ENTROPY, INFORMATION THEORY AND SPATIAL INPUT-OUTPUT ANALYSIS

by

David Batten

Umeå Economic Studies No.92
UNIVERSITY OF UMEÅ 1981

ENTROPY, INFORMATION THEORY, AND SPATIAL INPUT-OUTPUT ANALYSIS

av

David F. Batten

AKADEMISK AVHANDLING

vid

Samhällsvetenskapliga fakulteten

Umeå universitet

Framlägges för avläggande av filosofie doktorsexamen till offentlig granskning i Humanisthuset, hörsal E, torsdagen den 18 juni 1981, kl 1400

Department of Economics
University of Umeå
S-901 87 Umeå, SWEDEN

Batten, D.F., Entropy, Information Theory, and Spatial Input-Output Analysis, Umeå Economic Studies No. 92, 1981, 303 pages.

ABSTRACT. Interindustry transactions recorded at a macro level are simply summations of commodity shipment decisions taken at a micro level. The resulting statistical problem is to obtain minimally biased estimates of commodity flow distributions at the disaggregated level, given various forms of aggregated information. This study demonstrates the application of the entropy-maximizing paradigm in its traditional form, together with recent adaptations emerging from information theory, to the area of spatial and non-spatial input-output analysis. A clear distinction between the behavioural and statistical aspects of entropy modelling is suggested. The discussion of non-spatial input-output analysis emphasizes the rectangular and dynamic extensions of Leontief's original model, and also outlines a scheme for simple aggregation, based on a criterion of minimum loss of information. In the chapters on spatial analysis, three complementary approaches to the estimation of interregional flows are proposed. Since the static formulations cannot provide an accurate picture of the gross interregional flows between any two sectors, Leontief's dynamic framework is adapted to the problem. The study concludes by describing a hierarchical system of models to analyse feasible paths of economic development over space and time.

KEYWORDS: Commodity flows, dynamic analysis, entropy, information theory, input-output analysis, probability distributions, regional and interregional modelling.

ENTROPY, INFORMATION THEORY
AND SPATIAL INPUT-OUTPUT ANALYSIS

A Thesis

Presented to the Department of Economics

of Umeå University

in Partial Fulfillment for the Degree of

Doctor of Philosophy

by

David F. Batten

June 1981

ENTROPY, INFORMATION THEORY,
AND SPATIAL INPUT-OUTPUT ANALYSIS

Abstract

A great deal of theoretical and empirical attention has been paid to input-output analysis, since the appearance of Leontief's original model. This study focusses primarily on a number of *empirical* difficulties, arising from the fact that interindustry transactions which are recorded at the aggregated, or *macro*, level are nothing more than broad reflections (or summations) of individual commodity shipment decisions taken at a disaggregated, or *micro*, level. The resulting statistical problem involves making efficient use of various forms of information (including published input-output tables) which are available at the macro level, to obtain unbiased estimates of commodity flow distributions at the micro level.

The major purpose of the study is to examine the potential application of the *entropy-maximizing paradigm* in its traditional form, together with more recent adaptations emerging from *information theory*, to the general area of spatial and non-spatial input-output analysis. A clear distinction is maintained between the *behavioural* and the *statistical* aspects of entropy. It is suggested that the entropy-maximizing approach to the estimation of interindustry flows can more fruitfully adopt the first principles of information theory, in preference to various microstate descriptions derived by analogy with statistical mechanics.

The discussion of non-spatial input-output analysis emphazises the *rectangular* and *dynamic* extensions of Leontief's original static model. The information content of an input-output table is also discussed, and an ordering scheme for simple aggregation, based on the criterion of *minimum loss of information*, is outlined. Biproportional matrix adjustments are also examined.

In the chapters on spatial analysis three complementary approaches to the estimation of interregional and intersectoral flows are proposed, using a limited database of industrial and multiregional information. Unfortunately, the *static* formulations cannot provide an accurate estimate of the *gross* interregional flows between any two sectors. Leontief's *dynamic* input-output framework is therefore adapted to the interregional estimation problem.

The study concludes with a demonstration of how the various coefficient estimates, derived in earlier chapters, can be used to analyse feasible paths of economic development over space and time. A *hierarchical* system of models is proposed, and particular attention is focussed on relationships between, and within, the national and regional levels. Although, this system disaggregates the development problem, it also permits an autonomous, self-assertive tendency within each region, to counterbalance the integrative forces in the nation as a whole. It is concluded that information theory can play an extremely useful, complementary role in the analysis of hierarchical social systems.

BIOGRAPHICAL SKETCH

David F. Batten was born in Melbourne on 3 March 1947. He attended the University of Melbourne from 1965 to 1969. After receiving his Bachelor of Engineering in 1969, he joined the systems research group at the Division of Building Research, Commonwealth Scientific and Industrial Research Organization. He commenced graduate studies in economics, operations research, and sociology at Monash University in 1972. In 1976, he was admitted to the degree of Master of Administration. During a period of collaborative research at the University of Gothenburg in 1978, he enrolled for the PhD program. In 1980, he transferred to Umeå University.

To Jennifer Fay Wundersitz

who may have incurred excessive opportunity costs throughout the writing of this dissertation.

ACKNOWLEDGMENTS

The lack of parametric information currently available in spatial input-output modelling is the major *raison d'être* for this dissertation. My interest in this subject was first aroused during a visit to Australia by Anders Karlqvist in 1975. Further stimulation emanated from a later visit by Åke Andersson. During a period of collaborative research at the University of Gothenburg in 1978, I enrolled for the doctoral degree. Late in 1980, I followed the chosen route of my professor and transferred to Umeå University.

Åke Andersson has fulfilled the multiple roles of my professor, thesis adviser and inspirational mentor during this work. His creative insights, and boundless energy and enthusiasm, have made dramatic impressions on my own thinking and writing. He has been ably supported in the supervision of the thesis by Börje Johansson, a close colleague at the University of Gothenburg, whose perceptive comments and forthright suggestions have contributed greatly to the refinement of the entire manuscript. I am grateful to Jan Eriksson for providing me with a working version of his entropy-maximizing algorithm. Among many other Swedish colleagues who have assisted me considerably, my particular thanks go to Lars Lundqvist, Bertil Marksjö, Håkan Persson and Tönu Puu.

I owe an enormous debt of gratitude to all my colleagues and friends at the CSIRO Division of Building Research in Melbourne. Special thanks are due to John Brotchie, Ray Toakley and Ron Sharpe for their sustained encouragement and creative leadership during my years at the Division. For valuable comments on specific chapters, I also wish to thank Ron, Paul Lesse and John Roy. I am particularly grateful to the latter for many of the original ideas underlying Section 4.3.2. Chris Tremelling ably carried out the computations described in Appendix E.

There are many other colleagues outside Sweden and Australia who have aided me in various ways. In particular, I would like to thank Peter Burridge, Rolf Funck, Leif Johansen, William Miernyk, and various journal referees.

I have also been fortunate to receive considerable financial and administrative assistance during the period of this research. For their generous support, I am grateful to the CSIRO Division of Building Research, the Swedish Council of Building Research and the Universities of Gothenburg and Umeå. The typing has been capably shared by Maureen Wishart, Julie Siedses and Silvia Diiorio (in Melbourne), together with Ann-Christine Söderlund and Ing-Marie Nilsson (in Umeå).

An investigation of this kind, stretched over a number of years, puts considerable strain on those closely associated with the researcher and his environment. During my lengthy stays in Sweden, I have been treated to generous hospitality by the Andersson, Johansson, Karlqvist, Lundqvist and Marksjö families. I am also indebted to my persevering parents, who have tolerated many inconveniences during my periods in and away from Melbourne. Finally, but foremost, my fondest thanks for her support, encouragement, and patience go to Jennifer Fay Wundersitz, who has fulfilled, albeit simultaneously, the demanding roles of wife, travelling companion, secretary, and friend.

An enigmatic salute is directed to Bacchus, without whom this dissertation might have been completed many years ago, or might never have begun!

David Batten

Umeå, June 1981

The gaps in the subject as treated here are too enormous for this work ever to be regarded as presenting my system, or even my theory. It is to be regarded rather as an interconnected group of suggestions which it is hoped will be of some practical use to critics and students of literature. Whatever is of no practical use to anybody is expendable.

Northrop Frye, *Anatomy of Criticism.*

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

A great deal of theoretical and empirical attention in economics has been devoted to the subject of input-output analysis since the appearance of Leontief's original national model.[1] There are several good reasons for this. Firstly, input-output analysis is a *theoretically simple* technique for recognizing the interdependent nature of an economic system. By grouping the productive activities of firms into various industries or sectors, it is possible to describe the overall balance between the supply of, and demand for, various products in terms of a simple set of linear equations.

Secondly, input-output tables provide a *practically appealing* means of representing economic interdependencies. Considerable effort has already been devoted in most developed countries to the task of either constructing or updating intersectoral transaction tables, thereby assuring the empirical tractability of input-output analysis at the national level.

Thirdly, input-output models are being adopted more frequently for *short to medium term economic forecasting*. Future levels of production can be predicted, given known or exogenously determined levels of final demand, by assuming constant technical (input-output) coefficients. Moreover, the multiplier principle permits the quantitative intersectoral effects of prescribed changes in the production levels of one or more sectors to be estimated directly.

Although much has also been written about the problems and weaknesses of input-output analysis from an economic viewpoint,[2] we shall not dwell on these theoretical problems at any length. Instead, we shall focus primarily on a number of empirical difficulties arising from the fact that inter-industry transactions which are recorded on the aggregate, or macro, level are nothing more than broad reflections (or summations) of individual commodity shipment decisions taken at a disaggregate, or micro, level.

1.1 The Problem and its Importance

The statistical problem which is central to this dissertation involves making efficient use of various forms of information (including published input-output tables) which are available on an aggregate level, to obtain unbiased estimates of commodity flow distributions on a disaggregate level. A simple example may help to clarify the nature of the problem. Consider a case where the national input-output table furnishes information about the flows between different sectors, but is incapable of providing any information about the geographical origins and destinations of these flows. In this rather common situation, we require some means of estimating the distribution pattern of commodities over space given their flow pattern between sectors. In other words, we wish to derive a full interregional input-output table by spatial disaggregation of the national table.

The obvious drawback to any survey-based development of an interregional input-output model, covering the whole national economy, is the considerable cost and effort involved in its empirical implementation. A simple form of interregional model can be derived given the availability of a single regional input-output table, and another for the same period relating to the remaining wider aggregate. While this type of model makes small demands for data, it inevitably understates the true extent of interregional *feedbacks* and *spillovers*. In a genuine interregional system, the basic requirement is that each component region should be treated equivalently and directly, leading normally to consideration of a large number of regions.

Recent research into the formulation of models describing the spatial distribution of goods and people within an *urban* environment has largely been characterized by the adaptation of theories based on the laws of large numbers and of probabilities. These theories originated in the physical world of interacting particles and gravitational force, and have since inspired the development of many *entropy-based* models of spatial interaction. More recently, the probabilistic methods evolving from a seemingly independent field, namely *information theory*, have been used for similar types of model building.

Since the initial adaptation of these statistical techniques to the field of urban distribution modelling by Wilson,[3] there have been many refinements and extensions to the basic methodology.[4] Common to many of these developments is the adoption of the *entropy-maximizing paradigm* to derive new model formulations for various spatial distributions. To illustrate the adaptation of this approach to our interregional estimation problem, we shall consider two seemingly different interpretations of the entropy concept.

Firstly, we can characterize the entropy function in its traditional form as a measure of the probability of a physical system of particles being found in a particular state. The entropy of the system is logarithmically proportional to the number of possible microstates which correspond, or give rise to, that particular macrostate. In elementary statistical mechanics, this view typecasts the entropy-maximizing procedure as the process of determining the most probable macrostate which corresponds to the largest number of microstates. The fundamental assumption inherent in this approach is that all microstates are equally probable.

The potential analogy between such physical assemblies containing large numbers of particles and the system components describing the spatial distribution of goods is readily demonstrated. Suppose we wish to estimate x_{ij}^{rs}, the shipment of commodities from sector i in region r for use in the production of other commodities by sector j in region s. If we know X_i^r, the total production by sector i in region r, we have

$$\sum_j \sum_s x_{ij}^{rs} = X_i^r \quad . \tag{1.1}$$

Adopting the assumption that each commodity unit is *distinguishable*, the number of ways in which X_i^r units can be distributed into ℓ (= m x n) groups, with x_{ij}^{rs} (j = 1,.....,n; s = 1,.....,m) commodities in each group, is given by the combinatorial formula

$$W_i^r = \frac{X_i^r!}{\prod_{js} x_{ij}^{rs}!} \quad . \tag{1.2}$$

Considering all regions and sectors simultaneously, the complete microstate description becomes

$$W = \prod_{ir} \left[\frac{X_i^r!}{\prod_{js} x_{ij}^{rs}!} \right] . \qquad (1.3)$$

We can then determine the most probable commodity distribution by maximizing W subject to a *known* system of constraints.[5]

The second and more recent interpretation of the entropy concept characterizes it as a measure of the amount of *uncertainty* or *lack of information* associated with a probability distribution. In theory, this approach necessitates a transformation of variables such as x_{ij}^{rs} into probabilities, wherein the elementary event is the shipment of commodities from sector i in region r to sector j in region s; p_{ij}^{rs} is the probability of such an event, and is defined by

$$p_{ij}^{rs} = \frac{x_{ij}^{rs}}{X} \qquad (1.4)$$

where $X = \sum_i \sum_j \sum_r \sum_s x_{ij}^{rs}$. The most probable commodity distribution is then determined by maximizing the entropy function in the form

$$S = - \sum_i \sum_j \sum_r \sum_s p_{ij}^{rs} \log p_{ij}^{rs} \qquad (1.5)$$

subject to a set of constraints containing whatever flow information is available.

This information-theoretical approach provides a constructive criterion for estimating probability distributions on the basis of partial knowledge, and characterizes the maximum-entropy estimate as a type of *statistical inference*. It is simply the least biased estimate possible with the given information; and is maximally non-committal with regard to missing information.[6]

If we consider statistical mechanics to be a form of statistical inference, instead of just a purely physical theory, these two seemingly different views of entropy are essentially reconcilable. It has consequently been argued that the perspective which should be adopted may simply be a matter of *taste*.[7] On the contrary, we believe that it is both useful and important to maintain a clear distinction between the *behavioural* and the *statistical* aspects of entropy. The former consists of the correct enumeration of the feasible states of the system, whereas the latter is a straight-forward example of statistical inference.

1.2 Scope of the Present Investigation

The author is not currently aware of any comprehensive attempts to straddle the *general* area of input-output analysis and information-theoretical techniques. Wilson's early work included some interesting speculation, whereas Theil's initial research focussed mainly on information decomposition, and on the use of entropy as a measure of the information content of an input-output table.[8] A number of interesting papers have emerged during the seventies, but each has treated rather specific examples of commodity movements, in preference to a more general theoretical investigation of the area. We shall discuss these recent developments in the appropriate chapters.

The main purpose of this dissertation is to examine the potential application of the entropy-maximizing paradigm in its traditional form, together with more recent adaptations emerging from information theory, to the general area of spatial and non-spatial input-output analysis. A fundamental decision, which will be made as early as possible in the dissertation, is whether the entropy-maximizing approach to the estimation of interindustry flows may more fruitfully adopt the first principles of information theory, in preference to various microstate descriptions derived by analogy with statistical mechanics. There is clearly a stage beyond which such analogies can become misleading or inappropriate. This point may be reached when we enter the multisectoral world of input-output analysis. The rather restrictive assumption that all microstates are equally probable may prevent any of the traditional statistical distributions from reproducing empirical flows accurately.

The author's intellectual debt to various social scientists will become evident as the dissertation unfolds. To an eminent economist, Wassily Leontief, goes the credit for establishing input-output analysis as such a valuable tool for the investigation of economic interdependencies. To an eminent geographer, Alan Wilson, goes the credit for introducing the entropy-maximizing technique into the world of spatial analysis and model-building. Many other notable scientists, such as Walter Isard and Henri Theil, have made significant contributions to our multidisciplinary area of investigation. We shall certainly attempt to acknowledge them all within the following pages.

1.3 Plan of Each Chapter

In Chapter 2, we begin by examining the course of theoretical ideas which has led to the suggested nexus between the physical concept of entropy, and measures of uncertainty and information. Previous applications of the entropy concept, and related measures of information, to the analysis of spatial and economic phenomena are selectively reviewed. Particular attention is paid to earlier analyses of commodity movements, and the estimation of interindustry flow coefficients.

Chapter 3 explores the potential similarities between various statistical representations of physical systems, and the system components describing the spatial distribution of a single commodity. Quite different microstate descriptions can be derived, depending upon whether each commodity unit is regarded as identical or distinguishable. The chapter concludes by recommending that entropy-maximizing approaches to the estimation of inter-industry distributions should adopt the first principles of information theory, in preference to various microstate descriptions derived by analogy with statistical mechanics.

Whereas Chapter 3 considers each commodity in isolation, Chapter 4 demonstates how the flow patterns of various commodities can be linked together by subdividing the economy into a system of mutually inter-dependent industries. After examining Leontief's classic formulation, it is apparent that some fundamental problems exist. In particular, it is clear that more realistic representations of economic interdependencies may be developed by adopting the rectangular and dynamic extensions of his original static model. A major part of this chapter is therefore devoted to a demonstration of the use of various measures and methods, based on information theory, for the estimation of key parameters in these and other input-output models.

Throughout Chapter 4, the discussion is essentially non-spatial. The marriage of certain concepts derived in this chapter, with related work on spatial analysis, is celebrated in Chapters 5 and 6. To clarify the discussion in these chapters, some fundamental distinctions between *popular* terms are necessary. Although they may not be appropriate on all occasions, the following definitions are adopted for the purposes of this dissertation:

REGIONAL - a *general* term referring to the behaviour of a *single* region, with no detailed distinction between the internal and external flow relationships.

INTRAREGIONAL - a *specific* term referring to the behaviour inside a *single* region, with a detailed focus on the *internal* flow relationships.

MULTIREGIONAL - a *general* term referring to the behaviour of a *group* of regions, with no detailed distinctions between the internal and external flow relationships.

INTERREGIONAL - a *specific* term referring to the behaviour of a *group* of regions, with a detailed focus on the flow relationship *between* each pair of regions.

Chapter 5 begins with an examination of various non-survey techniques, both *intra-* and *inter-regional*, which have been adopted for the spatial estimation of intersectoral flow coefficients. It is concluded that no acceptable non-survey method for deriving intraregional coefficients from their national counterparts has been published. We then propose three different information-theoretical approaches to the estimation of interregional and intersectoral flows, using a limited database of industrial and multiregional information.

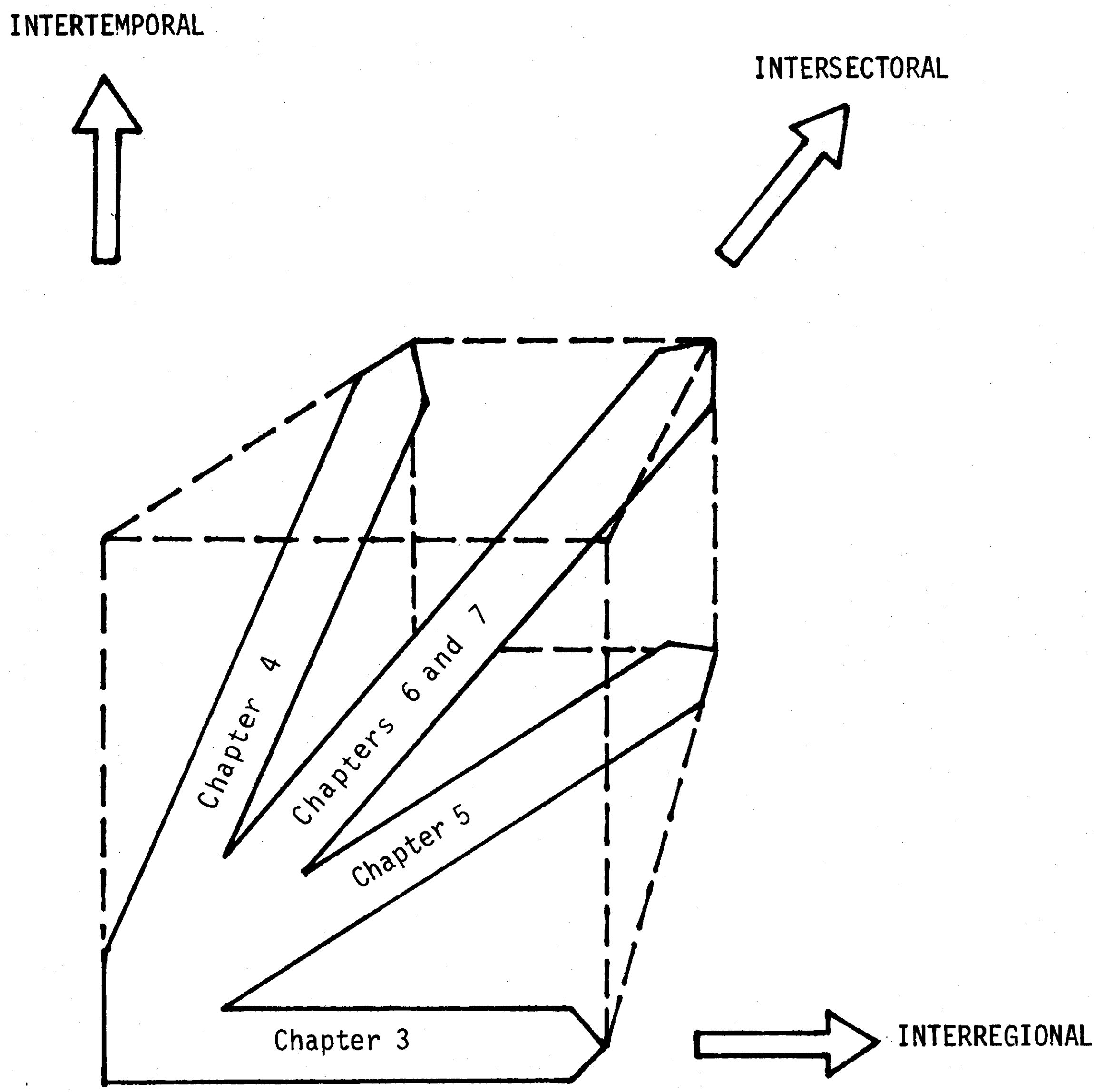
Unfortunately, none of these static formulations is capable of providing an accurate estimate of the *gross* interregional flows between any two sectors. Since the exact distribution of capital flows is unknown, they are also unable to analyse the repercussions of regional and industrial growth or decline. In Chapter 6, Leontief's dynamic input-output model is adapted to the interregional estimation problem. Through the use of a simple accelerator principle, a clear distinction can be made between the intermediate flows, which are described by the usual input-output coefficients, and productive capital flows, which are specified by an interregional matrix of capital coefficients. Thus the analysis in Chapter 6 is formulated in terms of coefficients, in contrast to the flow estimates described in Chapter 5.

The final chapter demonstrates how the various coefficient estimates, derived in Chapter 6 and earlier chapters, could be used to analyse feasible paths of economic development over space and time. In particular, a search is made for a plausible system of models to integrate national and regional development. The exploration begins by reviewing some existing approaches, which have adopted a dynamic interregional framework of the interindustry type, and concludes by suggesting a *hierarchical* system of models. Although this system disaggregates the development problem, it also permits an autonomous, self-assertive tendency within each region, to counterbalance the integrative forces in the nation as a whole.

As we move down our hierarchy, from the national to the regional level, we progress to a subsystem in which behaviour is more spatially disaggregated. In doing so, we face an increasingly difficult data problem: that of making efficient use of the information which is available at the national level, to coordinate, but not completely control, the pattern of behaviour within and between each region. But this is, in fact, the very information problem which is central to this dissertation. Clearly, information theory can play an extremely useful, complementary role in the development of our hierarchical modelling system.

As a guide to the nature of each chapter, Figure 1.1 depicts their relative positions in a three-dimensional system of spatial, industrial and temporal coordinates.

Figure 1.1


INTERTEMPORAL
INTERSECTORAL
Chapter 4
Chapters 6 and 7
Chapter 5
Chapter 3
INTERREGIONAL

FOOTNOTES FOR CHAPTER 1

1 See Leontief (1951; 1953).

2 See, for example, Richardson (1972).

3 See Wilson (1967; 1970 a, b).

4 See, for example, Fisk and Brown (1975 b), Snickars and Weibull (1977), and Lesse et al. (1978).

5 It is usually more convenient to maximize log W rather than W.

6 See Jaynes (1957, p 620).

7 See, for example, Williams and Wilson (1979).

8 See Wilson (1970 a, b) and Theil (1967).

Chapter 2

A REVIEW OF ENTROPY AND INFORMATION THEORY

2.1 Introduction

Despite its long history, which stretches back in excess of one hundred years, to many the term *entropy* still appears esoteric. In the early days of classical thermodynamics, perhaps it was; even though its original meaning was grounded in a bedrock of physical facts. But nowadays it is becoming increasingly popular in one field after another. To some extent, these more recent adaptations are related in a purely formal way to a simple algebraic formula which is the cloak under which entropy grows more familiar to an increasing number of social scientists.[1]

The purpose of the present chapter is twofold:

(i) to examine the course of theoretical ideas which has led to the suggested nexus between the entropy concept of the physicists and measures of uncertainty or information; and

(ii) to selectively review previous applications of the entropy concept and measures of information in spatial and economic analyses.

In section 2.2, we begin by defining the original thermodynamic concept of entropy, and then demonstrate how the emergence of statistical mechanics heralded an important redefinition of entropy as a measure of the *degree of disorder* existing within a system. At this stage, the concepts of *macrostate* and *microstate* are introduced. We then examine the course of ideas which has led to the suggestion that statistical entropy is equivalent to a probabilistic measure of uncertainty.

As early as 1928, some scientists argued that the entropy of a system constitutes an index of our *degree of ignorance* about the microstructure of that system. But another twenty years were to pass before the same entropy measure derived in statistical mechanics was formally proposed as a measure of the *degree of uncertainty*, or *missing information*, in a probability distribution.

Much more recently, the rather restrictive views of entropy adopted in practice have been widely challenged. Some authors have sought to establish other microstate descriptions, in most cases by analogy with statistical mechanics. Other changes have been prompted by the desire to incorporate *a priori* information (in the form of non-uniform prior probabilities) into measures of uncertainty. Some of these current issues are discussed near the end of Section 2.2.

Not long after Shannon's classic work on information theory,[2] the terms entropy and information were becoming magic words in a variety of disciplines. Towards the middle sixties, they made their first few appearances in the literatures of economists and geographers. Section 2.3 traces the ensuing applications in three fields, namely economics, spatial analysis, and their interdisciplinary focus, known as spatial economics.

In these three fields, we find that measures of uncertainty or information have been used for two quite different purposes. On the one hand, there are those who use information measures as a descriptive statistic to provide a statistical summary of a distribution. In these applications, information statistics are employed to measure the dividedness or degree of concentration or specialization existing in a distribution. On the other hand, there is the work pioneered by Wilson, which is largely concerned with generating the most likely probability distribution given a certain set of constraints. This body of work, using the entropy maximizing paradigm, demonstrates the importance of information theory as a flexible tool of estimation or statistical inference for practical model-building.

Although these two procedures appear somewhat different, when seen as steps in the process of generating and testing hypotheses, they can in fact be reconciled.[3] However, for the majority of the work undertaken in this thesis, the use of information theory as a versatile vehicle for empirical estimation will be emphasized.

2.2 Theory

2.2.1. Thermodynamics and Statistical Mechanics

The entropy concept emerged from a memoir of Sadi Carnot dealing with the efficiency of steam engines.[4] At the time, it was disguised under the name *calorique*. By 1865, Rudolf Clausius was able to give to the first two laws of thermodynamics their classic formulation:

> "The *energy* of the universe remains constant;
> the *entropy* of the universe at all times moves
> towards a maximum".[5]

Entropy was defined as a ratio relating the quantity of heat exchanged isothermally in a heat engine to the absolute temperature during the exchange. Clausius seems to have understood the importance of recognizing the evolutionary nature of the entropic process, for he coined the word entropy from a Greek word, equivalent in meaning to *evolution* (τροπη), by adding the prefix en- to resemble energy.

It was quite difficult not only for physicists but also for other men of science to reconcile themselves to the blow inflicted on the supremacy of mechanics by the science of heat. No wonder then that ever since thermodynamics appeared on the scene, physicists fervently strove to reduce heat phenomenon to locomotion. The result was a new thermodynamics, better known by the name of statistical mechanics. Within this new theoretical framework, entropy came to be redefined as a measure of the *degree of disorder* existing within a system.[6]

Statistical mechanics circumvents the difficulty of actually defining disorder by means of two basic principles:

(a) The disorder of a *microstate* is ordinally measured by that of the corresponding *macrostate*.

(b) The disorder of a macrostate is proportional to the *number* of corresponding microstates.[7]

A microstate is a state the description of which requires that each individual particle be identifiable. A macrostate corresponds to a group of microstates. The degree of disorder, computed according to rule (b), depends on the manner in which microstates are grouped into macrostates. Since statistical thermodynamics is concerned only with the mechanical co-ordinates of particles, all particles are treated as qualityless individuals distinguishable only by their names. The concept of macrostate, in which no particle names are used, corresponds to the obvious fact that the physical properties of an assembly of particles do not depend on which particle occupies a certain *state*. Each arrangement of particles in a given macrostate constitutes a microstate. However, the criterion according to which two such arrangements constitute two *different* microstates is an additional convention which varies from one approach to another. So does the criterion for what constitutes an acceptable macrostate.

In the earliest but still the basic approach,[8] two arrangements constitute two different microstates if and only if the names of the particles in some state(s) are not the same. Furthermore, no restriction is imposed upon the number of particles having the same state. In general, if there are m states and N particles, the number of microstates, W, corresponding to each macrostate $(N_1, N_2, \ldots, N_m)$, $\Sigma N_i = N$, is given by the familiar formula of combinatorial calculus

$$W = \frac{N!}{N_1!N_2!\ldots\ldots N_m!} \quad . \tag{2.1}$$

Boltzmann's famous formula for entropy, S, viewed as a measure of disorder is

$$S = k \log W \tag{2.2}$$

where log W is the natural logarithm of W and k is a *physical* constant known as Boltzmann's constant. For *large values* of N_i, we can use Stirling's approximation

$$\log N! \cong N \log N - N \tag{2.3}$$

to estimate the factorial terms in (2.1). We have

$$\log W = - \sum_i N_i \log (N_i/N) . \tag{2.4}$$

We can now rewrite (2.2) as follows:

$$S = - kNH \tag{2.5}$$

where

$$H = \sum_i p_i \log p_i$$

and (2.6)

$$p_i = N_i/N .$$

H is the function adopted by Boltzmann to formulate his famous H-theorem.[9] Clearly, -kH represents the average entropy per particle. For later reference, it is worth noting that H and S vary in opposite directions.

The H-theorem provides a bridge between the phenomenological investigations of Carnot and Clausius, and the atomistic views underlying the kinetic theory of gases. The proof that -kH and thermodynamic entropy are identical led to the unification of statistical mechanics and equilibrium thermodynamics. However, it is now quite clear that Boltzmann's formula (2.1) does not fit all conditions. Quantum mechanics provides at least three other known statistics, namely Maxwell-Boltzmann, Fermi-Dirac and Bose-Einstein, which lead to quite different combinatorial definitions for W, and hence to different values of S.[10] For the same macrostate, W is often greatest for Maxwell-Boltzmann, and smallest for Fermi-Dirac, statistics.

Ever since its conception, statistical entropy has been the object of serious criticism; and it still is.[11] It should be instructive, therefore, to closely examine the course of ideas which has gradually led to the conclusion that statistical entropy may be equivalent, or at least related, to various measures of uncertainty and information.

2.2.2 Information theory

As early as 1928, some scientists argued that the entropy of a system constitutes an index of our *degree of ignorance* or *lack of information about* the microstructure of that system.[12] We shall take a very simple example to illustrate this idea. Suppose we have four particles labelled P1, P2, P3, P4 and two states A and B. Let us consider the microstate P1, P2, P3 in A and P4 in B, and denote by S the entropy of this microstate. Since macrocoordinates do not depend on which particular particles are in each state, every microstate in which any three particles are in A and the other one in B must possess the same entropy S. From our knowledge of S, we therefore know the macrostate; i. e. we know that there are three particles in A and one in B, but not which particular particle is in each state. However, we also know that there are four microstates that are compatible with S. Now if we consider the microstate in which P1 and P2 are in A and P3 and P4 in B, then from our knowledge of the corresponding entropy S', we would know that there are six microstates compatible with S'.

Boltzmann's idea is that S = k log 4 and S' = k log 6. Knowing S, we wonder which of the four compatible microstates is actually the case. Knowing S', the spectrum of possibilities increases to six microstates.

Clearly, as the entropy of our system increases from S to S', our degree of ignorance - or our *degree of uncertainty* - about the actual microstate increases as well. As Lewis puts it: "The increase in entropy comes when a known distribution is converted into an unknown distribution. The loss, which is characteristic of an irreversible process, is loss of information".

A specific definition of the *amount of information* in relation to a probability distribution was introduced in 1948 by Norbert Wiener. He adopted a decision-oriented approach by suggesting that if we knew *a priori* that a variable lies between 0 and 1, and *a posteriori* that it lies on the interval (a, b) inside (0, 1), it is reasonable to regard any positive and monotonically decreasing function F as an ordinal measure of the *a posteriori* information, namely

$$\text{Amount of information} = F\left\{\frac{\text{measure of }(a,b)}{\text{measure of }(0,1)}\right\} \qquad (2.7)$$

where F (x) is strictly decreasing with x. But since it is reasonable to expect that (2.7) should yield the same value for all intervals equal to (a, b), it is necessary to assume that the variable related to (2.7) is uniformly distributed over (0, 1) in which case

$$\frac{\text{measure of }(a,b)}{\text{measure of }(0,1)}$$

is the probability that the variable lies within (a, b).

The general principle now becomes obvious: the amount of information, $I(E)$, that the event E of probability p has occurred is ordinally measured by the formula

$$I(E) = F(p) \tag{2.8}$$

where F is a strictly decreasing function which may be assumed to satisfy the condition $F=0$ for $p=1$. Wiener chose the negative logarithmic function, namely

$$I(E) = - \log p \tag{2.9}$$

which was suggested originally by Hartley.[13]

The choice has obvious merits. If in (2.7) we assume that $a = b$, then the information is extremely valuable because it completely determines the variable. According to (2.9), the value of (2.7) is then infinite. On the other hand, if $(a, b) = (0, 1)$, then the information tells us nothing we did not already know. The value of (2.9) is zero in this case. But perhaps the greatest advantage of choosing the logarithm arises from its ability to treat successive amounts of information additively.

All this is in order. But Wiener, using a rather obscure argument,[14] concludes that "a reasonable measure of the amount of information" associated with the probability density f(x) is

$$\int_{-\infty}^{+\infty} [\log f(x)]f(x)dx \tag{2.10}$$

He further suggests that this expression is the negative of the quantity usually defined as entropy in similar situations. In (2.9) we have the logarithm of a probability, whereas in (2.10) the logarithm is applied to the probability density. We shall demonstrate shortly why (2.10) cannot be regarded as the continuous form of Boltzmann's H-function.

The most celebrated way of connecting entropy with information is due to C. E. Shannon who, in the same year as Wiener, presented it in a classical memoir on *communication theory*.[15] Unlike Wiener, Shannon sought a measure of the capacity of a code system to transmit or store *messages*. Shannon was not concerned with whether the message contains any valuable information.[16] A basic problem in communication theory is which code has the largest capacity to transmit information. The shift in the meaning of information is accentuated by Shannon from the outset: "the number of messages.....or any monotonic function of this number, can be regarded as a measure of information produced when one message is chosen from the set". The case of messages transmitted in some ordinary language is a little complicated, since not all sequences of signs constitute messages. A long sequence of the same letter, for example, has no meaning in any language; hence it must not be counted in measuring the information capacity of a language. To arrive at a formula for this situation, Shannon sought a function that would fulfil prescribed analytical conditions. However, the same formula can be reached by a direct manner which has the merit of demonstrating why this formula is identical to Boltzmann's original definition of entropy.

Accepting that the relative frequency with which every sign (letter, punctuation or blank space) appears in any language has an *ergodic* limit, we can denote these frequency-limits by $(p_1, p_2, \ldots\ldots p_s)$.[17] A *typical* message of N signs must contain $(N_1 = p_1N, N_2 = p_2N\ldots\ldots, N_s = P_sN)$ signs of each type. The total number of typical messages is therefore given by the combinatorial formula

$$W = \frac{N!}{N_1!N_2!\ldots\ldots N_s!} \tag{2.11}$$

This is the same formula as (2.1), from which Boltzmann derived his H-function. Since

$$\log W = - N \sum_i p_i \log p_i \tag{2.12}$$

the Shannon information per signal is given by

$$\frac{\log W}{N} = - \sum_i p_i \log p_i \tag{2.13}$$

which is found to be independent of N.

Like Wiener, Shannon noted the identity between (2.3) and Boltzmann's formula, and proposed to refer to it as the *entropy* of the set of probabilities.[18] In the case of *typical* messages, Wiener's formula (2.9) yields

$$- \log (1/W) = \log W \tag{2.14}$$

which is identical to Shannon's result. For Shannon, this represents the number of binary units in typical messages of length N; while for

Wiener the same formula represents the amount of information. To this extent, we can see that information theory was founded independently in 1948 by Shannon *and* Wiener, although Shannon introduced the expression

$$U[P] = - \sum_{i=1}^{n} p_i \log p_i \tag{2.15}$$

as a measure of the *missing information,* or *uncertainty* in a probability distribution $P = (p_1, p_2, \ldots\ldots, p_n)$.

It is also worth noting that Shannon suggested an identical formula to Wiener for measuring the amount of information in a probability density (2.10), but with the opposite sign. Unfortunately, expression (2.15) is not invariant to the interval size over which the distribution is defined, so it is not a measure which can be defined consistently in both discrete and continuous terms. Both Wiener's and Shannon's treatments of the continuous case are inadequate.

Some ground between this newly founded information theory and statistical mechanics was established only after the research efforts of Brillouin and Jaynes were published.[19] Brillouin's information measure assumes that information on relative frequencies is obtained solely from observations, and is given by

$$I_B\,[N] = \frac{1}{N}\,(\log N! - \sum_i \log N_i!)\ . \tag{2.16}$$

The strong link between Brillouin's measure and statistical mechanics is evidenced by observing that (2.16) reduces to Boltzmann's combinatorial definition of entropy (2.2) if $k = 1/N$. Consequently, Brillouin's measure does not rely on Stirling's approximation; it is defined in terms of relative frequencies rather than probabilities.

A short time later, Jaynes demonstrated that Shannon's entropy measure is identical to statistical entropy if we consider statistical mechanics to be a form of statistical inference rather than simply a physical theory. He went on to formulate a principle of maximum uncertainty stating that a minimally prejudiced probability function can be estimated by maximizing Shannon's measure subject to related facts which are treated as constraints. His argument rests on the fact that the most probable distribution is the one which can occur in the maximum number of ways, and this corresponds to the state of maximum entropy. To argue that this distribution is the one which will occur in reality is a statement that is "maximally non-committal with regard to missing information".[20]

Later, Jaynes realized that Gibbs had already given a similar interpretation of the maximum entropy estimate as early as 1902.[21] Both Gibbs' and Jaynes' ideas support Liouville's Theorem in so much as they recognize that

(i) in the long run, a system is most likely to be found in the state which has maximum entropy; and

(ii) if the system is still evolving, its most likely direction of evolution is towards the state of maximum entropy.

The operational nature of Jaynes' principle heralded a plethora of applications, some of which shall be discussed later in this chapter. Although applications were limited originally to thermodynamics, the constraints may embody relationships among variables which describe vastly different systems.

In contrast to the Shannon and Brillouin measures, Kullback suggested a measure of *information gain* which rests on the assumption that information is a relative quantity, and compares probabilities before and after an observation.[22] Information gain is defined when a *posterior* distribution p_i is compared with a known *prior* distribution q_i. This gain, I_k [P:Q] , is given by

$$I_k\,[P;Q] = \sum_i p_i \log(p_i/q_i) \tag{2.17}$$

Similar definitions have been suggested by other writers.[23] Hobson has also demonstrated that I_k [P:Q] is an unique measure of the information content in a posterior probability assignment p_i when the prior probabilities are q_i. The conditions used by Hobson are essentially simple extensions of the original Shannon - Weaver conditions.[24]

Equation (2.17) is perhaps the most general of all information measures. It represents a relative measure which

(i) is independent of the number of events or observations;

(ii) is always positive;

(iii) has more reasonable additive properties than Shannon's measure;

(iv) can be extended to continuous sample spaces; and

(v) allows for non-uniform prior probabilities.

More recently, the use of I_k [P:Q] for various applications has received strong support. It is possible to derive a measure of uncertainty in terms of the difference between two information gains, namely

$$U_k [P] = I_k [P_{max};Q] - I_k [P;Q] \tag{2.18}$$

where P_{max} denotes the probability distribution characterizing the state of maximum knowledge (certainty). If each element of Q is deemed equiprobable, U_k [P] reduces to Shannon's measure U_s [P] . Furthermore, if the posterior distribution is uniform, (2.18) becomes

$$U_k = \log W \tag{2.19}$$

which is Hartley's definition of classical entropy known from statistical mechanics.

Kerridge defined a measure of information inaccuracy,[26] U [PQ] , as

$$U [PQ] = - \sum_i p_i \log q_i \tag{2.20}$$

which, together with Shannon's measure, can be used to derive I_k [P:Q] , since

$$I_k\ [P;Q] = U\ [PQ] - U_s[P]\ . \tag{2.21}$$

Other writers have derived similar functions, like the following definition of *spatial entropy*:

$$I[P;dX] = - \sum_i p_i \log(p_i/\Delta x_i) \tag{2.22}$$

in which Δx_i is the interval size over which p_i is defined. This function was prompted by the realization that Shannon's measure is not invariant to the interval size over which the distribution is defined. It is apparent, however, that (2.22) does not correspond exactly to Kullback's notion of information as a relative quantity, since the interval size (Δx_i) is simply a *property* of the distribution q_i rather than its *a priori* probability.

In conclusion, it now seems clear that Shannon's measure is really a special case of Kullback's information gain. The Kullback formula gereralizes Shannon entropy and also contains Hartley's form as a special case. However, one can readily conceive of the other ways in which both expressions could be generalized.

For example, Rênyi entropy of order α is defined by

$$U_R[P] = \frac{1}{1-\alpha} \log \sum_i p_i^\alpha \tag{2.23}$$

At the limit where α approaches one, U_R [P] approaches Shannon's formula.[28]

Taneja has suggested a generalization of Kullback's measure, namely

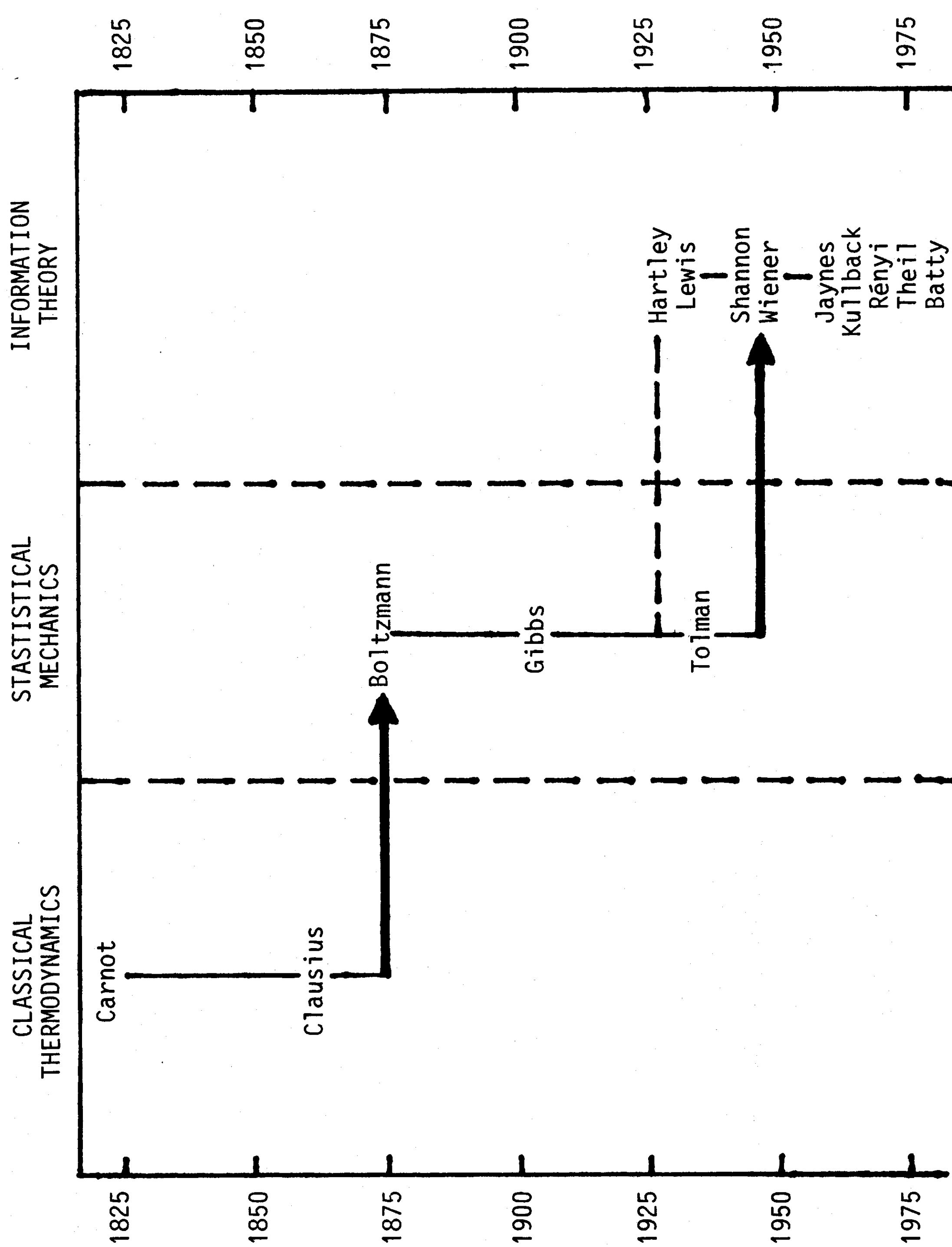
$$I_T[P;Q] = (2^{\alpha-1} - 1)\,(\sum_i p_i^{\alpha}\, q_i^{1-\alpha} - 1) \tag{2.24}$$

which approaches I_k [P;Q] as α approaches one.[29]

These generalizations indicate that there may be whole families of information measures merely awaiting discovery. However, we shall regard further generalizations as being of little practical importance for the forthcoming applications. Nevertheless, their existence is certainly acknowledged.

As a final salute to the remarkable historical developments associated with the entropy concept, Figure 2.1 summarizes the most significant theoretical constributions made to classical thermodynamics, statistical mechanics and information theory.

Figure 2.1


CLASSICAL THERMODYNAMICS
STASTISTICAL MECHANICS
INFORMATION THEORY
1825
1850
1875
1900
1925
1950
1975
Carnot
Clausius
Boltzmann
Gibbs
Tolman
Hartley
Lewis
Shannon
Wiener
Jaynes
Kullback
Rényi
Theil
Batty

2.3 Selected applications

Soon after Shannon's derivation of entropy as a measure of uncertainty, information theory was becoming a magic word in a number of disciplines. It was adopted by biologists in 1953, sociologists in 1954, psychologists in 1956 and ecologists in 1958. In some cases, the pioneers in some of these fields were a little optimistic. Nevertheless, applications continued to spread.

Economists and geographers were more apprehensive. But towards the middle sixties, the concepts of entropy and information made their first few appearances in the literature of both these disciplines. We shall firstly trace the relevant developments in two specific fields, namely economics and spatial analysis. Then we shall briefly review the use of entropy measures in the youthful interdisciplinary field known as spatial economics.

2.3.1 Economic Analysis

Perhaps the first invitation to include the theory of information in the economist's tool box came in 1967. In that year, Theil devoted a whole volume to the presentation of this idea, with many introductory examples falling within the economist's field of vision. Two of the earliest areas of interest were industrial concentration and input-output analysis.

Proposals to use Shannon's expression as an inverse measure of industrial concentration were made independently by a number of economists.[30] If we consider an industry that consists of n firms, we can write $p_1, p_2, \ldots\ldots. p_n$

for the annual sales (or number of workers, etc) of these firms measured as fractions of the total annual sales (or number of workers) in the industry. The entropy, S, is then interpreted as an inverse measure of concentration, where

$$S = - \sum_{i=1}^{n} p_i \log p_i \qquad (2.25)$$

S will be zero in a monopoly situation, and reach a maximum of log n when all n firms are of equal size. This maximum will increase as the number of firms increases.

One of the major attractions of the entropy measure for industrial concentration, and indeed for many other applications, is its ability to handle problems of aggregation and disaggregation. If we combine n firms into N groups of firms, with G_j (j=1,....,N) firms in each group, the entropy at that level of groups is then

$$S' = - \sum_{j=1}^{n} P_j \log P_j \qquad (2.26)$$

where

$$P_j = \sum_{i \in G_j} p_i \qquad (2.27)$$

for j=1,.....,N. Theil refers to S' as the *between-group entropy,* and then develops the following relationship between S and S':

$$S = S' + \sum_{j=1}^{n} P_j S_j \qquad (2.28)$$

where

$$S_j = \sum_{i \in G_j} \frac{p_i}{P_j} \log \frac{P_j}{p_i} \tag{2.29}$$

for j=1,.....,N. S_j is interpreted as the entropy within group G_j, and so the term $\sum_j P_j S_j$ is the *average within-group entropy.*[31] In the context of industrial concentration, S' and S_j measure deconcentration at the group level, and within group G_j, respectively. Clearly, these ideas could be extended to measure and compare industrial concentration in different geographical regions.

Theil was also instrumental in applying some information concepts to input-output analysis. In collaboration with Tilanus, he adopted the expression I[P;Q] in Equation (2.17) as a measure of the *information inaccuracy* of decomposition forecasts.[32] The forecasts relate to the input structure of certain industries in a future year. Using a similar approach to the one adopted in the earlier studies of industrial concentration, they developed expressions to monitor the predictive achievements of subgroups of input coefficients. The information measures so defined were then applied to a ten year time series of input-output tables for the Netherlands.

Building on this initial input-output work, Theil then examined the question of aggregation bias in the consolidation of individual firms into industrial groups. He introduced a measure for the information content of an input-output table,[33] namely

$$I = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{p_{i.}p_{.j}} \tag{2.30}$$

where p_{ij} is defined in terms of the interindustry flow between sectors i and j measured as a fraction of the total output of all sectors, and $p_{i.}$ and $p_{.j}$ are the corresponding row and column sums, respectively. By looking at the difference between the information contents before and after aggregation, he formulates an information decomposition equation for input-output aggregation. This equation is then applied to the same time-series of Dutch input-output tables to identify, in particular, the relative contributions of input and output heterogeneity to the information lost by aggregation.

