer*, 24, 335-337.

Britton, J.H., 1971. "The Interaction Model and Relative Location Variables", *Geografiska Annaler*, B, 53, 129-137.

Carroll, J.D., 1955. "Spatial Interaction and the Urban-Metropolitan Regional Description", *Regional Science Association (Papers and Proceedings)*, 1, D1-D14.

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

Cesario, F.J., 1975. "Linear and Non-linear Regression Models of Spatial Interaction", *Economic Geography*, 51, 69-77.

Chisholm, M. and P.O'Sullivan, 1973. *Freight Flows and Spatial Aspects of the British Economy*, Cambridge University Press: Cambridge, Mass.

Christ, C.F., 1966. *Econometric Models and Methods*, John Wiley and Sons, Inc.: New York.

Clark, G.L. and K.P. Ballard, 1979. "Modelling Out-Migration from Depressed Regions: The Significance of Origin and Destination Characteristics", *Discussion Paper D79-17*, Department of City and Regional Planning, Harvard University, Cambridge.

Clark, J.R., 1979. "Measuring the Flow of Goods with Archaelogical Data", *Economic Geography*, 55, 1-17.

Cliff, A.D., R.L. Martin and J.K. Ord, 1974. "Evaluating the Friction of Distance Parameter in Gravity Models", *Regional Studies*, 8, 281-286.

Cliff, A.D., R.L. Martin and J.K. Ord, 1975. "Map Pattern and Friction of Distance Parameters: Reply to Comments by R.J. Johnston, and by L. Curry, D.A. Griffith and E.S. Sheppard, *Regional Studies*, 9, 285-288.

Cliff, A.D., R.L. Martin and J.K. Ord, 1976. "A Reply to the Final Comment", *Regional Studies*, 10, 341-342.

Curry, L., 1972. "A Spatial Analysis of Gravity Flows", *Regional Studies*, 6, 131-147.

Curry, L., D.A. Griffith and E.S. Sheppard, 1975. "Those Gravity Parameters Again", *Regional Studies*, 9, 289-296.

Domencich, T.A. and D. McFadden, 1975. *Urban Travel Demand: A Behavioural Analysis*, North-Holland: Amsterdam.

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

Farrar, D.E. and R.R. Glauber, 1967. "Multicollinearity in Regression Analysis: The Problem Revisited", *Review of Economics and Statistics*, 49, 92-107.

Finney, N.D., 1972. "Trip Distribution Models", Chapter 3 in A.J. Catanese (ed.) *New Perspectives in Urban Transportation Research*, D.C. Heath: Lexington, Mass.

Fotheringham, A.S. 1979. "Polarised Growth within a Multi-Growth Centre Environment: A Case Study of the United States 1920-1970" *Environment and Planning*, 11, 193-208.

Fotheringham, A.S. and M.J. Webber, 1980. "Spatial Structure and the Parameters of Spatial Interaction Models", *Geographical Analysis*, 12, 33-46.

Frost, M., 1969. "Distribution Costs as a Factor in the Location of Industry Policy", *Discussion Paper 34*, Department of Geography, London School of Economics and Political Science.

Gordon, I.R., 1976. "Gravity Demand Functions, Accessibility and Regional Trade", *Regional Studies*, 10, 25-37.

Gordon, I.R., 1978. "Distance Deterrence and Commodity Values", *Environment and Planning*, 10, 889-900.

Gould, P., 1975. "Acquiring Spatial Information", *Economic Geography*, 51, 87-99.

Greenwood, M.J. and D. Sweetland, 1972. "The Determinants of Migration between Standard Metropolitan Statistical Areas", *Demography*, 9, 665-681.

Gregory, S. 1963. *Statistical Methods and the Geographer*, Longman: London.

Griffith, D.A. and K.G. Jones, 1980. "The Relationship between Spatial Structure and Spatial Interaction", *Environment and Planning*, 12, 187-201.

Hägerstrand, T., 1957. "Migration and Area" in *Migration in Sweden: A Symposium*, (eds.) D. Hannerberg, T. Hägerstrand and B. Odering. C.W.K. Gleerup: Lund, Lund Studies in Geography, Series B, 13, 27-158.

Haitovsky, Y., 1969. "Multicollinearity in Regression Analysis: A Comment", *Review of Economics and Statistics*, 51, 486-489.

Hansen, W.G., 1959. "How Accessibility Shapes Land Use", *Journal of the American Institute of Planners*, 25, 73-76.