Theil's exploratory work on this important aspect of input-output analysis was full of valuable insights, but remains unfinished. He concentrated on measuring the information lost by aggregating the input-output coefficients of a larger table into a smaller table of *predetermined* industrial groups. However, a scheme for choosing the most appropriate set of industrial groups, from the original table of disaggregated industrial information, can also be formulated with the aid of information theory. We shall outline such an approach in Chapter 4.

But the most common application of information theory to input-output analysis appeared initially in disguise. Studies of changes in input-output relations over time have been constantly frustrated by lack of time-series data. In 1962, Stone suggested that the problem of estimating a normalized transaction matrix $\{p_{ij}\}$ for a specific year can be simplified by adjusting a known matrix $\{q_{ij}\}$ to fit known row sums u_i and column sums v_j for that specific year. He proposed a sequence of biproportional matrix adjustments, generally referred to as the RAS method, to find this estimate.[34] Elements in $\{p_{ij}\}$ are given by

$$p_{ij} = r_i \, s_j \, q_{ij} \tag{2.31}$$

where the coefficients r_i and s_j are defined by an iterative adjustment process.

However, several authors have noted that the same solution can be achieved by formulating the problems as one of minimum information gain,[35] namely

$$\text{Minimize } I = \sum_i \sum_j p_{ij} \log(p_{ij}/q_{ij}) \tag{2.32}$$

subject to

$$\sum_j p_{ij} = u_i \tag{2.33}$$

$$\sum_i P_{ij} = v_j \tag{2.34}$$

and the usual non-negativity conditions. The original RAS method is actually equivalent to a special application of the principle of minimum information gain, in which the constraints are of the form specified in (2.33) and (2.34).

A detailed comparision of various methods which have been used for the adjustment of input-output tables over time is included in Chapter 4.

2.3.2 Spatial Analysis

At much the same time as Theil was advocating the use of information measures as descriptive statistics for the analysis of economic behaviour, geographers were beginning to analyse spatial phenomena in a similar manner. In most early geographical applications, it was assumed that Shannon's measure was the appropriate one to ascertain the degree of concentration or diversity existing in an observed spatial distribution.[36] The main thrust towards new information measures for spatial analysis has only occurred recently.[37]

Batty's concern was prompted by a realization that Shannon's measure is not invariant to the interval size over which the distribution is defined. He therefore developed a number of new measures for the analysis of spatial distributions, including one defined earlier as *spatial entropy*.[38] His primary aim was to find a spatial measure which can be defined in either discrete or continuous terms. Walsh and Webber have also discussed the appropriateness of various information measures for the analysis of spatial problems.[39]

In contrast to those geographers who use information measures to provide a statistical summary of spatial distributions, there is the work pioneered by Wilson which is largely concerning with generating the most likely probability distributions to describe patterns of spatial behaviour.[40]

This body of work, using the entropy maximizing paradigm, was largely stimulated by the ideas of Jaynes.

Entropy maximizing models may be distinguished according to the nature of the spatial phenomenon which they analyse. Two main classes appear in the literature. *Location* models specify the probability that an individual is located in a given area. Their purpose is to explain the distribution of people or households in various zones of a city or region. The earliest solutions were derived in the absence of any distance consideration.[41] In later models, the mean cost of distance from some centre or set of centres is specified as a constraint, and the pattern of trips to these centres is regarded as fixed.[42] Typically, a negative exponential decline of probability with distance from the centre is obtained.

Suggested improvements and alternatives to the original residential location models have been frequent. Dacey and Norcliffe suggest methods for incorporating zonal capacity constraints in a consistent fashion.[43] Their research appears to build upon earlier work by Webber, in which entropy-maximizing location models for nonindependent events are derived.[44] Webber has also developed some location models in which the distribution whose entropy is maximized is one of expenditures rather than items, and others based on the location and allocation problems of the urban consumer.[45] More recently, attempts to broaden the conventional entropy maximizing framework for location modelling have been made by considering other basic distributions in addition to the familiar Boltzmann statistics.[46]

In reality, residential location models are closely related to the second broad class of entropy-maximizing models, namely *trip-distribution* models. Originally formulated by Murchland and Wilson, [47] these models define the elementary event as a trip between an origin and a destination. The problem is one of estimating the complete distribution of trips between a set of origins r and a set of destinations s. Let T_{rs} be the total number of trips from r to s, O_r be the number of trips originating at r, and D_s be the number ending at s. The cost of a trip from r to s is c_{rs}, whereas the total cost of all trips is C. The most probable distribution is given by the matrix T_{rs} which maximizes entropy S, where

$$S = \log W, \tag{2.35}$$

$$W = \frac{T!}{\prod_{rs} T_{rs}!}, \tag{2.36}$$

and

$$T = \sum_r \sum_s T_{rs} \; ; \qquad (2.37$$

subject to the following constraints:

$$\sum_s T_{rs} = O_r \tag{2.38}$$

$$\sum_r T_{rs} = D_s \tag{2.39}$$

and

$$\sum_r \sum_s c_{rs} T_{rs} = C \; . \tag{2.40}$$

With the assistance of Stirlings' approximation (2.3), the solution can be determined, and is

$$T_{rs} = \exp(-\mu_r - \eta_s - \beta c_{rs}) \tag{2.41}$$

where μ_r, η_s and β are Lagrangian multipliers associated with the constraints (2.38), (2.39) and (2.40) respectively. The solution is often expressed in the form

$$T_{rs} = A_r B_s O_r D_s \exp(-\beta c_{rs}) \tag{2.42}$$

where

$$A_r = [\sum_s B_s D_s \exp(-\beta c_{rs})]^{-1} \tag{2.43}$$

and

$$B_s = [\sum_r A_r O_r \exp(-\beta c_{rs})]^{-1} . \tag{2.44}$$

The constants μ_r and η_s (or A_r and B_s) can be found by substitution, whereas β is normally the subject of calibration. However, if C is known, Equation (2.40) can be solved numerically for β. The most probable distribution of trips given by (2.41) or (2.42) is identical to the gravity distribution, developed originally by analogy with Newtonian mechanics.

Many refinements to this general interaction model are possible. In transport modelling, Wilson extended the basic model to incorporate person types, modal and route split, and traffic assignment.[48] Halder derived a similar interaction model in terms of expenditures rather than trip statistics, which underwent further refinement by Webber,[49] Fisk and

Brown derived new model formulations by revising the microstate descriptions of the tripmakers and their destinations.[50] This work has been extended by Roy and Lesse.[51] Dacey and Norcliffe developed a flexible doubly-constrained model incorporating inequality constraints.[52] Finally, both Fisk and Brown, and Snickars and Weibull, suggested approaches based upon the inclusion of *historical* trip distributions.[53] Snickars and Weibull demonstrated that the classical gravity model is considerably less powerful as a tool for describing changes in trip patterns than models based on *a priori* trip distributions. Their investigations constituted the first practical planning applications of the principle of minimum information gain. The implications of some of these recent innovations for the estimation of commodity distributions will be explored in Chapter 3.

It has been argued that the most successful planning applications of the entropy-maximizing paradigm have been based upon interaction models.[54] A plethora of formulations have certainly emerged since Wilson's fundamental paper appeared. Nevertheless, a number of important issues remain unsolved, including the necessary distinctions between statistical entropy and any *behavioural* theory of human interactions.

It is perhaps no surprise to learn that certain difficulties have also arisen amongst the school of geographers who subscribe to the use of entropy as a descriptive statistic. It now appears that these social scientists are beginning to recognize the potential relevance of other measures for various spatial applications.

2.2.3 Spatial Economics

It is now appropriate to discuss the use of information theory in an evolving area which falls somewhere between geography and economics. We shall refer to this area of interdisciplinary focus as spatial economics. The desire to integrate geographic and economic approaches has been hindered by the preoccupation of Anglo-Saxon economists with the introduction of the time element into their analyses. However, it is encouraging to observe that the spatial aspects of economic developement are now the subject of vigorous enquiry in many parts of the world.

Attempts to introduce information theory into this interdisciplinary arena have been very limited. Perhaps the earliest contribution came from Uribe, de Leeuw and Theil.[55] They suggested the adoption of an information-minimizing solution to constrained matrix problems dealing with interregional and international trade. More particularly, if p_{ij} is the (i,j)th element in an intersectoral, interregional or international trading matrix which has been normalized, then the complete matrix $\{p_{ij}\}$ can be regarded as a set of contingent probabilities. If q_{ij} is then regarded as an estimate or *a priori* probability of the contingency (i,j), the function I defined in (2.32) can be viewed as the expected information value of the message that the probabilities are actually $\{p_{ij}\}$.

They proposed to solve this problem by minimizing I over $\{q_{ij}\}$ subject to the constraints

$$\sum_j q_{ij} = u_i \qquad (2.45)$$

and

$$\sum_i q_{ij} = v_j \tag{2.46}$$

plus the usual non-negativity conditions. As $\{p_{ij}\}$ is also unknown, however, a solution to this system for $\{q_{ij}\}$ still requires some estimates of the elements p_{ij}.[56]

Theil persisted with this two-stage approach when he formulated a gravity model for interrregional commodity flows.[57] Defining x_i^{rs} as the total flow of commodity i from region r, X_i^{r*} and X_i^{*r} as the total production and useage, respectively, of commodity i in region r , and X_i as the total production in the system, he postulated that

$$x_i^{rs} = \frac{X_i^{r*} \; X_i^{*s}}{X_i} \, Q_i^{rs} \tag{2.47}$$

where

$$Q_i^{rs} = \frac{\hat{x}_i^{rs} \; \hat{X}_i}{\hat{X}_i^{r*} \; \hat{X}_i^{*s}} \quad . \tag{2.48}$$

$\hat{x}_i^{rs}$, $\hat{X}_i^{r*}$, $\hat{X}_i^{*s}$ and $\hat{X}_i$ are known values of the variables in some base year. Theil then discovered that his estimate of x_i^{rs} in (2.47) did not satisfy

$$\sum_r \sum_s x_i^{rs} = X_i \quad . \tag{2.49}$$

He suggested multiplication by a normalizing factor to remedy this, but discovered that the new estimate did not satisfy the origin and destination constraints, namely

$$\sum_s x_i^{rs} = X_i^{r*} \tag{2.50}$$

and

$$\sum_r x_i^{rs} = X_i^{*s} \ . \tag{2.51}$$

His final solution was to replace the normalized estimate of x_i^{rs} by $\bar{x}_i^{rs}$, obtained by minimizing the quantity

$$I = \sum_r \sum_s x_i^{rs} \ln \frac{x_i^{rs}}{\bar{x}_i^{rs}} \tag{2.52}$$

which, he argues, is a measure of information inaccuracy, subject to Equations (2.50) and (2.51) as constraints.

Wilson recognized the similarities between Theil's gravity model and a similar model proposed by Leontief and Strout.[58] The latter model does not assume that independent estimates of the regional totals X_i^{r*} and X_i^{*s} are available. It is of the form

$$x_i^{rs} = \frac{x_i^{r*} \ x_i^{*s}}{x_i^{**}} Q_i^{rs} \tag{2.53}$$

where x_i^{r*}, x_i^{*s} and x_i^{**} are the unknown equivalents of X_i^{r*}, X_i^{*s}, and X_i respectively.

Wilson's integration of the gravity and input-output models, using *entropy maximizing principles*, serves as a fundamental focus for this thesis. He builds on the multiregional framework developed by Leontief and Strout to achieve one of the most promising examples of model integration yet accomplished in spatial economics. His approach rests on one rather elegant assumption: that the ultimate destination of goods is irrelevant to producers, and that the origin of goods is irrelevant to consumers. Consequently, x_i^{rs} is now the quantity of commodity i produced in region r and shipped to a (notional) demand pool in region s.

The integrated model based on Leontief-Strout assumptions corresponds to an *unconstrained* input-output version of the gravity model. The only constraints included relate to freight costs and commodity balances. Firstly, we have

$$\sum_r \sum_s c_i^{rs} \; x_i^{rs} = C_i \qquad (2.54)$$

where c_i^{rs} is the unit cost of deliverying commodity i from region r to region s, and C_i is the total freight cost for commodity i. Secondly, we have the fundamental relationship:

$$\sum_s x_i^{sr} = \sum_j a_{ij}^r \sum_s x_j^{rs} + y_i^r \qquad (2.55)$$

where a_{ij}^r is an input-output coefficient for region r, and y_i^r is the final demand for commodity i in region r. Since there are no separate supply or demand constraints, the model is classified as unconstrained.[59] A solution is obtained by maximizing entropy S defined as

$$S = - \sum_i \sum_r \sum_s x_i^{rs} \log x_i^{rs} \qquad (2.56)$$

subject to Equations (2.54) and (2.55).[60]

In reality, this model is very difficult to implement empirically. It is more likely that historical estimates of X_i^{r*} (and even X_i^{*r}) will be available than estimates of the input-output coefficients (a_{ij}^r) or the final demands (y_i^r). For this reason, the development of an empirically operational entropy-maximizing model for the estimation of intraregional and interregional flows seems more appropriate. Such models are the subject of detailed discussion in Chapters 5 and 6.

An alternative approach to the question of interregional commodity flows has been suggested by Fisk and Brown.[61] They utilize Wilson's trip distribution framework to obtain a new model formulation for the distribution of commodity flows. Their revised expression for the entropy of the distribution system is derived using a statistical analogue emanating from quantum mechanics. This type of approach to the estimation of interregional flows will be discussed in detail in Chapter 3.

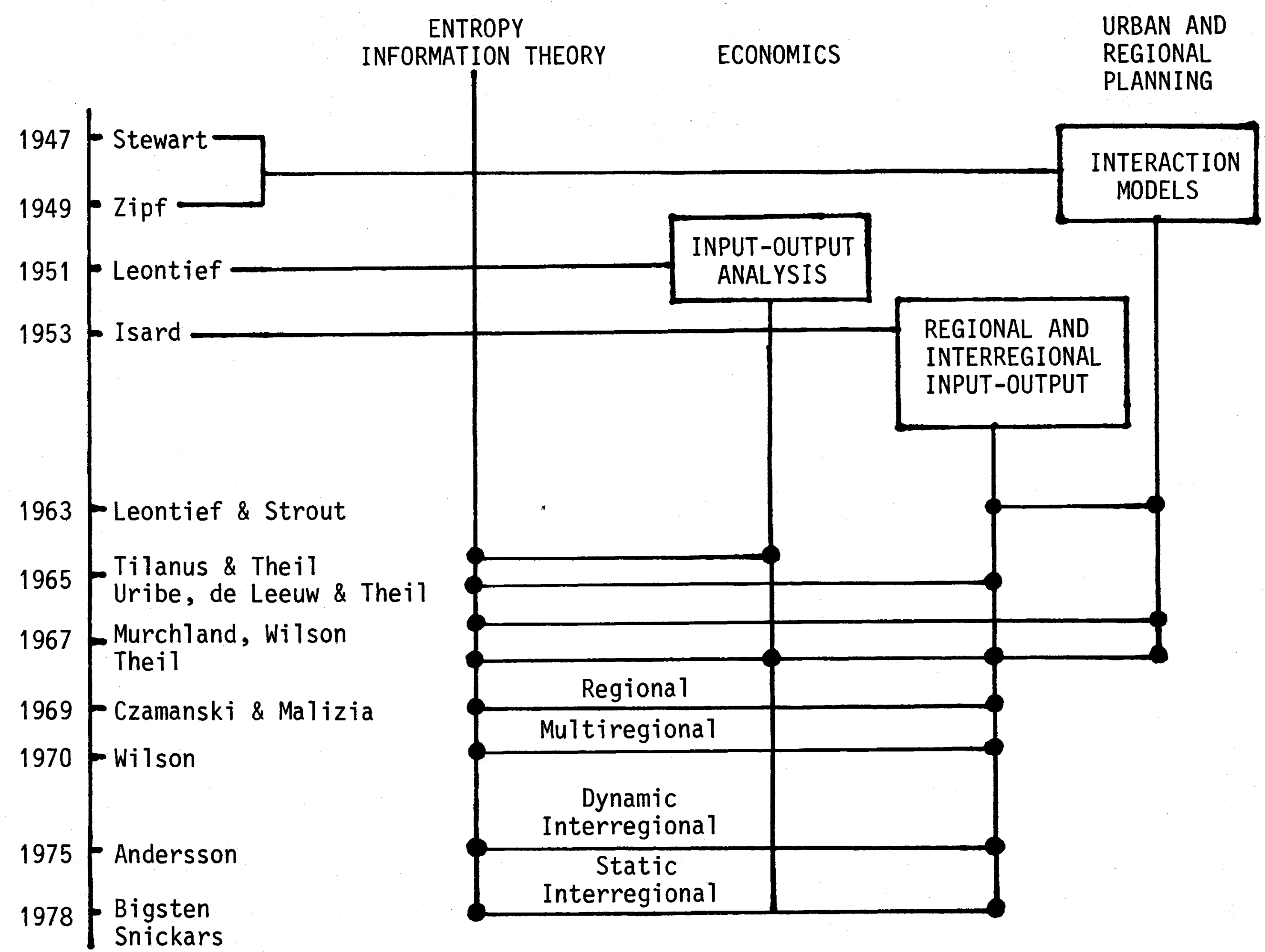
The abovementioned work has provided little more than an elementary background to the eventual integration of spatial models within a comprehensive economic framework. A major objective of the present study is to develop a production-interaction model in which the elementary event is the movement of commodities from sector i in region r for use in the production of other commodities by sector j in region s; p_{ij}^{rs} is the probability of such an event. This fully interregional input-output model can be regarded as a generalization of earlier work on location and interaction models, since $\sum_{ij} p_{ij}^{rs}$ is an interaction probability, $\sum_{ijs} p_{ij}^{rs}$ is a location probability, and $\sum_{jr} p_{ij}^{rs}$ is the production of sector i in region r.

To the authors's knowledge, the only *static* input-output interaction models to be formulated using information-theoretical concepts are attributable to Bigsten and Snickars.[62] Making use of different types of initial information (to produce prior estimates of the intermediate flows), they both apply the minimum information principle to derive estimates of full interregional input-output tables for Kenya and Sweden, respectively. Their respective approaches will be examined closely in Chapter 5.

Rather surprisingly, various authors have already proposed entropy-maximizing solutions for different kinds of *dynamic* input-output interaction models.[63] Some of these models deal with closed economies in which interregional trade, location and growth are integrated within a general equilibrium framework. In others, the approach is one of mathematical optimization. In each case, the interregional flows are estimated using techniques which are equivalent to entropy-maximizing solutions. Some of these models will be reviewed in the early pages of Chapter 7.

An historical summary of the more significant applications of information theory to input-output analysis and interaction modelling is depicted in Figure 2.2.

Figure 2.2


ENTROPY
INFORMATION THEORY
ECONOMICS
URBAN AND
REGIONAL
PLANNING
INTERACTION
MODELS
INPUT-OUTPUT
ANALYSIS
REGIONAL AND
INTERREGIONAL
INPUT-OUTPUT
1947 Stewart
1949 Zipf
1951 Leontief
1953 Isard
1963 Leontief & Strout
1965 Tilanus & Theil
Uribe, de Leeuw & Theil
1967 Murchland, Wilson
Theil
1969 Czamanski & Malizia
1970 Wilson
1975 Andersson
1978 Bigsten
Snickars
Regional
Multiregional
Dynamic
Interregional
Static
Interregional

FOOTNOTES FOR CHAPTER 2

1 See Georgescu-Roegen (1971, p 4).

2 See Shannon (1948).

3 As demonstrated by Sheppard (1976, p 742-3).

4 For a full translation of Carnot (1824), see Magie (1899).

5 See Clausius (1865).

6 For an authoritative discussion of this point, see Bridgman (1941).

7 See Margenau (1950).

8 Introduced by Boltzmann (1872).

9 Which simply states that H does not increase in a closed system.

10 For definitions of each of these statistical forms, see Gurney (1949) or Fast (1962).

11 Georgescu-Roegen (1971, p 147-8) suggests that the root of the difficulty lies in the step by which statistical entropy is endowed with additional meaning beyond that of a disorder index. He strongly objects to the numerical values of expressions such as (2.2) or (2.5) being described as measures of information.

12 See, for example, Hartley (1928) and Lewis (1930).

13 See Hartley (1928).

14 In which he acknowledges a suggestion from John von Neumann.

15 See Shannon (1948).

16 In contrast to Shannon, Jacob Marschak has been concerned primarily with the economics of information systems and the real value of coded messages. Useful references in this area are Marschak (1971; 1973; 1975 a, b).

17 The reason why these coefficients are not referred to as probabilities is related to the nature of linguistics. Although the letters in a language do not follow each other according to a fixed rule, neither do they occur completely at random like the numbers in a die tossing. For further insights into *ergodicity*, see Halmos (1956) or Georgescu-Roegen (1971, p 153-9).

18 On the advice of John von Neumann.

19 See Brillouin (1956) and Jaynes (1957).

20 See Jaynes (1957, p 620).

21 In 1968, Jaynes (1968) acknowledged the earlier suggestions of Gibbs (1902).

22 See Kullback (1959).

23 Notably Rényi (1966) and Theil (1967).

24 See Hobson (1969), and Shannon and Weaver (1949).

25 For example, Hobson (1971), Charnes, Raike and Bettinger (1972), March and Batty (1975), Snickars and Weibull (1977), and Webber (1979) have all suggested that I_K [P;Q] be used to define our uncertainty about the state of the system. Others, such as Hobson and Cheng (1973) and Batten and Lesse (1979), have claimed that I_K [P;Q] is a generalization of the Shannon measure, by allowing for non-uniform prior probabilities.

26 See Kerridge (1961).

27 The term "spatial entropy" was first suggested by Curry (1972), although Batty (1974 a, b; 1978 a, b) proposed the formula in (2.22).

28 See Rényi (1960,1961) for an explanation of this convergence.

29 See Taneja (1974).

30 Notably Hildebrand and Paschen (1964), Finkelstein and Friedberg (1967), Theil (1967) and Horowitz and Horowitz (1968).

31 See Theil (1967).

32 See Tilanus and Theil (1965).

33 See Theil (1967, p 333).

34 See Stone (1962).

35 Notably Bacharach (1970) and Macgill (1977).

36 Geographical examples based on this assumption include the work of Medvedkov (1967), Chapman (1970), and Semple and Golledge (1970) concerning settlement patterns, Ya Nutenko (1970) on zonal partitioning, and Garrison and Paulson (1973) on spatial concentration. Berry and Schwind (1969) followed Theil (1967) in analyzing migration flows using an algebra of probability.

37 Pioneered predominantly by Batty (1974 a, b; 1976; 1978 a, b).

38 See Equation (2.22).

39 See Walsh and Webber (1977).

40 For a broad spectrum of his ideas, see Wilson (1970 b).

41 See Curry (1963) and Berry (1964).

42 See, for example, Wilson (1969 b), Bussière and Snickars (1970), and Scott (1970).

43 See Dacey and Norcliffe (1976).

44 See Webber (1975).

45 See Webber (1976 b; 1977 a).

46 See, for example, Lesse et al. (1978).

47 Wilson (1967) is generally credited with the original formulation, but he himself acknowledges the earlier work of Murchland (1966).

48 See Wilson (1969 a).

49 See Halder (1970) and Webber (1976 b).

50 See Fisk and Brown (1975 a).

51 See Roy and Lesse (1980).

52 See Dacey and Norcliffe (1977).

53 See Fisk and Brown (1975 a, b) and Snickars and Weibull (1977).

54 See Webber (1977 b, p 263).

55 See Uribe, de Leeuw and Theil (1966).

56 Bacharach (1970, p 84) noted this back-to-front method of analysis, and suggested a modification based on the RAS method. His approach regarded the final estimate as the posterior message instead of the prior, thereby overcoming the difficulties associated with two-stage estimation.

57 See Theil (1967).

58 See Wilson (1970 a, b) for his elegant entropy-maximizing formulation of the multiregional framework developed by Leontief and Strout (1963).

59 In the terminology of Wilson (1970 b), four cases are possible: (i) the unconstrained model; (ii) the production-constrained model; (iii) the attraction-constrained model; and (iv) the production-attraction-constrained model.

60 There is no need to normalize x_i^{rs} in order to obtain the correct solution to this problem. The reader may care to confirm that the definition of S in (2.56) gives an identical solution to that obtained using the normalized form (x_i^{rs}/X_i). The latter is, of course, strictly correct if S is to be defined as the entropy of a probability distribution.

61 See Fisk and Brown (1975a).

62 See Bigsten (1978) and Snickars (1979).

63 See, for example, Andersson (1975), Sharpe and Batten (1976), Karlqvist et al. (1978), Andersson and Karlqvist (1979), and Andersson and Persson (1979).

Chapter 3

PROBABILITY DISTRIBUTIONS FOR COMMODITY FLOWS

3.1 Introduction

In elementary statistical mechanics, the entropy of a physical system of particles can be determined statistically by counting the number of possible microstates which correspond to a given macrostate. Although all of the early microstate descriptions were based upon Boltzmann statistics, the advent of quantum mechanics focused attention on three other statistical forms, namely Maxwell-Boltzmann, Fermi-Dirac and Bose-Einstein distributions. Each of these forms corresponds to a particular assumption about the manner in which microstates are grouped into macrostates.

The potential similarities between these three statistical representations of physical systems and the system components describing spatial distributions of goods and people have recently been recognized.[1] In the case of commodity flows, quite different microstate descriptions can be derived depending on whether each commodity unit is regarded as identical or distinguishable. The purpose of this short chapter is to explore some of these rather speculative physical analogies, and to decide whether entropy-maximizing approaches to spatial estimation problems should adopt the first principles of information theory in preference to various microstate descriptions derived by analogy with statistical mechanics.

Section 3.1 outlines various microstate descriptions and the corresponding entropy formulae which could be adopted to describe the spatial distribution of a single commodity in a production-constrained economy. The entropy-maximizing procedure is presented in its more traditional form, namely as the process of determining the most probable macrostate which corresponds to the largest number of possible microstates.[2] The assumption that each microstate is equi-probable may prevent any of these standard distributions from reproducing empirical flows accurately. The inclusion of non-uniform prior probabilities in the entropy formulations is therefore discussed.

Section 3.2 examines the doubly-constrained model, and demonstrates how two different microstate descriptions can produce the same solution. This convergence may have important implications for later attempts to derive more complex microstate descriptions relating to spatial input-output analysis. The presence of supply and demand constraints could lead to various simplifications which support the use of Shannon uncertainty, or Kullback information gain, as the most appropriate objective function for a range of constraint situations.

3.2 Production-constrained Distributions

3.2.1 Microstate Descriptions

In the early pages of Chapter 2, we introduced the concepts of macrostate and microstate. It was stated that the entropy of a system can be determined statistically by counting the number of possible microstates which correspond to a given macrostate. In other words, the entropy depends on the manner in which microstates are grouped into macrostates. Since the maximum number of microstates corresponding to any given macrostate depends largely on the degree of aggregation specified in the macrostate definition, it is clear that entropy is intrinsically related to the degree of aggregation or diversity inherent in the system definition.

For convenience, we shall now repeat Boltzmann's original formula for entropy S as a function of the number of microstates W, namely

$$S = k \log W \ . \tag{3.1}$$

Quantum mechanics provides three elementary statistical forms for the determination of W. In the standard approach, based on Maxwell-Boltzmann statistics, the counting of microstates is carried out as if each individual particle is *distinguishable*. Alternatively, systems of *identical*, independent particles are described using Bose-Einstein or Fermi-Dirac statistics.[3] The former allows each quantum state to contain an unlimited number of particles, whereas the latter does not permit more than one particle to occupy the same state.

The potential similarities between these physical assemblies, containing large numbers of particles, and the system components describing the spatial distribution of goods are readily demonstrated. Suppose we wish to estimate x_{rs}, the interregional distribution of flows between regions r and s for some type of commodity. We know X_r, the total production of that commodity in each region r, such that

$$\sum_s x_{rs} = X_r \ . \tag{3.2}$$

Adopting the assumption that each commodity unit is *distinguishable*, the number of ways, W_r in which the total number of commodity units, produced in region r can be distributed into m groups,[4] with x_{rs} (s=1,....,m) commodities in each group, is given by

$$W_r = \frac{X_r!}{\prod_s x_{rs}!} \ . \tag{3.3}$$

This is seen to be based on the same familiar formula of combinatorial calculus (2.1) which corresponds to Boltzmann's original definition of entropy.[5] Considering all regions simultaneously, the complete micro-state description becomes

$$W = \prod_r \left[\frac{X_r!}{\prod_s x_{rs}!} \right] \ . \tag{3.4}$$

However, expression (3.4) does not take into account the microstate space associated with the *individual destination* of each commodity unit entering region s. In general, Maxwell-Boltzmann statistics imply that x_{rs} individual commodity units arriving in a region s may be assigned to D_s individual destinations (depots) in $(D_s)^{x_{rs}}$ different ways, when more than one commodity unit per destination can be accomodated. Thus expression (3.4) should be modified to read

$$W = \prod_r \left[\frac{X_r!}{\prod_s x_{rs}!} \prod_s (D_s)^{x_{rs}} \right] . \tag{3.5}$$

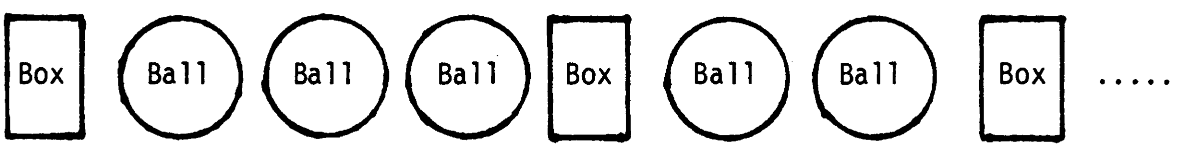
Up to this point, the commodity units have been treated as completely distinguishable. In reality, this may not always be the case. A distributor may not consider the sending of a commodity unit x_1 to a destination in region s_1, and an *identical* unit x_2 to region s_2, as being distinct from sending x_1 to s_2 and x_2 to s_1. In this case, either Fermi-Dirac or Bose-Einstein statistics may be relevant.

Fermi-Dirac statistics are only suitable for commodity flow analysis if no more than one commodity unit can be assigned to each destination. Since more than one commodity unit per destination is allowed, we shall restrict our attention to Bose-Einstein statistics. In this case, the microstate description relates to the number of ways in which x_{rs} identical commodity units may be assigned to D_s distinct destinations.

For a fixed origin r, a *given* set of x_{rs} commodity units may be assigned to each region s in only one distinct way. Using Bose-Einstein statistics, we can now calculate the number of ways in which the x_{rs} identical units can be distributed among the D_s destinations within region s. The problem is analogous to the following situation:

Suppose x_{rs} balls are to be arranged in D_s boxes, allowing any number of balls in a box. The procedure is to lay out the $(x_{rs} + D_s)$ objects (balls and boxes) in a straight line, but in a random order.[6] This linear arrangement is chosen to start with a box and might continue as depicted in Figure 3.1.

Figure 3.1.



Then, move those balls which are immediately to the right of a given box into that box. For the arrangement depicted in Figure 3.1, there are three balls in box 1, two in box 2 and so on. Since $(x_{rs} + D_s - 1)$ objects actually move,[7] there are $(x_{rs} + D_s - 1)!$ possible arrangements. But many of these arrangements are identical. For example, we must divide by 3! to take account of the fact that the balls in box 1 are all identical. In general, there are $x_{rs}!$ such permutations. Similarly, there are $(D_s - 1)!$ permutations of the boxes. The total number of distinguishable arrangements, W_{rs}, is thus given by

$$W_{rs} = \frac{(x_{rs}+D_s-1)!}{x_{rs}!\,(D_s-1)!} \quad . \tag{3.6}$$

This is indeed the number of ways in which x_{rs} identical commodity units from a given region of origin r may be distributed among D_s destinations or depots in region s. If we now consider all origin and destination regions simultaneously, the complete microstate description is of the form

$$W = \prod_{rs} \frac{(x_{rs}+D_s-1)!}{x_{rs}!(D_s-1)!} \quad . \tag{3.7}$$

Equation (3.7) is similar to the microstate description suggested originally by Fisk and Brown.[8] Since D_s is normally large compared with unity, a reasonable approximation of Equation (3.7) is given by

$$W \cong \prod_{rs} \frac{(x_{rs}+D_s)!}{x_{rs}!\cdot D_s!} \tag{3.8}$$

Returning to our analogous arrangement of balls into boxes, Equation (3.8) corresponds to a circular instead of a linear arrangement of objects.[9]

A summary of the microstate descriptions for each of the production-constrained statistical analogues discussed above is given in Table 3.1. To facilitate further understanding of the four basic microstate descriptions, a simple example pertaining to industrial location is discussed in Appendix A. A number of hybrid models could emerge from combinations of these microstate descriptions to define multi-indexed variables such as interregional, interindustry flows. It is obvious, however, that it becomes increasingly difficult to define the appropriate microstates in more complex situations, and standard assumptions such as homogeneity and independence become more difficult to fulfil.

TABLE 3.1

STATISTICAL FORM	W
Old Boltzmann	$\prod_r \left[\frac{X_r!}{x_{rs}!}\right]$
Maxwell-Boltzmann	$\prod_r \left[\frac{X_r!}{\prod_s x_{rs}!} \prod_s D_s^{x_{rs}}\right]$
Bose-Einstein	$\prod_{rs} \left[\frac{(x_{rs}+D_s-1)!}{x_{rs}!\,(D_s-1)!}\right]$

3.2.2 Entropy Maximands

By assuming that each microstate is equi-probable, we can determine the most probable commodity distribution by finding that solution which corresponds to the maximum number of microstates, subject to a known system of constraints. Jaynes suggests that "in making inferences on the basis of partial information, we must use that probability distribution which has maximum entropy subject to whatever is known. This is the only unbiased assumption we can make".[10] Before we discuss the appropriate constraints, it is informative to define the entropy maximand for the various statistical forms introduced in the previous section.

Maximizing the number of microstates W is equivalent to maximizing Boltzmann's entropy formula (3.1). To take advantage of Stirling's approximation for factorials, namely

$$\log n! \cong n(\log n-1) \tag{3.9}$$

we shall consider log W as the entropy maximand. If expression (3.4) is simplified using Stirling's approximation, we find that

$$\log W = X \{\sum_r p_r \log p_r - \sum_r \sum_s p_{rs} \log p_{rs}\} \tag{3.10}$$

where $X = \sum_r X_r$, $P_r = X_r/X$, and $P_{rs} = x_{rs}/X$. Neglecting all the constant terms, (3.10) reduces to Shannon's measure of the uncertainty in a probability distribution,[11] namely

$$S_1 = - \sum_r \sum_s p_{rs} \log p_{rs} \quad . \tag{3.11}$$

We shall refer to this reduced form of (3.10) as *Shannon entropy.*

If we start instead with Equation (3.5), use of Stirling's approximation leads to a different result, namely

$$\log W = X\{\log D + \sum_r p_r \log p_r + \sum_s p_{rs} \log p_s - \sum_r \sum_s p_{rs} \log p_{rs}\} \tag{3.12}$$

where $D = \sum_s D_s$ and $P_s = D_s/D$. Disregarding all the constant terms, (3.12) reduces to the following:

$$S_2 = - \sum_r \sum_s p_{rs} \log (p_{rs}/p_s) \tag{3.13}$$

It is noteworthy that Equation (3.13) resembles Batty's definition of *spatial entropy*, if D_s is regarded as a measure of regional size in terms of commodity useage. We shall therefore refer to S_2 as *Batty entropy.*[12]

If we now redefine p_{rs} as x_{rs}/D_s, use of Stirling's approximation to simplify Equation (3.8) produces the following result:

$$\log W = \sum_r \sum_s D_s \{(1+p_{rs}) \log (1+p_{rs}) - p_{rs} \log p_{rs}\} \tag{3.14}$$

Disregarding the constant terms, Equation (3.14) reduces to the following form

$$S_3 = - \sum_r \sum_s \{p_{rs} \log p_{rs} - (1+p_{rs}) \log (1+p_{rs})\} \tag{3.15}$$

Since this formula is based on Bose-Einstein statistics, we shall refer to S_3 as *Bose-Einstein entropy.*

The entropy formula for each of these production-constrained statistical analogues is included in Table 3.2.

TABLE 3.2

ENTROPY FORMS	S	DEFINITIONS
Shannon	$- \sum_r \sum_s p_{rs} \log p_{rs}$	$p_{rs} = x_{rs}/X$
Batty	$- \sum_r \sum_s \{p_{rs} \log p_{rs} - p_{rs} \log p_s\}$	$p_{rs} = x_{rs}/X$ $p_s = D_s/D$
Bose-Einstein	$- \sum_r \sum_s \{p_{rs} \log p_{rs} - (1+p_{rs}) \log (1+p_{rs})\}$	$p_{rs} = x_{rs}/D_s$

3.2.3 The Most Probable Distributions

In the previous section, a set of entropy formulae were derived for the case of production-constrained interregional flows, namely flows which satisfy the production constraint (3.2) for each region r. Some recent research has suggested that trip distribution models may be sensitive to the location of destinations about the region of origin.[13] The normal practice in these models is to then constrain the $\{x_{rs}\}$ distribution, so that it conforms to an origin-specific cost constraint, based on the *a priori* knowledge of average delivery costs between different regions.

In the case of commodity flows, binding constraints are more likely to be associated with the nodes (regions) themselves, rather than with the links between nodes. We shall therefore adopt a general *capacity* constraint of the form

$$\sum_s c_r \, x_{rs} \leqq C_r \qquad (3.16)$$

for each region r, where c_r represents the *physical* capacity requirements for delivering each commodity unit from region r, and C_r is the total capacity for handling outgoing deliveries from region r. This capacity constraint is included in the following formulations simply because it is analogous to the energy constraint imposed on systems of particles in quantum mechanics. Its inclusion in later formulations, derived using information-theoretical arguments, will require additional justification. We shall regard the commodity distributions resulting from any maximum entropy estimate, subject to relations (3.2) and (3.16), to be part of a set of basic solutions. In later chapters, further constraints will be introduced, and their effect on the basic solutions examined. The most probable commodity distribution is found by calculating the number of microstates associated with each distribution which conforms to constraints (3.3) and (3.16), and choosing the one which corresponds to the maximum number of microstates.

Formally, the problem is to maximize W (or S) subject to Equations (3.2) and (3.16), and the standard non-negativity conditions. This constrained-matrix estimation can be achieved using the method of Lagrange multipliers. We shall demonstrate this solution technique for the case of Batty entropy.

Firstly, we form the Lagrangian $\mathcal{L}$, namely

$$\mathcal{L} = \log W + \sum_r \mu_r (X_r - \sum_s x_{rs}) + \sum_r \beta_r (C_r - \sum_s c_r\, x_{rs}) \tag{3.17}$$

where μ_r and β_r are the Lagrangian multipliers associated with Equations (3.2) and (3.16) respectively. Note that it is more convenient to maximize log W rather than W, since it is then possible to use Stirling's approximation to simplify $\mathcal{L}$. We know that

$$\log W = \log X_r! - \sum_s \log x_{rs}! + \sum_s x_{rs} \log D_s \tag{3.18}$$

$$= X_r \log X_r - X_r - \sum_s [x_{rs} \log x_{rs} - x_{rs} - x_{rs} \log D_s] \ .$$

The $\{x_{rs}\}$ distribution which maximizes $\mathcal{L}$, and which therefore constitutes the most probable distribution of commodities, is the solution of

$$\frac{\partial \mathcal{L}}{\partial x_{rs}} \leq 0 \tag{3.19}$$

and the constraint equations (3.2) and (3.16). We have

$$\frac{\partial \mathcal{L}}{\partial x_{rs}} = - \log x_{rs} + \log D_s - \mu_r - \beta_r c_r \tag{3.20}$$

and

$$\frac{\partial^2 \mathcal{L}}{\partial x_{rs}^2} = - \frac{1}{x_{rs}} . \tag{3.21}$$

The first derivative vanishes when

$$x_{rs} = D_s \exp(-\mu_r - \beta_r c_r) \tag{3.22}$$

and the second order conditions are negative definite as long as all $x_{rs} \neq 0$. The constants μ_r and β_r can be computed from (3.3) and (3.16). Results for each of the basic cases are summarized in Table 3.3.

TABLE 3.3

ENTROPY FORM	BASIC SOLUTIONS
Shannon	$[\exp(\mu_r + \beta_r c_r]^{-1}$
Batty	$D_s[\exp(\mu_r + \beta_r c_r)]^{-1}$
Bose-Einstein	$D_s[\exp(\mu_r + \beta_r c_r)-1]^{-1}$

It is interesting to note that the "quantized" solutions can be expressed in the general form

$$x_{rs} = D_s[\exp(\mu_r + \beta_r c_r) + \alpha]^{-1} \tag{3.23}$$

where α takes on the value 0, 1 or -1 depending on whether the distribution is based on Maxwell-Boltzmann, Fermi-Dirac, or Bose-Einstein statistics respectively.

3.2.4 Prior Information

If any of these production-constrained models provides a good representation of the distribution system, the $\{x_{rs}\}$ matrix will correspond closely to survey results. If not, there is clearly something amiss with the initial model hypotheses. For example, the assumption (implicit in the model formulation) that all microstates satisfying (3.2) and (3.16) are equally probable may prove unrealistic owing, for example, to different average consumption rates in each region.

Although it is possible to apply various weightings to the entropy function to compensate for *a posteriori* knowledge (as described for the case of Batty entropy), it is also feasible to treat some or all of the *a priori* possibilities as being, at least in principle, empirically tractable. Suppose we already know, or can independently estimate, prior information in the form of an historical distribution $\{x^o_{rs}\}$. Designating by q_{rs} the *a priori* probability that a commodity unit from region r will be sent to region s, we have

$$q_{rs} = x^o_{rs}/X^o \qquad (3.24)$$

where

$$\sum_r \sum_s x^o_{rs} = X^o$$

How can we introduce this *a priori* information into our basic model formulations?

Gaseous particles are known to exhibit uniform prior probabilities of entering various quantum states. Consequently, there is little to be learned from the various statistical forms discussed earlier, since all these analogies are based on the *a priori* assumption of equi-probability. For non-uniform prior distributions, a different approach is needed.

Fisk and Brown have suggested that if q_{rs} designates the *a priori* probability that a tripmaker will travel to a destination in region s, the microstate description (for the case of work trip distributions) should be weighted by the factor

$$\pi_{s} (q_{rs})^{x_{rs}}$$

for each region r.[14] The resulting maximization for each of the quantized commodity distributions would lead to a general solution of the form

$$x_{rs} = q_{rs} D_s [\exp(\mu_r + \beta_r c_r) + \alpha]^{-1} \qquad (3.25)$$

where α takes on the value 0, 1 or -1 depending on whether the distribution is based on Maxwell-Boltzmann, Fermi-Dirac or Bose-Einstein statistics respectively.

Note, however, that D_s biases the final distribution according to the *a posteriori* number of destinations (or depots) in each region s, whereas q_{rs} biases the same distribution according to the *a priori* probability that a commodity unit from region r will be sent to region s. Solutions based on (3.25) therefore imply that D_s and q_{rs} warrant equal weighting in their joint influence on the final distribution. Such an assumption appears to have neither theoretical nor empirical foundation, so Equation (3.25) will not be adopted as the means of introducing prior probabilities in the context of commodity distributions. Consideration of conditional probabilities and Bayesian likelihood ratios may provide a more acceptable theoretical underpinning for such an analysis, but these avenues cannot be explored in the space of this dissertation.

If instead we return to the first principles of information theory, we could adopt Kullback's measure of information gain,[15] which rests on the assumption that information is a relative quantity; it compares probabilities before and after an observation. For the case of commodity flows, our information gain, I [P;Q] , is given by

$$I[P;Q] = \sum_r \sum_s p_{rs} \log (p_{rs}/q_{rs}) \tag{3.26}$$

Adopting Jaynes' suggested extension of the maximum entropy paradigm,[16] we can choose that distribution $\{p_{rs}\}$ which minimizes I [P;Q] subject to related facts about P, which are treated as constraints. The principle of minimum information gain is equivalent to choosing the most probable distribution given that the prior probabilities are non-uniform. Its application emphasizes the inertia of the *a priori* distribution, since the final solution is the one which most closely resembles the original distribution in an informative sense.

If we minimize I [P;Q] subject to constraints (3.2) and (3.16) and the usual non negativity conditions, we obtain the following revised solution for the most probable commodity distribution (x_{rs}) :

$$x_{rs} = x^{o}_{rs}[\exp(\mu_r + \beta_r c_r)]^{-1} . \tag{3.27}$$

The important question concerning acceptable means by which the *a priori* probabilities may be determined remains partly unanswered. If a complete historical flow matrix $\{x^{o}_{rs}\}$ is available, there is no difficulty in calculating the prior probabilities. If partial information about historical flow patterns can be obtained, this information can be used to derive a complete matrix of *a priori* flows, so long as the total flows are known. The procedure involved represents a compromise between the entropy-maximizing and information-minimizing paradigms. To demonstrate its application, a simple example is presented in Appendix B.

If we do not have any historical information available, it is doubtful whether *a priori* probabilities should be "guestimated" by adopting some ad hoc assumptions (such as regional consumption levels) to bias the $\{q_{rs}\}$ matrix. In this situation, it seems preferable to return to the entropy-maximizing principle (with its implicit assumption of uniform prior probabilities), which offers the additional descriptive flexibility afforded by the various statistical forms. An entropy-maximizing solution could certainly be adopted as the *a priori* distribution in cases where historical information (in the form of constraints) is superior to our current knowledge about constraining influences on the distribution of commodities.

A further consideration is the degree to which the *a priori* probabilities contain distributional information partly related to that contained in the c_r coefficients which bias capacity constraints. Following the discovery that the classical gravity model is much less powerful as a tool for describing historical changes in trip patterns than models based on *a priori* trip patterns,[17] it is not all certain whether any constraints like (3.16) should be included in any model which is based on *a priori* probabilities. The inertia embodied in the distribution system may ensure that historical patterns are a superior guide to distributional behaviour in the short to medium term. In this situation, the solution given by (3.27) could be modified by dropping the ($\beta_r c_r$) term.

3.3 Doubly-constrained Distributions

Having derived a full set of basic solutions for the production-constrained models, it is now appropriate to examine some doubly-constrained models. To do this we introduce the additional constraint

$$\sum_r x_{rs} = D_s \tag{3.28}$$

for each region, where D_s now represents the demand for (or useage of) commodities in region s, such that

$$\sum_s D_s = \sum_r X_r = X \ . \tag{3.29}$$

Equation (3.29) implies that demand equals supply at the aggregate level, so much so that the doubly-constrained model is simply an equilibrium model. If we now maximize W subject to a constraint system which includes supply (3.2) and demand (3.28) equations, the resulting $\{x_{rs}\}$ distribution is therefore an equilibrium solution.

Returning to the three statistical analogues (see Tables 3.1 and 3.2) discussed in the previous section, a rather interesting convergence occurs if Maxwell-Boltzmann statistics are applied. When each commodity unit is regarded as distinguishable, Maxwell-Boltzmann statistics imply that

$$W = \prod_r \left[\frac{X_r!}{\prod_s x_{rs}!}\right] \prod_{rs} \left[D_s^{x_{rs}}\right] . \tag{3.30}$$

Since $\sum_r x_{rs} = D_s$, the second term in Equation (3.30) can be simplified as follows:

$$\begin{aligned}
\prod_{rs} \left[D_s^{x_{rs}}\right] &= \prod_s \left[D_s^{\sum_r x_{rs}}\right] \\
&= \prod_s \left[D_s^{D_s}\right] \\
&= \text{a constant} .
\end{aligned}$$

This constant term will disappear when the Lagrangian is differentiated, yielding an identical solution to the one obtained using Shannon entropy. In other words, the double-constrained solution for Batty entropy reduces to that of Shannon entropy.

This convenient equivalence feature emphasizes the importance of Shannon entropy for doubly-constrained models. Unfortunately, the same equivalence does not extend to the Bose-Einstein distribution, which retains a similar solution to the production-constrained case. Solutions for the three doubly-constrained models are given in Table 3.4, wherein η_s is the Lagrange multiplier associated with Equation (3.28).

As mentioned in Chapter 2, there has been particular interest in the doubly-constrained model. A number of writers have indicated that the doubly-constrained information-minimizing model produces identical solutions to those of a group of techniques classified as biproportional matrix adjustments.[18] The RAS method and two-dimensional contingency table analysis belong to this group.

At first sight, this wealth of investigation into models which have a similar structure looks very promising. From the viewpoint of a spatial analyst, however, methods based on biproportional matrix adjustments may sometimes be too restrictive, since they ignore all information beyond the standard origin (supply) and destination (demand) constraints. For this reason, biproportional matrix adjustment may be regarded as a *special case* of minimum information gain in which the constraints are of a particular form. The latter approach promises to provide a very flexible approach to various matrix estimation problems which may be confronted in spatial economics. For a detailed discussion of biproportional matrix adjustments, see Section 4.3.3 in the following chapter.

3.4 Concluding Remarks

The analysis of commodity flows in a production-constrained economy (Section 3.1) reveals that certain analogies based on Maxwell-Boltzmann or Bose-Einstein statistics may be useful, depending on whether each

TABLE 3.4

ENTROPY FORM	SOLUTIONS
Shannon	$[\exp(\mu_r + \eta_s + \beta\ c_r)]^{-1}$
Batty	$[\exp(\mu_r + \eta_s + \beta\ c_r)]^{-1}$
Bose-Einstein	$D_s[\exp(\mu_r + \eta_s + \beta\ c_r)-1]^{-1}$

commodity unit can be regarded as distinguishable or identical respectively. The appropriate entropy maximands for these statistical forms are summarized in Table 3.2. It is noteworthy that each maximand envolves Shannon's measure of uncertainty as the first term in its entropy expression. This common feature suggests that it may be possible to reproduce all the "quantized " distributions by adopting Shannon's measure as the objective function for each case, and recasting the conditions imposed by each microstate description into the form of additional constraints.

Further support for the use of Shannon's measure as a general entropy maximand is provided by the doubly-constrained model. In the presence of both supply and demand constraints, the equilibrium solution for Batty entropy reduces to that of Shannon entropy. This has particularly important implications for input-output analysis, since Leontief's multisectoral world cannot boast a homogenous one-to-one relationship between each commodity and the sector to which it is allocated. Consequently, the need to distinguish different commodities in any one sector implies the adoption of Maxwell-Boltzmann statistics, and either the Shannon or Batty entropy form.

The entropy maximizing assumption that each microstate is equi-probable may prevent any of these statistical representations from reproducing empirical flows accurately. To overcome this, adoption of Kullback's measure of information gain is recommended to compare prior and posterior probability distributions. The principle of minimum information gain searches for the most probable distribution, given that the prior probabilities are non-uniform. To this degree, the need for additional distributional assumptions, such as Wilson's familiar cost constraints, may be reduced.

There is clearly a stage beyond which analogies drawn from statistical mechanics become misleading and perhaps even inappropriate. This point may be reached as we enter the multisectoral world of input-output analysis. Jaynes' entropy-maximizing paradigm requires no physical analogy, since its use can be justified by mathematical reasoning and a logical principle. For this reason, together with the arguments outlined above, we shall adopt the first principles of information theory in the remaining chapters of this dissertation.

Henceforth we shall restrict our attention to Shannon's measure of uncertainty and Kullback's measure of information gain. The former will be used as the entropy maximand for situations in which no satisfactory *a priori* information about the flows can be ascertained, or in times of change when the reproduction of historical trends is to be avoided. The latter measure will be adopted as the information minimand for those cases where adequate *a priori* information is available, and resistance to change is sustainable.

FOOTNOTES FOR CHAPTER 3

1 See, for example, Fisk and Brown (1975 a), Dacey and Norcliffe (1976), Lesse et al. (1978), or Roy and Lesse (1981).

2 In this instance, each macrostate corresponds to an interregional commodity distribution.

3 These particles are known as *bosons* or *fermions*, respectively.

4 In physical systems, these groups correspond to a set of energy states.

5 See Equation (2.1).

6 Achieved, for example, by tossing a coin.

7 Box 1 is a fixed reference point.

8 See Fisk and Brown (1975 a).

9 In which there is no specific reference point.

10 See Jaynes (1957, p 620).

11 Shannon's measure is given originally in Equation (2.15). For only two events (say E and not E), it can be expressed as

$$U\,[P] = -\,[p_1 \log p_1 + (1 - p_1) \log (1 - p_1)]$$

which resembles the familiar *Fermi-Dirac* statistical form with i = 1. Differentiating U with respect to the probability p_1, we get

$$\frac{dU}{dp_1} = -\log \left(\frac{p_1}{1 - p_1}\right)$$

which is the inverse of the *logit* function. Theil (1972) concludes that the logit model measures the sensitivity of uncertainty (or entropy) to variations in probabilities.

12 Spatial entropy is defined by Equation (2.22), which is the form proposed by Batty (1974 a, b; 1978 a). A similar derivation based on Equation (3.5) can be found in Snickars and Weibull (1977), but their interpretation of the $\{p_s\}$ distribution differs. They describe $\{p_s\}$ as the *a priori* most probable distri-

bution, which should be of the same order as the *a posteriori* $\{p_{rs}\}$ distribution. Equation (3.13) is actually biased according to the proportion of destinations in each region.

13 See, for example, Wilson (1973) or Fisk and Brown (1975 b).

14 See Fisk and Brown (1975 a).

15 See Equation (2.17).

16 See Jaynes (1968).

17 See Snickars and Weibull (1977).

18 See, for example, Bacharach (1970), Macgill (1977) or Hewings and Janson (1980).

Chapter 4

NON SPATIAL INPUT-OUTPUT ANALYSIS

4.1 Introduction

In Chapter 3, we discussed shipments of goods by considering each commodity in *isolation*. In this chapter, we shall link these commodity flows together by describing the economy as a system of *interdependent* activities being carried out by many mutually interrelated industries. The traditional method of describing these economic interdependencies is known as input-output analysis, and was developed initially for the American economy by Leontief.[1]

In examining Leontief's original static model, it is possible to identify some fundamental problems associated with his classic formulation. Firstly, there is the question of units. In order to distinguish between published input-output tables which normally record interindustry transactions in *value* terms, and tables which express the *volumes* of interindustry flows, we must take account of prices explicitly. The conversion from value units to physical units is quite straightforward if the appropriate price data are available. Unfortunately, this is not often the case in practice, so much so that tables recorded in inappropriate units are used in many studies.

Secondly, the fundamental Leontief form assumes a square matrix of technical coefficients, thereby implying that each industry produces just one homogeneous product. Although each commodity may be regarded as homogeneous for practical purposes, a one-to-one sector product assumption is clearly unrealistic. Fortunately, the development of rectangular input-output tables has necessitated a specific distinction between sectors and products. The salient features of rectangular input-output models are discussed in Section 4.2.2.

A third difficulty with Leontief's original static model arises because the input coefficients do not reflect the capital stock requirements of the economy. Leontief recognized this problem and proposed a dynamic extension of his original model to introduce the fundamental characteristics of a growth model, in which the influence of time is recognized explicitly. A major obstacle to the implementation of his dynamic input-output model, however, has been the scarcity of reliable capital coefficients. We shall describe his original dynamic formulation in Section 4.2.3, and then demonstrate a means by which this empirical difficulty might be alleviated.

Having introduced the essential characteristics of pertinent non-spatial input-output models, the rest of this chapter demonstrates the use of methods based on information theory for the estimation of key parameters in these models. In Section 4.3.1, Theil's original measure of the information content of an input-output table is replaced by a more general expression, namely Shannon's measure of uncertainty.[2] The discussion in Section 4.3.2 embraces some speculation on an ordering scheme for simple aggregation by adopting the criterion of minimum loss of information. The objective function is expressed in terms of an information loss criterion between the original array,and an array of the same size acting as a surrogate for the reduced array.

In Section 4.3.3, we examine biproportional matrix adjustments for updating input-output tables. It is found that standard iterative solution procedures like the RAS method and the IPFP algorithm are equivalent to certain restrictive applications of the principle of minimum information gain.[3] The latter principle is more flexible, since it is capable of including additional or alternative information to that demanded by these standard approaches. Although this flexibility may not be regarded as essential for the adjustment of input-output tables over time, it is certainly important when attempts are made, in later chapters, to adjust these matrices over space.

Finally, two information-theoretical models are formulated for the estimation of unknown flow coefficients. In Section 4.3.4, the assumption of one-to-one correspondence between sectors and commodities (which is central to the classic Leontief model) is abandoned in an attempt to simplify the aggregation problem. The resulting model makes use of information contained in rectangular make and absorption matrices, as well as the square intersectoral matrix, to estimate the elements of a new array $\{x_{ijk}\}$, in which each element, x_{ijk}, defines the flow of commodity k from industry i to industry j. Since published flow matrices are usually out of date by the time of their release, the initial estimates may be updated in a second stage of estimation using the principle of minimum information gain.

A similar two-stage approach is suggested in Section 4.3.5 to improve the method of estimating capital coefficients in dynamic models. First-stages estimates are computed using a standard two-factor model of independence, which assumes that *a priori* coefficients depend more on investments in the base period than on existing capacities. In the second stage, base period coefficients are adjusted to satisfy *a posteriori* capacity constraints using the same principle.

Throughout this chapter the discussion is primarily non-spatial. The marriage of certain non-spatial concepts derived from the analyses in this chapter with related work on spatial analysis will be celebrated in Chapters 5 and 6.

4.2 Basic Model Characteristics

4.2.1 The Original Static Input-Output Model

Consider an economy which is (figuratively) divided into n production sectors. Denote by x_i the total output of sector i, and by x_{ij} the intermediate demand by sector j for goods produced by sector i. Further denote by y_i the demand by final users for goods produced by sector i. The overall input-output balance of this n-sector economy can be described in terms of n linear equations:

$$x_i = \sum_{j=1}^{n} x_{ij} + y_i \; . \qquad (i = 1,\ldots,n) \qquad (4.1)$$

The input-output structure of any particular sector can be described by a vector of technical coefficients, a_{ij}, each of which states the amount of a particular input from sector i which is absorbed by sector j per unit of its own output. Thus the commodity flows included in Equations (4.1) are subject to the following set of structural relationships:

$$x_{ij} = a_{ij}\, x_j \quad . \qquad (i,j = 1,\ldots,n) \qquad (4.2)$$

Substituting (4.2) in (4.1), we have

$$x_i = \sum_{j=1}^{n} a_{ij}\, x_j + y_i \quad . \qquad (i = 1,\ldots,n) \qquad (4.3)$$

This is a system of n linear equations in which final demand can be regarded as the steering mechanism. The essential concern of input-output analysis in this most basic form is to determine future levels of production (x_i) in the endogenous sectors, given known or exogenously determined levels of final demand (y_i), and assuming constant technical coefficients (a_{ij}). This is done by forming the familiar Leontief inverse, achieved by rearranging terms in (4.3) to give (in matrix notation):

$$X = (I-A)^{-1}Y, \qquad (4.4)$$

where X is a vector of endogenous sectoral outputs, Y is a vector of exogenous final demands, A is the matrix of technical coefficients, and I is the unit matrix.

It is worthwhile noting that there is a complementary row sum relationship for the primary and intermediate inputs. If v_j denotes the sum of all primary inputs to sector j, the complementary equation to (4.1) is

$$x_j = \sum_{i=1}^{n} x_{ij} + v_j \qquad (j = 1,\ldots,n) \qquad (4.5)$$

It is convenient to represent these intersectoral relationships in tabular form, as depicted in Table 4.1 .

An augmented set of input-output coefficients can be formed by replacing v_j by $x_{n+1,j}$. In this case, it is obvious that

$$\sum_{i=1}^{n+1} a_{ij} = 1 \qquad (j = 1,\ldots,n) \qquad (4.6)$$

This particular feature, together with the knowledge that all coefficients are non-negative, implies that any augmented column of input-output coefficients can easily be transformed into a vector which has the same technical features as a probability distribution. As we shall discover later in this chapter, this characteristic enables us to apply concepts derived from information theory.

There has been no attempt above to offer anything more than a rudimentary version of the basic input-output methodology. In so doing, a clear passage is left to identify some of the difficulties associated with Leontief's original formulation.

Firstly, there is the question of *units*. Published tables normally record intersectoral transactions in value terms, thereby allowing row and column sums to be expressed on a comparable scale. To study the effect of technological conditions upon flow relationships, however, we prefer to know flow volumes. In order to distinguish between tables which are expressed in physical and value units, we must take account of prices explicitly.

Table 4.1

Supply \ Useage	Sector 1	2	...	...	...	n	Total Intermediate Demand	Final Demand	Total Useage
Sector 1	x_{11}	x_{12}	...	...	...	x_{1n}	$\sum_{j=1}^{n} x_{1j}$	y_1	x_1
2	x_{21}	x_{22}	...	...	...	x_{2n}	$\sum_{j=1}^{n} x_{2j}$	y_2	x_2
⋮	⋮	⋮				⋮	⋮	⋮	⋮
n	x_{n1}	x_{n2}	...	...	...	x_{nn}	$\sum_{j=1}^{n} x_{nj}$	y_n	x_n
Total Intermediate Inputs	$\sum_{i=1}^{n} x_{i1}$	$\sum_{i=1}^{n} x_{i2}$	...	...	...	$\sum_{i=1}^{n} x_{in}$			
Primary Inputs	v_1	v_2	...	...	...	v_n			
Total Supply	x_1	x_2	...	...	...	x_n			

To eliminate the price effect, we can convert the technical coefficients from value units (a_{ij}) to physical units (a'_{ij}). Mathematically, this conversion can be easily accomplished by the following multiplication:

$$a'_{ij} = \frac{p_i}{p_j} a_{ij} \tag{4.7}$$

where p_i is the average price of products from sector i. In practice, however, we must overcome the difficulty of obtaining the appropriate data on prices. Consequently, many studies which deal with interindustry flows have actually employed technical matrices computed in value terms.

For our present purpose, a particularly limiting feature of the fundamental Leontief form is that Equations (4.4) assume a square matrix of technical coefficients, thereby implying that each sector produces just one homogenous product. This assumption ensures (amongst other things) that matrix inversion and an exact algebraic solution are possible. Although no commodity is homogenous in the strict sense, it may be regarded as such for most practical purposes. However, the one-to-one sector-product assumption central to the classic Leontief model is clearly unrealistic, and greatly confounds the aggregation problem. We must eventually abandon this assumption if realistic models of commodity flows are to be developed.

Fortunately, some of the theoretical and applied work on input-output tables has acknowledged a specific distinction between sectors and products or, more generally, between industries and commodities. This body of work has led to the development of rectangular input-output models, discussion of which follows.

4.2.2 Rectangular Input-Output Models

Rectangular input-output tables have been published in Canada and Norway for some years in response to the 1968 recommendations by the Statistical Commission of the United Nations for a new system of national accounts.[4] These tables exhibit a fundamental distinction between industries and commodities, which is illustrated schematically in Table 4.2.[5] In this diagram , capital letters denote matrices, lower case letters denote vectors, and primes indicate transposed column vectors (row vectors). The complete notation is as follows:

U is an absorption matrix; a typical element u_{kj} represents the amount of commodity k used up by industry j,

V is a make matrix; a typical element v_{ik} represents the amount of commodity k produced by industry i,

q is a vector of total commodity outputs,

g is a vector of total industry outputs,

y is a vector of final demands for commodities,

e is a vector of primary inputs.

The square partitions in Table 4.2 (namely the commodity-by-commodity and industry-by-industry partitions) have no entries at all. Any attempt to use a model as simple as Equation (4.4) requires either one, and only one, non-zero element in each row of the absorption matrix so as to ensure a one-to-one correspondence between aggregated commodities and industries, or must manipulate the data from both matrices to provide

Table 4.2

	Commodities	Industries	Final Demands	Totals
Commodities		u_{kj}	y_k	q_k
Industries	v_{ik}			g_i
Primary Inputs		e_j		
Totals	q'_k	g_j		

entries for the empty square partitions. In the former case, the technical coefficients in matrix A could be given by u_{kj}/g_j, and the make matrix V would then be redundant. In the latter case, industries could produce any number of products, so matrices U and V need not to be square and assumptions regarding the internal structure of the square matrices are required.

In terms of Table 4.2, the fundamental input-output balance is given by:

$$q_k = \sum_j u_{kj} + y_k \tag{4.8}$$

with the corresponding column sums yielding

$$g_j = \sum_k u_{kj} + e_j \ . \tag{4.9}$$

Further relationships for the output balances are given by

$$q_k = \sum_i v_{ik} \tag{4.10}$$

and

$$g_i = \sum_k v_{ik} \ . \tag{4.11}$$

Various assumptions that relate to the technical conditions of production can be expressed as follows:

(a) The assumption that intermediate inputs of commodities are proportional to the industry outputs into which they enter, namely

$$u_{kj} = b_{kj} \, g_j \tag{4.12}$$

where $B = \{b_{kj}\}$ is a matrix of technical coefficients of dimensions commodity-by-industry.[6]

(b) The assumption that each industry makes commodities in its own fixed proportions, namely

$$v_{ik} = c_{ki}\, g_i \tag{4.13}$$

where $C = \{c_{ki}\}$ is a matrix of coefficients of dimensions commodity-by-industry.

(c) The assumption that commodity outputs are allocated among industries in fixed proportions, namely

$$v_{ik} = d_{ik}\, q_k \tag{4.14}$$

where $D = \{d_{ik}\}$ is a matrix of market share coefficients of dimensions industry-by-commodity.

Combining (4.8) and (4.12) leads to the following basic relationship:

$$q_k = \sum_j b_{kj}\, g_j + y_k \tag{4.15}$$

The assumption that matrix B describes technical structure implies that it is reasonably stable, irrespective of the levels of final demand and of the resulting levels of commodity and industry outputs. However, Equation (4.15) is not yet an input-output model. Such models define a linear transformation from final demands to either industry or commodity outputs. Consequently, a relationship linking industry and commodity outputs is required.

Assumption (b), corresponding to Equation (4.13), leads to a *commodity technology model* in which there is only one way of producing each commodity. A commodity produced by several industries will have the same input structure in each industry, so much so that the input structures of industries will be linear combinations of the input structures of the commodities they produce.

Combining (4.10), (4.13) and (4.15) results in the following input-output relationship for the commodity technology model (using matrix algebra):

$$q = B\,C^{-1}\,q + y \tag{4.16}$$

In theory, the rectangular matrix C cannot be inverted, implying that the assumption of a commodity technology is only permissible if the number of industries equals the number of commodities. In reality, however, a pseudo-inverse of C may exist under certain conditions.[7]

Assumption (c), corresponding to Equation (4.14), leads to an *industry technology model* in which the market shares of industries are assumed to be stable, and are regarded as independent of the levels of commodity or industry outputs. From Equations (4.11) and (4.15), we can write

$$\begin{aligned} q &= BD\,q + y \\ &= (1-BD)^{-1}y \end{aligned} \tag{4.17}$$

and

$$\begin{aligned} g &= D(1-BD)^{-1}y \\ &= (1-DB)^{-1}Dy\ . \end{aligned} \tag{4.18}$$

Thus, if we denote the square input-output coefficient matrix by A, we find that A = BD for a commodity-by-commodity table, whereas A = DB for an industry-by-industry matrix. With the assumption of an industry technology, there is no need for the number of industries to equal the number of commodities.

A case can be made for each of assumptions (b) and (c) in particular contexts. Moreover, mixed technology models have been proposed to allow for both assumptions, depending on the nature of each commodity produced. Distinctions between principal or secondary products, joint products and by-products can be implemented by dividing the V matrix into two parts, V_1 and V_2, so that

$$V = V_1 + V_2 \ . \qquad (4.19)$$

The elements of V_1 are outputs that more reasonably conform to the assumption of a commodity technology; those of V_2 are treated using the market share hypothesis. Appropriate matrices of coefficients can then be calculated.

To take full advantage of the basic relationships embodied in rectangular input-output theory,[9] it is necessary to introduce a new interindustry flow matrix $\{x_{ijk}\}$ which is capable of defining supply-demand interactions more realistically than the traditional matrix $\{x_{ij}\}$. A typical element, x_{ijk}, in this new array defines the flow of commodity k from industry i to industry j.[10] Subsets of elements in $\{x_{ijk}\}$ can be related directly to the elements of the absorption and make matrices (in Table 4.2), by simply noting that summation over i automatically defines elements in the make matrix, and summation over j defines elements in the absorption matrix.

This three-dimensional array formally recognizes that the functioning of an economy takes place in terms of specific commodity flows between producers and consumers. The form $\{x_{ijk}\}$ allows both producer's (i's) and receivers (j's) to be explicitly defined, as well as the specific item (commodity k) which is involved in any particular transaction between them. The make matrix then defines the sources (producers) of these interactions, whereas the absorption matrix defines the sinks (receivers). The natural focus of interest is in the complete interaction array $\{ x_{ijk} \}$, which is the subject of further discussion in Section 4.3.4 below.

4.2.3 Dynamic Input-Output Models

Before we discuss certain information-theoretical approaches which might be adopted for the estimation of some key parameters in the input-output models discussed above, there is a third difficulty arising from the classical Leontief form which warrants our attention. In its original static formulation, the Leontief model describes the mutual interdependence of various sectors in the economy by means of a set of technical coefficients, which represent the flows between producers and consumers during one time period. However, these input coefficients do not reflect the stock requirements of the economy. They do not, and cannot, explain the magnitude of those input flows which serve directly to satisfy the capital needs of each sector.[11]

In Equations (4.1), (4.2) and (4.3), all investment behaviour is consolidated into final demand (y_i). This means that although the effects of investment demand are considered, the actual magnitude of this particular demand cannot be ascertained. Explicit evaluation is only possible if the stock requirements of all economic sectors are included in the structural representation of the system, and distinguished from the flow requirements previously described. This dynamic extension of the static input-output model enables us to recognize the stock-flow relationship explicitly, and thereby introduce the fundamental characteristics of a growth model in which the influence of time is recognized explicitly. Leontief introduced this dynamic element in 1953. In the following section, we shall begin with his original formulation, and then demonstrate a means by which one of the major empirical difficulties associated with this model might be alleviated.

Let us reconsider our n-sector economy (initially discussed in Section 4.2.1) and introduce the time element into the model. Accordingly, $x_i(t)$ and $y_i(t)$ now represent, respectively, total output and final demand for the goods from sector i in period t. Final demand, however, now excludes the annual additions to the stocks of fixed capital (by way of productive investment) used by each sector to expand their productive capacity.

If $s_{ij}(t)$ represents the stock of a commodity produced by sector i and used by sector j at time t, the rate of change of this stock at any point in time can be written as

$$\frac{ds_{ij}(t)}{dt}$$

or simply as $\dot{s}_{ij}(t)$. Equation (4.1) can now be rewritten as follows:

$$x_i(t) = \sum_{j=1}^{n} x_{ij}(t) + \sum_{j=1}^{n} \dot{s}_{ij}(t) + y_i(t) \qquad (i = 1,\ldots,n) \; . \qquad (4.21)$$

In this formulation, all allocations of commodity i to the current replacement and maintenance requirements of existing capital stocks are included in the flow coefficients $x_{ij}(t)$. Accumulated stocks for future consumption are considered as another form of consumption, and are therefore included in final demand, $y_i(t)$. Consequently, the $\dot{s}_{ij}(t)$ term in Equation (4.21) refers only to those changes in stocks which are designed to increase the existing productive capacity.[12]

The set of structural equations (4.2) describing the input-output structure of each sector must now be supplemented by a corresponding set of capital-output relationships, namely

$$s_{ij}(t) \geq b_{ij}\, x_j(t) \qquad (i,j = 1,\ldots,n) \qquad (4.22)$$

where b_{ij} is the stock or capital coefficient for capital goods produced by sector i and used in sector j. Changes in stocks can be represented by differentiating both sides of (4.22) with respect to time, yielding

$$\dot{s}_{ij}(t) \geq b_{ij}\, \dot{x}_j(t) \quad . \qquad (i,j = 1,\ldots,n) \qquad (4.23)$$

Substitution of relations (4.2) and (4.23) into (4.21) leads to our fundamental system of dynamic input-output relationships

$$x_i(t) \geq \sum_{j=1}^{n} a_{ij}\, x_j(t) + \sum_{j=1}^{n} b_{ij}\, \dot{x}_j(t) + y_i(t) \qquad (i = 1,\ldots,n) \; (4.24)$$

which describe the dynamic interdependencies of the economy. We can further simplify relations (4.24) by assuming that capital goods produced in year t are installed and operable by the following year, t+1. Then the direct interdependence between the outputs of all sectors in two successive years can be described by the following balance equations:

$$x_i(t) = \sum_{j=1}^{n} a_{ij}\, x_j(t) + \sum_{j=1}^{n} b_{ij}\, [x_j(t+1)-x_j(t)] + y_i(t) \qquad (i = 1,\ldots,n) \tag{4.25}$$

if we assume full utilization of capital stocks in each period.

Methods for the solutions of this system of linear difference equations have been suggested,[13] and the model's inherent stability problem has received much attention.[14] However, a major reason for the lag between theoretical and empirical developments has been the scarcity of reliable capital coefficients, $\{ b_{ij} \}$.[15]

With this obstacle in mind, Brôdy assumed a direct relationship between the coefficients in the stock and flow matrices to obtain his solution for the American economy using a closed version of Leontief's dynamic model.[16] Although this implicit connection had been recognized earlier,[17] Brody estimated values for the quotients of the corresponding elements in his simplified matrices. The explicit relationship is

$$b_{ij} = a_{ij}\, t_{ij} \qquad (i,j = 1,\ldots,n) \tag{4.26}$$

where t_{ij} is defined as the turnover time, being the time required for capital goods from sector i to be used up in the production of goods in sector j.

Unfortunately, accurate estimates of the turnover time in each sector are usually unavailable. Approximations are possible by substituting physical life span for turnover time.[18] The former is more easily measured and is independent of the current price mechanism. But, in reality, the two concepts are not equivalent. Their exact relationship requires further investigation.

The purpose of the following discussion is to suggest a method for the direct estimation of capital coefficients which does not require prior knowledge of the turnover time in each sector.[19] However, the method does not preclude the existence of a relationship like (4.26) between the stock and flow coefficients. We shall begin with the system of linear difference equations (4.25) developed above.

The usual manner of investigating the growth possibilities in a system like (4.25) is to study the corresponding homogenous system for which $y_i(t) = 0$. For this closed system, there exists a balanced growth path which achieves the largest possible rate of expansion of the system.[20] For our immediate purpose, studying this homogenous system would be inappropriate. Instead we shall assume that each sector has a growth path of its own, with a corresponding growth rate λ_i, at which

$$x_i(t) = x_i(0)\ \{1-\lambda_i\}^t \quad . \qquad (i = 1,\ldots,n) \qquad (4.27)$$

Substituting (4.27) into (4.25), setting t=0, and writing x_i for $x_i(0)$, we have

$$x_i = \sum_{j=1}^{n} a_{ij}\ x_j + \sum_{j=1}^{n} b_{ij}\ \lambda_j\ x_j + y_i \ . \quad (i = 1,\ldots,n) \qquad (4.28)$$

We shall assume that base values of x_i, y_i, λ_i, and a_{ij} are available at our reference point in time ($t = 0$). Thus we can simplify Equations (4.28) to the following:

$$\sum_{j=1}^{n} \alpha_j b_{ij} = \beta_i \qquad (i = 1,\ldots,n) \tag{4.29}$$

where

$$\alpha_j = \lambda_j x_j \qquad (j = 1,\ldots,n)$$

and

$$\beta_i = x_i - y_i - \sum_{j=1}^{n} a_{ij} x_j \, . \quad (i = 1,\ldots,n)$$

We can also express Equations (4.26) in the proportional form

$$\frac{b_{ij}}{b_{i\ell}} = \frac{a_{ij}\, t_{ij}}{a_{i\ell}\, t_{i\ell}} \qquad (i,j = 1,\ldots,n) \tag{4.30}$$

where ℓ is a reference column in the appropriate matrices. Equations (4.30) express proportional relationships between the elements in columns j and ℓ of any row i. A more convenient form is

$$\frac{b_{ij}}{b_{i\ell}} = \gamma^{i}_{j\ell} \qquad (i,j = 1,\ldots,n) \tag{4.31}$$

where $\gamma^{i}_{j\ell}$ are constants yet to be determined.

Equations (4.31) form a system of n by (n-1) non-trivial equations.[21] Combining Equations (4.29) and (4.31) results in a system of n^2 equations which can be solved directly by simple substitution, giving

$$b_{ij} = \frac{\beta_i \ \gamma^i_{j\ell}}{\sum\limits_{j=1}^{n} \alpha_j \ \gamma^i_{j\ell}} \ . \qquad (i,j = 1,\ldots,n) \qquad (4.32)$$

In Equations (4.32), the α_j and β_i terms ensure consistency with other elements of the model.[22] As well as this, they determine the relative distribution of capital flows between rows in the matrix of capital stocks. In a complementary manner, the $\gamma^i_{j\ell}$ terms determine the relative distribution between the columns of the same matrix. Since the success of this approach rests entirely on the accuracy of these row and column proportionality terms, we shall refer to this technique as the *Biproportionality Method*.[23]

To complete the estimation, however, we require a suitable means of predetermining the $\gamma^i_{j\ell}$ terms. Two simple possibilities spring to mind. Firstly, we could follow Bródy's seemingly drastic simplification, and suppose that capital goods from sector i have a uniform turnover time, t_i, which is independent of destination sector j.[24] Thus we redefine our original square matrix $\{t_{ij}\}$ to be a diagonal matrix of n turn-over times (t_i), and Equations (4.30) simplify to the following:

$$\frac{b_{ij}}{b_{i\ell}} = \frac{a_{ij}}{a_{i\ell}} \ . \qquad (i,j = 1,\ldots,n) \qquad (4.33)$$

Assumption (4.33) implies that for each delivery sector i, the distribution of capital flows between different destination sectors j are in the same relative proportions as the distribution of intermediate flows. To implement this approach, no prior knowledge of turnover times is required.

Secondly, we could postulate an alternative relationship between the coefficients in each row, namely

$$\frac{b_{ij}}{b_{i\ell}} = \frac{k_j}{k_\ell} \qquad (i,j = 1,\ldots,n) \qquad (4.34)$$

where k_j is the average or marginal capital/output ratio in sector j. Equations (4.34) introduce the assumption that for each delivery sector i, the distribution of capital flows between different destination sectors j are proportional to the capital/output ratios in those destination sectors. If the sectoral capital/output ratios are known, the $\gamma^{i}_{j\ell}$ terms may be estimated using (4.34).

Matrices of capital coefficients may be approximated using Equations (4.32), together with assumption (4.33) or (4.34). The relative merits of each approach have been assessed in Appendix C, using a closed version of Leontief's model adopted originally by Brődy in a study of the American economy.[24] Equations (4.32) were found to produce superior results to Brődy's when the appropriate solutions were compared.

Although Equations (4.32) provide a convenient means of estimating a complete matrix of capital coefficients in an *ex post* fashion, the assumptions embodied in (4.33) and (4.34) are clearly unrealistic for any detailed model. Inaccuracy can partly be gauged by comparision with the existing capital stock, K_i, in each sector i. Under the assumption of full capacity utilization, the following equations should hold:

$$\sum_{j=1}^{n} b_{ij} x_j = K_i \quad . \qquad (i = 1,\ldots,n) \qquad (4.35)$$

In reality, most sectors are not operating at full capacity, so Equations (4.35) will not hold identically. It is more appropriate to regard the existing capital stock as a capacity constraint in each sector, namely

$$\sum_{j=1}^{n} b_{ij} x_j \leq K_i \quad . \qquad (i = 1,\ldots,n) \qquad (4.36)$$

Finally, it may sometimes be necessary to consider each input coefficient, a_{ij}, as the sum of two components: a current input coefficient, a^1_{ij}, and an input coefficient for the maintenance of fixed capital, so that

$$a_{ij} = a^1_{ij} + q_i\, b_{ij} \qquad (i,j = 1,\ldots,n) \qquad (4.37)$$

where q_i is the average rate of depreciation of goods produced by sector i.[25] If published input-output statistics record the current input coefficients (a^1_{ij}), Equations (4.28) can be modified to read

$$x_i = \sum_{j=1}^{n} a^1_{ij}\, x_j + \sum_{j=1}^{n} b_{ij}(q_i + \lambda_j)x_j + y_i \quad (i = 1,\ldots,n) . \qquad (4.38)$$

This can be simplified to the following:

$$\sum_{j=1}^{n} \alpha_{ij}\, b_{ij} = \beta_i \qquad (i = 1,\ldots,n) \qquad (4.39)$$

where

$$\alpha_{ij} = (q_i + \lambda_j)x_j \qquad (i,j = 1,\ldots,n)$$

and

$$\beta_i = x_i - y_i - \sum_{j=1}^{n} a_{ij}^{1}\, x_j \quad (i = 1,\ldots,n)\ .$$

Combining (4.26) and (4.37) gives

$$b_{ij} = \left\{\frac{t_{ij}}{1-q_i\, t_{ij}}\right\} a_{ij}^{1} \qquad (i,j = 1,\ldots,n) \qquad (4.40)$$

and the modified general solution becomes

$$b_{ij} = \frac{\beta_i\, \gamma_{j\ell}^{i}}{\sum_{j=1}^{n} \alpha_{ij}\, \gamma_{j\ell}^{i}} \ . \qquad (i,j = 1,\ldots,n) \qquad (4.41)$$

An alternative approach, which adopts techniques derived from information theory, is presented in Section 4.3.5.

4.3 Some Information-Theoretical Formulations

4.3.1 The Information Content of an Input-Output Table

The application of concepts derived from information theory to measure the information content of input-output tables was first attempted in the middle sixties.[26] Theil's work is adopted as a starting point for the reformulations discussed in this section.

Returning to our static multisectoral economy,[27] it is possible to normalize the original input-output matrix to form a bivariate probability distribution, p_{ij}, in two different ways. Firstly, we can express the intermediate flows as fractions of the *total* output of all sectors, namely

$$p_{ij} = x_{ij} / \sum_{k=1}^{n} x_k \quad . \qquad (i,j = 1,\ldots,n) \tag{4.42}$$

Secondly, we can express the same flows as fractions of the *intermediate* output of all sectors, thus

$$p_{ij} = x_{ij} / \sum_{k=1}^{n} \sum_{\ell=1}^{n} x_{k\ell} \quad . \qquad (i,j = 1,\ldots,n) \tag{4.43}$$

Theil adopted (4.42) to define his bivariate distribution, and then expressed the *intermediate* demand for products from sector i, namely $p_{i.}$, in the fractional form

$$p_{i.} = (x_i - y_i) / \sum_{k=1}^{n} x_k \qquad (i = 1,\ldots,n) \tag{4.44}$$

together with the *total* supply of products to sector j, namely $p_{\cdot j}$, as

$$p_{\cdot j} = x_j / \sum_{k=1}^{n} x_k \quad . \qquad (j = 1,\ldots,n) \qquad (4.45)$$

The curious feature of Theil's approach is that primary inputs are included on the supply axis of the bivariate array, whereas final demands are excluded from the receiving sectors. Although this omission may be justified if we simply wish to disaggregate primary inputs, it does little to clarify the choice between (4.42) and (4.43) as the most appropriate bivariate array $\{p_{ij}\}$ for defining the information content of the original table. To restore consistency, we shall treat primary inputs *and* final demands endogenously, and focus our attention on the flows between all sectors.

The column sums, $p_{\cdot j}$, are given by (4.45), with

$$x_n = \sum_{i=1}^{n} y_i \; .$$

The row sums, $p_{i\cdot}$, are redefined as

$$p_{i\cdot} = x_i / \sum_{k=1}^{n} x_k \qquad (i = 1,\ldots,n) \qquad (4.46)$$

where $x_n = \sum_{j=1}^{n} v_j$. The elements of the array $\{p_{ij}\}$ can now be viewed as part of a bivariate probability distribution for which

$$\sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} = 1 \qquad (4.47)$$

and $p_{ij} \geq 0$ for all i and j. Information-theoretical concepts can therefore be adopted.

Theil defines the information concept of the input-output table as the expected mutual information of his bivariate array $\{p_{ij}\}$, namely

$$I = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} \log \frac{p_{ij}}{p_{i.}\, p_{.j}} \quad . \tag{4.48}$$

Expression (4.48) actually measures the difference between the information contained in the complete input-output distribution, and that contained in the row ($p_{i.}$) and the column ($p_{.j}$) sums alone. In other words, (4.48) measures the deviation of the actual distribution from an independence model, namely $p_{ij} = p_{i.}\, p_{.j}$.[29]

A more plausible definition of the total information content involves Shannon's measure of uncertainty, U, given as

$$U = - \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} \log p_{ij} \quad . \tag{4.49}$$

Expression (4.49) implies that the information contained in the table is measured relative to the completely ignorant assumption of equiprobability.[30] Consequently, U will be adopted as a general measure of the *total* information content of an input-output table.

We now turn to a simple example for clarification of this new definition. In his 1967 analysis, Theil chose a 4 by 4 array which included primary inputs as an additional row vector, together with additional zero entries in the corresponding column vector. Since we have chosen to include final demands in our bivariate array, the need for dummy column entries does not arise. Accordingly, we have modified Theil's original array to represent the flows between four non-zero sectors. The resulting array $\{p_{ij}\}$ is reproduced in Table 4.3.

TABLE 4.3

Demand / Supply		Sector j 1	2	3	4	Row Sum $p_{i.}$
Sector i	1	.05	.05	0	.05	.15
	2	.05	.1	.05	0	.2
	3	.1	.05	.05	0	.2
	4	.2	.1	.1	.05	.45
Column Sum $p_{.j}$		.4	.3	.2	.1	1

Statistics describing Table 4.3 can be computed using both measures. For Theil's measure (I), the information content is 0.182 bits. Using Shannon's expression (U), the information content is 3.522 bits. The larger information content exhibited under the latter assumption reflects the greater degree of ignorance assumed beforehand. If we are interested in the *total* information content of the table, this latter definition is more appropriate.

4.3.2 Information Losses in Simple Aggregation

As a logical extension of the discussion concerning the information content of input-output tables, we now turn to the subject of aggregation. By aggregation, we refer to the process of combining individual firms into groups which are generally referred to as industries or sectors. This aggregation procedure has significant effects on any predictions of intermediate and final demand.

It is almost thirty years since the aggregation problem was first recognized as a significant issue in input-output analysis.[31] In the intervening period, many useful theorems concerning the conditions required for the absence of aggregation bias have been formulated.[32] In addition, various criteria have been suggested for selecting an appropriate aggregation scheme,[33] the two most popular of which attempt to preserve input or output homogeneity.

Our present aim is to improve computational efficiency by determining, with some scheme, the means of reducing the size of an input-output array by aggregation of complete rows and columns (i. e. "simple" aggregation), such that, in a special sense, the reduced array is the best substitute for the original in the required areas of application.

In contrast to the previous section, aggregation involves a *loss* of information which prompts the adoption of a relative, rather than an absolute measure. The ordering scheme will therefore be defined as that which results in an aggregated matrix suffering minimum loss from the original in the Kullback tradition,[34] representing a converse formulation of the standard minimum information gain approach.

As this approach implies comparison of arrays of the same size, it does not seem directly applicable to the case where the original matrix and a smaller aggregated matrix are being compared. In addition, an information loss, rather than a gain, is incurred in proceeding from the original matrix to the aggregated matrix. However, defining our original matrix as an n by n matrix Q, we seek an n by n surrogate matrix P which is equivalent in information to a given m by m candidate aggregated matrix $\bar{P}$ (where $m < n$). We therefore choose that aggregation strategy to form $\bar{P}$ (and thus P) which minimizes the information gain I(Q: P), defined as

$$I(Q:P) = \sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} \log (q_{ij}/p_{ij}) \tag{4.50}$$

between P and Q (or the information loss between Q and P). Matrix P is determined as that having maximum Shannon uncertainty, with just the values of $\bar{P}$ provided as constraint information. It is of interest to note that Fei uses a completely different argument to support a similar transformation from his aggregated matrix A* to an "equivalent" augmented matrix A^*_*.[35]

Consider that we have available the original matrix Q, with terms already normalized by some scheme, such that

$$\sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} = 1 \tag{4.51}$$

enabling Q to be regarded as a probability distribution. We wish to reduce this matrix by combining complete rows and the corresponding columns to form an m by m matrix $\bar{P}$. The transformation from Q to any particular $\bar{P}$ is given as

$$\bar{P} = S'QS \tag{4.52}$$

where the elements of the m by m matrix S are either zero or one, with a single one in each row located in the column corresponding to the position of that row of Q in the reduced matrix $\bar{P}$.

In order to determine the surrogate matrix P, we maximize the Shannon uncertainty of P, defined as U(P), in the form

$$U(P) = - \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} \log p_{ij} \tag{4.53}$$

with the only information provided being the terms of $\bar{P}$ itself, which enter into additivity constraints. In this way, an *information equivalent* version P of $\bar{P}$ is obtained, which has the dimensions of the original matrix Q. Letting I and J stand for a typical row and column of the aggregated matrix $\bar{P}$, the constraints are given in matrix form as

$$S' \; P \; S = \bar{P} \quad . \tag{4.54}$$

Upon differentiating Equation (4.53) with respect to each term p_{ij} under the relevant constraint from Equation (4.54), P may then be written as

$$P = S \; W \; S' \tag{4.55}$$

where the m by m matrix W contains terms (I, J) of $e^{-\lambda IJ}$, and λ_{IJ} is the Lagrange multiplier on the $(I, J)^{th}$ constraint in Equation (4.54). Now, defining a diagonal matrix D as

$$D = S' S \tag{4.56}$$

where a typical term d(I) represents the number of original sectors in aggregated sector I, the expression for P in Equation (4.55) may be substituted into Equation (4.54) to yield W. This is then resubstituted back into Equation (4.55) to finally give P as

$$P = S D^{-1} \bar{P} D^{-1} S' . \tag{4.57}$$

Then, Equations (4.52) and (4.56) enable Equation (4.57) to be given directly in terms of the original matrix Q as

$$P = V Q V \tag{4.58}$$

where the symmetrical matrix V is given as $S(S' S)^{-1} S'$.

Returning to Equation (4.50), the objective may be written in terms of the unknown transformation matrix S as

$$\min_{S} \sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} \log [q_{ij}/(S_{i.}(S'S)^{-1} S' Q S(S.S)^{-1} S_{j.}')] \tag{4.59}$$

where $S_{i.}$ and $S_{j.}$ are the i^{th} and j^{th} rows respectively of matrix S. Now, Equation (4.50) may be simplified by firstly using Equations (4.52) and (4.54) to write

$$S'_{.I} P S_{.J} = S'_{.I} Q S_{.J} = \bar{P}_{IJ} \tag{4.60}$$

where $S_{.I}$ and $S_{.J}$ are the I^{th} and J^{th} columns of matrix S respectively. If, in Equation (4.57), all terms in matrix $\bar{P}$ except $\bar{P}_{IJ}$ are set to zero, the contributions to matrix P are seen to comprise $[d(I).\ d(J)]$ identical terms of value $\bar{P}_{IJ}/\ [d(I).\ d(J)]$. The combination of these facts allows Equation (4.50) to be written as

$$I(Q:P) = \sum_{i=1}^{n} \sum_{j=1}^{n} (q_{ij} \log q_{ij} - p_{ij} \log p_{ij}) \tag{4.61}$$

or, alternatively, in the form

$$I(Q:P) = \sum_{i=1}^{n} \sum_{j=1}^{n} (q_{ij} \log q_{ij}) - \sum_{I=1}^{m} \sum_{J=1}^{m} \bar{P}_{IJ} \log [\bar{P}_{IJ}/(d(I).\ d(J))] \tag{4.62}$$

where Equations (4.52), (4.56) and (4.57) allow the expressions to be given entirely as a function of the original matrix Q and the sought transformation matrix S.

From Section 4.3.1, the expected information content U(x) for a probability distribution x is defined as

$$U(x) = - \Sigma\, x_i \log x_i \tag{4.63}$$

enabling Equation (4.61) to be given as

$$I(Q:P) = U(P) - U(Q)\ . \tag{4.64}$$

A coefficient of information loss C_I may then be defined as

$$C_I = [U(P) - U(Q)]/U(Q) \tag{4.65}$$

which is to be minimized in terms of S, via minimization of I(Q:P) in Equations (4.59), (4.61) or (4.62). The spread of C_I values for trial aggregation orderings S will provide some measure of the advantage of having a systematic approach to the choice of aggregation scheme, rather than merely relying on chance.

Theil proposed a clever separation of the information loss in his model into a column effect, a row effect and a cell effect, in order to analyse the effects of input and output heterogeneity.[36] An analogous separation of I(Q:P) has been reported elsewhere.[37] If we are primarily interested in minimizing aggregation bias, we should concentrate on the input heterogeneity criterion.

Two alternative approaches may be adopted to solve the above aggregation problem in practice. Firstly, when one has a good intuitive feel of the economic sectors which seem to belong together, one may form a list of preferred alternatives, and merely run them through a function evaluation via the relevant objective, choosing that which produces the minimum value. The information loss coefficients C_I should also be computed to check if they are reasonable. Secondly, in larger problems or cases when aggregation strategies are far from obvious, a more formal procedure must be employed. One obvious possibility would be to employ a random search method.

Although further work is required to see if any algorithms which work directly on minimization of I(Q:P) are feasible, one promising approach warrants attention. In a recent unpublished report, Marksjö outlines an algorithm which involves a multi-knapsack formulation.[38] An interesting feature of this approach is the capacity to handle rectangular input-output tables as well as the usual square matrices.

In order to clarify the manual application of the criteria developed in this section, two small test problems are described in Appendix D.

4.3.3 Adjusting Input-Output tables over time

We shall now return to the subject of biproportional matrices and changes in input-output coefficients over time.[39] In Chapter 2, we suggested that studies of changes in input-output relations over time have been frustrated by lack of comparable time-series data. To overcome this problem, Stone suggested a method of biproportional matrix adjustments for estimating the transactions matrix $\{p_{ij}\}$ for a specific year, by adjusting a known matrix $\{q_{ij}\}$ to fit the known row sums, $p_{i.}$, and column sums, $p_{.j}$, for that specific year.[40] This approach, termed the *RAS method*, determines elements in $\{p_{ij}\}$ by an equation of the form:

$$p_{ij} = r_i \, s_j \, q_{ij} \tag{4.66}$$

where the coefficients r_i and s_j are defined by the following iterative adjustment process (in which t numbers the iteration):

$$p_{ij}^{2t-1} = r_i^t \, p_{ij}^{2t-2} = r_i^t \, s_j^t \, p_{ij}^{2t-3} = \prod_{k=1}^{t} r_i^k \prod_{\ell=1}^{t-1} s_j^\ell \, p_{ij}^o \tag{4.67}$$

as

$$r_i^t = p_{i.} \left(\sum_j p_{ij}^{2t-2}\right)^{-1} \tag{4.68}$$

$$s_j^t = p_{.j} \left(\sum_i p_{ij}^{2t-1}\right)^{-1} \tag{4.69}$$

where $p_{ij}^{o} = q_{ij}$.

This adjustment process turns out to be identical to an iterative proportional fitting procedure (IPFP) proposed by Deming and Stephan to estimate cell probabilities $\{p_{ij}\}$ in a contingency table for which the marginal totals $p_{i.}$ and $p_{.j}$ are known.[41] This latter procedure operates as follows:

(1) Suppose we have n_{ij} observations in the (i,j) cell, where

$$\sum_i \sum_j n_{ij} = n \quad . \tag{4.70}$$

(2) At the (2t) step, we take

$$p_{ij}^{2t-1} = p_{ij}^{2t-2} \frac{p_{i.}}{p_{i.}^{2t-2}} \quad . \tag{4.71}$$

(3) At the (2t + 1) step, we take

$$p_{ij}^{2t} = p_{ij}^{2t-1} \frac{p_{.j}}{p_{.j}^{2t-1}} \quad . \tag{4.72}$$

(4) These iterations are continued until two successive sets of values for the cell probabilities agree closely.

Various proofs of convergence on a solution for $\{p_{ij}\}$ can be found in the literature.[42] From their respective formulations, it is a straightforward task to demonstrate the equivalence of the RAS and IPFP algorithms.

The IPFP was suggested originally as a means of arriving at estimates which minimized

$$\sum_i \sum_j \frac{(p_{ij} - q_{ij})^2}{q_{ij}} \tag{4.73}$$

subject to the known marginal totals, $p_{i.}$ and $p_{.j}$. Expression (4.73) is a modified Chi-square statistic, which is normally used to measure *goodness of fit*. Several authors have already noted that identical solutions to those obtained by minimizing (4.73) can be achieved by formulating the problem as one of minimum information gain,[43] namely

$$\text{minimize } I = \sum_i \sum_j p_{ij} \log (p_{ij}/q_{ij}) \tag{4.74}$$

subject to

$$\sum_j p_{ij} = p_{i.} \tag{4.75}$$

$$\sum_i p_{ij} = p_{.j} \tag{4.76}$$

$$p_{ij} \geq 0 .^{44} \tag{4.77}$$

The equivalence of these forms is evident from the observation that a classical Lagrangian derivation of a solution to the system (4.74) through (4.77) can be made to produce (4.66) directly. The relevant Lagrangian, , is given by

$$\mathcal{L} = \sum_i \sum_j p_{ij} \log (p_{ij}/q_{ij}) + \sum_i \alpha_i (\sum_j p_{ij} - p_{i.})$$
$$\sum_j \beta_j (\sum_i p_{ij} - p_{.j}) \tag{4.79}$$

where α_i and β_j denote the multipliers associated with constraints (4.75) and (4.76), respectively. After differentiation of $\mathcal{L}$, the solution follows as

$$p_{ij} = q_{ij} \exp (-1-\alpha_i-\beta_j) \ . \tag{4.80}$$

It is obvious that Equation (4.80) reduces to (4.66) if we make the simple substitutions

$$r_i = \exp (-1-\alpha_i) \tag{4.81}$$

and

$$s_j = \exp (-\beta_j) \ . \tag{4.82}$$

Explicit values of r_i and s_j may be determined using the constraint equations (4.75) and (4.76).

The RAS method is therefore equivalent to (i) the estimation of cell probabilities in two-way contingency tables given marginal totals, and

(ii) a doubly constrained model of minimum information gain.[45] We shall now attempt to identify those conditions under which one approach may prove superior to another. Since the RAS method can be regarded as a special case of contingency table analysis,[46] we shall restrict our comparision to two approaches:

(i) the estimation of cell probabilities in contingency tables, and

(ii) the family of models employing the principle of minimum information gain.

For the popular situation in which both row and column sums are known, or can be estimated independently, both approaches are feasible. Standard computer packages exist for both techniques.[47] Although comparative tests of both approaches have yet to be reported, we suspect that convergence would be rapid for either approach.

However, standard packages based on the iterative proportional fitting procedure are currently unsuitable for problems involving ad hoc sets of constraints. This does not imply that such procedures cannot be devised to handle different conditions. Modified versions of the RAS method have been developed for this very purpose. The problem is simply that contingency table analysis is designed to test hypotheses which conform to standard sets of equality constraints. In marked contrast, a standard information-minimizing algorithm is designed to accommodate any set of equality constraints. Accordingly, various sets of constraint information can be handled by the same routine *without any algorithmic modifications*. Algorithms have also been developed to handle a mixture of equality and inequality constraints.[48] The flexibility of these algorithms extends far beyond the boundaries of standard procedures devised for multidimensional contingency table analysis.

In conclusion, it appears that the minimum information principle has more general application because it is capable of incorporating additional or alternative information to that assumed in methods of biproportional matrix adjustment. Although this flexibility may not always be necessary when making comparative static adjustments, it is certainly needed when attempts are made to adjust these matrices over space. Further comparisons of the two approaches will therefore be undertaken, when we discuss the estimation of intra- and interregional flows in Chapter 5.

4.3.4 The Estimation of Intersectoral Commodity Flows

The publication of one or more historical matrices recording total intersectoral flows $\{x_{ij}\}$ within the national economy is now commonplace among most developed countries. These historical tables can be regularly updated using methods like those discussed in the previous section. Much less common, however, is the existence of a matrix describing the intersectoral flow of each commodity $\{x_{ijk}\}$. A typical element in this array defines the intermediate flow of commodity k from sector i to sector j. We shall now consider possible means by which information theory may assist in the estimation of this three-dimensional flow matrix.

Assuming that no *a priori* flow matrix is available, a simple approach to the generation of $\{x_{ijk}\}$ estimates would match directly the terms in a known intersectoral matrix $\{x_{ij}\}$, to their most likely counterparts in the absorption matrix $\{u_{kj}\}$ and the make matrix $\{v_{ik}\}$, taking due account in the make matrix of outputs destined for final demand. The intersectoral matrix immediately furnishes the constraints

$$\sum_k x_{ijk} = x_{ij} \qquad (i,j = 1,\ldots,n) \tag{4.83}$$

a corresponding absorption matrix the constraints

$$\sum_i x_{ijk} = u_{kj} \qquad (j = 1,\ldots,n;\ k=1,\ldots,m) \tag{4.84}$$

and a corresponding make matrix the constraints

$$\sum_j x_{ijk} = v_{ik} - y_{ik} \qquad (i = 1,\ldots,n;\ k = 1,\ldots,m) \tag{4.85}$$

where y_{ik} denotes the amounts of commodity k produced by sector i for final demand. Equation (4.85) simply signifies that a typical make matrix includes outputs that are destined for final as well as intermediate consumption.