Heanue, K.E. and C.E. Pyers, 1966. "A Comparative Evaluation of Trip Distribution Procedures", *Highway Research Record*, 114.

Hua, C. and F. Porell, 1979. "A Critical Review of the Development of the Gravity Model", *International Regional Science Review*, 4, 97-126.

Huang, D.S., 1970. *Regression and Econometric Methods*, John Wiley and Sons Inc.: New York.

Huff, D.L., 1963. "A Probabilistic Analysis of Shopping Center Trade Areas", *Land Economics*, 53, 81-90.

Hutchinson, B.G. and D.P. Smith, 1979. "Modelling Work Trip Distribution Patterns in Urban Ontario with Census Data", Ontario Ministry of Transportation and Communications, Research and Development Division, Final Report, Project No. 8510, TS-79-106.

Hyman, G.M.,1969. "The Calibration of Trip Distribution Models, *Environment and Planning*, 1, 105-112.

Hyman, G.M. and D. Gleave, 1976. "A Reasonable Theory of Migration", *CES-RP-22*, Centre for Environmental Studies, London.

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

Iklé, F.C., 1954. "Sociological Relationship of Traffic to Population and Distance", *Traffic Quarterly*, 8, 123-136.

Isard, W., 1960. *Methods of Regional Analysis: An Introduction to Regional Service*, The M.I.T. Press: Cambridge, Mass.

Johnston, R.J., 1973. "On Frictions of Distance and Regression Coefficients", *Area*, 5, 187-191.

Johnston, R.J., 1975. "Map Pattern and Friction of Distance Parameters: A Comment", *Regional Studies*, 9, 281-283.

Johnston, R.J., 1976. "On Regression Coefficients in Comparative Studies of the Friction of Distance", *Tijdschrifte voor Economishe en Sociale Geografie*, 67, 15-28.

Klein, L.R., 1962. *An Introduction to Econometrics*, Prentice-Hall: Englewood Cliffs.

Kullback, S., 1959. *Information Theory and Statistics*, John Wiley and Sons, Inc.: New York.

Leinback, T.R., 1973. "Distance, Information Flows and Modernisation: Some Observations from West Malaysia", *The Professional Geographer*, 25, 7-11.

Linneman, H.V., 1966. *An Econometric Study of International Trade Flows*, North-Holland Publishing Co.: Amsterdam.

Long, W.H., 1970. "The Economics of Air Travel Gravity Models", *Journal of Regional Science*, 10, 353-363.

Lowry, I.S., 1964. *A Model of Metropolis*, RM-4035-RC, RAND Corporation, Santa Monica.

Mackay, J.R., 1958. "The Interaction Hypothesis and Boundaries in Canada: A Preliminary Study", *Canadian Geographer*, 11, 1-8.

Macgill, S., 1975. "A Convergence Theorem for the Multipliers in Biproportional Matrices", *Working Paper 113*, Department of Geography, University of Leeds.

Merrill, W.C. and K.A. Fox, 1970. *Introduction to Economic Statistics*, John Wiley and Sons, Inc.: New York.

Olsson, G., 1965a. "Distance and Human Interaction: A Migration Study", *Geografisker Annaler*, Series B, 47, 3-43.

Olsson, G., 1965b. *Distance and Human Interaction - A Review and Bibliography*, Regional Science Research Institute (Philadelphia), University of Pennsylvania.

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

Olsson, G., 1976. *Birds in Egg*, Michigan Geographical Publication, University of Michigan, Department of Geography, Ann Arbor, Chapter 13, 455-458.

Openshaw, S., 1976. "An Empirical Study of Some Spatial Interaction Models", *Environment and Planning*, 8, 23-41.

Openshaw, S., 1977. "Optimal Zoning Systems for Spatial Interaction Models", *Environment and Planning*, 9, 169-184.

Openshaw, S. and C.J. Connolly, 1977. "Empirically Derived Deterrence Functions for Maximum Performance Spatial Interaction Models", *Environment and Planning*, 9, 1067-1079.

Penn-Jersey Transportation Study, Volumes I and II, Penjerdel: Philadelphia.

Pirie, G.H., 1979. "Measuring Accessibility: A Review and Proposal", *Environment and Planning*, 11, 299-312.

Porter, R., 1956. "Approach to Migration Through its Mechanism", *Geografiska Annaler*, B, 38, 317-343.