The estimation of elements in $\{x_{ijk}\}$ is analogous to that of estimating the expected frequencies in a three-way contingency table with three sets of one-way marginal constraints. Standard results for three-way tables indicate that the problem can be solved directly, with the elements in $\{x_{ijk}\}$ being given by

$$x_{ijk} = x_{ij}\, u_{kj}(v_{ik} - y_{ik})/Q^2 \tag{4.86}$$

where

$$Q = \sum_i \sum_j x_{ij} = \sum_k \sum_j u_{kj} = \sum_i \sum_k v_{ik} \; . \qquad (4.87)$$

An identical result could be obtained from a standard Lagrangian solution to the entropy problem:

$$\text{maximize } U = - \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{m} x_{ijk} \log x_{ijk} \qquad (4.88)$$

subject to Equations (4.83), (4,84) and (4.85) as constraints.[50]

In the absence of any additional information, the direct solution afforded by Equation (4.86) is very convenient. However, if the set of available information includes suitable data on freight costs, or the physical capacities for handling deliveries, an additional set of constraints should be included. If we know freight costs, for example, we may have

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ik} \, x_{ijk} = C_k \qquad (k = 1,\ldots,m) \qquad (4.89)$$

where c_{ik} is the unit cost of delivering one unit of commodity k from sector i, and C_k is the total cost of delivering commodity k. In this instance, *standard* iterative procedures devised for three-way contingency tables are no longer suitable because of the appearance of weighted coefficients in the system of constraints. An entropy-maximizing formulation is then required to obtain the solution.

Having determined an historical estimate for each element in a base year array (which we shall henceforth refer to as $\bar{x}_{ijk}$), it is interesting to speculate on how this particular array might be updated to reflect existing or future intersectoral commodity flows. Macgill proposes a second stage of modelling in which the cells in the base year array $\{\bar{x}_{ijk}\}$ are updated, to derive new estimates $\{x_{ijk}\}$, by minimizing the information gain between the old and new arrays, namely

$$\text{minimize } I = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} x_{ijk} \log (x_{ijk}/\bar{x}_{ijk}) \tag{4.90}$$

subject to the following constraint:[51]

$$\sum_{i} x_{ijk} = b_{kj}(\sum_{i}\sum_{k} x_{jik} + y_j) \ . \tag{4.91}$$

In this formulation, b_{kj} are coefficients derived from the absorption matrix,[52] and y_j is the output of sector j which is destined for final consumption.

Equation (4.91) supposedly defines the assumed constancy of commodity inputs to total sectoral outputs. Rather curiously, elimination of the absorption coefficients from (4.9) and (4.12), using (4.84), leads to a slightly different relationship, namely

$$\sum_{i} x_{ijk} = b_{kj}(\sum_{i}\sum_{k} x_{ijk} + e_j) \tag{4.92}$$

A notable omission from Macgill's system is a set of constraints derived from the make matrix. Macgill argues that adoption of either the commodity technology assumption or the industry technology assumption should be avoided,[53] because

(i) inclusion of *a priori* values $\{\bar{x}_{ijk}\}$ adequately replaces assumptions embodied in the c_{ki} and d_{ki} coefficients; and

(ii) inclusion of either assumption determines an unique solution for $\sum_j x_{ijk}$, by reducing the system of make equations to one of (m by n) equations in (m by n) unknowns.

Both these arguments appear rather tenous. Firstly, there is no empirical evidence available to support her implicit suggestion that *a priori* flow values can more aptly replace assumptions embodied in the base-year c_{ki} and d_{ik} coefficients, than those embodied in the corresponding technical coefficients, b_{kj}. Secondly, inclusion of an assumption like (4.92) determines an unique solution for $\sum_i x_{ijk}$, by reducing the system of absorption equations to a similar system of (m by n) equations in (m by n) unknowns. In other words, the inclusion of assumptions based originally on either technology assumption appears no more restrictive than those embodied in Equations (4.92).

There is some evidence to suggest that the industry-commodity coefficients (b_{kj}, c_{ki} and d_{ik}) may be less stable than their intersectoral counterparts (a_{ij}). This stems mainly from the aggregate nature of intersectoral flows, which disguises the overall effects of individual changes in commodity technologies. It may therefore be wiser to revert to entropy-maximizing models for all $\{x_{ijk}\}$ estimates, rather than to speculate on the use of a two-stage model which attempts to perpetuate the base-year technologies.

4.3.5 The Estimation of Capital Coefficients

In Section 4.2.3 of this chapter, we introduced the basic elements of dynamic input-output theory. To overcome the scarcity of survey-based capital coefficients, a direct method of estimation was proposed.[54] To implement this method, however, an assumption was required concerning the proportional relationships between any two elements in the same row (or delivery sector i) of the capital matrix. In this section, a simple information-minimizing model will be formulated wherein no such assumption is required.

We shall commence where we left off in Section 4.2.3, by considering the declining productivity of capital goods over time. Equation (4.38) will be repeated here for convenience:

$$x_i = \sum_{j=1}^{n} a_{ij}^{1} x_j + \sum_{j=1}^{n} b_{ij} (q_i + \lambda_j) x_j + y_i \qquad (i = 1,\ldots,n) . \qquad (4.38)$$

We also note that this can be simplified to

$$\sum_{j=1}^{n} \alpha_{ij} b_{ij} = \beta_i \qquad (i = 1,\ldots,n) \qquad (4.39)$$

where α_{ij} and β_i are coefficients which can be computed in advance.

Equations (4.39) will first be complemented by a description of capital useage, namely

$$\sum_{i=1}^{n} \alpha_{ij} b_{ij} = I_j \qquad (j = 1,\ldots,n) \qquad (4.93)$$

where I_j is the total investment demand for capital goods by sector j.

In the absence of any further information relating to the unknown coefficients $\{b_{ij}\}$, we could estimate them directly using a standard two-factor model of independence,[55] namely

$$b_{ij} = \frac{\beta_i \, I_j}{\alpha_{ij} \, I} \qquad (i,j = 1,\ldots,n) \qquad (4.94)$$

where $I = \sum_{i=1}^{n} \beta_i = \sum_{j=1}^{n} I_j$.

As we suggested earlier, the accuracy of any estimates can partly be gauged by comparisions with existing capital stocks in each sector. In reality, most sectors are not operating at full capacity, so Equations (4.35) will not strictly hold. It is more appropriate to regard the existing capital stock in each sector, K_i, as a capacity constraint, namely

$$\sum_{j=1}^{n} b_{ij} \, x_j \leq K_i \quad . \qquad (i = 1,\ldots,n) \qquad (4.95)$$

This introduction of inequality constraints (4.95) necessitates an entropy-maximizing approach. Algorithms have been developed recently to solve problems containing both equality and inequality constraints. Thus the capacity constraints can be included in the following formulation:

$$\text{maximize } U = - \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} \log b_{ij} \qquad (4.96)$$

subject to the constraints (4.39), (4.93) and (4.95), together with the usual nonnegativity conditions.

Unfortunately, this maximum entropy estimate is only satisfactory if we ignore the real influence of time. In reality, current investments are usually made well in advance of the resultant production equipment being ready for use. In other words, our assumed β_i and I_j values do not necessarily relate to the same time period as the capacities, K_i. A number of models which allow for various gestation lags (together with finite durabilities and declining productivity of capital equipment) have been proposed.[56] We do not intend to review these nonlinear models at this present stage. Instead we shall describe a simple modification to our maximum entropy estimate which introduces some temporal reality.

In the first stage of our modified approach, we determine the elements of an array $\{\bar{b}_{ij}\}$, which describes the capital flows (coefficients) during our base period, using Equation (4.94). This base period is conveniently chosen to precede our real period of interest by the average construction or gestation period for all sectors and types of capital equipment. The matrix $\{\bar{b}_{ij}\}$ respresents an *a priori* hypothesis that base period coefficients relate more to base period investments than to base period capacities.

In our second stage of estimation, the base period coefficients are adjusted to satisfy existing capacity constraints at the end of the gestation period. This is done using the principle of minimum information gain, namely

$$\text{minimize } I = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} \log (b_{ij}/\bar{b}_{ij}) \tag{4.97}$$

subject to the inequality constraints (4.95) and the usual nonnegativity conditions.

In practice, the assumption of a uniform gestation period for all sectors and types of capital equipment is unrealistic. Different construction periods may be very important, ranging from less than one year up to some seven or more years.[57] To introduce these variations, we could decompose b_{ij} into elements b_{ij1}, $b_{ij2}, \ldots, b_{ijT}$ such that

$$\sum_{\theta=1}^{T} b_{ij\theta} = b_{ij} \tag{4.98}$$

where $b_{ij\theta}$ is the capital input which must be delivered θ periods before the new production capacity, of which it is a part, is ready for use ($\theta=1, \ldots\ldots, T$). Accordingly, T is the longest gestation period. However, the introduction of this new subscript (θ) necessitates an evaluation of the complete time profile of each input into capital construction, which is a task considered beyond the scope of this dissertation.

FOOTNOTES FOR CHAPTER 4

1 See Leontief (1951, 1953).

2 The former measure can be found in Theil (1967, p 333), whereas the latter is based on Equation (2.15).

3 The Iterative Proportional Fitting Procedure (IPFP) was developed originally for contingency table analysis, but may also be applied to the adjustment of input-output tables.

4 Canadian tables have been published by the Dominion Bureau of Statistics (1969), whereas rectangular accounts of the Norwegian economy have appeared in Statistisk Sentralbyrå (1978). For an outline of the United Nations proposal, see either United Nations (1968) or Aidenoff (1970).

5 Similar tables have appeared in Aidenoff (1970), Gigantes (1970), and Macgill(1978).

6 An assumption of this type is central to the development of a Norwegian input-output model known as MODIS IV, which distinguishes between commodities, sectors and activities. For a detailed outline of this model, see Bjerkholt and Longva (1980). The assumption itself has been thoroughly investigated using Norwegian data by Sevaldson (1970; 1972; 1974).

7 This inverse is commonly referred to as the *Penrose inverse* in memory of its proponent. For further details, see Penrose (1956).

8 See, for example, Gigantes and Matuszewski (1968), Aidenoff (1970), or Gigantes (1970).

9 Namely, Equations (4.8) through (4.18).

10 This array is actually an aggregated form of an even larger array which appeared originally in Cripps, Macgill and Wilson (1974).

11 See Leontief (1953).

12 Lange (1957) defines this process as *productive investment*.

13 See Leontief (1970).

14 See, for example, Jorgenson (1960), Tsukui (1961; 1968) and Petri (1972).

15 Early efforts to develop capital coefficients were made by the Harvard Economic Research Project and the Interindustry Analysis Branch of the U. S. Bureau of Mines. The Harvard group's work dates back to at least 1948, and was terminated in 1972. The Bureau groups have made estimates of incremental capital coefficients for many sectors. Fixed capital coefficients for the U. S. economy in 1939 were computed by Grosse (1953). For a review of some of the problems encountered in these efforts, see Carter (1957).

16 Brődy (1966; 1970) developed this simplified form from the original dynamic model proposed by Leontief (1953). The first closed-model computations were made at Harvard in the early fifties. Their results, and Tsukui's (1966) earlier work, were never published.

17 In discussing a similar model, Hawkins (1948) noted that the quotient of the respective elements of the two matrices, b_{ij}/a_{ij}, is of the dimensionality of time.

18 Lange (1957) did not distinguish between durability, a physical characteristic of capital goods, and turnover time which is an economic characteristic. Brődy (1966; 1970) substituted physical life spans for turnover times, taking some of his guesses from Domar (1957). Lange's interpretation of the concept of turnover time is worthy of recollection. Let the durability of that part of the output of sector i which is allocated to sector j as additional means of production be t_{ij} units of time. In order to produce a unit of output from sector j during a unit period of time, the quantity a_{ij} of products from sector i must be used up during that period of time; a_{ij} is the technical coefficient. Thus to increase (in the next period) the output of sector j by one additional unit, a quantity of output $a_{ij} \cdot t_{ij}$ from sector i must be allocated to sector j. In this way, exactly a_{ij} of output from sector i will be used up in the next unit period in sector j, and this will produce one unit of output.

19 This method was originally proposed and evaluated in Batten (1979), and an extension was suggested in Batten (1981).

20 This rate of expansion corresponds to the von Neumann growth rate. The associated solution vector represents a unique equilibrium structure which is often referred to as the *turnpike* solution. For a discussion of turnpike theorems, and their application to multisectoral models, see Tsukui (1966, 1979), Murakami et al. (1970), and Tokoyama et al. (1976). See also Appendix C.

21 Since there are n trivial equations in which $j=1$.

22 As defined in Equations (4.28).

23 Stone (1962) suggested a method of biproportional matrix adjustment to modify a known matrix in order to fit new row and column sums. Bacharach (1970) elaborated further on this approach. Equation (4.32) is really an *independence* model, since *a priori* coefficients are not available. For further information on the latter, see Fienberg (1977).

24 See Bródy (1966).

25 As suggested, for example, by Wurtele (1960) and Morishima (1964). Wurtele divided the economy into two industries: consumption goods and capital goods. He then proposed that $a_{ij} = q_i b_{ij}$ for the capital goods industries alone. Equations (4.37) are more general, in that they allow the goods produced in any sector to be used for intermediate consumption, capital maintenance, or both.

26 Theil (1967, p 331) acknowledges Skolka (1964) as the first proponent. Skolka's work appeared in Czech with only a summary in English. For other examples, see Theil and Uribe (1965) or Tilanus and Theil (1965).

27 Which was introduced originally in Section 4.2.1.

28 See Theil (1967, p 332).

29 This latter expression represents the contingency table solution.

30 Under the assumption of complete ignorance, S attains its maximum value, namely log n^2.

31 See Hatanaka (1952).

32 See, for example, Balderston and Whitin (1954), Theil (1957), Ara (1959), and Morimoto (1970; 1971).

33 See, for example, Balderston and Whitin (1954), Fischer (1958), and Kossov(1970).

34 See Kullback (1959).

35 See Fei (1956).

36 See Theil (1967, pp 331-8).

37 See Roy, Batten and Lesse (1981).

38 See Marksjö (1981).

39 See Bacharach (1970).

40 See Stone (1962).

41 See Deming and Stephan (1940).

42 See, for example, Ireland and Kullback (1968), or Fienberg (1970).

43 Notably Ireland and Kullback (1968), Bacharach (1970), Kādas and Klafszky (1976), and Macgill (1977). The proposed measure of information gain is equivalent, at its first-order approximation, to the sum of the squares of the relative deviations. The following proof is due to Kādas and Klafszky :

For $|x| \ll 1$, we have

$$\log (1+x) \cong x .$$

The following simplification is therefore possible:

$$\begin{aligned}
\sum_i \sum_j p_{ij} \log \left(\frac{p_{ij}}{q_{ij}} \right) &= \sum_i \sum_j p_{ij} \log \left(\frac{p_{ij}-q_{ij}}{q_{ij}} + 1 \right) \\
&\cong \sum_i \sum_j p_{ij} \left(\frac{p_{ij}-q_{ij}}{q_{ij}} \right) \\
&= \sum_i \sum_j \frac{(p_{ij}-q_{ij})^2}{q_{ij}} + \sum_i \sum_j (p_{ij}-q_{ij}) \\
&= \sum_i \sum_j \frac{(p_{ij}-q_{ij})^2}{q_{ij}} .
\end{aligned}$$

44 Note that $p_{ij} \log(p_{ij}/q_{ij}) = 0$ whenever $p_{ij} = 0$, and $p_{ij} = 0$ whenever $q_{ij} = 0$.

45 Which may alternatively be described as an entropy-maximizing model incorporating non-uniform prior probabilities.

46 In which the marginal probabilities, $p_{i.}$ and $p_{.j}$, are known.

47 See Haberman (1973) or Fay and Goodman (1975) for algorithms based on the Iterative Proportional Fitting Procedure. Eriksson (1978; 1980) has developed a versatile information-minimizing algorithm.

48 See, in particular, Eriksson (1980).

49 See Fienberg (1977, p 25).

50 A normalization constraint, corresponding to Equation (4.47), is redundant here, and has therefore been omitted.

51 Macgill (1978) actually maximizes the function

$$S''' = -\sum_i \sum_j \sum_k x_{ijk} \log(x_{ijk}/\bar{x}_{ijk})$$

which is, of course, identical to (4.90).

52 See Equation (4.12).

53 See Equations (4.13) and (4.14), respectively.

54 See Equations (4.32) and (4.41).

55 See, for example, Fienberg (1977, p 12).

56 See, for example, Johansen (1978), and Åberg and Persson (1980).

57 See Johansen (1978).

Chapter 5

INTERSECTORAL FLOWS IN SPACE: STATIC FORMULATIONS

5.1 Introduction

Intraregional and interregional input-output models both represent spatial disaggregations of the standard Leontief model introduced in Chapter 4. Leontief's original model was developed primarily for a national economy. A regional economy is, however, much more open or trade-oriented than its national counterpart. Interregional transactions therefore play a significant role in determining the structure of economic activity within any single region or system of regions.

The purpose of the present chapter is essentially twofold:

(i) to examine various non-survey techniques, both *intra-* and *inter-regional*, which have been adopted for the spatial estimation of intersectoral flow coefficients; and

(ii) to suggest certain alternative approaches to these estimation problems based on elements of information theory.

In Section 5.2, we begin by re-examining Leontief's original *static* model from an intra- and interregional viewpoint. Although the intraregional input-output model is of a comparable form to the national model, important structural differences exist between the two formulations. For example, because each region is largely dependent on other regions for its economic survival, detailed account must be taken of the trading patterns between regions. To analyse these regional interdependencies meaningfully, an interregional approach is necessary. Section 5.2.2 introduces the basic flow relationships of the static interregional input-output model.

In Section 5.3, various non-survey techniques which have been adopted for the spatial estimation of input-output coefficients are discussed. For the estimation of intraregional coefficients, the methods outlined may be broadly categorized as

(i) location quotient approaches;
(ii) commodity balance approaches;
(iii) biproportional matrix adjustments; and
(iv) semi-survey techniques.

Because of their inability to make allowances for cross-hauling, it appears that the first three methods underestimate regional imports and exports. Consequently, no acceptable non-survey method can be cited for deriving intraregional coefficients from their national counterparts.

For the estimation of interregional coefficients, two approaches are reviewed, namely (i) quotient methods; and (ii) efficient information adding. The latter approach, which can involve two stages of estimation, appears quite promising for a number of reasons. Not the least of these is its ability to cope with cross-hauling. It represents a practical compromise between two extremes, namely full survey methods and theoretical trade models.

Section 5.4 proposes three alternative approaches to the estimation of interregional, intersectoral flows from a limited database of industrial and regional information. For each method, a distinction is made between flows to intermediate and to final demand. In contrast to earlier approaches which have adopted various *a priori* flow assumptions, the present study investigates four different cases describing the extent to which information is available on intraregional demands.

In the first approach, each case is treated as a simple form of hypothesis testing, in which the expected frequencies in a four-dimensional contingency table are estimated using various sets of marginal constraints. In these formulations, flow capacities at the different nodes are not taken into account, and supplies and demands are regarded as pooled. By dealing with this simple interregional accounting system, it is possible to solve all four cases using a standard iterative proportional fitting procedure. Furthermore, an explicit solution for two of these cases is possible, since their maximum-likelihood estimates can be expressed in closed form. It is concluded that the extent to which information about intraregional demands is available determines the viability of an explicit solution.

If we now add a set of *capacity constraints* to our interregional accounting system, the iterative procedures devised for multidimensional contingency table analysis are no longer appropriate. The introduction of these constraints prompts our second approach, namely adoption of the entropy-maximizing paradigm. This approach is sufficiently flexible to accommodate all types of linear constraints, including those on handling capacities at the nodes, which involve inequalities. It is also possible to specify some of the flows in advance without any need to modify the solution algorithms. Explicit solutions are found numerically using iterative methods.

The existence of an *a priori* flow pattern allows us to relax the equi-probability assumption associated with the entropy-maximizing paradigm. In modifying our objective function to allow for non-uniform prior probabilities, we introduce our third approach, which employs the principle of minimum information gain. Although this technique enjoys most of the advantages possessed by the entropy-maximizing paradigm, it also embodies some important differences. These contrasting features are compared near the end of the chapter.

5.2 Basic Model Characteristics

5.2.1 Intraregional Input-Output Models

Consider a single region r whose economy is divided into n production sectors. Denote by x_i^r the total output of sector i in region r, and by x_{ij}^r the intermediate demand by sector j in region r for goods produced by sector i in the same region. Further denote by y_i^r the local final demand for goods produced by sector i in region r. The overall input-output balance of this regional economy can be described in terms of our characteristic set of n linear equations:

$$x_i^r = \sum_{j=1}^{n} x_{ij}^r + y_i^r \quad . \qquad (i = 1,\ldots,n) \tag{5.1}$$

The input-output structure of any particular sector j in region r can be described by a vector of technical coefficients, a_{ij}^{r}, each of which states the amount of a particular input produced by sector i in region r which is absorbed by sector j per unit of its own regional output. This implies the following set of structural relationships:

$$x_{ij}^{r} = a_{ij}^{r} \; x_{j}^{r} \qquad (i,j = 1,\ldots,n) \qquad (5.2)$$

and leads to the basic equilibrium equation:

$$x_{i}^{r} = \sum_{j=1}^{n} a_{ij}^{r} \; x_{j}^{r} + y_{i}^{r} \quad . \qquad (i = 1,\ldots,n) \qquad (5.3)$$

In this familiar form, input-output analysis is often used to determine future levels of production in the endogenous sectors (x_{i}^{r}), given exogenously determined levels of final demand (y_{i}^{r}), and assuming constant technical coefficients (a_{ij}^{r}). This is done by forming the standard Leontief inverse.[1]

Although the comparable relationships are identical, a number of important structural differences exist between this regional model and its national counterpart. Because intraregional tables are more open than the national table to which they correspond, exports and imports account for a larger share of total transactions in the region than in the nation. So, the size of the import coefficient in any given column of the intraregional matrix may be quite large, causing local input coefficients in the same column to fall below those in the national table. For this reason alone, it is easy to understand why the adoption of national coefficients in regional models can be misleading. Clearly, there are wide variations in export and import patterns from region to region.

Miernyk acknowledged the importance of these variations when he suggested extensions to the basic model designed to separate the technical requirements of local industries from the interregional trade patterns of the economy.[2] The result was the so-called *dog-leg* input-output model, the accounting system for which is depicted in Figure 5.1. This variant of Leontief's original model recognizes explicitly that the regional economy is not a self-contained production-consumption entity. Consequently, the coefficients underlying the intraregional transactions matrix in Figure 5.1 can be redefined as intraregional requirements coefficients.

Further concern often centres on the assumption of coefficient stability. Several factors can cause the coefficients to alter over time: technological change, variations in product mix, price changes, input substitutions, and shifts in trade patterns.[3] The question of coefficient instability, however, is essentially an empirical one. Since we are not concerned with long-term forecasting using static models, the question of coefficient stability will not trouble us here. On the other hand, the interindustry structure does appear to be sensitive to short-run disturbances in a region's propensity to import,[4] so an accurate picture of the trading patterns existing between regions is essential. To examine these regional interdependencies, a full interregional approach is required.

Figure 5.1

	Intermediate demands	Final demands		
Intermediate inputs	Intraregional transactions matrix	Exports to other regions	Other final demands	Gross output
Primary inputs	Imports from other regions			
	Other primary inputs			
	Gross input			

5.2.2 Interregional Input-Output Models

The original static interregional model was formulated by Isard, whereas the scheme proposed by Leontief was designed specifically as a theory of intranational relationships.[5] Since those early days, a number of other approaches have emerged in an attempt to develop an empirically workable model. We shall examine some of these proposals a little later. The present introduction will be limited to the basic structural and flow relationships suggested by Isard.

Consider a system of regional economies containing n production sectors which are distributed over m regions. Denote by x_{ij}^{rs} the intermediate flow from sector i in region r to sector j in region s, and by y_i^{rs} the deliveries from sector i in region r to final demand in region s. The basic set of interregional flow relationships for intersectoral balance can be specified in the form

$$x_i^r = \sum_{j=1}^{n} \sum_{s=1}^{m} x_{ij}^{rs} + \sum_{s=1}^{m} y_i^{rs} \quad . \quad \begin{pmatrix} i=1,\ldots\ldots,n; \\ r=1,\ldots\ldots,m \end{pmatrix} \tag{5.4}$$

We can add to this the corresponding inflow of factors needed to satisfy regional production requirements in each sector, namely

$$x_j^s = \sum_{i=1}^{n} \sum_{r=1}^{m} x_{ij}^{rs} + v_j^s \qquad \begin{pmatrix} j=1,\ldots\ldots,n; \\ s=1,\ldots\ldots,m \end{pmatrix} \tag{5.5}$$

where v_j^s is the value added to sector j in region s. The input-output structure of a particular sector j in any individual region s is then described by a matrix of interregional technical coefficients. Each element, a_{ij}^{rs}, states the amount of goods produced by sector i in region r which are absorbed by sector j per unit of its own output in region s. This implies the following set of structural relationships:

$$x_{ij}^{rs} = a_{ij}^{rs} \; x_j^s \qquad \begin{pmatrix} i,j=1,\ldots\ldots,n; \\ r,s=1,\ldots\ldots,m \end{pmatrix} \tag{5.6}$$

Substituting (5.6) into (5.4) yields

$$x_i^r = \sum_{j=1}^{n} \sum_{s=1}^{m} a_{ij}^{rs} x_j^s + \sum_{s=1}^{m} y_i^{rs} \quad . \qquad \begin{matrix}(i=1,\ldots\ldots,n;\\ r=1,\ldots\ldots,m)\end{matrix} \qquad (5.7)$$

If we assume constant technical coefficients and a given pattern of final demands, this system of equations can be solved for regional outputs. The solution is obtained by forming the standard Leontief inverse. The model is regarded as an *ideal interregional input-output archetype,*[6] whose empirical implementation has been restricted severely by inadequate data describing the interregional coefficients (a_{ij}^{rs}) themselves.

As we shall see shortly, elements of information theory can assist us to obtain reasonable estimates of these coefficients. But before turning to these, it is appropriate to review some of the earlier attempts to estimate intra- and interregional input-output tables. We shall begin with intraregional estimation.

5.3 Existing Nonsurvey Techniques for Estimating Spatial Flows

5.3.1 Intraregional Coefficients

The various approaches to the estimation of intraregional input-output tables can be categorized rather broadly as *survey* or *non-survey* techniques. Although it is generally agreed that there is still no acceptable substitute for a survey-based study, there have been numerous articles promoting or assessing the feasibility of various non-survey methods.[7] Most of these non-survey methods attempt to adapt national coefficients for regional purposes, an approach which has much in common with similar attempts to adjust for *temporal* changes in the national tables.[8]

Some authors have been extremely critical of the use of national coefficients in regional estimation.[9] We are generally in agreement with much of this criticism, since it is rather unlikely that adjustment of national figures can take all the pertinent regional influences into account. Nevertheless, many of these non-survey techniques seem likely to proliferate on purely pratical grounds, so we shall briefly examine the following methods which already exist in the literature:

(i) location quotient approaches;
(ii) commodity balance approaches;
(iii) biproportional matrix adjustments; and
(iv) semi-survey techniques.

The first three procedures have been applied and compared extensively. All three attempt to adjust the national coefficients to the regional level, by assuming that intraregional trade coefficients differ from their national counterparts only by the magnitude of the regional import coefficient. More explicitly,

$$r_{ij} = a_{ij} - m_{ij} \tag{5.8}$$

where a_{ij} is the national input-output coefficient, r_{ij} the regional trade coefficient, and m_{ij} is a regional import coefficient. None of these techniques can therefore account for region-specific product mixes or local production funtions.[10]

5.3.1.1 Location Quotients

The *location quotient* is a measure which compares the relative regional importance (usually measured in terms of output or employment) of an industry to its standing at the national level. In its simplest form, this quotient is defined for industry i as

$$LQ_i = (X_i^r/X^r)/(X_i/X) \tag{5.9}$$

where X_i denotes output (or employeemnt) in sector i, X denotes total output (or employment), and r denotes regional values. When $LQ_i > 1$, the assumption is that intraregional requirements for commodity i can all be met locally. The national technical coefficients are thus used in row i of the regional trade coefficients matrix, that is, $r_{ij} = a_{ij}$ for all j. If $LQ_i < 1$, then $r_{ij} = a_{ij}(LQ_i)$ for all j.

To overcome some of the deficiencies in the simple location quotient approach, various other quotients have been suggested. *The purchases-only* location quotient ensures that the summation of total output (or employment), in the calculation of the quotient for industry i, is confined to those industries which make purchases from industry i. The *cross-industry* location quotient compares the national output proportion of selling industry i, in the region, to that of purchasing industry j. The *logarithmic cross-industry* location quotient refines the cross-industry approach, by considering the relative size of the region compared to the nation.

5.3.1.2 Commodity balances and supply-demand pooling

The commodity balance approach was first suggested by Isard in 1953, and relies upon calculating the balance, b_i, between the local output of commodity i, namely x_i^r, and the local demand for commodity i, d_i; the latter being estimated using national coefficients. The net surplus (deficit), or *commodity balance,* for each industry is obtained by simple subtraction:

$$b_i = x_i^r - d_i \quad . \tag{5.10}$$

The resulting table indicates whether each product should be imported into or exported from the region. It does not, however, indicate the origins of inputs, so further assumptions are required to complete the intraregional table.

Extensions to Isard's work were accomplished by a technique know as supply-demand pooling.[11] When b_i is positive, the national coefficients are substituted into the appropriate row of the regional matrix, imports are set at zero, and exports are then calculated. If b_i is negative, the national coefficients of row i are reduced by the factor (x_i^r/d_i), exports are set at zero, and imports may then be computed. Further modifications of supply-demand pooling have also been suggested, to allow for predetermined estimates of final demand. Schaffer and Chu introduced an additional refinement by employing an iterative approach. Their RIOT (Regional Input-Output Table) simulation attempts to distribute local production according to both the national sales pattern and local needs.

5.3.1.3 Biproportional matrix adjustments

Various biproportional matrix methods have been used to adjust national coefficients until they conform to predetermined sectoral totals of intermediate inputs and outputs for each region. Of all these techniques, the RAS method has received the most attention in the literature.[12] This latter method has already been described in the context of matrix adjustments over time,[13] where it was suggested that the approach is equivalent to both the estimation of cell probabilities in two-way contingency tables given marginal totals, and a doubly-constrained model of minimum information gain.[14]

The virtues of the RAS method lie in its simplicity and economy. Whenever it has been used for the estimation of regional coefficients, it has consistently outperformed the two approaches discussed earlier.[15] Yet the biproportional assumption has no special economic meaning; there is no reason to believe that input coefficients should change in a uniform manner along each row and column. Consequently, a number of writers have opposed the use of this method for the estimation of intraregional coefficients.[16]

5.3.1.4 Semi-survey techniques

Attempts have been made to improve on the basic RAS method by incorporating a limited amount of survey information. Other semi-survey techniques have also been proposed. Su proposed a technique which relies on the direct determination of regional imports by survey.[17] Schaffer outlined a framework within which a nonsurvey procedure could be improved by direct estimation of regional exports, and insertion of known local transactions.[18] But, perhaps the most ambitious attempt to develop a semi-survey approach has been completed recently by Jensen and his associates.[19] Their method, known as the GRIT technique, applies various adjustments to the national table to allow for prices, international trade, and regional imports; and then advocates the systematic insertion of *superior data* whenever reliable flow statistics are available. Their system has been adopted in a number of empirical studies for Australian regions.

5.2.1.5 Discussion

A major weakness of the location quotient and commodity balance approaches is that balances are calculated in net terms. The possibilities of accounting for cross-hauling (the import and export of similar goods by the same sector) are denied, thereby implying that intraregional trade is maximized.[20] Consequently, these two approaches (along with the RAS method) tend to underestimate regional imports and exports.

An acceptable method of deriving intraregional input coefficients from their national counterparts has yet to be formulated. To develop a realistic representation of the technical structure existing inside any single region, proper cognizance must be taken of various development patterns occuring *outside* the region. To this extent, an acceptable hybrid model should contain a certain amount of interregional survey material to supplement any nonsurvey technique. This interregional data would enable the model to accurately define regional imports and exports, which are the key to successful estimation. If this trade data could be supported by survey information about industries in which the regional economy specializes, a promising hybrid model might emerge. To the author's knowledge, such an approach has not yet been reported.

To take proper account of the interdependent trading patterns which exist between regions, a full interregional analysis is required. Faced with the daunting task of finding sufficient data to implement an interregional model accurately, it is no surprise to find that most regional economists have concentrated on less demanding problems. In the long run, however, it is the author's firm belief that the interregional approach is the only one which can provide satisfactory results in the spatial context. Accordingly, we shall now evaluate the few attempts which have been made to develop non-survey estimates of interregional tables.

5.3.2 Interregional Coefficients

The key to the successful calibration of any interregional model is the extent to which information on the interindustry flows *between* regions is available, toghether with the accuracy and consistency of these interregional statistics. In the absence of survey information, various models have been used to estimate the interregional trade flows for aggregate commodity groups. Models incorporating fixed column coefficients were the forerunners of a linear programming approach, but the empirical results were not very impressive.[21] In cases where heterogeneous products are combined within sectors, the effects of cross-hauling appear to rule out linear programming approaches.

An important contributionoccurred in 1963, when Leontief and Strout proposed their gravity trade model; which requires the bare minimum of information, but allows for cross-hauling between regions.[22] This elegant multiregional model has been adopted in a number of empirical studies,[23] and served as useful vehicle for Wilson's classic integration of the gravity and input-output assumptions using entropy-maximizing principles. In later formulations, further use will be made of their notion of *supply-demand pooling*.

Few other analysts have enjoyed comparable access to the comprehensive data on interstate commodity shipment available to Polenske. As a result, many other modelling exercises have been forced to rely on crude or inadequate data, together with some rather arbitrary assumptions.[24] The results of these exercises must therefore be treated cautiously.

Among the more recent contributions to the estimation of interregional trade flows, two nonsurvey approaches to the static problem warrant special attention. The first is based upon multiregional extensions of location quotient methods, whereas the second involves efficient information adding. We shall begin with the quotient approach.

5.3.2.1 Quotient methods

It has been argued that quotient methods are typical representatives of a general class of techniques which include all iterative and commodity balance methods.[25] Building upon earlier work which basically followed a commodity balance approach,[26] Round defines five different quotients to determine alternative sets of trading coefficients. The general procedure is to consider a quotient, α_{ij}^{r}, defined for every pair of sectors i and j in region r. The value of α_{ij}^{r} identifies import or export orientation of the sector i in supplying the needs of sector j. A value less than one indicates import orientation, whereas a value greater than one implies exports.

The method devised relates this quotient, α_{ij}^{r}, to the trading coefficient, t_{ij}^{r}. If α_{ij}^{r} is less than one, t_{ij}^{r} is set equal to it. If, on the other hand, α_{ij}^{r} indicates self-sufficiency (export orientation), then this is reflected by a trading coefficient of unity; implying that all the local demand from sector j for the products of sector i are satisfied by local production. Herein lies a major weakness of the quotient approach. Commodity exports are invariably ascertained as a residual, after final and intermediate sales have been deducted from gross sales.

In extending the quotient approach to a system of two regions, Round develops a two-stage procedure, which allows initial estimates of the trading coefficients to be adjusted to satisfy certain consistency constraints imposed by the full interregional system. The result is a consistent and fully balanced set of intra- and interregional intermediate transactions. This two-stage approach represents a significant conceptual improvement to the estimation process. Not surprisingly, the validity of the quotient approach for nonsurvey estimation cannot be demonstrated using Round's simple two-region model. What is amply demonstrated, however, is the value of *consistency constraints* describing the full interregional system. A single region input-output framework has none of these advantages. Substantial improvements to intraregional estimation may therefore be possible by simply considering each region as part of a two-region model: the region itself and the rest of the world.

Although two-region models make small demands for data, they inevitably understate the true extent of interregional feedbacks and spillovers. In a genuine interregional model, the basic requirement is that all regions in the system be treated equivalently and directly, leading normally to consideration of a large number of regions. There is a fundamental need for a reliable non-survey method to estimate trading relationships within the full interregional system which is also consistent with the information already available on trade flows. Certain techniques based on information theory show considerable potential, so a critical review of two earlier formulations follows.

5.3.2.2 Efficient information adding

In order to adopt the principle of minimum information gain, a complete *a priori* matrix of interregional flows is required.[27] The collection of flow statistics generally published does not furnish this complete matrix. In the presence of incomplete information, some assumptions are needed to estimate the unknown elements.

A complete *a priori* matrix may be estimated in a number of ways. For example, Bigsten adopted an assumption which closely resembles the gravity hypothesis, whereas Snickars regressed certain log-linear models which, he argued, contain gravity-type models as particular cases.[28] As a result, Snickars derived his *a priori* estimates using the contingency table approach. In a manner similar to the second stage of Round's estimation procedure, Bigsten and Snickars both used the principle of minimum information gain to achieve consistency between their *a priori* matrices, and exogenously determined data expressed in the form of constraints. This consistency was achieved by an efficient addition of information to complete the *a posteriori* matrix.

This two-stage process of *efficient information adding* warrants further investigation. Firstly, there is a fundamental need for consistency between the first and second stages of the estimation. Adoption of particular assumptions, like the gravity hypothesis, for the estimation of *a priori* values, is hardly consistent with the *unbiased* nature of the second stage. In the absence of an historically complete *a priori* matrix of flows, we shall later advocate use of the entropy-maximizing paradigm for the first stage of estimation. In so doing, we then ensure that both stages remain "maximally non-committal with regard to missing information".[29]

Secondly, there is a need for consistency with respect to the treatment of interregional deliveries to satisfy intermediate and final demands. Both Bigsten and Snickars derived their *a priori* estimates of final demand deliveries using assumptions which were neither consistent with those adopted to estimate intermediate deliveries, nor suitably unbiased. We shall attempt to overcome these deficiencies by treating intermediate and final demands simultaneously in the same model.

Finally, there is a need to recognize that regional production levels in various sectors may differ vastly from the region's internal demand for the products from those same sectors. This difference assumes considerable importance owing to the greater propensity of regions to import a high percentage of their needs, and to export goods from only a few sectors. These intraregional imbalances have not been considered directly in either Bigsten's or Snickars' formulations. We shall attempt to examine their effects by investigating four different assumptions concerning intraregional demands.

It should be emphasized from the outset that the reservations expressed above apply only to the *first* stage of the estimation process. We have no theoretical objections to the second phase, namely efficient information adding. Although adoption of the principle of minimum information gain does correspond to a tendency towards maintenance of the *status quo*, this merely emphasizes the need to devote careful attention to the first stage. In the following section, we begin our discussion of this important stage with a brief description of the basic elements of an interregional accounting system.

5.4 A Consistent Approach to Interregional Estimation

5.4.1 Interregional Accounts

We now return to the interregional system introduced in Section 5.1.2. Our system of regional economies may correspond to a nation, a state, or some other aggregate for which certain industrial statistics are available. It contains n productive sectors which are distributed over m regions. Each region exports (imports) some of its products to (from) other regions and some to (from) the outside world. We shall treat interregional trade endogenously, but regard international trade as exogenous data. Assuming that all transactions are measured in real terms, and that the location pattern of production is predetermined, the interregional shipments can be related directly to the aggregate structure of the economy, and movements generated by imbalances between production and demand inside the different regions. The basic relationships may be viewed as a collection of accounts in which the following notations are used:

a_{ij} = amount of sector i absorbed per unit of output sector j;

a_{ij}^{*s} = amount of sector i absorbed per unit of output by sector j in region s;

a_{ij}^{rs} = amount of sector i from region r absorbed per unit of output by sector j in region s;

c_i^r = physical capacity requirements per unit delivery from sector i in region r;

C^r = total capacity for handling outgoing deliveries from region r;

e_i = exports (to abroad) from sector i;

e_i^r = exports (to abroad) from sector i in region r;

f_{ij}^{rs} = gross deliveries from sector i in region r to sector j in region s;

m_i = imports (from abroad) to sector i;

m_i^r = imports (from abroad) to sector i in region r;

v_i = value added to sector i;

v_i^r = value added to sector i in region r;

x_i = gross production in sector i;

x_i^{r*} = gross production of sector i in region r;

x_i^{*s} = gross deliveries from sector i to region s;

x_{ij} = intermediate deliveries from sector i to sector j;

x_{i*}^{*s} = intermediate deliveries from sector i to region s;

x_{ij}^{rs} = intermediate deliveries from sector i in region r to sector j in region s;

y_i = final demand (excluding exports) for sector i;

y_i^{r*} = final demand deliveries from sector i in region r;

y_i^{*s} = final demand deliveries from sector i to region s;

y_i^{rs} = final demand deliveries from sector i in region r to region s.

In a manner similar to Snickars, we can write down the basic elements of our interregional accounting system, commencing with the standard equations encountered in the literature.[30] We begin with the outflow relationship:

$$x_i^{r*} + m_i^r = \sum_{j=1}^{n} \sum_{s=1}^{m} x_{ij}^{rs} + \sum_{s=1}^{m} y_i^{rs} + e_i^r \qquad (i=1,.....,n;\ r=1,.....,m) \qquad (5.11)$$

and then define the equivalent inflow of factors needed to support the same intraregional production levels in each sector, namely

$$x_j^{s*} = \sum_{i=1}^{n} \sum_{r=1}^{m} x_{ij}^{rs} + v_j^s \ . \qquad (j=1,.....,n;\ s=1,.....,m) \qquad (5.12)$$

To maintain consistency with industrial statistics, the following additional relationships apply:

$$\sum_{r=1}^{m} \sum_{s=1}^{m} x_{ij}^{rs} = x_{ij} \qquad (i,j=1,.....,n) \qquad (5.13)$$

$$\sum_{r=1}^{m} \sum_{s=1}^{m} y_i^{rs} = y_i \qquad (i=1,.....,n) \qquad (5.14)$$

$$\sum_{r=1}^{m} x_i^{r*} = \sum_{s=1}^{m} x_i^{*s} = x_i \qquad (i=1,.....,n) \qquad (5.15)$$

$$\sum_{r=1}^{m} e_i^r = e_i \qquad (i=1,.....,n) \qquad (5.16)$$

$$\sum_{r=1}^{m} m_i^r = m_i \qquad (i=1,.....,n) \qquad (5.17)$$

$$\sum_{r=1}^{m} v_i^r = v_i \ . \qquad (i=1,.....,n) \qquad (5.18)$$

Equations (5.15) specify that total supply equals total demand at the aggregate level. Although a similar equality holds at the regional level, it is worthwhile stressing that interregional exports and imports are a major part of total transactions in the region, whereas they play a minor role in the nation's accounts. Owing to wide variations in the export and import patterns from region to region, regional production levels in the various sectors may differ vastly from each region's internal demand for products from those same sectors. These intraregional imbalances have not been explored in earlier work. Since they are of fundamental importance, we shall examine the sensitivity of various estimation procedures to the extent to which these imbalances can be quantified in advance.

5.4.2 The Estimation Problem

The primary aim of the present analysis is to estimate the interregional pattern of intermediate flows $\{x_{ij}^{rs}\}$ from a limited database of industrial and multi regional information. To do this, a distinction must be made between intermediate and final demand. Consequently, a secondary but simultaneous exercise is to estimate the interregional pattern of deliveries to final demand $\{y_i^{rs}\}$.

In terms of Equations (5.11) to (5.18), aggregate information from industrial statistics is certainly available concerning v_i, x_i, y_i and x_{ij}. We shall further assume that regional disaggregation of industrial statistics can provide information about e_i^r, m_i^r, v_i^r and x_i^{r*}. The following four assumptions concerning intraregional demands will be investigated:

Case (i) - that neither intermediate demand (x_{i*}^{*s}) nor final demand (y_i^{*s}) for each sector's products are known (or can be estimated independently) for each region;

Case (ii) - that final demand (y_i^{*s}) for each sector's products is known for each region;

Case (iii) - that total demand (x_i^{*s}) for each sector's products is known for each region;

Case (iv) - that both intermediate demand (x_{i*}^{*s}) and final demand (y_i^{*s}) for each sector's products are known for each region.

In the following discussion, estimates of unknown parameters will be denoted by a tilde (~). Thus an estimated pattern of deliveries to final demand would be written as $\tilde{y}_i^{rs}$. In the next section, we shall ignore both transportation costs, and any physical capacity constraints at the various nodes, and simply assume that supplies and demand are pooled. The nature of this class of problem can be amply understood by regarding each case as one of estimating the expected frequencies in a multidimensional contingency table containing various sets of marginal constraints. In a later section (5.4.4), consideration of capacity constraints renders this approach infeasible, so we then turn to the entropy-maxmimizing paradigm for assistance.

5.4.3 Contingency Table Analysis

In Section 4.3.3, we introduced an iterative procedure, sometimes known as the Deming-Stephan algorithm, to estimate the cell probilities in a two-way contingency table for which the marginal totals are known. As we shall demonstrate shortly, this type of iterative algorithm can also be applied to the three and four-way tables which arise in interregional analysis. However, standard tests of multidimensional contingency table analysis indicate that an iterative approach is not always necessary, because certain problems can be solved explicitly.[32] If the model formulations are such that the maximum-likelihood estimate can be expressed in *closed form*, an explicit solution is immediately possible.

5.4.3.1 Case (i): intraregional demands unknown

The given information for this situation can be summarized in the following four sets of constraints:

$$\sum_{j=1}^{n} \sum_{s=1}^{m} \tilde{x}_{ij}^{rs} + \sum_{s=1}^{m} \tilde{y}_{i}^{rs} = x_i^{r*} + m_i^r - e_i^r \quad (i=1,\ldots\ldots,n;\ r=1,\ldots\ldots,m) \tag{5.19}$$

$$\sum_{i=1}^{n} \sum_{r=1}^{m} \tilde{x}_{ij}^{rs} = x_j^{s*} - v_j^s \quad (j=1,\ldots\ldots,n;\ s=1,\ldots\ldots,m) \tag{5.20}$$

$$\sum_{r=1}^{m} \sum_{s=1}^{m} \tilde{x}_{ij}^{rs} = x_{ij} \quad (i,j=1,\ldots\ldots,n) \tag{5.21}$$

$$\sum_{r=1}^{m} \sum_{s=1}^{m} \tilde{y}_{i}^{rs} = y_i \quad (i=1,\ldots\ldots,n) \tag{5.22}$$

This is essentially Snickars' constraint set. The estimation problem can be regarded as one of estimating the expected frequencies in a four-dimensional contingency table with no third or fourth order interaction terms, but three sets of two-way constraints.[33] For this purpose, it is convenient to rewrite y_i^{rs} as $x_{i,n+1}^{rs}$. The problem can then be solved directly, since intermediate deliveries by sector of origin (i.e, $\sum_r x_{ij}^{rs}$) are constrained only by (5.20) and (5.21). The solution process is decomposed into two almost trivial parts: (a) estimate intermediate deliveries to region s, and (b) given these, estimate interregional deliveries.
From (5.20) and (5.21), it follows that

$$\sum_{r=1}^{m} \tilde{x}_{ij}^{rs} = \frac{x_{ij}\,(x_j^{r*} - v_j^s)}{\sum\limits_{s=1}^{m} (x_j^{s*} - v_j^s)} \tag{5.23}$$

whereupon supply and demand pooling leads to the solutions

$$\tilde{x}_{ij}^{rs} = \left(\frac{x_i^{r*} + m_i^r - e_i^r}{x_i + m_i - e_i}\right)\left(\frac{x_j^{s*} - v_j^s}{x_j - v_j}\right) x_{ij} \tag{5.24}$$

and

$$\tilde{y}_i^{rs} = \left(\frac{x_i^{r*} + m_i^r - e_i^r}{x_i + m_i - e_i}\right)\left(\frac{y_i}{m}\right) \quad . \tag{5.25}$$

Consider the estimates of intraregional input-output coefficients emanating from this formulation. By definition we have

$$\tilde{a}_{ij}^{*s} = \frac{\sum_{r=1}^{m} \tilde{x}_{ij}^{rs}}{x_j^{s*}} \tag{5.26}$$

which, from (5.23), becomes

$$\tilde{a}_{ij}^{*s} = a_{ij}\left(\frac{x_j^{s*} - v_j^s}{x_j - v_j}\right)\left(\frac{x_j}{x_j^{s*}}\right) \quad . \tag{5.27}$$

It is evident that $\tilde{a}_{ij}^{*s}$ is obtained by proportional column-wise adjustment of a_{ij} for each region.

5.4.3.2 Case (ii): final demands (y_i^{*s}) known

For this case, the given information consists of three of the four previous sets of constraints, namely (5.19), (5.20) and (5.21), together with the following additional set:

$$\sum_{r=1}^{n} y_i^{rs} = y_i^{*s} \quad . \qquad (i=1,\ldots\ldots,n;\ s=1,\ldots\ldots,m) \tag{5.28}$$

Once again, the estimation process can be decomposed into two trivial parts, whereupon supply and demand pooling yields the following solutions:

$$\tilde{x}_{ij}^{rs} = \left(\frac{x_i^{r*} + m_i^r - e_i^r}{x_i + m_i - e_i}\right)\left(\frac{x_j^{s*} - v_j^s}{x_j - v_j}\right) x_{ij} \tag{5.29}$$

and

$$\tilde{y}_i^{rs} = \left(\frac{x_i^{r*} + m_i^r - e_i^r}{x_i + m_i - e_i} \right) y_i^{*s} \tag{5.30}$$

The estimates of intraregional coefficients are identical to those given in (5.27) above.

5.4.3.3 Case (iii): total demands (x_i^{*s}) known

To quantify the total demands by all intermediate and final consumers in each region, an interregional version of the Leontief-Strout relationship is added to the four original sets of constraints, namely

$$\sum_{j=1}^{n} \sum_{r=1}^{m} \tilde{x}_{ij}^{rs} + \sum_{r=1}^{m} y_i^{rs} = x_i^{*s} \quad . \qquad (i=1,\ldots\ldots,n; \; s=1,\ldots\ldots,m) \tag{5.31}$$

Whereas Equations (5.19) suggest that intraregional production levels may be influenced predominantly by demands in other regions, Equations (5.31) imply that levels of intraregional demand may be satisfied largely by inflows from other regions.

We can view this problem as one of estimating the expected frequencies in a four-way table with four sets of two-way marginal constraints. It is again convenient to write $x_{i,n+1}^{rs}$ for y_i^{rs}. The awkward feature of this formulation is that (5.31) does not reduce to a constraint on intraregional coefficients because final demands appear on the left hand side. The problem cannot therefore be solved explicitly, but the following iterative proportional fitting procedure will converge to the maximum-likelihood solution:

(1) Assign arbitrary initial values (ones, for example) to all those elements of x_{ij}^{rs} which are nonzero,[34] and denote these values by $\overset{(o)}{\tilde{x}}{}_{ij}^{rs}$. Set t=0.

(2) At the (4t+1)th step, we take

$$\tilde{x}_{ij}^{rs\,(4t+1)} = \tilde{x}_{ij}^{rs\,(4t)} \cdot x_i^{r*} / \{ \sum_{j=1}^{n+1} \sum_{s=1}^{m} \tilde{x}_{ij}^{rs\,(4t)} \} \qquad (i=1,\ldots\ldots,n+1;\ r=1,\ldots\ldots,m)$$

(3) At the (4t+2)th step, we take

$$\tilde{x}_{ij}^{rs\,(4t+2)} = \tilde{x}_{ij}^{rs\,(4t+1)} \cdot (x_j^{s*} - v_j^s) / \{ \sum_{i=1}^{n} \sum_{r=1}^{m} \tilde{x}_{ij}^{rs\,(4t+1)} \} \qquad (j=1,\ldots\ldots,n;\ s=1,\ldots\ldots,m)$$

(4) At the (4t+3)th step, we take

$$\tilde{x}_{ij}^{rs\,(4t+3)} = \tilde{x}_{ij}^{rs\,(4t+2)} \cdot x_{ij} / \{ \sum_{r=1}^{m} \sum_{s=1}^{m} x_{ij}^{rs\,(4t+2)} \} \qquad (i,j=1,\ldots\ldots,n+1)$$

(5) At the (4t+4)th step, we take

$$\tilde{x}_{ij}^{rs\,(4t+4)} = \tilde{x}_{ij}^{rs\,(4t+3)} \cdot x_i^{*s} / \{ \sum_{j=1}^{n+1} \sum_{r=1}^{m} x_{ij}^{rs\,(4t+3)} \} \qquad (i=1,\ldots\ldots,n;\ s=1,\ldots\ldots,m)$$

(6) Repeat sequences (2) through (5) for t=1,2,....., etc. until the process converges satisfactorily.

When convergence has occured, the solutions are of the form

$$\tilde{x}_{ij}^{rs} = A_i^r \, B_j^s \, C_{ij} \, D_i^s \tag{5.32}$$

and

$$\tilde{y}_i^{rs} = A_i^r \, D_i^s \,, \tag{5.33}$$

where A_i^r, B_j^s, C_{ij} and D_i^s are coefficients associated with the constraints (5.19), (5.20), (5.21) and (5.31) respectively.

5.3.3.4 Case (iv): both intermediate (x_i^{*s}) and final (y_i^{*s}) demands known

In this case, the given information consists of three of the original sets of constraints, namely (5.19), (5.20), and (5.21), together with the following two additional sets:

$$\sum_{j=1}^{n} \sum_{r=1}^{m} \tilde{x}_{ij}^{rs} = x_{i*}^{*s} \qquad (i=1,\ldots\ldots,n;\ s=1,\ldots\ldots,m) \tag{5.34}$$

and

$$\sum_{r=1}^{m} \tilde{y}_i^{rs} = y_i^{*s} \ . \qquad (i=1,\ldots\ldots,n;\ s=1,\ldots\ldots,m) \tag{5.35}$$

From (5.19) and (5.35), it is immediately evident that part of the problem is decomposable. Final demands are obtained by pooling, namely

$$\tilde{y}_i^{rs} = \left(\frac{x_i^{r*} + m_i^r - e_i^r}{x_i + m_i - e_i} \right) y_i^{*s} \ . \tag{5.36}$$

The remainder of the problem is essentially one of finding $\sum_r x_{ij}^{rs}$ to satisfy (5.20), (5.21) and (5.34). This is analogous to a three-way (i, j, s) contingency table with no third-order interaction, but all three second-order terms. No direct solution is possible, but application of a three-step procedure, similar to the four-step approach outlined for the previous case, yields the following result:

$$\sum_r x_{ij}^{rs} = B_j^s \ C_{ij} \ F_i^s \tag{5.37}$$

where B_j^s, C_{ij} and F_i^s are coefficients associated with the constraints (5.20), (5.21) and (5.34) respectively.

5.3.3.5 Comparison of results for each case

Although the contingency table solutions for $\tilde{x}_{ij}^{rs}$ and $\tilde{y}_{i}^{rs}$ could theoretically involve up to six two-way interaction terms,[35] the prevailing constraints dictate that general expressions for the estimates of intermediate deliveries can be restricted to the following four two-way interactions terms:

$$\tilde{x}_{ij}^{rs} = A_i^r \, B_j^s \, C_{ij} \, F_i^s \tag{5.38}$$

and those for the final demand estimates need only involve two second-order terms:

$$\tilde{y}_i^{rs} = A_i^r \, D_i^s \quad . \tag{5.39}$$

Using these expressions as a guide, it is possible to summarize and contrast the solution techniques and resulting estimates for each of the four cases studied. Table 5.1 contains this summary.

TABLE 5.1

	$\tilde{x}_{ij}^{rs}$	$\tilde{y}_i^{rs}$
Case (i): demands unknown	$A_i^r\ B_j^s\ C_{ij}$ by direct solution +	$A_i^r\ D_i$ by direct solution +
Case (ii): y_i^{*s} known	$A_i^r\ B_j^s\ C_{ij}$ by direct solution +	$A_i^r\ D_i^s$ by direct solution +
Case (iii): x_i^{*s} known	$A_i^r\ B_j^s\ D_i^s$ by four-way contingency tables	$A_i^r\ D_i^s$ by four-way contingency tables
Case (iv): x_{i*}^{*s} and y_i^{*s} known	$A_i^r\ B_j^s\ C_{ij}\ F_i^s$ by three-way contingency tables	$A_i^r\ D_i^s$ by direct solution +

$$^{+}\ A_i^r = \frac{x_i^{r*} + m_i^r - e_i^r}{x_i + m_i - e_i}$$

$$B_j^s = \frac{x_j^{s*} - v_j^s}{x_j - v_j}$$

$$C_{ij} = x_{ij}$$

$$D_i^s = y_i^{*s}$$

$$D_i = \frac{y_i}{m}$$

Although a direct solution is impossible for some of the formulations proposed above, an explicit solution for each of the four cases can always be found using standard iterative algorithms devised for multi-dimensional contingency table analysis.[36] Such an algorithm has been included in the INTEREG package, which transforms the mathematical sets of constraints described above into the FORTRAN programming language, for computer estimation of the interregional flows. This package has been used to investigate each of the four cases using a simple system of two regions. The results are summarized in Appendix E. We shall confine our immediate analysis of these results to a brief discussion of the extent to which the information about intraregional demands influences the estimated pattern of interregional flows.

Firstly, the results contained in Appendix E confirm the validity of all the expressions included in Table 5.1. Secondly, Table 5.1 indicates that the type of available information on intraregional demands determines whether or not a direct solution is possible. Cases (i) and (ii), which assume that no aggregate demand information on intermediate deliveries is available, are easily solved by decomposition into two trivial stages. Cases (iii) and (iv), which introduce this additional information on intermediate demands, suffer by comparison since they require iterative solution techniques. We are therefore entitled to ask what exactly can be gained by collecting this costly information.

Tables E.4 to E.7 do not enlighten us greatly, except by suggesting (somewhat weakly) that the degree of cross-hauling is quite sensitive to the level of information available. Tables E.8 to E.11, however, reveal that the total levels of interregional trade are *reduced* by the inclusion of more detailed information on intraregional demands.[37] There is a complementary increase in the levels of intraregional deliveries. This may simply correspond to an increase in accuracy achieved by the tightening up of certain supply-demand interactions. More elaborate studies are needed to analyse these effects in greater detail.

5.4.4 Maximimum Entropy Formulations

In each of the four cases discussed above, the components of our interregional accounting system take the form of consistency constraints, which simply define the sums of certain subsets of flows. In other words, only *unit coefficients* have been included in the flow constraints. This type of information has been called an *accounting* constraint by some writers, bacause it is independent of any theory concerning distribution or dispersion.[38] Accounting constraints refer, in general, either to restrictions which must be satisfied by partial sums of the flow matrices $\{x_{ij}^{rs}\}$ and $\{y_i^{rs}\}$, or to *logical* relationships which exist between different elements of these two matrices.

Apart from accounting constraints, there are other sets of restrictions which have been adopted to influence the distribution of interregional shipments across possible states (i, j, r, s) of the system. The well-known cost constraint is a restriction of this kind. Information about freight or delivery costs has been used in a variety of ways to constrain the set of feasible shipments, and therefore influence the most probable distribution of flows.

A typical constraint of this type might assume that delivery costs are a matter for calculation on the part of the producer,[39] whereby the flows are expected to satisfy a production-specific cost constraint of the form

$$\sum_{j=1}^{n} \sum_{r=1}^{m} \sum_{s=1}^{m} c_i^{rs} x_{ij}^{rs} = C_i \qquad (i=1,\ldots,n) \qquad (5.40)$$

where c_i^{rs} is the mean transport cost per unit delivery from sector i in region r to region s, and C_i is the total cost of deliveries from sector i. This type of constraint assumes that transportation costs either remain fixed (at prescribed levels), or conform to some expected value of C_i.[40]

There are some serious objections to the use of constraints like (5.40). Firstly, it seems premature to try to define the cost coefficients, c_i^{rs}, prior to the estimation process. These costs are themselves derivable from some Lagrangian procedure, and are thus intrinsically associated with any chosen formulation. To specify them *in advance* is therefore potentially contradictory. Secondly, specifications of C_i, which assume fixed or expected values, may be applying an artificial restriction on the assumed behavior of the transportation system.[41] Finally, and perhaps foremost, it is now apparent that binding constraints on the transportation system are more likely to be associated with congestion at the nodes (i.e.within the regions) themselves, rather than on the links between nodes.

We shall therefore disregard the traditional cost constraint, in favour of a general *capacity* constraint of the form

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{s=1}^{m} c_i^r x_{ij}^{rs} \leq C^r \tag{5.41}$$

where c_i^r describes the physical capacity requirements per unit delivery from sector i (using a particular transport mode) in region r, and C^r defines the total capacity for handling outgoing deliveries from region r. A similar set of capacity constraints could also be formulated for *incoming* deliveries, but we shall not consider these in our current formulations.

If we add a set of capacity constraints, based on node restraints,[42] to our earlier formulations which involved only accounting constraints, standard iterative procedures devised for multidimensional contingency table analysis are no longer appropriate. Constraints like (5.41) no longer involve unit coefficients, but include coefficients (c_i^r) which result in a weighted sum of flows. Consequently, we shall adopt the entropy-maximizing paradigm to make further progress. We now define the following entropy function:

$$S = - \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{r=1}^{m} \sum_{s=1}^{m} \tilde{x}_{ij}^{rs} \log \tilde{x}_{ij}^{rs} - \sum_{i=1}^{n} \sum_{r=1}^{m} \sum_{s=1}^{m} \tilde{y}_i^{rs} \log \tilde{y}_i^{rs} \tag{5.42}$$

and proceed to reformulate our four basic cases so as to include a set of capacity constraints.

5.4.4.1 Case (i): intraregional demands unkown

This problem can be reformulated in the following manner:

Maximize S subject to the four basic accounting constraints (5.19), (5.20), (5.21) and (5.22), together with the following node constraints:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{s=1}^{m} c_i^r x_{ij}^{rs} + \sum_{i=1}^{n} \sum_{s=1}^{m} c_i^r y_i^{rs} \leq C^r \qquad (r=1,\ldots.,m) \tag{5.43}$$

and the usual non-negativity conditions. A standard Lagrangian derivation leads to the solutions

$$\tilde{x}_{ij}^{rs} = \exp\,(-\,\alpha_i^{r*} - \beta_j^{s*} - \gamma_{ij} - \varepsilon_r\, c_i^r) \tag{5.44}$$

and

$$\tilde{y}_i^{rs} = \exp\,(-\,\alpha_i^{r*} - \delta_i - \varepsilon_r\, c_i^r) \tag{5.45}$$

where α_i^{r*}, β_j^{s*}, γ_{ij}, δ_i and ε_r are the Lagrange multipliers associated with the constraints sets (5.19), (5.20), (5.21), (5.22) and (5.43) respectively.

The above solutions can be expressed in a more convenient (and familiar) form, namely

$$\tilde{x}_{ij}^{rs} = A_i^r\, B_j^s\, C_{ij} \exp\,(-\,\varepsilon_r\, c_i^r) \tag{5.46}$$

and

$$\tilde{y}_i^{rs} = A_i^r\, D_i \exp\,(-\,\varepsilon_r\, c_i^r) \tag{5.47}$$

where A_i^r, B_j^s, C_{ij} and D_i are coefficients which are exponentially related to the corresponding Lagrange multipliers.[43]

Returning to our earlier contingency table solutions (which are summarized in Table 5.1), we notice that the coefficients in (5.46) and (5.47) are identical to those for case (i) contingency table analysis. In other words, the first four sets of Lagrange multipliers defined above are simply negative logarithms of the corresponding coefficients included in our contingency table solutions.[44] This means that these coefficients may be specified explicitly. We have

$$\alpha_i^{r*} = \log \left(\frac{x_i + m_i - e_i}{x_i^{r*} + m_i^r - e_i^r} \right) \tag{5.48}$$

$$\beta_j^{s*} = \log \left(\frac{x_j - v_j}{x_j^{s*} - v_j^s} \right) \tag{5.49}$$

$$\gamma_{ij}^s = - \log x_{ij} \tag{5.50}$$

$$\delta_i = - \log \left(\frac{y_i}{m} \right) \tag{5.51}$$

To obtain explicit solutions for $\tilde{x}_{ij}^{rs}$ and $\tilde{y}_i^{rs}$, the final set of Lagrange multipliers (ε_r), associated with the capacity constraints, must also be limited. This can be done numerically using various iterative methods. For the INTEREG package (see Appendix E), we have adopted Eriksson's routine.[45]

5.4.4.2 Case (ii): final demands (y_i^{*s}) known

In this case, our reformulation takes the form:

Maximize S subject to the constraints (5.19), (5.20), (5.21), (5.28), (5.43) and the usual non-negativity conditions. The solutions are given by (5.44) and

$$\tilde{y}_i^{rs} = \exp\left(-\alpha_i^{r*} - \delta_i^{*s} - \varepsilon_r c_i^r\right) \tag{5.52}$$

with the Lagrange multipliers being defined by (5.48), (5.49), (5.50) and

$$\delta_i^{*s} = -\log y_i^{*s} \quad . \tag{5.53}$$

5.4.4.3 Case (iii): total demands (x_i^{*s}) known

For this reformulation, the constraint equation (5.28) is replaced by (5.31). The set of multipliers (δ_i^{*s}) now appears in both solutions, namely

$$\tilde{x}_{ij}^{rs} = \exp\left(-\alpha_i^{r*} - \beta_j^{s*} - \gamma_{ij} - \delta_i^{*s} - \varepsilon_r c_i^r\right) \tag{5.54}$$

and

$$\tilde{y}_i^{rs} = \exp\left(-\alpha_i^{r*} - \delta_i^{*s} - \varepsilon_r c_i^r\right) \quad . \tag{5.55}$$

The corresponding contingency table solution cannot be solved directly, so all the Lagrange multipliers in (5.54) and (5.55) must be eliminated using iterative techniques.

5.4.4.4 Case (iv): both intermediate (x_{i*}^{*s}) and final (y_i^{*s}) demands known

The replacement of (5.31) by two sets of constraints, namely (5.34) and (5.35), necessitates the replacement of δ_i^{*s} in (5.54) by a new set of multipliers λ_{i*}^{*s}, so that

$$\tilde{x}_{ij}^{rs} = \exp\,(-\,\alpha_i^{r*} - \beta_j^{s*} - \gamma_{ij} - \lambda_{i*}^{*s} - \varepsilon_r\; c_i^r) \tag{5.56}$$

Once again, an iterative procedure is needed to converge on an explicit solution.

5.4.3.5 Discussion of results

General expressions for the delivery estimates to both intermediate and final demand can be derived, and take the following exponential form:

$$\tilde{x}_{ij}^{rs} = A_i^r\; B_j^s\; C_{ij}\; F_i^s\; \exp\,(-\,\varepsilon_r\; c_i^r) \tag{5.57}$$

and

$$\tilde{y}_i^{rs} = A_i^r\; D_i^s\; \exp\,(-\,\varepsilon_r\; c_i^r)\;. \tag{5.58}$$

The resulting estimates for each of the four cases are summarized in Table 5.2. A quick glance back at Table 5.1 reveals that the exponential terms, associated with the capacity constraints, are the distinguishing feature of the maximum entropy estimates; in comparison with the correspon-ding contingency table solutions.

Explicit solutions for each case can be found numerically using the INTEREG package, which was also adopted earlier for the contingency table solutions. The package includes an iterative algorithm which can solve entropy-maximizing problems containing equality or inequality constraints. It has been applied to a simple system of two regions in order to investi-gate the influence of certain constraints on the flow estimates. The initial results are summarized in Appendix E. The marked difference between a typical contingency table solution (Tables E.4 and E.8), and the corresponding maximum entropy estimate (Tables E.13 and E.14), accentuates the sensitivity of the solution to constraints containing weighted coefficients like (5.40) or (5.41).

Owing to the influential part played by capacity constraints in our entropy-maximizing formulations, it is desirable to specify their precise role in this first stage of our estimation process. The imposition of constraints like (5.43) ensures that the solution algorithm generates a pattern of flows wherein the capacity requirements per commodity unit are consistent with observed nodal behaviour. These unit capacities (c_i^r) are defined to be the means of what is really a range of handling capacities, so much so that if additional information about their variance, V_r^2, is known, an additional set of constraints of the form

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{s=1}^{m} (c_i^r - C^r)^2 \, x_{ij}^{rs} = V_r^2 \qquad (5.59)$$

could be included in the formulation. In this case, the maximum entropy estimates would no longer be of exponential form, but would result in a normal Gaussian or truncated normal distribution depending on the range of c_i^r.[46]

For the present, we shall assume that the available information is limited to the *average* sectoral capacities within each region. This approach allows for modal split, to the extent that each sector may adopt a different transport mode (or vehicle type) for the delivery of their own commodities.

If historical information about these capacities is available for the period corresponding to our first stage of estimation, then it is appropriate to include a set of these constraints in the formulation. The resulting distributions are those presented in Table 5.2.

Historical, or *a priori* capacity constraints, are an important ingredient for the first stage of our estimation procedure involving the entropy-

TABLE 5.2

	$\tilde{x}_{ij}^{rs}$	$\tilde{y}_i^{rs}$
Case (i): demands unknown	$A_i^r\ B_j^s\ C_{ij}\ \exp\ (-\varepsilon_r\ c_i^r)^+$	$A_i^r\ D_i\ \exp\ (-\varepsilon_r\ c_i^r)^+$
Case (ii): y_i^{*s} known	$A_i^r\ B_j^s\ C_{ij}\ \exp\ (-\varepsilon_r\ c_i^r)^+$	$A_i^r\ D_i^s\ \exp\ (-\varepsilon_r\ c_i^r)^+$
Case (iii): x_i^{*s} known	$A_i^r\ B_j^s\ C_{ij}\ D_i^s\ \exp\ (-\varepsilon_r\ c_i^r)$	$A_i^r\ D_i^s\ \exp\ (-\varepsilon_r\ c_i^r)$
Case (iv): x_i^{*s} and y_i^{*s} known	$A_i^r\ B_j^s\ C_{ij}\ F_i^s\ \exp\ (-\varepsilon_r\ c_i^r)$	$A_i^r\ D_i^s\ \exp\ (-\varepsilon_r\ c_i^r)^+$

$$^+\ A_i^r = \frac{x_i^{r*} + m_i^r - e_i^r}{x_i + m_i - e_i}$$

$$B_j^s = \frac{x_i^{s*} - v_j^s}{x_j - v_j}$$

$$C_{ij} = x_{ij}$$

$$D_i^s = y_i^{*s}$$

$$D_i = \frac{y_i}{m}$$

maximizing paradigm. By including handling capacities which are actually known (rather than expected), we are ensuring that our estimation procedure minimizes statistical bias. Expected value constraints may be considered in a second stage of estimation, where an existing flow pattern requires modification to comply with expected changes in flow information. This second stage is more predictive, and may require certain assumptions about changes in the distributional mechanism (c_i^r) and total capacities (C^r). These assumptions will be explored in the following section, which deals with information adding by application of the principle of minimum information gain.

5.4.5 Information Adding

The results of the first stage calculations are a minimally biased pattern of deliveries prevailing in a period for which some information is available. These first-stage results are obvious candidates as prior values for a second stage of estimation designed to calculate the expected future pattern of flows in some later period. The task of this second stage, then, is to forecast the flows for a later period of interest, using the *a priori* flow pattern in the base period, together with partial information about the flows in the later period. To do this, we must modify our objective function to allow for non-uniform prior flow probabilities. In other words, we now progress from Jaynes' entropy-maximizing paradigm (which assumes uniform prior probalities) to the principle of minimum information gain.

Complete *a priori* flow matrices, which we shall denote by $\{x_{ij}^{rs}\}$ and $\{y_i^{rs}\}$, can be used to derive a conservative estimate of the expected flow pattern in a later period. We simply minimize the information gain, I, between the prior and posterior distributions, where

$$I = \sum_i \sum_j \sum_r \sum_s \tilde{x}_{ij}^{rs} \log (\tilde{x}_{ij}^{rs}/x_{ij}^{rs}) + \sum_i \sum_r \sum_s \tilde{y}_i^{rs} \log (\tilde{y}_i^{rs}/y_i^{rs}) \qquad (5.60)$$

subject to a set of linear constraints on $\tilde{x}_i^{rs}$ and y_i^{rs}, containing all the available information about the flows in the later period. For comparative purposes, we shall investigate two possibilities concerning available information. Firstly, we shall assume that it is limited to the set of accounting constraints discussed in Section 5.4.3. Secondly, our information will be assumed to contain some *expected* pattern of physical node capacities, thereby introducing an additional set of expected value constraints, which complete a similar set of restrictions to those discussed in Section 5.4.4.

It may be shown that there is always a unique solution to the above type of minimization problem, which can be expressed in terms of a set of monotonic transformations of the Lagrange multipliers pertaining to the full set of constraints. For the set of constraints discussed in Section 5.4.3, the following general expressions for the estimated deliveries to intermediate and final demand can be derived:

$$\tilde{x}_{ij}^{rs} = A_i^r \; B_j^s \; C_{ij} \; F_i^s \; x_{ij}^{rs} \tag{5.61}$$

and

$$\tilde{y}_i^{rs} = A_i^r \; D_i^s \; y_i^{rs} \; . \tag{5.62}$$

It is obvious that these two expressions reduce to our earlier results if all *a priori* flows considered equiprobable.[47]

As we discussed in Section 5.3.4, an iterative approach is not always needed to obtain an explicit solution. For cases (i) and (ii), a direct solution is possible because the maximum-likelihood estimate can be expressed in closed form. Cases (iii) and (iv), do however, require an iterative approach. The resulting estimates for each of our four cases studied are summarized in Table 5.3.

The inclusion of physical capacity constraints, such as Equations (5.43), results in the appearance of an additinal exponential term in each of the solutions. These revised estimates are also included in Table 5.3. Unfortunately, inclusion of capacity constraints, which restrict the interregional freight flows to an *expected* value of total node capacity, is not as easily justified in this second stage, as the inclusion of similar historical

TABLE 5.3

		$\tilde{x}_{ij}^{rs}$	$\tilde{y}_i^{rs}$
Case (i): demands unknown	Excluding cost constraints	$A_i^r B_j^s C_{ij} x_{ij}^{rs}$	$A_i^r D_i^s y_i^{rs}$
	Including cost constraints	$A_i^r B_j^s C_{ij} x_{ij}^{rs} \exp(-\varepsilon_r c_i^r)$	$A_i^r D_i^s y_i^{rs} \exp(-\varepsilon_r c_i^r)$
Case (ii): y_i^{*s} known	Excluding cost constraints	$A_i^r B_j^s C_{ij} x_{ij}^{rs}$	$A_i^r D_i^s y_i^{rs}$
	Including cost constraints	$A_i^r B_j^s C_{ij} x_{ij}^{rs} \exp(-\varepsilon_r c_i^r)$	$A_i^r D_i^s y_i^{rs} \exp(-\varepsilon_r c_i^r)$
Case (iii): x_i^{*s} known	Excluding cost constraints	$A_i^r B_j^s C_{ij} D_i^s x_{ij}^{rs}$	$A_i^r D_i^s y_i^{rs}$
	Including cost constraints	$A_i^r B_j^s C_{ij} D_i^s x_{ij}^{rs} \exp(\varepsilon_r c_i^r)$	$A_i^r D_i^s y_i^{rs} \exp(\varepsilon_r c_i^r)$
Case (iv): x_{i*}^{*s} and y_{i*}^{*s} known	Excluding cost constraints	$A_i^r B_j^s C_{ij} F_i^s x_{ij}^{rs}$	$A_i^r D_i^s y_i^{rs}$
	Including cost constraints	$A_i^r B_j^s C_{ij} F_i^s x_{ij}^{rs} \exp(-\varepsilon_r c_i^r)$	$A_i^r D_i^s y_i^{rs} \exp(-\varepsilon_r c_i^r)$

information in the first stage.

The influence of earlier capacity constraints will be embodied in the *a priori* patterns $\{x_{ij}^{rs}\}$ and $\{y_i^{rs}\}$. If the physical capacities do not vary greatly over time, one may doubt the wisdom of including such constraints in the second stage of estimation. Under such conditions, the *a priori* flow patterns could at least partly reflect these restrainting influences.

Before we compare and contrast the salient features of all three approaches discussed in this chapter, it is worth noting one other feature of our two-stage process of estimation. The function $\tilde{x}_{ij}^{rs} \log (\tilde{x}_{ij}^{rs}/x_{ij}^{rs})$ is defined to be zero wherever $\tilde{x}_{ij}^{rs} = 0$, and $\tilde{x}_{ij}^{rs}$ will automatically be zero whenever $x_{ij}^{rs} = 0$. To this extent, application of the principle of minimum information gain suffers from a similar drawback to methods of biproportional matrix adjustment; namely unadjustable zero entries.

The principle does not, however, require *a priori* knowledge or specification of zero entries for them to appear in the final solution. If certain expected value constraints (which involve the same two-factor interaction terms as some of the accounting constraints) are included in the formulation, the boundaries of the feasible region may prevent particular shipments from assuming positive values. In this situation, the solution is no longer preconditioned by the internal structure of the interindustry matrix $\{x_{ij}\}$, but depends more on interactions within the complete set of constraints.

5.3.6 Comparative Conclusions

In this chapter, three approaches to the estimation of interregional, intersectoral flows from a limited database of industrial and regional information have been described. In all three, a distinction has been made between deliveries to intermediate and final demands. Four separate assumptions concerning intraregional demands have been investigated. It is useful to summarize what we have learnt.

In the contingency table approach, each case is regarded as one of estimating the expected frequencies in a multidimensional contingency table with various sets of marginal constraints. In these formulations, nodal capacity constraints are not taken into account; regional supplies and demands are regarded as pooled. By dealing with an interregional system containing only accounting constraints, it is possible to solve all four cases using a standard iterative proportional fitting procedure. Furthermore, a direct solution to two of these cases is possible, since the maximum-likelihood estimates for these formulations can be expressed in closed form. The extent to which information about intraregional demands is available determines the viability of a direct solution.

If we now add a set of capacity constraints to our interregional accounting system, the iterative procedures devised for multidimensional contingency table analysis are no longer appropriate. The introduction of these constraints prompts our second approach, namely adoption of the entropy-maximizing paradigm, and highlights some important properties of this versatile approach to statistical estimation.

Firstly, the entropy-maximizing procedure can accommodate additional linear constraints, which may even contain the same interaction terms as those included in some of the accounting constraints. Secondly, if some of the interregional flows can be ascertained in advance, this limited survey data can be included in the constraint set. This feature is not normally possible using the contingency table approach, since the only entries which are certain to retain their original values, are those which are assigned zeros at the outset of the iterative scheme. Thirdly, methods based on multi-dimensional contingency tables suffer from the disadvantage that no zero element can be assigned a positive value during the iterative procedure, and no positive entry can be reduced uniquely to zero. This limitation is severe, since a number of positive entries in any national matrix may have zero values in the corresponding cells of some regional tables.

In general, the entropy-maximizing frame-work seems more suitable for the first-stage estimation of interregional flows than methods devised for multidimensional contingency table analysis. In cases where capacity constraints are excluded from the formulation, however, iterative procedures devised for the latter may prove more economical. In all other respects, the entropy approach offers greater flexibility.

For the second stage of estimation (if flow modifications are warranted), we have advocated application of the principle of minimum information gain. Using this approach, *a priori* flow estimates are modified to satisfy known constraints at some later date. The information gain approach enjoys most of the advantages inherent in the entropy-maximizing paradigm. In particular, it is possible to solve formulations involving both equality and inequality constraints. This feature is needed if we wish to introduce capacity constraints.

There are, nevertheless, certain features of this information-adding principle which differ greatly from those possessed by the entropy-maximizing paradigm. Of major importance is the property of *inertia,* which perpetrates a tendency to retain the existing (relative) pattern of flows to the largest possible extent. Of lesser importance is the fact that *a priori* flow patterns may reflect historical capacity constraints. This may undermine the usefulness of updated node handling information. These constrasting features emphasize an urgent need for some empirical research measuring the degree of inertia existing in interregional trade patterns.

At this stage, it is important to recognize that none of the static formulations described in Chapter 5 can provide an accurate estimation of the *gross* interregional flows between any two sectors. The aggregate nature of the deliveries to final demand makes it impossible to distinguish between consumption and capital investment. Since the exact distribution of capital flows is unknown, the gross flows between any two sectors cannot be determined. We can express the gross flows in the form

$$f_{ij}^{rs} = \tilde{x}_{ij}^{rs} + p_{ij}^{rs}\, \tilde{y}_{i}^{rs} \tag{5.62}$$

where p_{ij}^{rs} is the proportion of final demand deliveries from sector i in region r which are used for capital investment by sector j in region s. It is clear that no estimates of f_{ij}^{rs} are possible unless p_{ij}^{rs} is known. However, we shall not explore the question of capital expansion in this form, but instead turn to a more familiar framework: Leontief's dynamic input-output model. In the following chapter, the estimation of interregional flows is reexamined using this dynamic approach.

FOOTNOTES FOR CHAPTER 5

1 See Equations (4.4).

2 See Miernyk (1966).

3 For a thorough discussion of these factors, see Conway (1980).

4 See, for example, Emerson (1976) or Conway (1980).

5 See Isard (1951) and Leontief (1953).

6 A term proposed by Riefler (1973, p 136).

7 A strong preference for survey-based approaches has been shared, amongst others, by Czamanski and Malizia (1969, p 73), Schaffer and Chu (1969, p 96), and Miernyk (1972, p 267; 1976, p 54). For a thorough discussion of the problems associated with a survey-based study, see Isard and Langford (1971). For assessments of various non-survey methods, see Moore and Petersen (1955), Schaffer and Chu (1969), Czamanski and Malizia (1969), Hewings (1969: 1971), Morrison and Smith (1974), and Schaffer (1976).

8 Methods of biproportional matrix adjustment, such as the RAS method devised by Stone (1962), were initially designed to adjust input-output tables over time rather than space.

9 See, in particular, Tiebout (1957) and Miernyk (1968; 1976).

10 See, for example, Johansson and Stromqvist (1980).

11 Moore and Petersen (1955) initiated this extension, which was later formalized by Schaffer and Chu (1969).

12 See, for example, Stone (1962), Bacharach (1970), or Allen and Lécomber (1975).

13 See Section 4.3.3.

14 For discussions of the former approach, see Ireland and Kullback (1968) or Fienberg (1970); on the latter, see Bacharach (1970), Macgill (1977) or Hewings and Janson (1980).

15 See, for example, Czamanski and Malizia (1969) or Smith and Morrison (1974).

16 See, amongst others, Malizia and Bond (1974) and Miernyk (1976).

17 See Su (1970).

18 Schaffer's (1976) *exports-only* approach resembles the *rows-only* technique introduced by Hansen and Tiebout (1963).

19 See Jensen, Mandeville and Karunaratne (1979).

20 See Jones, Sporleder and Mustafa (1973).

21 Fixed coefficient models were tested independently by Chenery (1953) using Italian data, and Moses (1955) using American data. Moses (1960) also tested a linear programming model, in an attempt to explain the shipments of all goods within the United States.

22 For a brief theoretical introduction to this model, see Section 2.2.3.

23 See, for example, Polenske's (1970; 1972) multiregional input-output analyses of the Japanese and American economies.

24 For example, Courbis and Vallet (1976) relied on crudely aggregated transportation data, whereas Vanwynsberghe (1976) integrated features from several earlier studies to decompose the Belgian national table. This latter approach (known as the *Rococo method*) suffered from some very arbitrary assumptions, which may simply correspond to trade minimization.

25 See Round (1978 a, b).

26 See Nevin, Roe and Round (1966).

27 This *a priori* matrix include flows to both intermediate and final demand.

28 Snickars' *a priori* estimates of the intermediate flows are of the form

$$x_{ij}^{rs} = A_{ij}^{r} \; B_{i}^{rs}$$

where A_{ij}^{s} is termed the *technology effect*, since it describes a typical input structure for each receiving sector and region, and B_{i}^{rs} is called the *interregional transport effect*, since it reflects a typical interregional delivery pattern for the output of each sector. This model is of the type proposed by Chenery (1953) and Moses (1955). Bigsten (1978) and Snickars (1979) both derived estimates of final demand deliveries independently of intermediate shipments.

29 See Jaynes (1957, p 620).

30 See, for example, Hartwick (1971) or Richardson (1972).

31 The notion of *supply and demand pooling* implies that the ultimate destination of goods is irrelevant to producers, and that the origin of goods is irrelevant to consumers. For further elaboration, see Leontief and Strout (1963) or Wilson (1970 a, b).

32 See, for example, Fienberg (1977) or Haberman (1978).

33 There are three other possible sets which do not appear in this formulation, namely i-s, j-s and r-s.

34 Wherein $x^{rs}_{n+1,j} = 0$ for all j.

35 Together with up to four three-way terms and four one-way terms, if such information was available and not redundant.

36 Based, for example, on the Deming-Stephan (1940) algorithm or the Newton-Raphson technique. For further information, see Fienberg (1977) or Haberman (1978). See also Section 5.4.3.3.

37 Total imports into region r ($\sum_j \sum_s m^{sr}_j$), and total exports out of region r ($\sum_i \sum_s e^{rs}_i$), are both affected in the same direction.

38 A view expressed, for example, by Williams and Wilson (1980).

39 This approach has been suggested , for example, by Macgill and Wilson (1979).

40 Williams and Wilson (1980) define expected value constraints as those which embody a known *or* expected value of total freight expenditure, together with some assumptions regarding the distributional behaviour of individual sectors or producers.

41 A skepticism expressed, for example, by Puu (1979).

42 As specified in Equation (5.41).

43 With the aid of the constraint equations, it is always possible to derive general expressions for each coefficient in terms of the exogenously-given information and the other coefficients. For this reason, these coefficients are sometimes referred to as *balancing factors*.

44 The fifth set of multipliers, ε_r, are associated with the capacity constraints (5.43).

45 See Eriksson (1978; 1980).

46 Entropy-maximizing procedures can produce a variety of well-known distributions, depending on the nature of the set of constraints included in any formulation. For example, higher-order moments of the capacities lead to normal distributions, instead of the standard exponential distribution. Inclusion of logarithmic terms can result in a solution closely resembling the gamma distribution. For some examples of the range of distributions possible, see Reza (1961), Tribus (1969), Dowson and Wragg (1973), or Lee (1974b).

47 If all $x^{rs}_{ij} = y^{rs}_i = 1$, then Equations (5.60) and (5.61) reduce to Equations (5.38) and (5.39), respectively.

Chapter 6

INTERSECTORAL FLOWS IN SPACE: DYNAMIC FORMULATIONS

6.1 Introduction

The static formulations introduced in Chapter 5 provide a useful introduction to the various approaches which can facilitate the estimation of intersectoral flows over space. They do, however, suffer from two major deficiencies. Firstly, their failure to distinguish between the various components of final demand renders them incapable of estimating the *gross* intersectoral flows. Secondly, their inability to deal directly with repercussions of regional and sectoral growth restricts their relevance to analyses in the short run. For the purposes of medium to long term forecasting, the need for a dynamic model is obvious.

In this chapter, Leontief's dynamic input-output framework is adapted to the interregional estimation problem. The basic difference is that flows of capital goods are now treated endogenously, instead of being relegated to final demand. By quantifying individual capital flows between sectors, it is possible to describe the expansive capabilities of our multiregional economy. At the same time, this additional information enables the interregional pattern of gross intersectoral flows to be estimated.

To the extent that productive capital flows are treated endogenously, Leontief's dynamic model is partially closed by definition. Nevertheless, open and closed versions of his model are usually distinguished by their treatment of *all* final demand. In Section 6.2, we analyse open and closed versions in their basic national form, before deducing the fundamental relationships embodying the interregional extension.

Section 6.3 describes the approaches which enable the interregional pattern of gross intersectoral flows to be estimated from a limited database of national and regional information. Through the use of Leontief's dynamic framework and a simple accelerator principle, a clear distinction can be made between the intermediate flows, which are described by the usual input-output coefficients, and productive capital flows, which are specified by an interregional matrix of capital coefficients. Thus the present analysis is formulated in terms of coefficients rather than absolute flows.

Although the separable nature of the formulation allows each matrix of interregional coefficients to be estimated simultaneously, it also limits our choice of solution techniques. The objective function, and the associated system of constraints, no longer relate to simple flow sums, but rather to expressions involving weighted summation. Thus, iterative procedures devised for multidimensional contingency table analysis are now unsuitable. The methodology therefore draws exclusively on information theory for its solutions.

To complement the four elementary cases introduced in Capter 5, a similar set of four assumptions concerning the information available on intraregional demands is investigated. The resulting interregional accounting systems described in Section 6.3 are all formulated as fully closed models. It is left to the reader to ponder the appropriate simple adjustments which are needed to derive comparable open versions of each model.

6.2 Basic Model Characteristics

6.2.1 The Open Dynamic Leontief Model

In Chapter 4, we established a fundamental system of differential equations which,[1] at the national level, took the form

$$x_i(t) = \sum_j a_{ij}\, x_j(t) + \sum_j b_{ij}\, \dot{x}_j(t) + y_i(t) \qquad (6.1)$$

and describe the dynamic interdependencies of the economy.[2] According to the acceleration principle of capital formation, the investment terms can be expressed as a linear function of the change in production. namely

$$x_i(t) = \sum_j a_{ij}\, x_j(t) + \sum_j b_{ij}\, [x_j(t+1) - x_j(t)] + y_i(t) \ . \qquad (6.2)$$

If we further assume that each economic sector is following an independent growth path, with a corresponding average growth rate, g_i, then we also have

$$x_i(t) = x_i(0)[1 + g_i]^t \ . \qquad (i = 1,\dots,n) \qquad (6.3)$$

Substituting (6.3) into (6.2), setting t = 0, and rewriting $x_i(0)$ as x_i, we have

$$x_i = \sum_{j=1}^{n} (a_{ij} + b_{ij}\, g_j)x_j + y_i \ . \qquad (i = 1,\dots,n) \qquad (6.4)$$

This is our basic system of equations for the open national model.

The logic of this open system is simply that by treating final demands exogenously, we also make them decisive. As a set of independent variables, the final demand vector may then be determined by relationships other than those described in these equations themselves. Such an open system provides an analytical tool which is particularly well suited for investigating the implications of alternative policy decisions. For this purpose, the final demand vector may be viewed as the objective function of the economic process.

An open model like (6.1) can therefore be transformed, or rather interpreted, from the viewpoint of optimal processes. The generic model used in the theory of optimal control is also an open one, in which the behaviour of the system is determined by exogenous variables. In such formulations, the course of economic development may be steered by external agents such as consumers, overseas traders or governments, rather than by producers or investors.

A number of authors have already discussed the instability problem in national and regional economic systems, arising from the assumption of equilibrium in the dynamic Leontief formulation.[3] Such systems may be stabilized by allowing excess supply or demand conditions to occur, and then introducing a suitable process of control to promote stabilization.[4] Although most applications of control theory to economic systems have been based on macroeconomic management, a small number of regional applications have appeared recently.[5] These applications emphasize the importance of the open model.

It is not our purpose here to formulate a control theory approach to the problem of multiregional instability, but merely to develop an interregional extension of Leontief's open model as a basis for further analysis. To do this, we firstly return to the interregional system introduced in

Chapter 5. Defining b_{ij}^{rs} as the amount of capital goods required from sector i in region r per unit increase of output by sector j in region s, and g_i^r as the average rate of growth in output from sector i in region r, we can derive a modified version of Isard's original outflow relationship,[6] namely

$$x_i^r = \sum_{j=1}^{n} \sum_{s=1}^{m} (a_{ij}^{rs} + b_{ij}^{rs} g_j^s) x_j^s + \sum_{s=1}^{m} y_i^{rs} \quad . \tag{6.5}$$

$$(i = 1,\ldots,n;\ r = 1,\ldots,m)$$

Equations (6.5) define our basic set of relationships for balanced interregional commodity flows in Leontief's open model.

6.3.2 The Closed Dynamic Leontief Model

An interesting feature of systems such as (6.1) is that it is much easier to study the corresponding homogenous system in which $y_i(t) = 0$, namely

$$x_i(t) = \sum_j a_{ij}\, x_j(t) + \sum_j b_{ij}\, \dot{x}_j(t) \quad . \tag{6.6}$$

This closed formulation does not actually eliminate final demand activites from the model; it is not a consumptionless model. Instead, (6.1) is closed by the introduction of an additional equation with coefficients reflecting the structural properties of this particular sector.[7] For example, the behaviour of households can be viewed as an industry supplying its output, mainly labour, to other industries and receiving consumer goods from them. Household activities would then appear in the enlarged system as two sets of additional parameters: the flow coefficients, $a_{1,n+1}$, $a_{2,n+1}$,..... and

$a_{n+1,1}$, $a_{n+1,2}$,....., and the capital coefficients, $b_{1,n+1}$, $b_{2,n+1}$,..... and $b_{n+1,1}$, $b_{n+1,2}$,.....; the latter describing consumer's investments in housing, automobiles, appliances, and other kinds of household durables.

Interest in the closed model has been restricted mainly to theoretical economists who aim to study a system which is as endogenous as possible.[8] For the case in which there is only a single output, (6.6) reduces to the familiar Harrod-Domar growth model.[9] The input-output matrix corresponds to the marginal propensity to consume, whereas the capital matrix corresponds to the accelerator coefficient.

By adopting the acceleration principle of capital formation, together with the assumption of fully utilized capital stocks in each period, it is possible to calculate a balanced growth path which attains the largest possible rate of expansion of this closed system. The resulting proportionate growth path corresponds to the von Neumann trajectory,[10] and yields an unique equilibrium structure known as the turnpike solution.

Far from being only of academic interest, the turnpike solution may be used to assess the long-term growth potential of an existing technological structure. It may also be used to evaluate the development potential (or efficiency) of alternative technologies. Further discussion of its inherent properties, as well as various solution algorithms, is postponed to Chapter 7 and Appendix A. For the present, we shall restrict our discussion of the closed national model to a statement of the complementary system of equations to the open version (6.4). Following similar arguments to those adopted for the open model, we arrive at

$$x_i = \sum_{j=1}^{n+1} (a_{ij} + b_{ij}\ g_j)x_j \qquad (i = 1,\ldots,n+1) \qquad (6.7)$$

where (n+1) stands for the additional household sector.

An interregional extension of (6.7) will be developed in the following section, prior to the application of information theoretical techniques for the simultaneous estimation of both intermediate and capital flows on an intersectoral and interregional basis.

6.3 A Closed Model Approach to Interregional Estimation

6.3.1 Interregional Accounts

We now return to the interregional system developed in Section 5.4.1 of the previous chapter. As before, our system of regional economies is divided into n sectors and distributed over m regions. We shall now close this system by treating capital investment, consumption, and overseas trade as endogenous components of the model.

Capital flows which serve to maintain or revive existing prodictive capacity will be treated as intermediate consumption, by including them in the input-output matrix. Capital flows which are designed to expand existing productive capacity will be represented by a matrix of capital coefficients describing the interregional distribution of productive investment between various sectors. Household behaviour and international trade will be treated as additional industries, and are therefore included in the original n sectors. Once again, the basic relationships can be viewed as a collection of accounts in which the following additional notations are introduced:

b_{ij} = amount of capital goods required from sector i per unit increase of output by sector j,

b_{ij}^{rs} = amount of capital goods required from sector i in region r per unit increase of output by sector j in region s,

g_i = average rate of growth in output from sector i,

g_i^r = average rate of growth in output from sector i in region r,

k_{i*}^{*s} = investment demand in region s for capital goods from sector i,

u_i = gross operating surplus in sector i,

u_i^r = gross operating surplus in sector i of region r.

If we assume full utilization of capital stocks in each period, the non-spatial input-output equilibrium for our economy is derived from Equation (6.7). We shall now assemble the basic elements of our interregional accounting system for this closed model. It is left to the reader to make the appropriate simple adjustments to derive a comparable system for the open model. We begin again with a modified version of Isard's original outflow relationship, namely

$$x_i^{r*} = \sum_{j=1}^{n} \sum_{s=1}^{m} (a_{ij}^{rs} + b_{ij}^{rs} g_j^s) x_j^{s*} \qquad (i = 1,\ldots,n;\ r = 1,\ldots,m) \qquad (6.8)$$

and add to this the corresponding inflow of factors required to satisfy regional production in each sector, that is

$$x_j^{s*} = \sum_{i=1}^{n} \sum_{r=1}^{m} a_{ij}^{rs} x_j^{s*} + u_j^s \quad . \qquad (j = 1,\ldots,n; \; s = 1,\ldots,m) \qquad (6.9)$$

To ensure consistency with industrial statistics, the following additional relationships apply:

$$\sum_{r=1}^{m} \sum_{s=1}^{m} a_{ij}^{rs} x_j^{s*} = a_{ij} x_j \qquad (i,j = 1,\ldots,n) \qquad (6.10)$$

$$\sum_{r=1}^{m} \sum_{s=1}^{m} b_{ij}^{rs} g_j^s x_j^{s*} = b_{ij} g_j x_j \qquad (i,j = 1,\ldots,n) \qquad (6.11)$$

$$\sum_{r=1}^{m} g_i^r x_i^{r*} = g_i x_i \qquad (i = 1,\ldots,n) \qquad (6.12)$$

$$\sum_{r=1}^{m} u_i^r = u_i \qquad (i = 1,\ldots,n) \qquad (6.13)$$

and as before

$$\sum_{r=1}^{m} x_i^{r*} = \sum_{s=1}^{m} x_i^{*s} = x_i \quad . \qquad (i = 1,\ldots,n) \qquad (6.14)$$

This completes the specification of our interregional accounting system for the closed model.

6.3.2 The Estimation Problem

The main aim of the present analysis is to estimate the interregional pattern of *gross* intersectoral flows $\{f_{ij}^{rs}\}$ from a limited database of industrial and regional information. To do this, we have expressed these flows in the form

$$f_{ij}^{rs} = (a_{ij}^{rs} + b_{ij}^{rs}\, g_j^s)x_j^{s*} \quad . \tag{6.15}$$

In contrast to the formulations derived in Chapter 5, we have now made a distinction between intermediate and capital flows. Consequently, it is possible to estimate interregional matrices of input-output and capital coefficients simultaneously. The present analysis is therefore formulated in terms of coefficients rather than actual flows.

In terms of Equations (6.8) to (6.14), aggregate information from industrial statistics is certainly available concerning g_i, u_i, x_i and a_{ij}. Some countries are fortunate enough to have compiled a matrix of capital coefficients (b_{ij}) at the national level. Others may be forced to estimate these coefficients by alternative means. A number of non-survey approaches to this problem have been proposed in Chapter 4,[11] so we shall treat b_{ij} as a matrix which can be estimated independently at the aggregate level. We shall further assume that regional disaggregation of industrial data can provide information about g_i^r, u_i^r and x_i^{r*}. Our closed model investigations will be limited to the following four assumptions concerning intra-regional demands:

Case (i) - that neither intermediate demand (x_{i*}^{*s}) nor investment demand (k_{i*}^{*s}) for each sector's products are known (or can be estimated independently) for each region;

Case (ii) - that intermediate demand (x_{i*}^{*s}) for each sector's products is known for each region;

Case (iii) - that total demand (x_i^{*s}) for each sector's products is known for each region;

Case (iv) - that both intermediate demand (x_{i*}^{*s}) and investment demand (k_{i*}^{*s}) for each sector's products are known for each region.

Although the separable nature of the gross flows (f_{ij}^{rs}) allows interregional tables of input-output and capital coefficients to be estimated simultaneously, it also limits our choice of solution techniques. The inclusion of growth rates (g_i^r) in the present formulation alters the nature of the constraint system from one of simple flow sums to one of weighted summation. This means that standard solution procedures devised for multidimensional contingency table analysis are no longer suitable.[12] Consequently, we shall draw exclusively on information-theoretical procedures for our solutions. In so doing, we are now at liberty to include a set of nodal capacity constraints, similar to Equations (5.41), in all our model formulations.

6.3.3 Maximum Entropy Formulations

In the absence of any prior knowledge about the individual flows (that is, assuming that each shipment is equi-probable), we can define the following entropy function:

$$S = - \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{r=1}^{m} \sum_{s=1}^{m} (\tilde{a}_{ij}^{rs} + \tilde{b}_{ij}^{rs} \; g_j^s) x_j^{s*} \log (\tilde{a}_{ij}^{rs} + \tilde{b}_{ij}^{rs} \; g_j^s) x_j^{s*} \quad (6.16)$$

which serves as our objective function for the four cases discussed below.

6.3.3.1 Case (i): intraregional demands unknown

The relevant information from our interregional accounts can be written as the following four constraint sets:

$$\sum_{j=1}^{n} \sum_{s=1}^{m} (\tilde{a}_{ij}^{rs} + \tilde{b}_{ij}^{rs} \; g_j^s) x_j^{s*} = x_i^{r*} \qquad (i = 1,\ldots,n;\ r = 1,\ldots,m) \quad (6.17)$$

$$\sum_{i=1}^{n} \sum_{r=1}^{m} \tilde{a}_{ij}^{rs} \; x_j^{s*} = x_j^{s*} - u_j^s \quad (j = 1,\ldots, n;\ s = 1,\ldots,m) \quad (6.18)$$

$$\sum_{r=1}^{m} \sum_{s=1}^{m} \tilde{a}_{ij}^{rs} \; x_j^{s*} = a_{ij} \; x_j \quad (i,j = 1,\ldots,n) \quad (6.19)$$

$$\sum_{r=1}^{m} \sum_{s=1}^{m} \tilde{b}_{ij}^{rs} \; g_j^s \; x_j^{s*} = b_{ij} \; g_j \; x_j \quad (i,j = 1,\ldots,n) \quad (6.20)$$

The resulting entropy-maximizing problem is then to maximize S subject to the constraints (6.17) through (6.20) and the usual non-negativity conditions. A standard Lagrangian derivation yields

$$\tilde{a}_{ij}^{rs}\, x_j^{s*} = \exp(-\alpha_i^{r*} - \beta_j^{s*} - \gamma_{ij}) \tag{6.21}$$

and

$$\tilde{b}_{ij}^{rs}\, g_j^s\, x_j^{s*} = \exp(-\,\alpha_i^{r*} - \delta_{ij}) \tag{6.22}$$

where α_i^{r*}, β_j^{s*}, γ_{ij} and δ_{ij} are the Lagrange multipliers associated with the constraint sets (6.17) to (6.20) respectively. These solutions may be expressed in a more familiar form, namely

$$\tilde{a}_{ij}^{rs}\, x_j^{s*} = A_i^r\, B_j^s\, C_{ij} \tag{6.23}$$

and

$$\tilde{b}_{ij}^{rs}\, g_j^s\, x_j^{s*} = A_i^r\, D_{ij} \tag{6.24}$$

where A_i^r, B_j^s, C_{ij} and D_{ij} are coefficients which are related exponentially to the corresponding Lagrange multipliers.