Porter, H., 1964. "Application of Intercity, Intervening Opportunity Models to Telephone, Migration and Highway Traffic Data", Northwestern University Department of Geography, Ph.D. dissertation.

Rushton, G., 1969. "Analysis of Spatial Behaviour by Revealed Space Preference", *Annals of the Association of American Geographers*, 59, 391-400.

Sayer, R.A., 1976. "A Critique of Urban Modelling", *Progress in Planning*, 6, 187-254.

Shannon, C.F., 1948. "A Mathematical Theory of Communication", *Bell System Technical Journal*, 27, 379-423, and 623-656.

Sheppard, E.S., 1979a. "Gravity Parameter Estimation", *Geographical Analysis*, 11, 120-132.

Sheppard, E.S., 1979b. "Geographic Potentials", *Annals of the Association of American Geographers*, 69, 438-447.

Sheppard, E.S., D.A. Griffith and L. Curry, 1976. "A Final Comment on Mis-specification and Autocorrelation in Those Gravity Parameters", *Regional Studies*, 10, 337-339.

Smith, D.A., 1973. Unpublished Study Reported in E.J. Taaffe and H.L. Gauthier, Jr. *Geography of Transportation*, Prentice-Hall Inc.: Englewood Cliffs, N.J.

Smith, D.P. and B.G. Hutchinson, 1978. "Work Trip Lengths and Spatial Structure in Canadian Urban Areas", Paper presented to the Canadian Association of Geographers.

Southworth, F., 1980. "Temporal Versus Other Effects on Spatial Interaction Model Parameter Values", Paper presented to the 76th Annual Meeting of the A.A.G., Louisville.

Stetzer, F., 1976. "Parameter Estimation for the Constrained Gravity Model: A Comparison of Six Methods", *Environment and Planning*, 8, 673-683.

Stillwell, J.C.H., 1977. "Some Historical Tests of Spatial Interaction Models for Inter-Area Migration: Part 1: Inter-Country Migration", *WP185*, School of Geography, University of Leeds.

Stillwell, J.C.H., 1978. "Interzonal Migration: Some Historical Tests of Spatial-Interaction Models", *Environment and Planning*, 10, 1187-1200.

Stouffer, S.A., 1940. "Intervening Opportunities: A Theory Relating Mobility to Distances", *American Sociological Review*, 5, 845-867.

Taaffe, E.J. and H.L. Gauthier, Jr., 1973. *Geography of Transportation*, Prentice-Hall: Englewood Cliffs, N.J.

Taylor, P.J., 1971. "Distance Transformation and Distance Decay Functions", *Geographical Analysis*, 3, 221-238.

Taylor, P.J., 1975. "Distance Decay in Spatial Interactions", *Concepts and Techniques in Modern Geography*, 2.

Thomas, R.W., 1977. "An Interpretation of the Journey-to-Work on Merseyside using Entropy-Maximising Methods", *Environment and Planning*, 9, 817-834.

Tomazinis, A.R., 1962. "A New Method of Trip Distribution in an Urban Area", *Highway Research Board Bulletin*, 347, 77-99.

Ullman, E.L., 1956. "The Role of Transportation and the Bases for Interaction", in W.L. Thomas (ed.) *Man's Role in Changing the Face of the Earth*, University of Chicago Press, Chicago, 862-880.

United States Civil Aeronautics Board, "Domestic Origin-Destination Survey of Airline Passenger Traffic, 1970, Summary Data". Compiled by the Bureau of Accounts and Statistics, Air Transport Association of America, Washington, D.C.[1]

Wilson, A.G., 1970a, "Advances and Problems in Distribution Modelling", *Transportation Research*, 4, 1-8.

Wilson, A.G., 1970b. *Entropy in Urban and Regional Modelling*, Pion Ltd.: London.

Wilson, A.G., 1971. "A Family of Spatial Interaction Models, and Associated Developments", *Environment and Planning*, 3, 1-32.

Wilson, S.R., 1976. "Statistical Notes on the Evaluation of Calibrated Gravity Models", Transportation Research 10, 343-345.

Young , W.J., 1975. "Distance Decay Values and Shopping Centre Size", *Professional Geographer*, 27, 304-309.

Yuill, R.S., 1967. "Spatial Behaviour of Retail Customers", *Geografiska Annaler*, B, 49, 105-114.

[1]Published by the Airline Finance and Accounting Conference, a division of the Air Transport Association of America, 1000 Connecticut Ave., N.W., Washington, D.C. 20036.