Explicit solutions for $\tilde{a}_{ij}^{rs}$ and $\tilde{b}_{ij}^{rs}$ in this and all succeeding formulations may be obtained by eliminating the balancing factors or their corresponding Lagrange multipliers. As suggested in Chapter 5, this can be done numerically using the INTEREG package described in Appendix E.

6.3.3.2 Case (ii): intermediate demands (x_{i*}^{*s}) known

For this case, the given information consists of the four previous sets of constraints,[13] together with the following additional set:

$$\sum_{j=1}^{n} \sum_{r=1}^{m} \tilde{a}_{ij}^{rs} x_j^{s*} = x_{i*}^{*s} \quad . \qquad (i = 1,\ldots,n;\ s = 1,\ldots,m) \qquad (6.25)$$

The entropy-maximizing problem is now to maximize S subject to the constraints (6.17) through (6.20), (6.25), and the usual non-negativity conditions. An extra multiplier (μ_{i*}^{*s}) appears in the solution to represent the new set of constraints (6.25). The results are:

$$\tilde{a}_{ij}^{rs} x_j^{s*} = \exp(-\alpha_i^{r*} - \beta_j^{s*} - \gamma_{ij} - \mu_{i*}^{*s}) \qquad (6.26)$$

and

$$\tilde{b}_{ij}^{rs} g_j^s x_j^{s*} = \exp(-\alpha_i^{r*} - \delta_{ij}) \quad . \qquad (6.27)$$

6.3.3.3 Case (iii): total demands (x_i^{*s}) known

To quantify the total demands by all intermediate and capital users in each region, a modified version of the Leontief-Strout relationship is added to the four original sets of constraints.[13] We have

$$\sum_{j=1}^{n} \sum_{r=1}^{m} (\tilde{a}_{ij}^{rs} + \tilde{b}_{ij}^{rs} g_j^s) x_j^{s*} = x_i^{*s} \quad . \qquad (i = 1,\ldots,n;\ s = 1,\ldots,m) \qquad (6.28)$$

The resulting entropy-maximizing problem is now subject to the constraints (6.17) through (6.20), together with (6.28) and the usual non-negativity conditions. A new multiplier (η_i^{*s}) is now included in the solution, and on this occasion it appears in both equations. The solutions come out as

$$\tilde{a}_{ij}^{rs} x_j^{s*} = \exp(-\alpha_i^{r*} - \beta_j^{s*} - \gamma_{ij} - \eta_i^{*s}) \tag{6.29}$$

and

$$\tilde{b}_{ij}^{rs} g_j^s x_j^{s*} = \exp(-\alpha_i^{r*} - \delta_{ij} - \eta_i^{*s}) \quad . \tag{6.30}$$

6.3.3.4 Case (iv): both intermediate demands (x_{i*}^{*s}) and capital demands (k_{i*}^{*s}) known

In this situation, the available information consists of the five sets of constraints appearing in case (ii),[14] together with the following additional set:

$$\sum_{j=1}^{n} \sum_{r=1}^{m} \tilde{b}_{ij}^{rs} g_j^s x_j^{s*} = k_{i*}^{*s} \quad . \qquad (i = 1,\ldots,n;\ s = 1,\ldots,m) \tag{6.31}$$

The replacement of (6.28) by the two independent sets of constraints, namely (6.25) and (6.31), necessitates the introduction of another new Lagrange multiplier (λ_{i*}^{*s}) to represent Equation (6.31). The resulting solutions are

$$\tilde{a}_{ij}^{rs} x_j^{s*} = \exp(-\alpha_i^{r*} - \beta_j^{s*} - \gamma_{ij} - \mu_{i*}^{*s}) \tag{6.32}$$

and

$$\tilde{b}_{ij}^{rs} g_j^s x_j^{s*} = \exp(-\alpha_i^{r*} - \delta_{ij} - \lambda_{i*}^{*s}) \quad . \tag{6.33}$$

6.3.3.5 Discussion of results

The resulting estimates for both intermediate and capital flow coefficients can be expressed in the general form:

$$\tilde{a}_{ij}^{rs} = A_i^r \; B_j^s \; C_{ij} \; F_i^s(x_j^{s*})^{-1} \tag{6.34}$$

and

$$\tilde{b}_{ij}^{rs} = A_i^r \; D_{ij} \; G_i^s(g_j^s \; x_j^{s*})^{-1} \tag{6.35}$$

where A_i^r, B_j^s, C_{ij}, D_{ij}, F_i^s and G_i^s are coefficients which can be related exponentially to the correspondingly subscripted Lagrange multipliers introduced for each case. Estimates for the gross intersectoral flows $(\tilde{f}_{ij}^{rs})$ can therefore be expressed as

$$\tilde{f}_{ij}^{rs} = A_i^r(B_j^s \; C_{ij} \; F_i^s + D_{ij} \; G_i^s) \quad . \tag{6.36}$$

Table 6.1 contains expressions for the balancing coefficients defined in terms of the appropriate Lagrange multipliers for each case.

6.3.3.6 Inclusion of capacity constraints

If we wish to incorporate a similar set of nodal capacity constraints to the ones introduced in Chapter 5, the result is an additional Lagrange multiplier in the corresponding solutions.[15] For example, we could include a set of capacity constraints in the form

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{s=1}^{m} c_i^r \; (\tilde{a}_{ij}^{rs} + \tilde{b}_{ij}^{rs} \; g_j^s)x_j^{s*} \leq C^r \quad (r = 1,\dots,m) \tag{6.37}$$

TABLE 6.1

For all four cases		
$A_i^r = \exp(-\alpha_i^{r*})$ $B_j^s = \exp(-\beta_j^{s*})$ $C_{ij} = \exp(-\gamma_{ij})$ $D_{ij} = \exp(-\delta_{ij})$		
Case	F_i^s	G_i^s
(i) demands unknown	1	1
(ii) intermediate demands known	$\exp(-\mu_{i*}^{*s})$	1
(iii) total demands known	$\exp(-\eta_i^{*s})$	$\exp(-\eta_i^{*s})$
(iv) intermediate and capital demands known	$\exp(-\mu_{i*}^{*s})$	$\exp(-\lambda_{i*}^{*s})$

in each of our earlier formulations. The result would be an additional term in the solution estimates (6.34) and (6.35). If we let ε^r represent the set of Lagrange multipliers associated with (6.37), the modified solutions would read

$$\tilde{a}_{ij}^{rs} = A_i^r\ B_j^s\ C_{ij}\ F_i^s(x_j^{s*})^{-1}\ \exp(-\ \varepsilon^r\ c_i^r\ ) \tag{6.38}$$

and

$$\tilde{b}_{ij}^{rs} = A_i^r\ D_{ij}\ G_i^s(g_j^s\ x_j^{s*})^{-1}\ \exp(-\ \varepsilon^r\ c_i^r\ ) \quad . \tag{6.39}$$

Historical information on nodal capacities may prove binding in cases where the information available concerning intraregional demands is very limited.[16] They may therefore be an important ingredient in the first stage of our estimation procedure, namely computation of historical flow estimates using the entropy-maximizing paradigm. Wherever an historical flow pattern requires modifications to comply with expected changes in flow information, however, we require a second stage of estimation. In the following section, we discuss this information-adding procedure by reference to the principle of minimum information gain.

6.3.4 Partial Information Adding

In the previous section, we assumed that no *a priori* information about the flows is available. The results so obtained are an obvious choice for prior values in any second stage of estimation designed to update the historical flow pattern to comply with changing supply or demand information. To facilitate these modifications to our original flow estimates, we must alter our objective function to allow for non-uniform prior flow probabilities. We now minimize the information gain, I, in terms of gross flows, namely

$$I = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{r=1}^{m} \sum_{s=1}^{m} \tilde{f}_{ij}^{rs} \log(\tilde{f}_{ij}^{rs}/f_{ij}^{rs}) \tag{6.40}$$

subject to a set of linear constraints on $\tilde{f}_{ij}^{rs}$ containing all the pertinent flow information.[17]

For the four elementary cases analysed in the previous section, the resulting modified estimates for the gross intersectoral flows take the general form

$$\tilde{f}_{ij}^{rs} = A_i^r(B_j^s\ C_{ij}\ F_i^s + D_{ij}\ G_i^s)f_{ij}^{rs} \tag{6.41}$$

where the unknown coefficients retain the same expressions(given originally in Table 6.1), in terms of the appropriate Lagrange multipliers, for each case. It is obvious that (6.41) reduces to (6.36) if all the a priori flows (f_{ij}^{rs}) are equiprobable.

Of particular interest is the case where we have *a priori* knowledge of the input-output coefficients (a_{ij}^{rs}), but no knowledge of the capital coefficients (b_{ij}^{rs}) on an interregional basis. This situation requires separate terms for the intermediate and capital flows in the objective function. Our revised function to be minimized takes the following form:

$$I = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{r=1}^{m} \sum_{s=1}^{m} (\tilde{a}_{ij}^{rs}\ x_j^{s*}) \log(\tilde{a}_{ij}^{rs}\ x_j^{s*}/\bar{a}_{ij}^{rs}\ \bar{x}_j^{s*}) \\ + (\tilde{b}_{ij}^{rs}\ g_j^s\ x_j^{s*}) \log(\tilde{b}_{ij}^{rs}\ g_j^s\ x_j^{s*}) \tag{6.42}$$

where $\bar{a}_{ij}^{rs}$ and $\bar{x}_j^{s*}$ define *a priori* values.[18] The modified solutions are of the general form:

$$\tilde{a}_{ij}^{rs} = A_i^r \; B_j^s \; C_{ij} \; F_i^s(\bar{a}_{ij}^{rs} \; \bar{x}_j^{s*})(x_j^{s*})^{-1} \tag{6.43}$$

and

$$\tilde{b}_{ij}^{rs} = A_i^r \; D_{ij} \; G_i^s(g_j^s \; x_j^{s*})^{-1} \; . \tag{6.44}$$

In situations where we have partial *a priori* knowledge of either flow matrix, this information can be used to deduce a complete *a priori* matrix of flows. The adopted approach involves the calculation of the net sum of all unknown flows, and then the distribution of these unknown shipments on an equiprobable basis. To demonstrate the procedure, a simple example is included in Appendix B.

6.3.5 Conclusions

In this chapter, two further approaches to the estimation of interregional, intersectoral flows from a limited database of information have been described. In contrast to the previous chapter, adoption of Leontief's dynamic framework allows the interregional pattern of *gross* intersectoral flows to be estimated. A clear distinction can be made between the intermediate flows, which are described by the usual input-output coefficients, and productive capital flows, which are specified using capital coefficients. All the cases studied are therefore formulated in terms of coefficients instead of actual flows.

The entropy-maximizing approach is adopted because the various relationships no longer describe simple sums involving unit coefficients. The inclusion of growth rates in the various formulations alters the nature of the constraint system to one of weighted summation. Consequently, standard solution algorithms devised for multidimensional contingency table analysis are no longer useful.

If modifications to any historical flow patterns are warranted, we have reiterated our support for the principle of minimum information gain. This second stage of estimation is largely complementary to the entropy-maximizing paradigm. It does, however, produce conservative estimates, to the extent that the solution corresponds to a pattern of least change. To eventually clarify its relevance to the field of spatial input-output analysis, further research into the degree of inertia existing in interregional trade patterns is obviously necessary.

FOOTNOTES FOR CHAPTER 6

1 See Equations (4.24).

2 In this open formulation, $y_i(t)$ now represents final *non-productive* demand for the goods from sector i in period t.

3 See, for example, Jorgenson (1960), Tsukui (1968), Brődy (1970) and Lesse and Sharpe (1980).

4 This suggestion appears in Andersson and Karlqvist (1979).

5 See, for example, Nijkamp (1978), Bennet and Tan (1979), or Lesse and Sharpe (1980).

6 Based on a simple reformulation of Equations (5.4), which correspond to an interregional extension of Equations (6.4).

7 If the final demand vector can be broken down into several separate categories, a corresponding number of additional sectors (and equations) may be needed.

8 Following the initial formulations by Leontief (1953), other analyses of the closed model can be found in Morishima (1958), Sraffa (1960), Wurtele (1960), Jorgenson (1960), Brődy (1966; 1970), and Johansen (1973; 1978).

9 See Harrod (1939) and Domar (1946).

10 See von Neumann (1945).

11 See Sections 4.2.3 and 4.3.5, as well as Appendix C.

12 As well as the obvious difficulty posed by the presence of weighted coefficients, there is also the possibility that two sets of constraints may involve the same interaction terms.

13 Namely Equations (6.17) through (6.20).

14 Namely Equations (6.17) through (6.20) and Equations (6.25).

15 It is recommended that the reader refer to Sections 5.4.4 and 5.4.5 of the previous chapter for a discussion of various problems associated with the inclusion of cost constraints.

16 Case (i) could benefit greatly from the inclusion of capacity constraints.

17 In (6.40), f_{ij}^{rs} defines the *a priori* flow value, which may be estimated historically using (6.36) if a survey is unavailable.

18 Expression (6.42) is derived assuming equiprobable capital *flows* rather than equiprobable coefficients.

Chapter 7

TOWARDS A SYSTEM OF MODELS FOR INTEGRATED NATIONAL AND REGIONAL DEVELOPMENT

7.1 Introduction

The preceding chapters of this dissertation have been concerned predominantly with the estimation of intersectoral and interregional commodity flows, from a limited database of industrial and spatial information. Their main purpose has been to demonstrate a practical means by which the economic analyst can reduce his survey needs, which can otherwise be both extensive and expensive in the case of intersectoral and interregional models.[1] It now remains to demonstrate how the various flow or coefficient estimates could be used to analyse feasible paths of economic development over space and time. This is the task of this final chapter.

Having already expressed a preference for a dynamic interregional framework of the interindustry type, Section 7.2 sets out in search of a plausible system of models to integrate national and regional development. The exploration begins by reviewing some existing approaches, which fulfil the basic requirements, and concludes by suggesting a *hierarchical system of models*. Although this system successively disaggregates the development problem, it also permits an autonomous, self-assertive tendency at each level, to counterbalance the integrative nature of the system as a whole.

By regarding each level in our five-level modelling hierarchy as being responsible for a certain degree of detail, a separability of focus is maintained, leading to an efficient specialization of function by each model in the hierarchy.[2] An *equilibrium* function is examined at the national level. Compromise solutions, based on Simon's notion of *satisficing*,[3] are proposed for the various regional models. At the lower levels, where the decisions of individuals can be recognized more easily, the logical function is one of *optimization*.

It is argued that the generation of various development scenarios is a fundamental feature of the planning process at both the national and regional levels. By adopting a closed version of Leontief's dynamic model at the national level, it is possible to evaluate the development potential (or efficiency) of alternative production technologies. The natural choice would be the set of techniques which minimizes costs, and maximizes growth of production, in the long run.

The generation of plausible development scenarios at the regional level allows different policies and goals to be quantified and compared. We assume that the predominant function of regional economic planning is to satisfy (in some way) a multiplicity of objectives or decision problems. Simultaneous optimization of a system involving multiple criteria is usually impossible, due to the conflicting nature of various objectives. We therefore derive a satisficing version of Leontief's model, designed to search for a simultaneous compromise solution between conflicting criteria. The information gain statistic is used to maximize the *inertia* of the *a priori* production pattern, thereby retaining as much of the region's current characteristics as the multiplicity of criteria allows.

We can now adopt the type of formulation discussed in Chapter 6 to estimate the vital structural properties of our national-regional interface. Information theory can be used to calculate an interregional pattern which is consistent with both the structure of production at the national level, *and* supply-demand imbalances within each region. The relative degree of control exerted by each set of constraints is closely related to the amount and type of information collected at each level.

At this point, we can easily recognize the significant role which information theory can play in hierarchical systems analysis. It allows us to make efficient use of the information which is available at a higher, more aggregated level, to coordinate and integrate the patterns of behaviour at more disaggregated levels below. To this extent, it can therefore contribute to the fundamental process of *decomposability* in hierarchical systems.

But hierarchical theory is not essential to demonstrate the complementary role of information theory to the analysis of national and regional economies. Nor is the system of models proposed in this chapter considered to be the only, or even the optimal, path towards an integrated spatial development system. In reality, there is ample scope to modify the model formulations at each level, or even to disregard any hierachical assumptions completely.

7.2 In Search of a Framework for Integration

7.2.1 Theoretical Background

Regional and interregional modelling presently lack firm theoretical foundations. The attempts to generalize neoclassical economic theory, so as to encompass the spatial dimension, have largely failed because of their simplistic approach to the determinants of interregional flows; possibly the most distinctive feature of regional development. Neoclassical economics has neglected spatial factors, such as distance and location, which may be of critical importance in explaining regional growth.[4]

Forecast-oriented techniques, such as regional input-output analysis and development planning models,[5] should not primarily be seen as a contribution to regional growth theory. Their usefulness is related to examining the consequences of specific changes in exogenous factors (via impact analyses or scenario generation), or determining the most likely or most desirable pattern of development; rather than to any improvment in our understanding of the regionalization process itself. It is very much in this latter tradition that the following search for a suitable modelling framework should be viewed.

Although input-output analysis provides an extremely flexible framework for regional and interregional modelling, we have stressed repeatedly that its regional role is quite different to that on the national level. The regional economy is extremely *open* in comparison with the nation to which it belongs. This has two very important consequences. Firstly, effective regional planning must take account of various development patterns occurring outside the region in question. Each region must recognize its interdependency on other surrounding regions, as well as its role in national development. Thus the model framework should include *interregional linkages*.

Secondly, regional rates of growth and decline are much more accentuated than on the national level. In any medium to long-term forecasting, the repercussion of different growth rates cannot be ignored. Thus the model framework should also be *dynamic*.[6]

Having established a fundamental need for a dynamic interregional framework of the interindustry type, at least two other important decisions remain. Within the class of permissible models, either *optimization* or *equilibrium* solutions are always available. Furthermore, either *open* or *closed* versions of each model may be explored. Our decisions regarding these properties will be deferred, however, until after we have reviewed some existing models which fulfil our basic requirements.

7.2.2 A Review of Some Earlier Models

Spatial versions of Leontief's dynamic model were first suggested in theory more than twenty years ago.[7] In the lengthy period following this theoretical underpinning, very few models have become fully operational. Some are summarized in Table 7.1. One intraregional model is included in the table, because of its early contribution to the advancement of dynamic modelling. The seven other models are all interregional.

TABLE 7.1

Model	Author	Year first Published	Model Structure	Objective Function	Method of Solution	National-Regional Linkages	Determination of Trade Flows
1. West Virginia[8]	Miernyk and others	1970	Open Equilibrium	--------	Dynamic Inverse	Bottom-up	-----------
2. Maryland[11]	Harris	1970	Open Optimization	Minimize transport costs	Linear programming	Top-down	Transportation LP
3. Indian[13]	Mathur	1972	Open Optimization	Minimize transport costs	Linear programming	Top-down	Based on national flow coefficients
4. Swedish[15]	Andersson	1975	Closed Equilibrium	Balanced growth	Algebraic eigenvalues	Mixed	Maximum Likelihood
5. TIM[18]	Funck, Rembold and others	1975	Open Equilibrium	Most probable development	Sequential regressions	Mixed	Gravity model
6. DREAM[21]	Karlqvist, Sharpe Batten & Brotchie	1976	Open or closed Optimization	Multiple objectives	Iterative LP or Entropy	Top-down or Mixed	Modified gravity model(max.likelih.)
7. Dutch[22]	Hafkamp and Nijkamp	1978	Closed Optimization	Multiple objectives	Compromise method	Two regions only	Assumed known
8. MORSE[24]	Lundqvist	1980	Open Optimization	Multiple objectives	Linear programming	Mixed	Chenery-Moses Assumptions

Table 7.1 is not intended to be an exhaustive summary, since other models have centrainly appeared. The models included therein are considered simply to be representative of the chronological pattern of advancement in this area. A brief discussion of each model follows.

7.2.2.1 The West Virginia Model

Miernyk and his associates made the first attempt to implement a dynamic regional input-output model in the late sixties.[8] The West Virginia model is not an interregional model, but makes a useful distinction between replacement and expansion capital. The fundamental system of first-order difference equations, outlined in the previous chapter,[9] is modified to read

$$x_i(t) = \sum_j a_{ij}x_j(t) + \sum_j d_{ij}x_j(t) + \sum_j b_{ij}[x_j(t)-x_j(t-1)] + y_i(t) \quad (7.1)$$

where d_{ij} is the flow of capital goods from sector i required to maintain stocks of capital equipment in sector j at existing levels. In other words, $\{d_{ij}\}$ is a matrix of capital replacement coefficients. This modification is not unlike the modification to input coefficients suggested in Chapter 4, to allow for maintenance of fixed capital.[10]

A slightly modified form of the Leontief dynamic inverse is used to project capital requirements. The viability of this approach depends on the assumptions embodied in the capital coefficient matrices $\{d_{ij}\}$ and $\{b_{ij}\}$. When tested by Miernyk, the dynamic model produced forecasts that were only marginally different from a series of comparative-static forecasts with a fairly simple Leontief-type model. The West Virginian example demonstrates that the analyst must choose carefully between the costs of additional data collection, and the strategic returns to be gained from a more detailed specification of the relationships between investment and growth.

7.2.2.2 The Maryland Model

At much the same time as Miernyk's work, Harris attempted to embed Almon's national model into an interregional framework.[11] His objective was to forecast industrial activity at the regional level, together with other regional variables including population, income, workforce and unemployment. He used linear programming to solve the transportation problem for shadow prices, rather than to estimate the optimum trade flows. His interest in trade flows was peripheral.[12] The shadow prices are used as independent variables in the industry location equations which, together with a set of population migration equations, attempt to explain the migration of capital and labour between various regions.

By using small geographical areas, Harris hoped to minimize the problems of cross-hauling. The Maryland model therefore embraced a system of 3112 regions (counties) and 100 industrial sectors. The resulting data requirements were enormous, and the variation in quality of the data series which were used cast some doubt on the reliability of the overall model.

7.2.2.3 The Indian model

Mathur implemented a transport cost-minimizing model for optimal regional allocation in India.[13] His open model combines linear programming techniques with dynamic input-output analysis. The Indian economy is divided into 5 regions and 27 sectors, for which three average growth trajectories (zero, 10% and 15%) are examined. Constraints may be imposed on regional trade balances and resource exploitation.

The results indicate that the optimum pattern of production is highly sensitive to rates of growth, as well as to trade balance constraints. In other words, we cannot talk of balanced regional growth, and an optimum pattern of location, without a specific reference to the dynamic aspects of the economy.[14] Nevertheless, heavy industries appear to be locationally inelastic.

7.2.2.4 The Swedish model

An interregional model which postulates balanced growth in a closed system of regional economies has been proposed by Andersson.[15] The model is of the equilibrium type, and adopts a dynamic interregional growth and allocation model as an organizing mechanism for spatial flows. The allocation of regional production is organized in such a way that demands and supplies are equilibrated at the various nodes in the transportation network. The rate of capacity use is then maximized for any given expectations of growth in product demand. Alternatively, at full capacity, the rate of expansion of the whole production system is maximized.[16] The latter approach corresponds to the turnpike model presented in Appendix C.

Andersson argues that the transportation system is in equilibrium if it preserves a balanced situation on each of the regionally differentiated commodity markets, and is consistent with goals like full employment and a given level of resource conservation. He adopts the principle of maximum likelihood or maximum entropy to estimate the interregional flows, and also proposes use of the minimum information principle if *a priori* flows are available.[17]

7.2.2.5 The TIM model

Since July 1970, six German research groups have been striving towards the completion of a Total Interregional Model (TIM) for the Federal Republic of Germany. An interim report explains that the model has four components, namely (i) a demand submodel, (ii) an input-output model, (iii) a production submodel, and (iv) a resource submodel.[18] Interregional, sector-specific commodity flows are derived using a modified version of the gravity model. Sector-specific incestment demand functions are also formulated. Unfortunately, this research has now been abandoned owing to insurmountable difficulties with data collection.

7.2.2.6 The DREAM model

A Dynamic Regional Economic Allocation Model (DREAM) has been developed primarily for Australian conditions, but is sufficiently general for applications elsewhere.[19] This optimization model has an input-output framework, with constraints on labour, migration, population distribution, production, investments, exports and imports, and consumption. The temporal structure is represented by a simple dynamic multiplier principle, which relates capital investment to output in various sectors, during the same time period, by a set of linear investment coefficients.

A useful distinction is made between products from *national* sectors, which are transferable between regions (footloose), and *regional* products which are not transferable. The flow-stock relationships for regional sectors take a closed form, similar to the usual balanced dynamic Leontief model. A dummy region may be used to absorb excess supply or demand within national sectors. A gravity model is used for calculating the interregional flows for all national sectors. This gravity model can be derived by entropy-maximizing methods.

An initial objective of maximizing net surplus (exports less imports less transportation costs) was chosen. Other objectives have been investigated by including production, labour, population distribution, investment, consumption, intermediate demand, export and import terms (all linear), plus transportation cost terms (quasi-quadratic), in the objective function. Various combinations have been explored by weighting each term, and discounting between time periods has been used to give greater importance to initial time periods. Thus the objective function, and the choice of constraints, may be manipulated to reflect various community goals.

The mathematical programming formulation can be solved using iterative linear programming techniques, or entropy-maximizing methods.[20] The computer program, which is fully operational, has been implemented in a wide variety of Australian studies.[21]

7.2.2.7 The Dutch model

Hafkamp and Nijkamp have developed and interregional model which links production, investment, employment and pollution on an intersectoral basis.[22] The welfare profile of each region is assumed to contain three elements (production, employment, and pollution), which form the basis of a multi-objective decision framework. Simultaneous optimization of all objective functions is impossible, owing to the conflicting nature of each objective. Hafkamp and Nijkamp suggest a compromise method, based on a distance metric, which minimizes the discrepancy between the set of efficient solutions and the *ideal* solution. A learning procedure is outlined to improve the welfare specifications of the model, until an ultimate compromise solution is reached. The notions of *satisficing* and *displaced ideals* are therefore implied.[23]

The full-information input-output model so adopted assumes that all interregional deliveries of intermediate and final products are known in advance. Its implementation has been restricted to a two-region (Rhine-delta area and the rest of the Netherlands) model containing 27 sectors, because of data limitations.

7.2.2.8 The MORSE model

A recent Swedish model employs a *mixed* approach to the task of achieving consistency between the national and regional level. The model (known as MORSE) is designed to view relationships between the energy sector and the rest of the economy in a multi-regional perspective.[24] All relationships concerning technologies and consumption are expressed at the regional level, whereas interregional trade is balanced nationally by means of Chenery-Moses assumptions. Other nationally specified constraints (on balance of payments, energy use and capital formations), together with regional commodity balances, income balances, and bounds on capital and labour utilization, combine to restrict the set of feasible solutions.

MORSE draws on achievements in multi-regional input-output theory, development modelling and mathematical programming. It is an open model, in order to allow explicit treatment of goals, resource constraints and capacity utlization. The multi-objective approach combines goals for economic, employment and energy planning into a linear programming framework. It has many similar features to the DREAM model, and is used to analyse the feasibility and consistency of regional developments, with respect to national ambitions in economic and energy policies.

7.2.2.9 Discussion

What insights can be gleaned from these dynamic interregional modelling experiences? Firstly, there is a clear need for internal consistency between economic behaviour at the national level and aggregate multi-regional behaviour. This does not mean that national and regional objectives must be identical, but simply that the various parameters must sum to the national totals over all regions. The pioneering interregional models achieved this consistency by employing a *top-down* approach, in which national aggregates are broken down into their regional components. Although this top-down approach represents a convenient means of extending national planning systems to facilitate regional forecasting, it suffers from a serious inability to quantify the effects on the national economy of changing regional conditions. The ideal interregional model requires a *mixed* approach, in which some entities are prescribed at the national level, while others are determined regionally.[24]

Secondly, traditional optimization models were based on the assumption of independent decision-making units striving for a unidimensional objective. In many of the early interregional models, this single objective was to minimize transportation costs. Fortunately, there is now a growing awareness that planners and policy-makers, at both the national and regional levels, must base their decisions on a *multiplicity of criteria* (for example, equity, efficiency, ecological balance, etc.). They must therefore consider a wide range of policy objectives (implying a *multidimensional goal function*), to reflect the different goals and aspirations which exist within their community.

The simultaneous consideration of multiple decision criteria, or multiple objectives, complicates the traditional programming methods. Two schools of thought have recently developed. The first approach is a simple extension of traditional optimization procedures, which applies certain *a priori* weightings to the various terms in the objective function. The second recognizes the conflicting nature of various goals, and searches for a compromise solution based upon interactive satisficing principles. A detailed discussion of these approaches is postponed to Section 7.5.

Thirdly, there is an increasing need to develop a flexible interregional framework which allows certain linkages and spillover effects to be explored in greater detail. Important issues, such as energy consumption, environmental pollution, and resource depletion, now require specific consideration within an integrated economic framework. A few of the models in Table 7.1 have explored some of these issues. Others have considered the intersections between energy, pollution and further economic issues, within a static interregional framework.[25] An extension of the latter approach into a dynamic setting would be extremely valuable.

Finally, but perhaps foremost, there is a formidable obstacle which is shared by all the interregional modelling exercises undertaken so far: that of limited availability of suitable data. This common difficulty seems likely to persist in the foreseen future, as modellers attempt to introduce additional dimensions to the planning process. In the face of these inevitable deficiencies in information, it is important to make progressive improvements to our methods of estimation. It is now clear that information theory can make a significant contribution to this area of endeavour.

To build further on these earlier modelling efforts, we shall now attempt to develop a more general modelling framework which

(i) provides a flexible mechanism for the integrated analysis of national and regional development options; and also

(ii) demonstrates the valuable and versatile role which information theory can play in this analysis.

7.2.3 A Hierarchical Modelling System

It is now clear that long-term economic planning cannot be based on a single goal function alone, but must encompass a multiplicity of goals at the various levels of the planning process. It must also allow for a mixture of variables, each of which may either be determined or constrained at quite different levels. Wherever there is organized economic activity, there appear to be multilevel, or *hierarchical* phenomena.

Yet hierarchical analysis is still practically nonexistent in traditional economic theory,[26] and has only recently been introduced into regional science.[27] We shall attempt to consolidate on these recent analyses, by describing a general hierarchical system which, for our purposes, will be simplified to consider only five different levels of modelling effort. This system has its foundations in Isard's global balanced regional input-output model, which identifies a hierarchical structure of political authorities and corresponding commodities.[28]

Our multilevel system is depicted in Figure 7.1. Although it successively disaggregates the development problem, it also permits an autonomous, self-assertive tendency at each level, to counterbalance the integrative nature of the system as a whole.[29] In reality, this hierarchy is open-ended in the downward, as it is in the upward direction.

The system of models corresponding to this five-level hierarchy is represented in Figure 7.2. At the uppermost level, decisions taken concerning international trade patterns provide important constraints on feasible

Figure 7.1

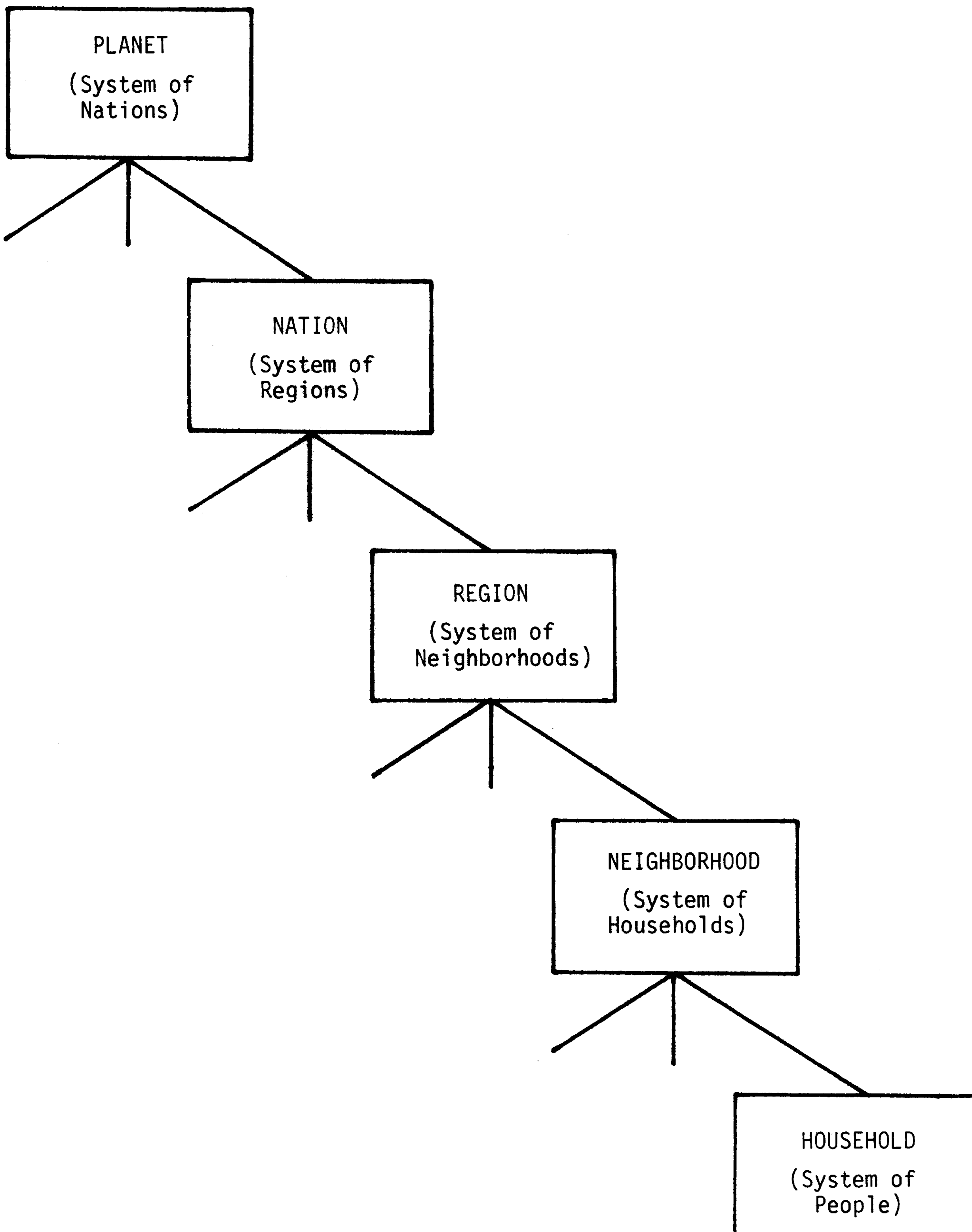
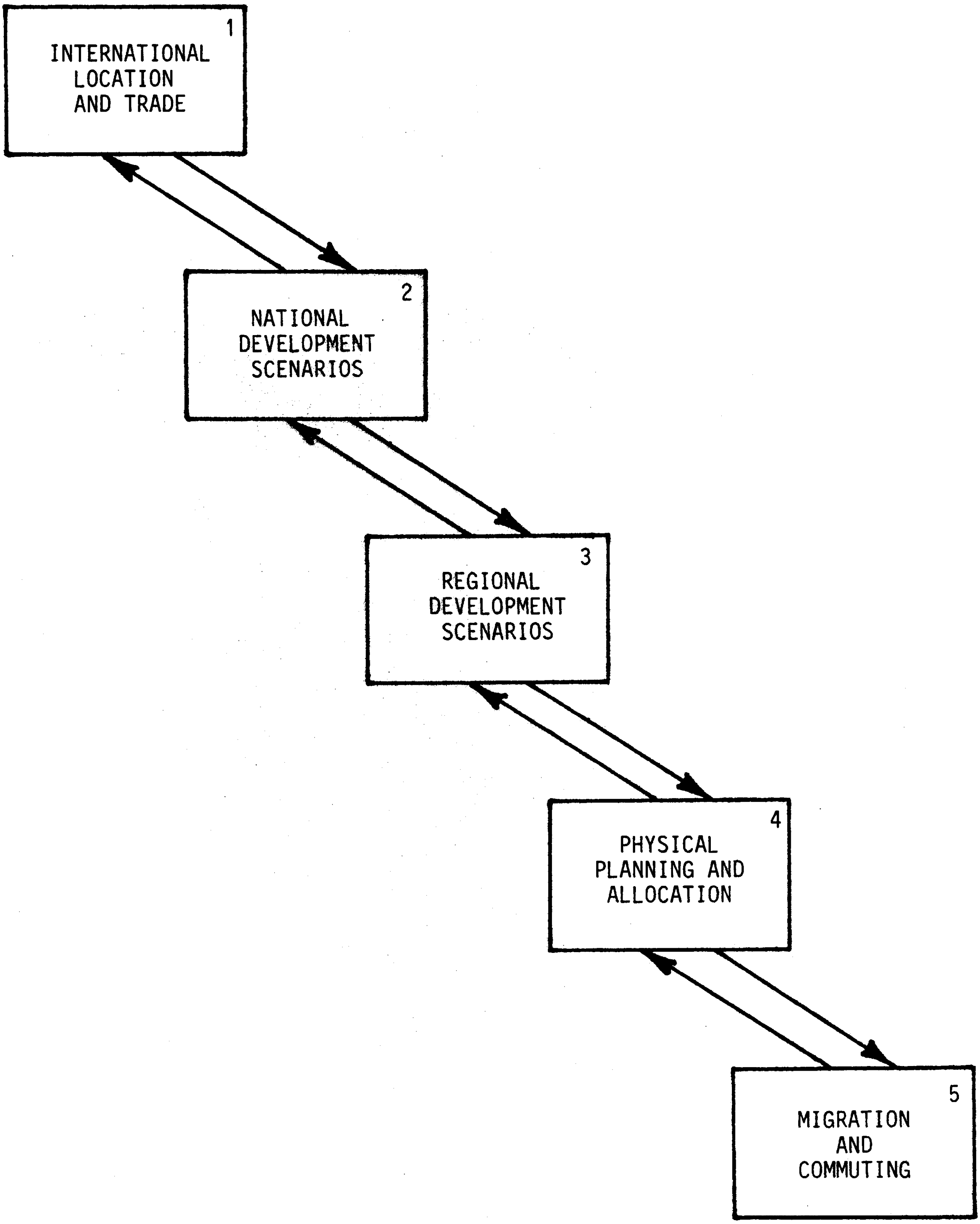

PLANET
(System of Nations)
NATION
(System of Regions)
REGION
(System of Neighborhoods)
NEIGHBORHOOD
(System of Households)
HOUSEHOLD
(System of People)

Figure 7.2


1
INTERNATIONAL
LOCATION
AND TRADE
2
NATIONAL
DEVELOPMENT
SCENARIOS
3
REGIONAL
DEVELOPMENT
SCENARIOS
4
PHYSICAL
PLANNING AND
ALLOCATION
5
MIGRATION
AND
COMMUTING

developments patterns in each nation. Similarly, decisions taken at both the international and national levels impose further constraints on the decision-maker at the regional level. However, it should be stressed that higher-level models can only co-ordinate, but not completely control, the goal-seeking activities at lower levels.[30]

We can associate this hierarchical structure of decision-making units with a similar commodity classification system. It is not only useful, but increasingly necessary, to recognize that some commodities are balanced (in terms of production and consumption) at the international level only. Others are balanced at the national, regional or neighbourhood levels. Let W denote world commodities (balanced on the world level only), N represent national commodities (whose production and consumption are balanced both at the world and the national level), R represent regional commodities (balanced at the world, national and regional levels), and L represent local commodities (with balance at all levels).

Similar distinctions are often made with respect to the *mobility* of industries.[31] World industries (more commonly referred to as *transnationals* or *multinationals*) are regarded as free to locate in any nation. National industries are free to locate in any region, and so on. World industries also tend to market their products to any nation; national industries to any region. Which goods or industries turn out to be world, and which remain national, regional and local, depends to a large extent on the structure and conditions of trade.

The advantage of this five-level hierarchical system lies in the ability to analyse each subsystem in a relatively independent fashion. This possibility arises because of the *near-decomposability* of subsystems.[32] Simon maintains that it is possible to focus on the dynamics of one level, by ignoring both higher and lower-level dynamics for the sake of simplification. "We can build a theory of a system at the level of dynamics that is observable, in *ignorance* of the detailed structure or dynamics at the next level down."[33]

The autonomy permitted on each level is, of course, accompanied by a set of constraints to coordinate and integrate the submodel's behaviour. The control exerted through these constraints is closely related to the amount and type of *information* collected at each level. Simon's point is that near-decomposability of hierarchies minimizes information flows between levels, and hence between submodels. At this point, the first clue to the role which information theory could play in hierarchical systems analysis emerges.

As we move down our five-level hierarchy, at each step we progress to a model in which behaviour is increasingly disaggregated on a spatial basis. In doing so, we face an increasingly difficult data problem: that of making efficient use of the information which is available at the higher, more aggregated level, to coordinate the patterns of behaviour at the more disaggregated levels below. But this is, in fact, the very problem which is central to this whole dissertation.[34] Information theory can therefore play a very useful role in our hierarchical modelling system.

In what follows, we shall try to distinguish between the *structural* and *functional* aspects of our hierarchical system. Koestler relates the former to the spatial properties of the system, and the latter to processes over time. Evidently, structure and function are not easily separated, and represent complementary aspects of an indivisible spatio-temporal process.[35] By regarding each model (level) in our hierarchy as being responsible for a certain degree of detail, a separability of focus is maintained, leading to an efficient *specialization of function* at each level in the hierarchy.

Since our major objective is the development of a flexible framework for the integrated analysis of national and regional development options, from hereon we shall concentrate on these two levels (2 and 3) in our hierarchy. The three other levels will not be examined explicitly, although any constraining influences on the national and regional levels will certainly be considered. In particular, the fundamental need to impose certain constraints on national development from the international level will be recognized.

In the following section, an *equilibrating* function is proposed for the national level. At the intermediate level of regional developments, a *satisficing* function is suggested, based on the need for compromise solutions. At even lower levels, where the decisions of individuals are more easily recognized, the logical function is one of *optimization*. If space and time permitted, alternative functional arrangements would certainly warrant investigation.

7.3 National Development Scenarios

It is argued in this section that the generation of different development scenarios is an important feature of the national planning process. In this way, the consequences of different technological assumptions and pricing policies can be quantified and compared. To do this efficiently, computer-aided procedures are not only useful, but generally necessary, since the number of fundamental dimensions and their degree of disaggregation can escalate very quickly.

For simplicity, we shall concentrate on equilibrium solutions to the various national development scenarios. While it is agreed that a number of other important, and perhaps conflicting, objectives may exist at the national level, we shall defer any discussion of multiple objectives until we reach the regional level. The current preoccupation with equilibrium solutions is motivated by a desire to avoid national balance of payments problems, and excessive reliance on volatile world markets. As we shall see shortly, it is possible to relax this equilibrium requirement to some extent, by classifying certain goods as world industries (W) which can remain unbalanced (individually) at the national level.

7.3.1 Model Structure

The basic form of our *hierarchical input-output model* is depicted in Table 7.2. The economy is divided into n_W world industries (or world commodities), n_N national industries, n_R regional industries, n_L local industries and, if necessary, n_H household classes.[36] *Strategic aggregation* of these industries will be an important feature of this modelling

TABLE 7.2

SUPPLY \ DEMAND		World Industries 1 2 ---- n_W	National Industries 1 2 ----- n_N	Regional Industries 1 2 ---- n_R	Local Industries 1 2 ----- n_L	Households 1 2 ----- n_H
World Industries	1 2 ⋮ n_W	X_{WW}	X_{WN}	X_{WR}	X_{WL}	X_{WH}
National Industries	1 2 ⋮ n_N	X_{NW}	X_{NN}	X_{NR}	X_{NL}	X_{NH}
Regional Industries	1 2 ⋮ n_R	X_{RW}	X_{RN}	X_{RR}	X_{RL}	X_{RH}
Local Industries	1 2 ⋮ n_L	X_{LW}	X_{LN}	X_{LR}	X_{LL}	X_{LH}
Households	1 2 ⋮ n_H	X_{HW}	X_{HN}	X_{HR}	X_{HL}	X_{HH}

system, since the degree of detail which should enter into the national model will differ from that needed at other levels.[37] The national input-output and capital matrices, which define the various technological scenarios at the national level, will be defined in terms of this strategic industrial classification.

Whereas Leontief's open dynamic model is well suited to investigating the implications of alternative policy decisions,[38] his closed version is more suitable for evaluating the development potential (or efficiency) of alternative technologies. The closed model is capable of investigating the effects of changes in the input structure of each industry (for example, the relative proportions of basic materials, energy, labour, and capital equipment), together with changes in the relative prices of these inputs, on the maximum attainable growth rate of the economy in the long run. It thus provides us with a convenient means of comparing various development scenarios based on alternative technologies at the national level.

We shall therefore return to the n-sector economy introduced in Chapters 4 and 6, and specify our closed dynamic input-output model in the form

$$x_i(t) = \sum_{j=1}^{n} a_{ij}\, x_j(t) + \sum_{j=1}^{n} b_{ij}[x_j(t+1) - x_j(t)] \tag{7.2}$$

which simply states that total production in each industry i (which includes world industries that are exported) is equal to its current use by all other industries, plus its investment use in the form of capital expansion to facilitate increased industrial production in the next period.

Before we discuss the mechanism by which the input-output (a_{ij}) and capital (b_{ij}) matrices may actually be determined, it is necessary to clarify the respective roles of *world industries* and *households* in this national model. World industries (W) are defined as those whose commodity flows remain significantly unbalanced (in terms of production and consumption) at the national level. They can only find balance on a world level. For the modelling exercise at the national level, these n_W industries will be aggregated to form *one* sector, which we can define as the *world trade sector*.

The advantage of combining all the world industries into one sector relates to the nature of the closed model. By restricting the activities of world industries to one sector, we ensure that an overall balance of trade is preserved, but still allow individual world industries to remain unbalanced. This is a typical example of strategic aggregation.

One simplified means of estimating the import and export coefficients associated with each industry i, would be to calculate each industry's share of total imports or exports in terms of a probability distribution. If plausible assumptions can be made concerning, say, the export potential for each industry's products on the world market, subject to certain constraints imposed by domestic demand, a probabilistic model for export demand could be developed. The result may be similar to one of the standard statistical forms available in entropy modelling.[39] Although a similar approach could be adopted for import coefficients, this probabilistic approach does disregard price movements. It would be preferable to make use of more sophisticated trade models to determine suitable import and export coefficients.[40]

In our national model, the behaviour of households will also be restricted to one sector. This sector can be viewed as an additional industry supplying its output, essentially labour, to other industries and, in return, receiving consumer goods and households durables from those industries. If certain assumptions are made concerning household demand for the products of each industry i, it it possible to formulate a similar probabilistic model to the one suggested for exports. Although this simplified approach to consumer behaviour could be adopted initially, a more sophisticated analysis of the household sector would be needed eventually. For the present, we shall treat patterns of migration as exogenously determined, and regard the supply of labour as unrestricted.

7.3.2 Choice of Production Techniques

The basic elements of our national model now consist of a single world trade sector, n_N national industries, n_R regional industries, n_L local industries, and a single household sector. The exogenous derivation of suitable coefficients for the world trade and household sectors has already been discussed. In this section, we shall describe a means by which the choice between various alternative production techniques might be made using equilibrium prices at the national level. In the following section, we shall demonstrate the connection between this equilibrium price solution and the balanced growth solution.

A meaningful choice between various possible production techniques (that is, columns of input-output and capital coefficients) for each industry i can be made by consideration of input prices at equilibrium.

Let us now introduce a non-arbitrary system of equilibrium prices,[41] known also as production prices according to Marxian terminology. These prices, to be denoted by $p_1,.......,p_n$, are defined by

$$p_j = \sum_{i=1}^{n} p_i a_{ij} + r \sum_{i=1}^{n} p_i b_{ij} . \quad (j = 1,...,n) \tag{7.3}$$

The price of each commodity is thus required to cover the costs of intermediate goods plus a capital charge, or rate of interest r, on the value of capital tied up in production. In matrix notation, we require the following balance for general equilibrium:

$$p = pA + rpB . \tag{7.4}$$

If our criterion of optimality for choosing between alternative production techniques is simply the minimization of costs in the long run, we must choose that set of techniques (A, B) which minimizes total costs computed according to (7.4). In the following section, we shall describe a scheme which enables us to find this least cost equilibrium solution.

7.3.3 Balanced Growth Solutions

Returning to our system of difference equations (7.2), which describe the development of production in the various sectors when we assume full utilization of capital stocks, we can find a balanced growth path for the system by trying proportionate growth at rate, λ , where

$$x_i(t) = x_i(0)[1 + \lambda]^t \tag{7.5}$$

Substituting (7.5) into (7.2), and writing x_i instead of $x_i(0)$, we obtain

$$x_i = \sum_{j=1}^{n} a_{ij} x_j + \lambda \sum_{j=1}^{n} b_{ij} x_j \qquad (i = 1,\ldots,n) \tag{7.6}$$

or, in obvious matrix notation

$$x = Ax + \lambda Bx \quad . \tag{7.7}$$

Only one particular growth rate, which we shall call λ^* , will yield a sustainable vector (x^*) which contains non-negative values in all time periods. This balanced growth path corresponds to the fastest attainable rate of expansion of the system. The corresponding outputs (x^*) represent a unique equilibrium structure which is commonly referred to as the *turnpike* solution.

It is obvious that Equations (7.7) form an eigensystem for any value of λ, and actually represent a particular solution to the general system

$$\mu x = Ax + \lambda' Bx \tag{7.8}$$

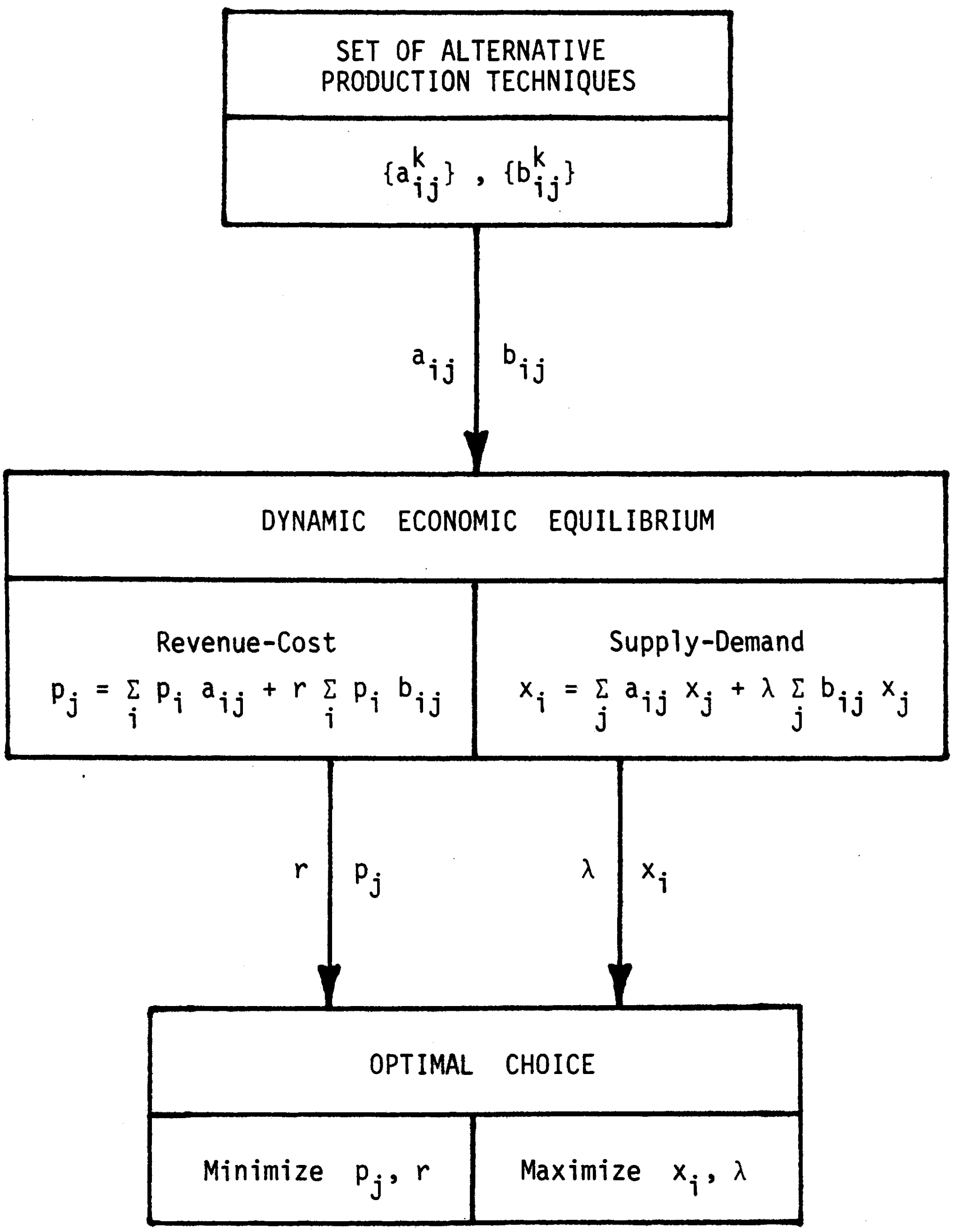
For any given value of λ', we can calculate the values of μ and x using a simple iterative approach (which is described in Appendix C). It is therefore possible to find the turnpike solution by adjusting the value of λ' until $\mu = 1$. So, for any given set of techniques (A, B), we can determine the maximal rate of balanced equilibrium growth, namely λ^* .

We also note that system (7.6) is of the same type as (7.3), only with the coefficient matrices transposed. Equations (7.3) do, in fact, represent the dual formulation corresponding to Equations (7.6) as the primal system.

Herein lies the vital connection between the prices and quantities in our closed system at equilibrium. Thus for each balanced growth rate λ^*, there is a corresponding interest rate r^* of the same value. Of major importance is this fact that the balanced growth solution is identical to our required least cost solution. This result leads us to the necessary *dynamic non-substitution theorem*, which can indicate correctly whether it is advantageous to substitute one technology for another in an industry, if the balanced growth rate (or overall profit rate) for the whole system is the main concern.[42]

Figure 7.3 depicts the general procedure for evaluating alternative development paths. The set of technological choices for each industry j consists of two groups of column vectors for each technique k. One contains alternative input-output coefficients, a_{ij}^k, and the other the corresponding capital coefficients b_{ij}^k. A complete set of coefficients $\{a_{ij}, b_{ij}\}$ is fed into the equilibrium model, which then calculates the balanced growth solution. The associated output and price vectors are retained. Further technology sets are analysed in an identical fashion, until all the relevant scenarios have been generated. The natural choice would be the set of techniques which minimizes costs, and maximizes growth of production, in the long run.

Figure 7.3


SET OF ALTERNATIVE PRODUCTION TECHNIQUES
$\{a^k_{ij}\}$, $\{b^k_{ij}\}$
a_{ij} b_{ij}
DYNAMIC ECONOMIC EQUILIBRIUM
Revenue-Cost
$p_j = \sum_i p_i a_{ij} + r \sum_i p_i b_{ij}$
Supply-Demand
$x_i = \sum_j a_{ij} x_j + \lambda \sum_j b_{ij} x_j$
r p_j
λ x_i
OPTIMAL CHOICE
Minimize p_j, r
Maximize x_i, λ

7.4 The National-Regional Interface

The general procedure discussed in Section 7.3 permits us to describe the interdependent pattern of economic activities at the national level. The production techniques for each industry i are defined in terms of a set of input-output and capital coefficients, which both include additional sectors for world trade and household activity. For the n_N national industries, these coefficients are assumed to describe the actual production functions employed. For the n_R regional industries, the coefficients simply act as co-ordinating constraints to integrate the productive activities of all R regions as a whole.

Traditionally, in hierarchical systems analysis we must allow these regional industries (and also the local industries) relative autonomy in choosing their own production techniques at the regional level. Such autonomy must be accompanied by certain constraints to coordinate the activities of individual regions.[43] As it turns out, in Chapters 5 and 6 we have already introduced a framework which permits the self-assertive, independent tendencies of each region to be tempered by the integrative tendencies of the nation as a whole.

7.4.1 Interregional Flows

In Section 6.3, we proposed various formulations to estimate the interregional pattern of gross intersectoral flows $\{f_{ij}^{rs}\}$ from a limited database of industrial and regional information. In each case, one set of constraints related to industrial information which ensured that the flows were consistent with activity at the national level. Another set described the imbalances between the regional production levels in various industries, and each region's demand for products from those same sectors. The results of this estimation process were a set of gross flows which allowed interregional tables of input-output and capital coefficients to be calculated simultaneously.

In our present hierarchical system, we can adopt this type of formulation to estimate the vital structural properties of our national-regional interface. Information theory can be used to calculate an interregional pattern which is consistent with both the structure of production at the national level, *and* supply-demand imbalances within each region. Using an identical notation to that introduced in Chapters 5 and 6, we shall derive an appropriate formulation for the case in which total demand (x_i^{*s}) for the products of each industry i is known for each region s.

The following two sets of constraints integrate regional activities at the national level:[44]

$$\sum_{r=1}^{m} \sum_{s=1}^{m} \tilde{a}_{ij}^{rs} \, x_j^{s*} = a_{ij} \, x_j \qquad (i = 1,\ldots,n;\ j \in n_R, n_L) \qquad (7.9)$$

$$\sum_{r=1}^{m} \sum_{s=1}^{m} \tilde{b}_{ij}^{rs} \, g_j^s \, x_j^{s*} = b_{ij} \, g_j \, x_j \qquad (i = 1,\ldots,n;\ j \in n_R, n_L) \qquad (7.10)$$

whereas the next three sets allow for supply-demand imbalances within each region:

$$\sum_{j \in n_R, n_L} \sum_{s=1}^{m} (\tilde{a}_{ij}^{rs} + \tilde{b}_{ij}^{rs} g_j^s) x_j^{s*} = x_i^{r*} - \sum_{j \in n_N'} \sum_{s=1}^{m} (a_{ij} + b_{ij} g_j) x_j^{s*} \tag{7.11}$$

$$(i = 1,\ldots,n;\ r = 1,\ldots,m)$$

$$\sum_{i=1}^{n} \sum_{r=1}^{m} \tilde{a}_{ij}^{rs} x_j^{s*} = x_j^{s*} - u_j^s \qquad (j \in n_R, n_L;\ s = 1,\ldots,m) \tag{7.12}$$

$$\sum_{j \in n_R, n_L} \sum_{r=1}^{m} (\tilde{a}_{ij}^{rs} + \tilde{b}_{ij}^{rs} g_j^s) x_j^{s*} = x_i^{*s} - \sum_{j \in n_N'} \sum_{r=1}^{m} (a_{ij} + b_{ij} g_j) x_j^{s*} \tag{7.13}$$

$$(i = 1,\ldots,n;s = 1,\ldots,m)$$

Equations (7.11) imply that intraregional production levels may be influenced largely by demands in other regions, whereas Equations (7.13) suggest that levels of intraregional demand may be satisfied mainly by inflows from other regions.

A standard Lagrangian derivation to the resulting entropy-maximizing problem yields identical expressions to those given earlier in (6.29) and (6.30). Explicit solutions for $\{\tilde{a}_{ij}^{rs}\}$ and $\{\tilde{b}_{ij}^{rs}\}$ may again be obtained using the INTEREG program, which is described in limited detail in Appendix E.

7.4.2 Intraregional Production Techniques

The results provided by our information-theoretical approach to interregional estimation are a nationally consistent set of input-output and capital coefficients, which could perhaps be used to generate various scenarios of multiregional development patterns. We do not intend, however, to pursue the interregional problem any further at this point. To extend our hierarchical analysis, we must now descend to the level of a single region.

An individual region's production techniques (in the regional and local industries) can be calculated from the full interregional matrices. Clearly, $\{a_{ij}^{rr}\}$ and $\{b_{ij}^{rr}\}$ describe intraregional input-output and capital structure. Total intranational exports, e_i^r, from sector i in region r are given by

$$e_i^r = \sum_j \sum_{s \neq r} f_{ij}^{rs} = x_i^{r*} - x_i^{*r} \qquad (7.14)$$

whereas intranational imports, m_j^s, used by sector j in region s appear as

$$m_j^s = \sum_i \sum_{r \neq s} f_{ij}^{rs} = x_j^{*s} - x_j^{s*} \qquad (7.15)$$

From these two relationships, import and export coefficients for each region may be calculated, thereby complementing the A and B matrices computed earlier. A complete structural description of intraregional production techniques can thus be assembled.

Figure 7.4

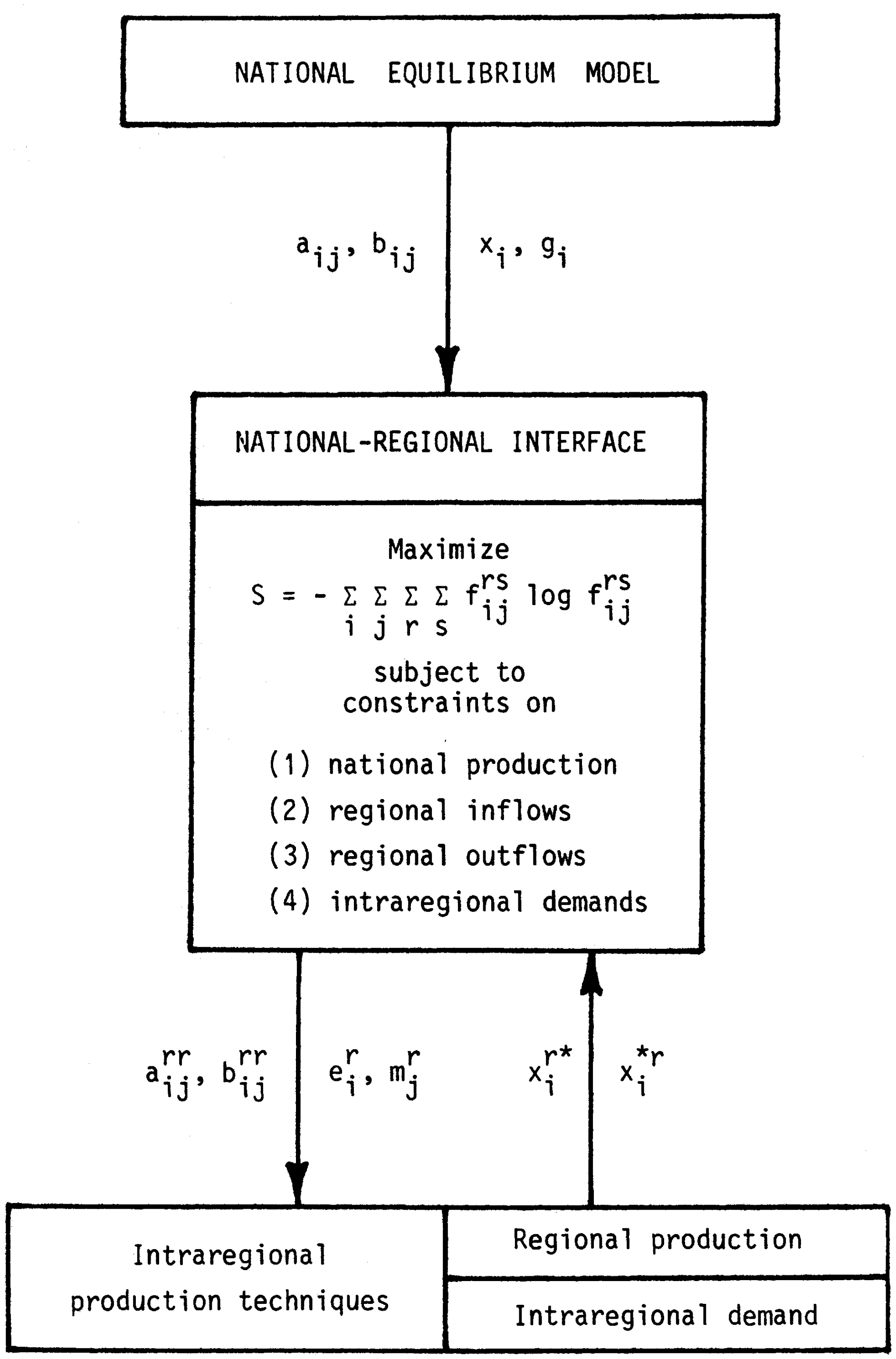

NATIONAL EQUILIBRIUM MODEL
a_{ij}, b_{ij}
x_i, g_i
NATIONAL-REGIONAL INTERFACE
Maximize
$S = -\sum_i \sum_j \sum_r \sum_s f_{ij}^{rs} \log f_{ij}^{rs}$
subject to
constraints on
(1) national production
(2) regional inflows
(3) regional outflows
(4) intraregional demands
a_{ij}^{rr}, b_{ij}^{rr}
e_i^r, m_j^r
x_i^{r*}
x_i^{*r}
Intraregional
production techniques
Regional production
Intraregional demand

Figure 7.4 depicts the overall framework, which permits each region some degree of autonomy. The coordinating role of the national level is apparent, together with a strong interdependency between the supply-demand imbalances within each region (the self-determined intraregional variables), and the estimation of intraregional production techniques. We shall now concentrate our discussion on the various forces which may be responsible for the development of these imbalances between each region's levels of production in various industries, and its demand for products from those same sectors.

7.5 Regional Development Scenarios

The generation of plausible development scenarios is also a significant feature of the regional planning process. In this way, the consequences of different policies or goals can be quantified and compared. Because of the open nature of regional economies, we shall not investigate their feasible development paths in a manner similar to the closed equilibrium approach adopted at the national level.[45] Instead, we shall assume that the predominant function of regional economic planning is to satisfy (in some way) a multiplicity of objectives or decisions problems.

Traditional planning and optimization techniques presupposed a clear and unambiguous specification of individual planning goals, their policy instruments, and all the constraints imposed upon the decision-making process. In most early formulations, a single "high-priority" goal (such as minimum cost or maximum consumption) was adopted as the objective function. More recently, optimizing principles for community welfare have been formulated by including various goals in the objective function, and then assigning arbitrary weights to each goal in order to simulate potential trade-offs.

Examples of both single and multiple-objective programming models are included in our earlier model review (see Table 7.1). The various objective functions employed by the optimization models included therein is summarized in Table 7.3. This historical progression from single to multiple objectives reflects a growing insight into economic decision-making, in which a broad spectrum of conflicting goals can often be observed.

TABLE 7.3

MODEL NAME	YEAR PUBLISHED	OBJECTIVE FUNCTION
Maryland	1970	Minimize transport costs
Indian	1972	Minimize transport costs
DREAM	1976	Maximize net surplus (exports - imports - transport costs)
Dutch	1978	Maximize production Maximize employment Minimize pollution
MORSE	1980	Maximize weighted sum of consumption plus employment less energy use

7.5.1 Conflicting Objectives

From a theoretical viewpoint, traditional optimization procedures are extremely elegant, since they provide a rigorous means of evaluating alternative strategies on the basis of their contribution to overall community welfare. As a practical tool in the regional planning process, however, their value is rather limited.[46] A major weakness relates to the specification of a regional welfare function *a priori*, which presupposes complete information about all possible actions, and the potential trade-offs between these actions. Furthermore, simultaneous optimization of a multiplicity of criteria is usually impossible (i.e. no feasible solution may exist), due to the conflicting nature of various objectives.

In view of these shortcomings, various writers have emphasized the need to adopt satisficing principles,[47] designed to converge on a satisfactory compromise solution. This view has led to the development of various multi-objective decision-making tools.[48] It is not our intention to review these approaches here. We shall simply outline an elementary satisficing formulation which takes advantage of our existing database of regional information. For a detailed review of alternative approaches to regional decision problems, the interested reader is directed elsewhere.[49]

7.5.2 A Simple Satisficing Formulation

Consider a regional economy in which the welfare profile contains five elements:

(i) consumption; (ii) employment; (iii) balance of trade; (iv) energy use; and (v) pollution.

The following set of mathematical relationships will be used to build a model which contains acceptable criteria for each of the above elements:

7.5.2.1 Minimum consumption criterion

A minimum per capita level of consumption is desirable to ensure an acceptable standard of living in the region, namely

$$\sum_i c_{it}^r \geq \omega^r P_t^r \tag{7.16}$$

where c_{it}^r is the intraregional consumption of products from sector i at time t, ω^r is the minimum wage rate per capita, and P_t^r is the population at time t.

7.5.2.2 Minimum employment criterion

Although the level of employment activity in any region has an upper limit which is dependent on available labour in the region, it is also desirable to control the level of unemployment by setting a lower limit on the activity rate. We thus have

$$\alpha_{t(min)}^r P_t^r \leq \sum_i \ell_{it}^r x_{it}^r \leq \alpha_{t(max)}^r P_t^r \tag{7.17}$$

where α_t^r is an activity rate (number of workers per unit population) in the region at time t, ℓ_{it}^r is the regional labour-output coefficient at time t for sector i, and x_{it}^r is the level of regional production for sector i at time t.

7.5.2.3 Balance of trade criterion

There is also a strong incentive for each region to achieve and maintain a positive net surplus or balance of trade

$$\sum_i (e^r_{it} - m^r_i \, x^r_{it}) \tag{7.18}$$

where m^r_i is the intraregional import coefficient for sector i, and e^r_{it} is the total level of exports from sector i in region r at time t.

7.5.2.4 Maximum energy use criterion

Goals in energy policy can be translated into maximum consumption levels for various forms of energy, namely

$$\sum_i (\mu_{ki} \, x^r_{it} + \eta_{ki} \, c^r_{it}) \leq U^r_{kt} \tag{7.19}$$

where μ_{ki} and η_{ki} are energy coefficients for production and consumption, respectively (the amount of energy of type k used in sector i), and U^r_{kt} is the upper bound on regional energy consumption of type k in period t.

7.5.2.5 Maximum levels of acceptable pollution

Emission levels of various pollutants are an important aspect of environmental policy, so upper bounds on these emissions are also required, namely

$$\sum_i \lambda_{ki} \, x^r_{it} < \beta^r_{kt} \tag{7.20}$$

where λ_{ki} is a matrix of pollution emission coefficients, which measure the amount of pollutant of type k per unit of output in sector i, and β_{kt} is the upper bound on regional emissions of pollutant k over time t.

7.5.2.6 Commodity balance

A final requirement is that a regionalized dynamic Leontief relationship of the form

$$(1 + m_i^r)x_{it}^r \geq \sum_j a_{ij}^{rr}\, x_{jt}^r + \sum_j b_{ij}^{rr}\, g_j^r\, x_{jt}^r + c_{it}^r + e_{it}^r \tag{7.21}$$

be satisfied for all sectors i in each region r at time t. Equations (7.21) simply ensure that the total supply of any sector in our regional economy exceeds its use by other sectors, whether for intermediate or final consumption, capital expansion, or export to other regions.

A convenient feature of the system of inequalities described in (7.16) through (7.21) is that they can all be reformulated in terms of our self-determining intraregional variables (x_i^{r*}, x_j^{*s}). Since the latter define the supply-demand imbalances, which are fundamental inputs to our inter-regional estimation problem, the determination of their feasible values is our prime concern. For convenience, we shall represent this set of all endogenous variables by a vector, p_i^r, and designate our base values by p_{i0}^r.

The salient feature of the system of inequalities (7.16) through (7.21) is the existence of a multiplicity of conflicting criteria (maximization of consumption *and* employment *and* net surplus, together with minimization of energy use *and* pollution). Simultaneous optimization of all five objectives is obviously impossible. However, the optimization of any single objective will also prevent all or most of the others from attaining their potential optima.

7.5.3 Compromise Solutions

Various approaches to this type of problem have been proposed. A central role in most of these is played by *Pareto-efficient* solutions. Ideal points, which are defined as the maximum feasible value of one objective function, are often used as reference points. Given a vector of ideal solutions, a compromise solution can be attained by selecting a point on the Pareto frontier (i.e. the set of efficient solutions) which minimizes the discrepancy (based, for example, on a distance metric) between the set of efficient solutions and the ideal set.[50]

In our present example, it is suggested that the five objectives first be ranked in hierarchical order. This approach corresponds to a *lexicographic* ordering of alternatives, which closely resembles the decision-making technique used by many policy-makers. An initial base period solution is provided by some known base value (p^r_{i0}), or by those values corresponding to the five ideal points. The first step in our compromise process is to combine the inequality constraints corresponding to the criterion of highest priority,[51] together with (7.21), to form the first constraint set. Our first compromise solution is obtained by minimizing the information gain statistic. Formally, we have

$$\text{Minimize} \sum_i p^r_{i1} \log(p^r_{i1}/p^r_{i0}) \qquad (7.22)$$

subject to one of (7.16) through (7.20), (7.21), and the usual non-negativity conditions.

The significance of using the information measure lies in its ability to maximize the resistance of our base period solution (p_{i0}) to any change. It thereby retains as much of the region's structural characteristics as the additional criteria permits. The minimization of (7.22) is approximately equivalent to the minimization of the Chi-squared statistic :

$$\sum_i (p_{i1} - p_{i0})^2 / p_{i0} \quad .$$

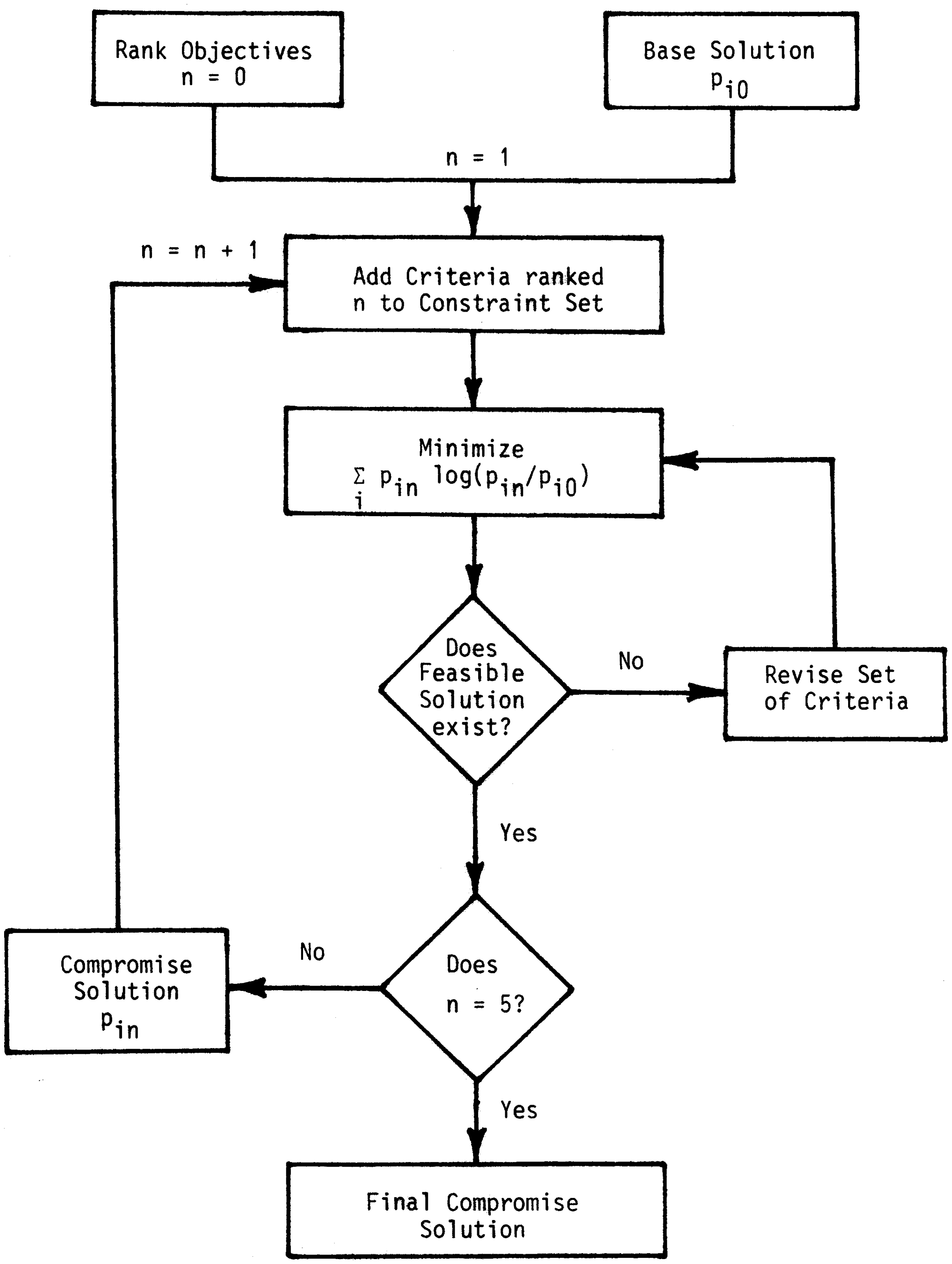
The second step in our sequential process consists of confronting the base solution (p_{i0}) with those constraints corresponding to the criterion of second highest priority, as well as the first set of constraints. A second compromise solution (p_{i2}) is then obtained by minimizing a similar information gain statistic. In general, step n in the compromise process consists of

$$\text{minimizing} \sum_i p_{in} \log(p_{in}/p_{i0})$$

subject to (7.21) and the inequality constraints corresponding to the first n (out of 5) criteria.

The various steps in this compromise process are depicted in Figure 7.5. The value of sequential introduction of various constraints (instead of simultaneous optimization) lies in the ability to identify infeasible solutions. It is quite possible that the simultaneous presence of all criteria precludes *any* feasible solution. However, the sequential inclusion of each criterion enables the transition to infeasibility to be detected. Iterative revision of one or more criteria is then possible.

Figure 7.5


Rank Objectives
n = 0
Base Solution
p_{i0}
n = 1
n = n + 1
Add Criteria ranked n to Constraint Set
Minimize
$\sum_i p_{in} \log(p_{in}/p_{i0})$
Does Feasible Solution exist?
No
Revise Set of Criteria
Yes
Compromise Solution
p_{in}
No
Does n = 5?
Yes
Final Compromise Solution

An interesting extension of this approach would be to search for a unique feasible solution, by successive revision of certain criteria.

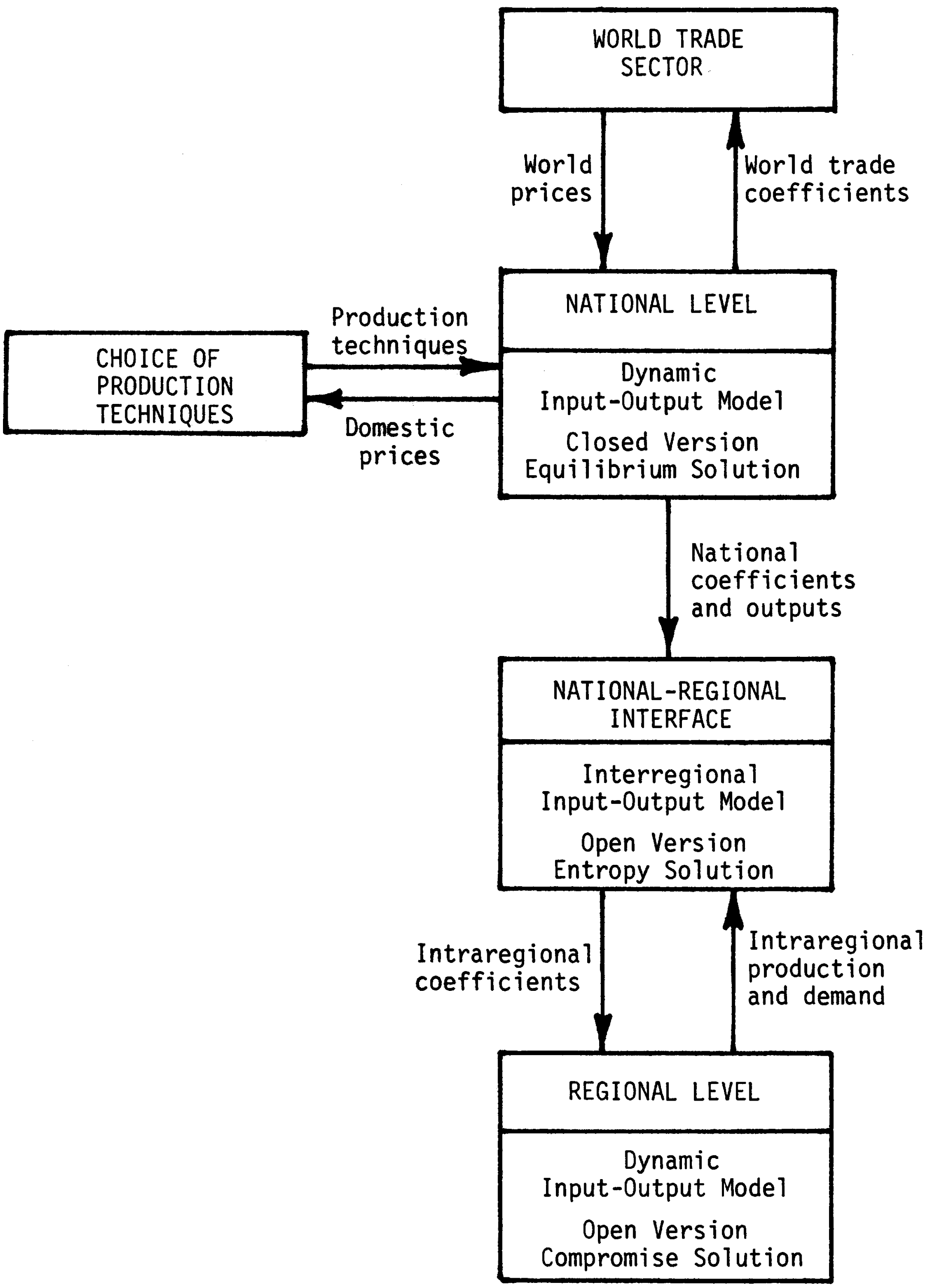
The proposed iterative procedure involves the adoption of a satisficing, rather than an optimizing concept, so that the final compromise solution complies with certain basic achievement levels specified by the various decision-makers. Such a notion of bounded rationality is relevant to regional development, which must confront different welfare goals, resource and energy constraints, consumption patterns, and demographic influences.

7.6 Concluding Remarks

A general summary of the interrelationships existing between our modelling system at the national and regional levels is depicted in Figure 7.6. It is abundantly clear that the role of information theory is not one of estimating the actual development path *at* each level, but rather of providing some structural connectivity *between* the two levels. By computing certain gaps in the information flow between these two levels, it does not restrict the self-determining principles adopted at each level. The control exerted through the constraints is closely related to the *amount* and *type* of information collected at each level.

The same procedure can be adopted between other levels in our hierarchical modelling system (see Figure 7.2). As we move down this hierarchy, the increasingly disaggregated information required at each level can be estimated using (i) a set of constraints to coordinate and integrate each

Figure 7.6


WORLD TRADE SECTOR
World prices
World trade coefficients
CHOICE OF PRODUCTION TECHNIQUES
Production techniques
Domestic prices
NATIONAL LEVEL
Dynamic Input-Output Model
Closed Version Equilibrium Solution
National coefficients and outputs
NATIONAL-REGIONAL INTERFACE
Interregional Input-Output Model
Open Version Entropy Solution
Intraregional coefficients
Intraregional production and demand
REGIONAL LEVEL
Dynamic Input-Output Model
Open Version Compromise Solution

subsystem's behaviour, and (ii) another counterbalancing set embodying the autonomous, self-assertive tendencies operating at each level. Information theory can therefore contribute to the fundamental process of *dissectibility* or *decomposition* in hierarchical systems analysis.

But hierarchical theory is not essential to demonstrate the complementary function of information theory in the analysis of national and regional economies. Nor is the system of models proposed in this chapter considered to be the only, or even the optimal, road towards an integrated spatial development system. In reality, there is ample scope to modify the model formulations at each level, or even to discard the hierarchical arrangements completely.

The main purpose of this and earlier chapters has been to use various forms of information which are available on an aggregate level to derive *unbiased* estimates of commodity flow behaviour on a disaggregate level. Information theory provides a number of versatile tools which are of particular relevance to this endeavour. Nevertheless, a proper realization of the full potential of information-theoretical methods has not yet eventuated. It is sincerely hoped that this dissertation may help to promote such an awareness.

FOOTNOTES FOR CHAPTER 7

1 As alluded to, for example, by Richardson (1973) and Riefler (1973).

2 This approach has been prompted by the ideas of certain hierarchical theorists, such as Koestler (1967), Mesarovic et al. (1970), Simon (1973) and Pattee (1973).

3 See Simon (1958).

4 A point emphasized by Richardson (1973).

5 For an introduction to the latter, see Tinbergen (1967), or Mennes, Tinbergen and Waardenburg (1969).

6 Models involving comparative statics will also be considered.

7 See Moses (1955; 1960).

8 See Miernyk et al. (1970).

9 See Equations (6.2).

10 In reality, $d_{ij} = q_i b_{ij}$. The assumption embodied in Equations (4.37) and (4.38) imposes more modest data requirements than Miernyk's approach.

11 Almon (1966) developed a sophisticated closed version of Leontief's dynamic model, which incorporated nonlinearities and changing technology, to complete a balanced growth rate for the American economy. See Almon et al. (1975) for later forecasts. Harris (1970) developed his interregional extension from Almon's early work.

12 See Harris (1970, p 174).

13 See Mathur (1972).

14 See Mathur (1972, p 220).

15 See Andersson (1975).

16 See Andersson and Persson (1979, p 40).

17 See Andersson and Karlqvist (1979), or Andersson and Persson (1979).

18 See Funck and Rembold (1975).

19 See Sharpe and Batten (1976), and Karlqvist et al. (1978).

20 For the original iterative scheme, involving piecewise linear approximations, see Brotchie, Toakley and Sharpe (1971) or Brotchie, Dickey and Sharpe (1980). For a discussion of the version containing Eriksson's (1980) entropy algorithm, see Sharpe, Batten and Anderson (1981).

21 See, for example, Sharpe and Batten (1976), Sharpe et al. (1977), Karlqvist et al. (1978), Sharpe, Ohlsson and Batten (1979), and Sharpe, Batten and Anderson (1981).

22 See Hafkamp and Nijkamp (1978; 1981).

23 See, for example, Simon (1958) or van Delft and Nijkamp (1977).

24 The MORSE model is described in Lundqvist (1980), and builds on earlier regional foundations laid by Snickars and Lundqvist (1978).

25 See, for example, Lesuis, Muller and Nijkamp (1980).

26 A point emphasized, amongst others, by Kornai (1971).

27 See Kaniss (1978) or Isard and Liossatos (1979, Chapter 10 and Section 12.4).

28 See Isard (1975) or Isard and Liossatos (1979, pp 285-287).

29 This notion of balance between the autonomous and constrained tendencies of subsystems in a hierarchy has been emphasized by Koestler (1967). In a well-adjusted subsystem, the self-assertive tendency and its opposite, the *integrative* tendency, are more or less equally balanced. Koestler coined the term *holon* to describe hierarchical subsystems, and to stress that they exhibit both the properties of independent wholes in certain domains, and those of dependent parts in other domains. For later presentation of similar ideas, see Mesarovic et al. (1970), Patte (1973), and Simon (1973).

30 A point emphasized by Mesarovic, Macko and Takahara (1970).

31 See, for example, Tinbergen (1967) or Karlqvist et al. (1978).

32 See Simon (1973).

33 See Simon (1973, pp 110-117).

34 See Section 1.1.

35 See Koestler (1967, p 59).

36 For this exploratory exercise, all commodities are assumed homogeneous, and the one-to-one correspondence between industries and commodities (implicit in Leontief's original model) is also retained.

37 Isard and Liossatos (1979, p 284 n) emphasized this need by taking an interregional linear programming model as an example. On a world level, this could be highly disaggregated, involving say 100 nations, each with 50 sectors; each sector could be an aggregate for a nation. The same type of model might also be adopted at a national level. Here there might be 20 regions with 100 sectors in each region, two or more sectors of this national model corresponding to a single sector of the international model, with national magnitudes of the international model being disaggregated by region in the interregional model. At the regional level, we might have a linear programming model involving 200 or more sectors for the region itself, two or more sectors of this model corresponding to a single sector in the national model. Furthermore, the distinction between skilled and unskilled labour in the household sector might be of critical importance in regional models, but irrelevant to national and international linear programming models.

38 The logic of this open system is simply that by treating final demands exogenously, they may be viewed as the objective function of the economic process. The generic model applicable to the theory of optimal control is also an open one, in which the behaviour of the system is determined by exogenous variables. Open systems may therefore be stabilized by allowing excess supply or demand conditions to develop, and then introducing a suitable process of control to promote stabilization.

39 See, for example, Fisk and Brown (1975 a, b).

40 Fairly extensive work on an international trade model has been undertaken jointly by the International Institute of Applied Systems Analysis and the University of Maryland. The resulting INFORUM model is based on an assumption of slowly changing trade shares which are regulated by movements in the prices of traded commodities in the producing countries and on the world market. The production and price possibilities in each country are predicted using dynamic input-output theory. For further details of the INFORUM model, see Almon (1966; 1975) and Nyhus (1980).

41 So named by Brôdy (1970) and Johansen (1973).

42 See Johansen (1973, p 83).

43 Hierarchical theorists, such as Koestler (1967), Mesarovic, Macko and Takahara (1970), and Simon (1973), emphasize repeatedly that multilevel organization requires both independence of decision-making, and integrative or coordinative constraint on this autonomy, in such a way that overall stability and harmony results.

44 In earlier formulations, a set of capacity or cost constraints have sometimes been included, when the appropriate information is available. Since this type of constraint normally includes coefficients which strongly influence the distribution pattern between all regions, it seems preferable to exclude it from our present hierarchical formulation.

45 For further discussion of turnpike theorems and optimal growth paths at the regional level, see Fujita (1978).

46 A view taken, for example, by van Delft and Nijkamp (1977, p 8).

47 Based on Simon's (1957) notion of bounded rationality.

48 For a general background to multi-objective decision methods, see, amongst others, Cochrane and Zeleny (1973), Wallenius (1975), and Cohon (1978). For the application of these methods to decision problems in spatial systems, see Nijkamp and Rietveld (1976), van Delft and Nijkamp (1977), Nijkamp (1977; 1978), Blair (1978), and Rietveld (1979).

49 See, for example, Nijkamp (1977) or Rietveld (1979).

50 An approach described, for example, by Hafkamp and Nijkamp (1978; 1980) and Lesuis, Muller and Nijkamp (1980).

51 Namely, one of (7.16) to (7.20).

Appendix A

BASIC MICROSTATE DESCRIPTIONS

A very simple example is presented to enable the reader to grasp the four basic classes of microstate descriptions introduced in Chapter 3. Suppose we wish to locate four industries in two regions, region A having two industrial sites available and region B four sites. No further definite information is available. It is required to find the most probable macrostate, that is, the most probable number of industries in each region. If a typical macrostate is defined as (a, b), where a is the number of industries in region A and b is the number in region B, then the possible macrostates are (0, 4), (1, 3), (2, 2), (3, 1) and (4, 0). The basic microstate descriptions corresponding to our four elementary distributions are now considered in turn.

(i) *Bolzmann entropy*

In the earliest and still the most popular analogy, two arrangements would constitute two different microstates if, and only if, each industry is distinguishable and there are no restrictions on the number of industries permitted in each region. The possible number of microstates, W, is given by

$$W = \frac{4!}{a!\ b!}$$

and represents the number of ways of arranging the 4 industries into 2 groups containing a and b industries respectively. Thus, for each macrostate, W is easily calculated as

Macrostate:	(0.4)	(1.3)	(2.2)	(3.1)	(4.0)
W:	1	4	6	4	1

The reader may readily confirm that macrostate (2, 2), which yields 6 microstates, is the most probable one.

(2) *Maxwell-Boltzmann entropy*

In this instance, the microstate definition is based on Boltzmann's original conditions, together with a finer-grained microstate space to cater for the allocation of individual industries to particular sites in each region. Here W is given as

$$W = \frac{4!}{a!\ b!}\,(2^a)(4^b)$$

indicating, firstly, the number of ways of arranging the 4 industries into 2 groups containing a and b industries respectively, times, secondly, the number of ways of arranging a individual industries in the 2 individual sites of region A times that of arranging b individual industries in the 4 sites of region B. The results emerge as

Macrostate:	(0.4)	(1.3)	(2.2)	(3.1)	(4.0)
W:	256	512	384	128	16

giving (1, 3) as the most probable macrostate.

(3) *Fermi-Dirac entropy*

In this case, two arrangements constitute two different microstates if, and only if, each industry is identical and only one industry per site is permitted. The expression for W is then

$$W = \frac{2!}{a!(2-a)!} \cdot \frac{4!}{b!(4-b)!}$$

which defines the number of ways of arranging a identical industries within the 2 individual sites of region A ($a \leq 2$) times that of arranging b industries within the 4 sites of region B ($b \leq 4$). The results are

Macrostate:	(0.4)	(1.3)	(2.2)
W:	1	8	6

Macrostate (1, 3) again turns out to be the most probable.

(4) *Bose-Einstein entropy*

In this location problem, each industry is also regarded as identical, but more than one industry per site is allowed. Here W is given by

$$W = \frac{(2+a-1)!}{a!(2-1)!} \cdot \frac{(4+b-1)!}{b!(4-1)!}$$

which is the number of ways of arranging a identical industries within the 2 individual sites of region A times that of arranging b industries within the 4 sites of region B, with no limits on the number of industries per site. The results are

Macrostate:	(0.4)	(1.3)	(2.2)	(3.1)	(4.0)
W:	35	40	30	16	5

which confirm (1, 3) as the most probable macrostate.

This very simple example has indicated that (1, 3) is the most probable distribution in three out of the four cases examined. However, in the much larger problems which occur in practice, quite different macrostates may emerge as the most probable for each case.

Appendix B

INCOMPLETE PRIOR INFORMATION: A SIMPLE EXAMPLE

Consider a normalized 4 by 4 flow matrix $\{p_{ij}\}$ which is to be estimated using information-theoretical principles. Prior to the estimation, we are given certain historical information concerning an earlier flow matrix $\{q_{ij}\}$. This incomplete *a priori* information is shown in Table B.1. Before we can apply the principle of minimum information gain,[1] some assumptions about the missing entries must be made. The entropy-maximizing assumption of equi-probability is the least biased statistical stance we can take.

To complete the entries in our *a priori* flow matrix $\{q_{ij}\}$, we first note that

$$\sum_j q_{1j} + q_{21} + q_{24} + q_{41} + q_{42} + q_{44} = 1.0 - 0.4 - 0.15 = 0.45$$

Distributing this subtotal to the nine appropriate elements on an equi-probable basis, we have

$$q_{11} = q_{12} = q_{13} = q_{14} = q_{21} = q_{24} = q_{41} = q_{42} = q_{44} = 0.05$$

Using a similar argument, each element in the third row assumes the value of 0. 1. Our complete *a priori* flow matrix is given in Table B.2.

FOOTNOTE

1 Namely, minimize $\sum_i \sum_j p_{ij} \log (p_{ij}/q_{ij})$ subject to various constraints which contain all the available information about the *a posteriori* matrix $\{p_{ij}\}$.

Table B.1.

$i \backslash j$	1	2	3	4	$\sum_j q_{ij}$
1					
2		0	0.15		
3					0.4
4			0		
$\sum_i q_{ij}$					1.0

Table B.2.

$i \backslash j$	1	2	3	4	$\sum_j q_{ij}$
1	0.05	0.05	0.05	0.05	0.2
2	0.05	0	0.15	0.05	0.2
3	0.1	0.1	0.1	0.1	0.4
4	0.05	0.05	0	0.05	0.15
$\sum_i q_{ij}$	0.25	0.2	0.3	0.25	1.0

Appendix C

COMPUTING CAPITAL COEFFICIENTS AND TURNPIKE SOLUTIONS:

The DYNIO Package

The DYNIO package transforms the mathematical formulations described in Section 4.2.3 into the FORTRAN programming language, thereby facilitating the computer estimation of capital coefficients, using the Biproportionality Method.[1] It can also be used to calculate the balanced growth solution for a closed dynamic input-output model. The turnpike (or von Neumann) growth rate, together with the associated output vector, are found iteratively as the solution to an algebraic eigenvalue problem.

The package has been tested extensively on a seven-sector model adopted originally by Brōdy in a study of turnpike solutions for the American economy.[2] The starting point for all these calculations is Brōdy's set of input-output coefficients for the American economy in 1947 and 1958.[3] These matrices describe a closed economy which excludes foreign trade. Estimates of the corresponding capital coefficients are made using Equation (4.32) together with assumption (4.33) or (4.34).

The two capital matrices resulting from the use of Brōdy's flow coefficients, and his own supposition that capital goods from any sector i have a uniform turnover time (t_i) irrespective of their destination sector j,[4] share one feature in common: the appearance of negative coefficients in some sectors. Re-examination of his flow matrices sheds some light on this unexpected result. In some sectors we find that

$$\sum_{j=1}^{7} a_{ij}x_j > x_i$$

Owing to the unbalanced nature of Bródy's flow matrices, arising from the exclusion of foreign trade and some approximations in the household sector, the appropriate rows in the corresponding capital matrix have negative coefficients simply for consistency.[5] Consequently, his own capital matrices are inconsistent with the flow matrices upon which they are based. The former have relevance in an open economy, but the latter relate to an unbalanced, closed economy.

To restore the appropriate balance to the flow matrices, the input coefficients in the household sector have been adjusted to make proper allowance for the net effects of foreign trade and household consumption. The matrices of flow coefficients resulting from these corrections are reproduced in Table C.1. Tables C.2 and C.3 contain the matrices of capital coefficients emanating from the revised flow matrices and Bródy's assumption of origin-specific turnover times. His original estimates are included for comparison. The absence of negative coefficients confirms that no improperly balanced sectors remain in the flow matrices.

It is now appropriate to return to our original set of equilibrium equations introduced in Chapter 4. Bródy's closed economy implies that all $y_i = 0$, so we can rewrite Equation (4.28) in the form

$$x_i = \sum_{j=1}^{n} a_{ij}x_j + \lambda_0 \sum_{j=1}^{n} b_{ij}x_j \qquad (i = 1,\dots,n) \qquad \text{(C.1)}$$

TABLE C.1

1947		1	2	3	4	5	6	7
1	Agriculture, food, textiles	3924	400	195	798	98	839	2810
2	Chemicals, plastics, rubber, metals	298	3758	1876	857	238	218	227
3	Machinery, fabricated metal products	245	303	2586	1176	373	561	473
4	Construction, cement, glass	128	219	147	684	462	541	25
5	Fuel, electric utilities	122	380	66	217	3420	212	418
6	Transport, services, undistributed	1237	998	751	1650	895	1825	5957
7	Households	5094	2112	3721	3342	1561	3338	389
1958		**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	Agriculture, food, textiles	4014	319	168	662	149	619	2892
2	Chemicals, plastics, rubber, metals	294	3024	1420	716	191	226	283
3	Machinery, fabricated metal products	232	493	2521	1300	487	576	534
4	Construction, cement, glass	112	240	141	913	546	403	1
5	Fuel, electric utilities	152	497	101	247	3311	346	557
6	Transport, services, undistributed	1244	1479	1112	2081	1070	2142	6462
7	Households	3269	1837	2451	3809	526	4511	267

TABLE C.2

	Brōdy's Method	1	2	3	4	5	6	7
1	Agriculture, food, textiles	3924	400	195	798	98	839	2335
2	Chemicals, plastics, rubber, metals	745	9395	4690	2143	595	545	645
3	Machinery, fabricated metal products	3675	4545	38790	17640	5595	8415	300
4	Construction, cement, glass	6400	10950	7350	34200	23100	27050	50
5	Fuel, electric utilities	12	38	7	22	342	22	35
6	Transport, services, undistributed	124	100	75	165	90	183	609
7	Households	0	0	0	0	0	0	0
	Biproportionality Method	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	Agriculture, food, textiles	31	3	2	6	1	7	22
2	Chemicals, plastics, rubber, metals	5	57	29	13	4	3	3
3	Machinery, fabricated metal products	3274	4050	34573	15715	4983	7497	6322
4	Construction, cement, glass	4333	7414	4794	23156	15638	18316	846
5	Fuel, electric utilities	0	0	0	0	0	0	0
6	Transport, services, undistributed	1327	1070	805	1770	960	1957	6389
7	Households	0	0	0	0	0	0	0

TABLE C.3

	Brődy's Method	1	2	3	4	5	6	7
1	Agriculture, food, textiles	4014	319	168	662	149	619	2778
2	Chemicals, plastics, rubber, metals	735	7560	3550	1790	478	565	443
3	Machinery, fabricated metal products	3480	7395	37815	19500	7305	8640	330
4	Construction, cement, glass	5600	12000	7050	45650	27300	20150	300
5	Fuel, electric utilities	15	50	10	25	331	35	57
6	Transport, services, undistributed	124	148	111	208	107	214	612
7	Households	0	0	0	0	0	0	0

	Biproportionality Method	1	2	3	4	5	6	7
1	Agriculture, food, textiles	21	2	1	5	1	5	22
2	Chemicals, plastics, rubber, metals	2	19	9	4	1	1	2
3	Machinery, fabricated metal products	1821	3869	19789	10200	3821	4520	4191
4	Construction, cement, glass	4668	10004	5878	38057	25759	16797	42
5	Fuel, electric utilities	0	0	0	0	0	0	0
6	Transport, services, undistributed	746	887	667	1249	642	1285	3877
7	Households	0	0	0	0	0	0	0

where λ_0 is the uniform rate of expansion of the system. Solving Equation (C.1) for λ_0 and the corresponding outputs, x_i, yields the maximum attainable growth rate and the optimal proportions of the von Neumann trajectory.

Rather than follow Brődy's iterative procedure,[6] we observe that Equations (C.1) form an eigensystem for any λ_0 , and thus represent a particular solution of the general system

$$\mu x_i = \sum_{j=1}^{n} a_{ij}x_j + \lambda \sum_{j=1}^{n} b_{ij}x_j \qquad (i = 1,\ldots,n) \tag{C.2a}$$

or, in obvious matrix notation

$$\mu x = (A + \lambda B)x \tag{C.2b}$$

in which $\mu = 1$. For any given value of λ, we can calculate the values of μ and x using a simple power method.[7] This technique uses an iterative improvement approach to adjust the value of λ until $\mu = 1$. The exact relationship takes the following form

$$\lambda^{m+1} = \lambda^m + k(1 - \mu^m) \tag{C.3}$$

where k is a chosen constant and m is the number of the iteration. Using this approach, it is relatively simple to find the turnpike solutions for extended reproduction.

Table C.4 compares the computed turnpike solutions with the actual outputs produced in 1947 and 1958. Brődy's solutions are included for comparison. Table C.5 presents some reliability estimates for each method. To the

TABLE C.4

1947		Actual output	Brődy's method	Biproportionality method
1	Agriculture, food, textiles	139.3	135.4	139.4
2	Chemicals, plastics, rubber, metals	53.0	61.0	53.2
3	Machinery, fabricated metal products	78.1	68.4	80.4
4	Construction, cement, glass	49.4	61.2	45.9
5	Fuel, electric utilities	28.2	26.8	28.2
6	Transport, services, undistributed	210.1	206.4	210.7
7	Households	210.4	209.2	210.5
1958		**Actual output**	**Brődy's method**	**Biproportionality method**
1	Agriculture, food, textiles	192.9	199.2	193.1
2	Chemicals, plastics, rubber, metals	62.0	66.0	62.3
3	Machinery, fabricated metal products	104.5	105.1	108.2
4	Construction, cement, glass	81.0	84.1	76.4
5	Fuel, electric utilities	55.3	56.2	55.4
6	Transport, services, undistributed	335.5	319.8	336.7
7	Households	293.0	294.9	298.2

TABLE C.5*

Year	Brődy's Method	Biproportionality Method
1947	0.0988	0.0059
1958	0.0092	0.0048

* The reliability estimate, R, has been computed in the following manner:

$$R = \sum_{i=1}^{7} \left(\frac{y_i - x_i}{x_i}\right)^2$$

where x_i is the actual output in sector i and y_i is the computed output in sector i.

TABLE C.6

Year	Actual average	Brődy's Method	Biproportionality Method
1947	3.53	3.92	3.66
1958		3.78	3.97

limited extent that deviations from the actual output can be interpreted as an approximate measure of empirical reliability, the Biproportionality Method appears to produce more reliable solutions than Brődy's approach.

Table C.6 compares the actual and computed growth rates. Gross domestic output (at 1958 prices) grew at an annual average of 3.53 per cent between 1947 and 1958. Brődy's computations produced a reduction in the turnpike growth rate during the same period. In contrast, the Biproportionality Method yielded an increase during this period. Subsequent events would lend some support to this upward trend.

Part of the explanation for these contradictions may be interpreted in terms of turnover times. Table C.7 reveals that the average turnover times in the two major capital-producing sectors (machinery and fabricated metal products; construction, cement and glass) altered significantly between 1947 and 1958. Brődy's solutions were based on constant turnover times, and his poor reliability in 1947 can be attributed to an overestimate of the average turnover time in the latter sector.

Although the results achieved using the Biproportionality Method and assumption (4.33) may appear quite promising, they should be treated with extreme caution. Much of their apparent success is attributable to the very high degree of sectoral aggregation adopted in Brődy's example. Part of the explanation may also be traced to the rather mature structure of the American economy.[8]

TABLE C.7

Sector		Bródy's Life Spans	Biproportionality	
			1947	1958
1	Agriculture, food, textiles	1.0	0.01	0.01
2	Chemicals, plastics, rubber, metals	2.5	0.02	0.01
3	Machinery, fabricated metal products	15	13.4	7.85
4	Construction, cement, glass	50	33.9	41.7
5	Fuel, electric utilities	0.1	0	0
6	Transport, services, undistributed	0.1	1.07	0.60
7	Households	0	0	0

TABLE C.8

Sector j	1	2	3	4	5	6	7
$\gamma^i_{j\ell}$	1	1.22	0.67	0.98	3.66	3.03	0

The implausible nature of Bródy's assumption about turnover times, for highly disaggregated capital matrices is amply demonstrated by the case of electric motors. Consider first the inputs of electric motors into the textile machine industry, where they are used both as an input to be consumed in the production process, and as a capital good. Consider second their input into the textile industry, where the motors are used exclusively as capital goods. At this level of sectoral detail, the assumption of uniform turnover times is clearly unrealistic.

A simple alternative to the Bródy assumption, and one which has considerable practical appeal, is that individual elements in each row of the capital matrix are proportional to the total capital output ratios in the corresponding sectors.[9] The additional data required to implement this method are the k_j terms which enable the $\gamma^i_{j\ell}$ coefficients to be estimated. The relative values given in Table C.8 have been computed from a *capital requirements* table for 1963, and are assumed for each row i.

The matrices of capital coefficients resulting from this simple assumption are reproduced in Table C.9. In this case, the iterative scheme outlined in the previous section is no longer needed to compute the turnpike growth rate. A direct solution is possible, namely

$$\lambda_0 = \frac{\sum_{j=1}^{n} k_j \lambda_j x_j}{\sum_{j=1}^{n} k_j x_j} \quad . \tag{C.4}$$

TABLE C.9

1947		1	2	3	4	5	6	7
1	Agriculture, food, textiles	8	10	6	8	31	25	0
2	Chemicals, plastics, rubber, metals	5	6	3	5	18	15	0
3	Machinery, fabricated metal products	5658	6903	3791	5545	20708	17144	0
4	Construction, cement, glass	6694	8167	4485	6560	24501	20284	0
5	Fuel, electric utilities	0	0	0	0	0	0	0
6	Transport, services, undistributed	1695	2068	1136	1662	6205	5137	0
7	Household	0	0	0	0	0	0	0
1958		**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	Agriculture, food, textiles	7	8	5	7	25	21	0
2*	Chemicals, plastics, rubber, metals	2	2	1	2	6	5	0
3	Machinery, fabricated metal products	3232	3943	2166	3167	11830	9793	0
4	Construction, cement, glass	7481	9127	5012	7331	27380	22667	0
5	Fuel, electric utilities	0	0	0	0	0	0	0
6	Transport, services, undistributed	971	1185	651	952	3554	2942	0
7	Households	0	0	0	0	0	0	0

The computed growth rates for 1947 and 1958 are 4.07 and 4.19 per cent respectively. These rates are higher than those resulting earlier, and are closer to Almon's forecast based on an open version of Leontief's dynamic model.[10] To this extent, the assumptions underlying the approach adopted herein may possibly be regarded as superior to Bródy's simplification.

Although the two approaches described above have considerable practical appeal because of their ease of implementation, they both suffer from a vastly oversimplified view of the structure of capital flows. The refinements suggested in Section 4.3.5 represent an initial attempt to develop a more realistic and yet flexible framework which is capable of making full use of all available information pertaining to the capital matrix. Although these later formulations based on the norms of information theory have yet to be tested, it is believed that they hold considerable promise.

FOOTNOTES FOR APPENDIX C

1 See Equation (4.32).

2 See Brōdy (1966, 1970).

3 These tables of flow coefficients are based on preliminary estimates contained in the files of the Harvard Economic Research Project.

4 Namely, Expression (4.33).

5 Under these circumstances, the computed matrices should be interpreted as containing coefficients of *residual demands*. Brōdy's exclusion of foreign trade from his flow matrix results in the automatic transfer of these trade effects to the capital matrix. The resulting coefficients can no longer be regarded simply as measures of capital/output ratios.

6 See Brōdy (1970, p 90).

7 The power method is an algorithm for finding the eigenvalue largest in absolute value, together with the corresponding eigenvector, of a given matrix. For a description of the actual procedure, see Åberg and Persson (1980). A formal proof is given in Wilkinson (1965).

8 It is possible to isolate the productive contribution of simple reproduction alone by setting $\lambda = 0$ in Equation (C.2). By solving the corresponding eigensystem of equations, namely (in matrix form)

$$\mu x = Ax$$

we obtain a maximum eigenvalue of $\mu = 0.896$ for the 1947 economy, and $\mu = 0.907$ for 1958. Thus the scope for expanding reproduction, when viewed as the amount of additional capacity required to increase μ to one, is seen to be very limited in these two economies. Put another way, the relatively low growth rates of about 4 % reflect a mature system in which the potential for higher rates of future expansion is severely constrained. Such an economy is hardly ideal for assessing the reliability of estimated capital coefficients, since the changes in output levels resulting from capacity expansion will not be greatly sensitive to the accuracy of these coefficients.

9 This does not imply that the total capital/output ratios must remain constant. On the contrary, it freely permits their variation over time, such that an aggregated form of substitution can be accommodated. In practical terms, we can often improve on

this approach by isolating the contributions from the two major capital-producing sectors, namely building and construction, and machinery and equipment. If available data permit the partitioning of the capital/output ratios, Equation (4.34) can easily be reformulated to take advantage of this additional information.

10 Almon (1966) forecast an overall growth rate of just over five per cent. His model computed a more sophisticated turnpike which incorporated nonlinearities and changing technology.

Appendix D

MINIMIZING INFORMATION LOSSES IN SIMPLE AGGREGATION:

Two Test Problems

PROBLEM 1

This problem has been adapted from the existing literature on aggregation in input-output analysis,[1] and involves the reduction of the following 4 by 4 matrix to a 3 by 3 array:

$$Q = \begin{bmatrix} .05 & .05 & 0 & .05 \\ .05 & .1 & .05 & 0 \\ .1 & .05 & .05 & 0 \\ .2 & .1 & .1 & .05 \end{bmatrix} .$$

By evaluating Equation (4.60) for all six possible combinations,[2] it is found that combining sectors 1 and 3 causes the lowest information loss. In this case, the resulting matrices are

$$\bar{P} = \begin{bmatrix} .2 & .1 & .05 \\ .1 & .1 & 0 \\ .3 & .1 & .05 \end{bmatrix} \qquad D = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$P = \begin{bmatrix} .05 & .05 & .05 & .025 \\ .05 & .1 & .05 & 0 \\ .05 & .05 & .05 & .025 \\ .15 & .1 & .15 & .1 \end{bmatrix} \qquad S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

PROBLEM 2

In this instance, we have chosen an example in which the first and second rows and columns are almost identical, and should therefore be the obvious candidates for aggregation. With the original Q matrix as

$$Q = \begin{bmatrix} .04 & .05 & .05 & .12 \\ .04 & .04 & .03 & .12 \\ .08 & .08 & .04 & 0 \\ .05 & .04 & .12 & .1 \end{bmatrix}$$

the reader may wish to confirm that the combination of sectors 1 and 2 is overwhelmingly preferred.

FOOTNOTES

1 See Theil (1967, p 333).

2 The number of ways in which any two sectors of our 4 by 4 matrix may be combined to produce a 3 by 3 array, is given by the combinatorial formula

$$^{4}C_{2} = \frac{4\,!}{(4-2)!\;2!}$$
$$= 6$$

Appendix E

COMPUTING INTERREGIONAL AND INTERSECTORAL FLOWS:
The INTEREG Package

The INTEREG package transforms the mathematical formulations described in Chapters 5 and 6 into the FORTRAN programming language, thereby facilitating computer estimation of the interregional flows. It is designed to handle either contingency table, maximum entropy or minimum information gain formulations. Explicit solutions are obtained using Eriksson's iterative algorithm, which transforms the original formulation into a system of nonlinear equations that are subsequently solved by Newton's method.[1] Each linear step of Newton's method is solved by the method of conjugate gradients. Eriksson's algorithm is used because it is both versatile and computationally efficient.

The INTEREG package has been tested on a simple system of two regions representing the state of Victoria in Australia. Recent nonsurvey data, derived primarily from the national input-output tables using the RAS and GRIT procedures,[2] provide an initial database. The flow tables originating from these two sources have been reconciled, and then aggregated into two regions (Melbourne and the rest of Victoria) and three sectors (primary, secondary and tertiary industries). The resulting data sets consist of the Victorian input-output table (Table E.1), together with flow tables for Melbourne (Table E.2) and the rest of Victoria (Table E.3).

Notations are identical to those defined in Chapter 5. However, the intersectoral flows for Melbourne (x_{ij}^{*1}) include imports from the rest of Victoria (m_j^{21}) and similarly, the intersectoral flows for the rest of Victoria (x_{ij}^{*2})

TABLE E.1

x_{ij}	1	2	3	x_{i*}	y_i	x_i
1	64	882	12	958	1096	2054
2	196	6660	996	7852	11823	19675
3	21	1825	287	2133	7259	9392
x_{*j}	281	9367	1295	10943	20178	31121
v_j	1773	10308	8097	20178		
x_j	2054	19675	9392	31121		

TABLE E.2

x_{ij}^{*1}	1	2	3	x_{i*}^{*1}	y_i^{*1}	x_i^{1*}
1	6	122	2	130	164	294
2	28	4229	761	5018	10963	15981
3	6	1706	356	1968	5458	7426
x_{*j}^{*1}	40	6057	1019	7116	16585	23701
v_j^1	254	9924	6407	16585		
x_j^{1*}	294	15981	7426	23701		

TABLE E.3

x_{ij}^{*2}	1	2	3	x_{i*}^{*2}	y_i^{*2}	x_i^{2*}
1	58	760	10	828	932	1760
2	168	2431	235	2834	860	3694
3	15	119	31	165	1801	1966
x_{*j}^{*2}	241	3310	276	3827	3593	7420
v_j^2	1519	384	1690	3593		
x_j^{2*}	1760	3694	1966	7420		

include imports from Melbourne (m_j^{12}). Imports from interstate and abroad are included in primary inputs (v_j^1 or v_j^2), and exports to interstate and abroad are included among final demands (y_i^{*1} and y_i^{*2}).

Preliminary estimates of $\{x_{ij}^{12}\}$, $\{x_{ij}^{21}\}$, $\{y_i^{12}\}$ and $\{y_i^{21}\}$ have focussed on the four basic assumptions concerning intraregional demands.[3] Results of the corresponding contingency table computations are summarized in Tables E.4 to E.7.[4] These interregional tables specify the flows between Melbourne (region 1) and the rest of Victoria (region 2). They provide ample evidence of the allowance for *cross-hauling*, which is an inherent characteristic of the chosen methodology.[5]

Further insight into the intraregional estimation problem may be gained by examining the estimates of $\{x_{ij}^{11}\}$, $\{x_{ij}^{22}\}$, $\{y_i^{1*}\}$ and $\{y_i^{2*}\}$. For example, the resulting dog-leg input-output tables for Melbourne under each of the four assumptions about intraregional demands are depicted in Tables E.8 to E.11. If we now refer back to the original flow table for Melbourne (Table E.2), some interesting differences arise. The original table, based on the RAS and GRIT procedures, fails to distinguish between the intra- and interregional inputs to the Melbourne economy. In contrast, the INTEREG philosophy ensures that the technical requirements of local industries are separated from the interregional trade patterns of the economy. As a result of this distinction, the coefficients underlying the intraregional transactions segments of Tables E.8 to E.11 can now be redefined as intraregional requirements coefficients; they denote the requirements for local products from sector i per unit of local output in sector j.

TABLE E.4

x_{ij}^{12}	1	2	3	y_i^{12}
1	8.	45.	Ø.	78.
2	137.	1912.	172.	48Ø2.
3	14.	51Ø.	48.	287Ø.
x_{ij}^{21}	1	2	3	y_i^{21}
1	8.	489.	8.	47Ø.
2	5.	8Ø9.	147.	111Ø.
3	1.	247.	47.	76Ø.

TABLE E.5

x_{ij}^{12}	1	2	3	y_i^{12}
1	8.	45.	Ø.	133.
2	137.	1912.	172.	699.
3	14.	51Ø.	48.	1424.
x_{ij}^{21}	1	2	3	y_i^{21}
1	8.	489.	8.	141.
2	5.	8Ø9.	147.	2Ø58.
3	1.	247.	47.	1143.

TABLE E.6

x_{ij}^{12}	1	2	3	y_i^{12}
1	9.	116.	1.	126.
2	130.	1508.	157.	1207.
3	14.	511.	58.	972.
x_{ij}^{21}	1	2	3	y_i^{21}
1	0.	64.	1.	186.
2	7.	902.	151.	1941.
3	1.	247.	45.	1262.

TABLE E.7

x_{ij}^{12}	1	2	3	y_i^{12}
1	9.	108.	1.	133.
2	137.	1960.	205.	699.
3	8.	112.	11.	1424.
x_{ij}^{21}	1	2	3	y_i^{21}
1	1.	108.	2.	141.
2	5.	797.	140.	2058.
3	2.	352.	57.	1143.

TABLE E.8

x_{ij}^{11}	1	2	3	e_i^{12}	y_i^{1*}	x_i^{1*}
1	1.	82.	1.	54	156	294
2	23.	3498.	637.	2219	9604	15981
3	2.	933.	179.	572	5740	7426
m_j^{21}	14	1544	202			
v_j^1	254	9924	6407			
x_j^{1*}	294	15981	7426			

TABLE E.9

x_{ij}^{11}	1	2	3	e_i^{12}	y_i^{1*}	x_i^{1*}
1	1.	82.	1.	54	156	294
2	23.	3498.	637.	2219	9604	15981
3	2.	933.	179.	572	5740	7426
m_j^{21}	14	1544	202			
v_j^1	254	9924	6407			
x_j^{1*}	294	15981	7426			

TABLE E.10

x_{ij}^{11}	1	2	3	e_i^{12}	y_i^{1*}	x_i^{1*}
1	Ø.	11.	Ø.	127	156	294
2	3Ø.	39Ø2.	652.	1793	9604	15981
3	2.	932.	169.	583	5740	7426
m_j^{21}	8	1212	198			
v_j^{1}	254	9924	6407			
x_j^{1*}	294	15981	7426			

TABLE E.11

x_{ij}^{11}	1	2	3	e_i^{12}	y_i^{1*}	x_i^{1*}
1	Ø.	18.	Ø.	120	156	294
2	23.	345Ø.	6Ø4.	2300	9604	15981
3	9.	1331.	216.	130	5740	7426
m_j^{21}	8	1258	199			
v_j^{1}	254	9924	6407			
x_j^{1*}	294	15981	7426			

Following these early computations, more recent estimates have concentrated on formulations which include constraints containing weighted sums of coefficients. Cost coefficients of the type mentioned in (5.40) have been examined. For example, Table E.12 contains a set of *hypothetical* cost coefficients which have been derived from the transport sector of the original flow tables, and then reconciled to satisfy various flow sums. These coefficients are based on average relative trip distances of 1 within Melbourne, 2 between Melbourne and the rest of Victoria, and 4 within the rest of Victoria.

Estimates of the interregional tables containing $\{x_{ij}^{12}\}$, $\{y_i^{12}\}$, $\{x_{ij}^{21}\}$ and $\{y_i^{21}\}$ have subsequently been repeated for these cost-sensitive formulations. A typical entropy-maximizing solution is given in Table E.13 for case (i), which assumes no *a priori* knowledge of any intraregional demands. The resulting dog-leg input-output table for Melbourne under these revised conditions is depicted in Table E.14. A comparison of Tables E.8 and E.14 confirms our expectations concerning the influence of the cost constraints: intraregional flows have now been maximized, at the expense of exports to, and imports from, the rest of Victoria.

Current research is focussed on the inclusion of nodal capacity constraints like (5.41) in our entropy-maximizing formulations. Owing to the versatility of Eriksson's algorithm, inequality constraints of this type can now be handled successfully. Our principal attention is therefore directed towards suitable methods for measuring the physical handling capabilities of each node or region.

TABLE E.12

c_i^{rs}		s = 1	s = 2
r = 1	i = 1	0.0073	0.0146
	i = 2	0.0115	0.0230
	i = 3	0.1165	0.233
r = 2	i = 1	0.0146	0.0292
	i = 2	0.0230	0.0460
	i = 3	0.233	0.466

C_i	
i = 1	14
i = 2	181
1 = 3	498
$\sum_i C_i$	693

TABLE E.13

x_{ij}^{12}	1	2	3	y_i^{12}
1	10.	5.	0.	40.
2	144.	2081.	175.	7366.
3	15.	398.	32.	3968.
x_{ij}^{21}	1	2	3	y_i^{12}
1	29.	755.	10.	845.
2	0.	302.	59.	160.
3	0	164.	31.	244.

TABLE E.14

x_{ij}^{11}	1	2	3	e_i^{12}	y_i^{1*}	x_i^{1*}
1	4.	110.	2.	22	156	294
2	5.	3601.	705.	2066	9604	15981
3	1.	1125.	212.	348	5740	7426
m_j^{21}	30	1221	100			
v_j^1	254	9924	6407			
x_j^{1*}	294	15981	7426			

FOOTNOTES FOR APPENDIX E

1 See Eriksson (1978) for the original details of his algorithm, and Eriksson (1980) for a later version developed to handle formulations which contain inequality constraints.

2 The RAS approach is described in Cruickshank (1979), whereas the GRIT technique is outlined by Jensen et al. (1979).

3 See Section 5.4.2.

4 These results correspond to the formulations given in Section 5.4.3.

5 For heterogeneous commodity classes like those in our present example, Nijkamp (1975) has suggested that entropy models, which lead to a considerable dispersion of flows within the system, may give more adequate results for cross-hauling amongst interregional shipments.

REFERENCES

Aczél, J. and Z. Daróczy (1975) *On Measures of Information and their Characterizations*, New York: Academic Press.

Aidenoff, A. (1970) "Input-Output Data in the United Nations System of National Accounts", in A.P. Carter & A. Bródy (eds) *Applications of Input-Output Analysis*, Amsterdam: North-Holland.

Allen, R.I.G. and J.R.C Lecomber (1975) " Some Tests on a Generalized Version of RAS" in R.I.G. Allen and W.F. Gossling, (eds), Estimating and *Projecting Input-Output Coefficients*, London: Input-Output Publ Co.

Almon, C. (1966) *The American Economy to 1975*, New York: Harper and Row.

Almon, C. et al. (1975) *1984 Interindustry Forecast of the American Economy*, Lexington: Lexington Books.

Andersson, Å.E. (1975) "A Closed Nonlinear Growth Model for International and Interregional Trade and Location", *Regional Science and Urban Economics* 5, pp 427-444.

Andersson, Å.E. and D.F. Batten (1980) "An Interdependent Framework for Integrated Sectoral and Regional Development", Paper presented at IFAC/IFORS Conference on Dynamic Modelling and Control of National Economies, Warsaw, June.

Andersson, Å.E. and A. Karlqvist (1979) " An Equilibrium Model for Integrated Regional Development", in D.A Henscher and P.R. Stopher (eds), *Behavioural Travel Modelling*, London: Croom Helm.

Andersson, Å.E. and H. Persson (1979) "Integration of Transportation and Location Analysis - A General Equilibrium Approach", *Papers, Regional Science Association*, vol 42, pp 39-55.

Ara, K. (1959) "The Aggregation Problem in Input-Output Analysis", *Econometrica*, vol 27, pp 257-262.

Bacharach, M. (1970) *Biproportional Matrices and Input-Output Change*, London: Cambridge University Press.

Balderston, J.B. and T.M. Whitin (1954) "Aggregation in the Input-Output Model", in O. Morgenstern (ed), *Economic Activity Analysis*, New York: John Wiley & Sons.

Batten, D.F. (1979) "The Estimation of Capital Coefficients in Dynamic Input-Output Models", Memorandum No 76, Department of Economics, University of Gothenburg.

Batten, D.F. (1981) "The Interregional Linkages between National and Regional Input-Output Models", *International Regional Science Review*, in press.

Batten, D.F. (1981) "A Note on the Estimation of Capital Structure in a Multisectoral Economy", *Economics of Planning*, in press.

Batten, D.F. and P.F. Lesse (1979) "Information-Theoretical Methods in Regional Science", *Proceedings, 7th Australian Conference on Urban and Regional Planning Information Systems*: 4.7-4.12.

Batten, D.F. and C.J. Tremelling (1980) "The Estimation of Interregional Input-Output Tables for Victoria", *Papers, Regional Science Association (Australia - New Zealand Section)*, vol 5.

Batty, M. (1974 a) "Spatial Entropy", *Geographical Analysis*, vol 6, pp 1-31.

Batty, M. (1974 b) "Urban Density and Entropy Functions", *Journal of Cybernetics*, vol 4, pp 41-55.

Batty, M. (1976) "Entropy in Spatial Aggregation", *Geographical Analysis*, vol 8, pp 1-21.

Batty, M. (1978 a) "Speculations on an Information Theoretic Approach to Spatial Representation", in Ian Masser & Peter J.B. Brown (eds), *Spatial Representation and Spatial Interaction*, pp 115-147. Leiden/ Boston: Martinus Nijhoff.

Batty, M. and R. Sammons (1978 b) "On Searching for the Most Informative Spatial Pattern", *Environment and Planning* A, vol 10, pp747-779.

Beck, A.H.W. (1976) *Statistical Mechanics, Fluctuations and Noise*, London: Edward Arnold.

Bennet, R.J. and K.C. Tan (1979) "Stochastic Control of Regional Economies", in C.P.A. Bartels & R.H Ketellapper (eds), *Exploratory and Explanatory Statistical Analysis of Spatial Data*, Boston: Martinus Nijhoff.

Berry, B.J.L. (1964) "Cities as Systems within Systems of Cities", *Papers, Regional Science Association*, vol 13, pp 147-164.

Berry, B.J.L. and P. Schwind (1969) "Information and Entropy in Migrant Flows", *Geographical Analysis*, vol 1, pp 5-14.

Bigsten, A. (1978) "Regional Inequality and Development", Doctoral Dissertation, Department of Economics, University of Gothenburg.

Bigsten, A. (1981) "A Note on the Estimation of Interregional Input-Output Coefficients", *Regional Science and Urban Economics*, vol 11, pp 149-153.

Bjerkholt, O. and S. Longva (1980) "MODIS IV: A Model for Economic Analysis and National Planning", Economic Study No 43, Central Bureau of Statistics of Norway, Oslo.

Blair, P. (1978) *Multi-objective Regional Energy Planning*. Leiden: Martinus Nijhoff.

Boltzmann, L. (1872) "Weitere Studien über das Warmgleichgewicht unter Gasmolekulen", *K. Acad. (Wein) Sitzb., II Abt.*, vol 66, pp 275.

Bridgman, P.W. (1941) *The Nature of Thermodynamics*, Cambridge, Massachusetts: Harvard University Press.

Brillouin, L. (1956) *Science and Information Theory*, New York: Academic Press.

Brődy, A. (1966) "A Simplified Growth Model", *Quarterly Journal of Economics*, vol 80, pp 137-146.

Brődy, A. (1970) *Proportions, Prices and Planning*, Amsterdam: North-Holland.

Brotchie, J.F., J.W. Dickey and R. Sharpe (1980) *TOPAZ: General Planning Technique and its Applications at the Regional, Urban and Facility Planning Levels*, Berlin: Springer-Verlag.

Brotchie, J.F., A.R. Toakley and R. Sharpe (1971) "A Model for National Development", *Management Science*, vol 18, pp 14-18.

Bussière, R. and F. Snickars (1970) "Derivation of the Negative Exponential Model by an Entropy Maximizing Method", *Environment and Planning A*, vol 2, pp 295-302.

Carnot, S. (1824) *Réflections sur la Puissance Matrice du Feu et sur les Machines Propres à Developer cette Puissance*, Paris.

Carter, A.P. (1957) "Capital Coefficients as Economic Parameters: the Problems of Instability", in National Bureau of Economic Research, *Problems of Capital Formation, vol 19*, Princeton: Princeton University Press.

Cesario, F.J. (1975) "A Primer on Entropy Modelling", *Journal of the American Institute of Planners*, vol 41, pp 40-48.

Chapman, G.P. (1970) "The Application of Information Theory to the Analysis of Population Distributions in Space", *Economic Geography*, (Supplement) vol 46, pp 317-331.

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

Chenery, H. (1953) "Regional Analysis", in H. Chenery and P. Clarc (eds), *The Structure and Growth of the Italian Economy*, Rome: U.S. Mutual Security Agency.

Clausius, R. (1865) "Über Verschiedene für die Anwendung Begüme Formen der Hauptgleichungen der mechanischen Wärmetheorie", *Ann. der Phys.*, vol 125, p 353.

Cochrane, J.L. and M. Zeleny (1973) *Multiple Criteria Decision-making*, Columbia: University of South Carolina Press.

Cohon, J.L. (1978) *Multiobjective Programming and Planning*, New York: Academic Press.

Conway, R.S. (1980) "Changes in Regional Input-Output Coefficients and Regional Forecasting", *Regional Science and Urban Economics*, vol 10, pp 153-171.

Courbis, R. and D. Vallet (1976) "An Interindustry Interregional Table of the French Economy", in K.R. Polenske and J.V. Skolka (eds), *Advances in Input-Output Analysis*, Cambridge, Massachusetts: Ballinger.

Cripps, E.L., S.M. Macgill and A.G. Wilson (1974) "Energy and Material Flows in the Urban Space Economy", *Transportation Research*, vol 8, pp 293-305.

Cruickshank, A. (1979) "An Urban Input-Output Model for Melbourne", Masters Thesis, Department of Town and Regional Planning, University of Melbourne.

Curry, L. (1963) "Explorations in Settlement Theory: The Random Spatial Economy, Part I", *Annals, Association of American Geographers*, vol 54, pp 138-146.

Curry, L. (1972) "Spatial Entropy", in W.P. Adams and F.M. Helleiner (eds), *International Geography 2*, Toronto: University of Toronto Press.

Czamanski, S. and E.E. Malizia (1969) "Applicability and Limitations in the Use of National Input-Output Tables for Regional Studies", *Papers, Regional Science Association*, vol 23, pp 65-77.

Dacey, M.P. and A. Norcliffe (1976) "New Entropy Models in the Social Sciences: 1. Elementary Residential Location Models", *Environment and Planning* A, vol 8, pp 299-310.

Dacey, M.P. and A. Norcliffe (1977) "A Flexible Doubly-Constrained Trip Distribution Model", *Transportation Research*, vol 11, pp 203-204.

van Delft, A. and P. Nijkamp (1977) *Multi-criteria Analysis and Regional Decision-making*, The Hague: Martinus Nijhoff.

Deming, W.E. and F.F. Stephan (1949) "On a Least Squares Adjustment of a Sampled Frequency Table when Marginal Totals are Known", *Annals of Mathematical Statistics*, vol 11, pp 427-444.

Domar, E. (1946) "Capital Expansion, Rate of Growth, and Employment", *Econometrica*, vol 14, pp 137-147.

Domar, E. (1957) *Essays in the Theory of Growth*, New York: Oxford University Press.

Dominion Bureau of Statistics (1969) *The Input-Output Structure of the Canadian Economy, 1961*, Ottawa: The Queens Printer for Canada.

Dowson, D.C. and A. Wragg (1973) "Maximum Entropy Distributions having Prescribed First and Second Moments", *IEEE Transactions on Information Theory*, IT-9, pp 689-693.

Emerson, M.J. (1976) "Interregional Trade Effects in Static and Dynamic Input-Output Models", in K.R. Polenske and J.V. Skolka (eds), *Advances in Input-Output Analysis*, Cambridge, Massachusetts: Ballinger.

Eriksson, J. (1978) "Solution of Large Sparse Maximum Entropy Problems with Linear Equality Constraints", Research Report 02, Department of Mathematics, University of Linköping.

Eriksson, J. (1980) "On Solving Linearly Constrained Maximum Entropy Problems", Research Report 14, Department of Mathematics, University of Linköping.

Eskelinen, H. and M. Suorsa (1980) "A Note on Estimating Interindustry Flows", *Journal of Regional Science*, vol 20, pp 261-266.

Fast, J.D. (1962) *Entropy*, Eindhoven: Phillips Technical Library.

Fay, R.E. and L.A. Goodman (1975) *ECTA Program: Description for Users*, University of Chicago: Department of Statistics.

Fei, J.C-H. (1956) "A Fundamental Theorem for the Aggregation Problem of Input-Output Analysis", *Econometrica*, vol 24, pp 400-412.

Fienberg, S.E. (1970) "An Iterative Procedure for Estimation in Contingency Tables", *Annals of Mathematical Statistics*, vol 41, pp 907-917.

Fienberg, S.E. (1977) *The Analysis of Cross-Classified Categorical Data*, Cambridge, Massachusetts: MIT Press.

Finkelstein, M.O. and R.M. Friedberg (1967) "The Application of an Entropy Theory of Concentration to the Clayton Act", *The Yale Law Review*, vol 76, pp 677-717.

Fisher , W.D. (1958) "Criteria for Aggregation in Input-Output Analysis", *Review of Economics and Statistics*, vol 40, pp 250-260.

Fisk, C. and G.R. Brown (1975 a) "A Note on the Entropy Formulation of Distribution Models", *Operational Research Quarterly*, vol 26, pp 755-758.

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

Fujita, M. (1978) *Spatial Development Planning*, Amsterdam, North-Holland.

Funck, R. and G. Rembold (1975) "A Multiregion, Multisector Forecasting Model for the Federal Republic of Germany", *Papers, Regional Science Association*, vol 34, pp 69-82.

Garrison, C.B. and A.S. Paulson (1973) "An Entropy Measure of the Geographic Concentration of Economic Activity", *Economic Geography*, vol 49, pp 319-324.

Georgescu-Roegen, N. (1971) *The Entropy Law and the Economic Process*, Cambridge, Massachusetts: Harvard University Press.

Gerking, S.D. (1976) *Estimation of Stochastic Input-Output Models*, Leiden: Martinus Nijhoff.

Gibbs, J.W. (1902) *Elementary Principles in Statistical Mechanics*, New York.

Gigantes, T. (1970) "The Representation of Technology in Input-Output Systems", in A.P. Carter and A. Bródy (eds), *Contributions to Input-Output Analysis*, Amsterdam: North-Holland.

Gigantes, T. and T. Matuszewski (1968) "Rectangular Input-Output Systems: Taxonomy and Analysis", Paper presented at the 4th International Conference on Input-Output Techniques, Geneva.

Good, I.J. (1963) "Maximum Entropy for Hypothesis Formulation, especially for Multidimensional Contingency Tables", *Annals of Mathematical Statistics*, vol 34, pp 911-934.

Grosse, R.M. (1953) "The Structure of Capital", in W.W. Leontief et al.(eds), *Studies in the Structure of the American Economy*, New York, Oxford University Press.

Gurney, R.W. (1949) *Introduction to Statistical Mechanics*, New York.

Habermann, S.J. (1973) *CTAB: Analysis of Multidimensional Contingency Tables by Loglinear Models: User's Guide*, Chicago: International Education Services.

Haberman, S.J. (1978) *Analysis of Qualitative Data Volume 1: Introductory Topics*, New York: Academic Press.

Hafkamp, W. and P. Nijkamp (1978) "Environmental Protection and Spatial Welfare Patterns", in H. Folmer and J. Oosterhaven (eds), *Spatial Inequalities and Regional Development*, The Hague: Martinus Nijhoff.

Hafkamp, W. and P. Nijkamp (1980) "An Integrated Interregional Model for Pollution Control", in T.R. Lakshmanan and P. Nijkamp (eds), *Economic-environmental-energy Interactions*, Boston: Martinus Nijhoff.

Halder, A. (1970) "An Alternative Approach to Trip Distribution", *Transportation Research*, vol 4, pp 63-69.

Halmos, P.R. (1956) *Lectures on Ergodic Theory*, Tokyo.

Hansen, W.L. and C.M. Tiebout (1963) "An Intersectoral Flows Analysis of the Californian Economy", *Review of Economics and Statistics*, vol 45, pp 409-418.

Harris, C.C. (1970) "A Multiregional, Multi-industry Forecasting Model", *Papers, Regional Science Association,* vol 25, pp 169-180.

Harrod, R. (1939) "An Essay in Dynamic Theory", *Economic Journal,* vol 49, pp 14-33.

Hartley, R.V.L. (1928) "Transmission of Information", *Bell System Technical Journal,* vol 7, pp 535-563.

Hartwick, J.M. (1971) "Notes on the Isard and Chenery-Moses Interregional Input-Output Models", *Journal of Regional Science,* vol 11, pp 73-86.

Hatanaka, M. (1952) "Note on Consolidation within a Leontief System", *Econometrica,* vol 20, pp 301-303.

Hawkins, D (1948) "Some Conditions of Macroeconomic Stability", *Econometrica,* vol 16 pp 309-322.

Hewings, G.J.D. (1969) "Regional Input-Output Models using National Data: West Midlands", *Annals of Regional Science,* vol 3, pp 179-191.

Hewings, G.J.D. (1971) "Regional Input-Output Models in the U.K: Some Problems and Prospects for the use of Non-survey Techniques, *Regional Studies,* vol 5, pp 11-22.

Hewings, G.J.D. and B.N. Janson (1980) "Exchanging Regional Input-Output Coefficients: A Reply and Further Comments", *Environment and Planning* A, vol 12, pp 843-854.

Hildenbrand, W. and H. Paschen (1964) "Ein axiomatisch begründetes Konzentrationsmass", *Statistical Information,* vol 3, pp 53-61.

Hobson, A. (1969) "A New Theorem of Information Theory", *Journal of Statistical Physics,* vol 1, pp 383-391.

Hobson, A. (1971) *Concepts in Statistical Mechanics,* New York, Gordon and Breach.

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

Horowitz, A. and I. Horowitz (1968) "Entropy, Markov Processes, and Competition in the Brewing Industry", *Journal of Industrial Economics,* vol 16, pp 196-211.

Ireland, C.T. and S. Kullback (1968)"Contingency Tables with Given Marginals", *Biometrika,* vol 55, pp 179-188.

Isard, W. (1951) "Interregional and Regional Input-Output Analysis: A Model of a Space Economy", *Review of Economics and Statistics,* vol 33, pp 318-328.

Isard, W. (1953) "Regional Commodity Balances and Interregional Commodity Flows", *American Economic Review,* vol 43, pp 167-180.

Isard, W. (1960) *Methods of Regional Analysis: An Introduction to Regional Science*, Cambridge, Massachusetts: MIT Press.

Isard, W. (1977) "On Hierarchical Dynamics", *London Studies in Regional Science*, vol 7, pp 125-133.

Isard, W. and T.W. Langford (1971) *Regional Input-Output Study: Recollections, Reflections, and Diverse Notes on the Philadelphia Experience*, Cambridge, Massachusetts: MIT Press.

Isard, W. and P. Liossatos (1979) *Spatial Dynamics and Optimal Space-Time Development*, Amsterdam: North-Holland.

Jaynes, E.T. (1957) "Information Theory and Statistical Mechanics", *Physical Review*, vol 106, pp620-630; vol 108, pp 171-190.

Jaynes, E.T. (1968) "Prior Probabilities", *IEEE Transactions on Systems Science and Cybernetics*, SSC-4, pp227-241,

Jefferson, T.R. and C.H. Scott (1979) "The Analysis of Entropy Models with Equality and Inequality Constraints", *Transportation Research B*, vol 13, pp 123-132.

Jensen, R.C., T. Manderville and N.D. Karunaratne (1979) *Regional Economic Planning*, London: Croom Helm.

Johansen, L. (1973) "The Rates of Growth in Input-Output Models: Some Observations along Lines Suggested by O. Lange and A. Brōdy", in E. Boettcher et al. (eds), *Yearbook of East-European Economics*, Munich-Vienna: Günter Olzog Verlag.

Johansen, L. (1978) "On the Theory of Dynamic Input-Output Models with Different Time Profiles of Capital Construction and Finite Lifetime of Capital Equipment", *Journal of Economic Theory*, vol 19, pp 513-533.

Johansson, B. and U. Stromqvist (1980) *Earnings and Employment in Swedish Industry: A Structural Analysis of Sweden's Industries, 1969 - 1977* (in Swedish), Stockholm: SIND.

Jones, L.L., T.L. Sporleder and G. Mustafa (1973) "A Source of Bias in Regional Input-Output Models Estimates from National Coefficients", *Annals of Regional Science*, vol 7, pp 67-74.

Jorgenson, D.W. (1960) "Stability of a Dynamic Input-Output System", *Review of Economic Studies*, vol 28, pp 105-116.

Kādas, S.A. and E. Klafsky (1976) "Estimation of the Parameters in the Gravity Model for Trip Distribution: a New Method and Solution Algorithm", *Regional Science and Urban Economics*, vol 6, pp 439-457.

Kaniss, P. (1978) "Evolutionary Change in Hierarchical Systems", Doctoral Dissertation, Cornell University.

Karlqvist, A. and B. Marksjö (1971) "Statistical Urban Models", *Environment and Planning*, vol 3, pp 83-98.

Karlqvist, A. et al. (1978) " A Regional Planning Model and its Application to South Eastern Australia", *Regional Science and Urban Economics*, vol 8, pp 57-86.

Kerridge, D.F. (1961) "Innaccuracy and Inference", *Journal of the Royal Statistical Society B*, vol 23, pp 184-194.

Khinchin, A.I. (1957) *Mathematical Foundations of Information Theory*, New York: Dover.

Koopmans, T.C. and M. Beckmann (1957) "Assignment Problems and the Location of Economic Activities, *Econometrica*, vol 25, pp 53-76.

Kornai, J. (1971) *Anti-Equilibrium*,Amsterdam: North-Holland.

Kossov, V. (1970) "The Theory of Aggregation in the Input-Output Models", in A.P. Carter and A. Brődy (eds), *Contributions to Input-Output Analysis*, Amsterdam: North-Holland.

Kullback, S. (1959) *Information Theory and Statistics*, New York: Wiley.

Kullback, S. and R.A. Leibler (1951) "On Information and Sufficiency", *Annals of Mathematical Statistics*, vol 22, pp 79-86.

Köstler, A. (1967) *The Ghost in the Machine*, London: Hutchinson.

Lange, O. (1957) "Some Observations on Input-Output Analysis", *Sankhya*, vol 17, pp 305-336.

Lee, R. (1974) "Entropy Models in Spatial Analysis", Discussion Paper 15, Department of Geography, University of Toronto.

Leontief, W.W. (1951) *The Structure of the American Economy, 1919 - 1939*, New York: Odford University Press.

Leontief, W.W. (1953) "Dynamic Analysis", in W.W. Leontief et al. (eds), *Studies in the Structure of the American Economy*, New York: Oxford University Press.

Leontief, W.W. (1970) "The Dynamic Inverse", in A.P. Carter and A. Brődy et al. (eds), *Contributions to Input-Output Analysis*, Amsterdam: North-Holland.

Leontief, W.W. et al. (1953) *Studies in the Structure of the American Economy*, New York: Oxford University Press.

Leontief, W.W. and A. Strout (1963) "Multiregional Input-Output Analysis", in T. Barna (ed), *Structural Interdependence and Economic Development*, London: St. Martin's Press.

Lesse, P. et al. (1978) "A New Philosophy for Regional Modelling", *Papers, Regional Science Association (Australia - New Zealand Section)*, vol 3, pp 165-178.

Lesse, P. and R. Sharpe (1980) "A Control Theory Approach to Regional Stagnation", in W. Buhr and P. Friedrich (eds), *Regional Development under Stagnation,* Baden-Baden: Nomos Verlag.

Lesuis, P., F. Muller and P. Nijkamp (1980) "An Interregional Policy Model for Energy-Economic-Environmental Interactions", *Regional Science and Urban Economics*, vol 10, pp 343-370.

Levine, R. D. and M. Tribus (eds), (1979) *The Maximum Entropy Formalism,* Cambridge, Massachusetts: MIT Press.

Lewis, G.N. (1930) "The Symmetry of Time in Physics", *Science,* June 6, p 573.

Lundqvist, L. (1980) "A Dynamic Multiregional Input-Output Model for Analyzing Regional Development, Employment and Energy Use", *Papers, Regional Science Association,* vol 46.

Macgill, S.M. (1977) "Theoretical Properties of Biproportional Matrix Adjustments", *Environment and Planning* A, vol 9, pp687-701.

Macgill, S.M. (1978) "Rectangular Input-Output Tables, Multiplier Analysis and Entropy Maximizing Principles", *Regional Science and Urban Economics*, vol 8, pp 355-370.

Macgill, S.M. and A.G. Wilson (1979) "Equivalence and Similarities between some Alternative Urban and Regional Models", *Sistemi Urbani,* vol 1, pp 9-40.

Macqueen, J. and J. Marschak (1975) "Partial Knowledge, Entropy and Estimation", *Proceedings, National Academy of Sciences,USA,* vol 72, pp 3819-3824.

Magie, W.F. (1899) *The Second Law of Thermodynamics,* New York.

Malizia, E.E. and D.L. Bond (1974) "Empirical Tests of the RAS method of Interindustry Coefficient Adjustment", *Journal of Regional Science,* vol 14.

March, L. and M. Batty (1975) "Generalized Measures of Information, Baye's Likelihood Ratio and Jaynes' Formalism", *Environment and Planning B,* vol 2, pp 99-105.

Margenau, H. (1950) *The Nature of Physical Reality,* New York.

Marksjö, B. (1981) "Simple Aggregation and Disaggregation subject to Minimal Information Loss and Other Criteria", Stockholms Läns Landsting.

Marschak, J. (1971) "Economics of Information Systems", *Journal of American Statistical Association,* vol 66, pp 192-219.

Marschak, J. (1973) "Limited Role of Entropy in Information Economics", in R. Conti and A. Ruberti (eds), *Fifth Conference on Optimization Techniques, Part II*, New York, Springer Valley.

Marschak, J. (1975 a) *Economics of Organizational Systems*, Amsterdam: North-Holland.

Marschak, J. (1975 b) "Entropy, Economics, Physics", in N.A. Chigier and E.A. Stern (eds), *Collective Phenomena and the Applications of Physics to Other Fields of Science*, Brain Research Publications, Inc.

Mathai, A.M. and P.N. Rathie (1975) *Basic Concepts in Information Theory and Statistics*, New Delhi: Wiley Eastern Ltd.

Mathur, P.N. (1972) "Multi-regional Analysis in a Dynamic Input-Output Framework", in A. Brōdy and A.P. Carter (eds), *Input-Output Techniques*, Amsterdam: North-Holland.

Medvedkov, Y. (1967) "The Concept of Entropy in Settlement Pattern Analysis", *Papers, Regional Science Association*, vol 18, pp 165-168.

Mennes, L.B.M., J. Tinbergen and J.G. Waardenburg (1969) *The Element of Space in Development Planning*, Amsterdam: North-Holland.

Mesarovic, M.D., D. Macko and Y. Takahara (1970) *Theory of Hierarchical, Multilevel Systems*, New York: Academic Press.

Miernyk, W.H. (1966) *The Elements of Input-Output Analysis*, New York: Random House.

Miernyk, W.H. (1972) "Regional and Interregional Input-Output Models: A Reappraisal", in M. Perlman, C.J. Leven and B. Chinitz (eds), *Spatial, Regional and Population Economics*, New York: Gordon and Breach.

Miernyk, W.H. (1976) "Comment on Recent Developments in Regional Input-Output Analysis", *International Regional Science Review*, vol 1, pp 47-55.

Miernyk, W.H. et al. (1970) *Simulating Regional Economic Development*, Lexington, Massachusetts: D.C. Heath and Company.

Mogridge, M.J.H. (1972) "The Use and Misuse of Entropy in Urban and Regional Modelling of Economic and Spatial Systems", Working Paper No 80, Centre for Environmental Studies, London.

Moore, F.T. and J.W. Petersen (1955) "Regional Analysis: an Interindustry Model of Utah", *Review of Economics and Statistics*, vol 37, pp 368-381.

Morimoto, Y. (1970) "On Aggregation Problems in Input-Output Analysis", *Review of Economic Studies*, vol 37, pp 119-126.

Morimoto, Y. (1971) "A Note on Weighted Aggregation in Input-Output Analysis", *International Economic Review*, vol 12, pp 138-143.

Morishima, M. (1958) "A Dynamic Leontief System", *Econometrica*, vol 26, pp 358-380.

Morishima, M. (1964) *Equilibrium, Stability and Growth: A Multi-sectoral Analysis*, Oxford: Clarendon Press.

Morrison, W.I. and P. Smith (1974) "Nonsurvey Input-Output Techniques at the Small Area Level: An Evaluation", *Journal of Regional Science*, vol 14, pp 1-14.

Moses, L.N. (1955) "The Stability of Interregional Trading Patterns and Input-Output Analysis", *American Economic Review*, vol 45, pp 803-832.

Moses, L.N. (1960) "A General Equilibrium Model of Production, Interregional Trade, and Location of Industry", *Review of Economics and Statistics*, vol 42, pp 209-224.

Murakami, Y. et al. (1970) "Efficient Paths of Accumulation and the Turnpike of the Japanese Economy", in A.P. Carter and A. Brōdy (eds), *Applications of Input-Output Analysis*, Amsterdam: North-Holland.

Murchland, J.D. (1966) "Some Remarks on the Gravity Model of Traffic Distribution and an Equivalent Maximization Formulation", LSE-TNT-38, Graduate School of Business, Transport New Theory Unit, London University.

McManus, M. (1956) "General Consistent Aggregation in Leontief Models", *Yorkshire Bulletin of Economic and Social Research*, vol 8, pp 28-48.

Nathanson, M. (1978) "Information Minimization, Markov Transition and Dynamic Modelling", *Environment and Planning A*, vol 10, pp 879-888.

von Neumann, J. (1945) "A Model of General Equilibrium", *Review of Economic Studies*, vol 13, pp 1-9.

Nevin, E., A.R. Roe and J.I. Round (1966) *The Structure of the Welsh Economy*, Cardiff: University of Wales Press.

Nijkamp, P. (1975) "Reflections on Gravity and Entropy Models", *Regional Science and Urban Economics*, vol 5, pp 203-225.

Nijkamp, P. (1977) *Theory and Application of Environmental Economics*, Amsterdam: North-Holland.

Nijkamp, P. (1978 a) "Competition among Regions and Environmental Quality", in W. Buhr and P. Friedrich (eds), *Competition among Small Regions*, Baden-Baden: Nomos Verlag.

Nijkamp, P. (1978 b) "Compromise Choices in Spatial Interaction and Regional Planning Models", in Karlqvist, A. et al. (eds), *Spatial Interaction Theory and Planning Models*, Amsterdam: North-Holland.

Nijkamp, P. and P. Rietveld (1976) "Multi-objective Programming Models: New Ways in Regional Decision-making", *Regional Science and Urban Economics*, vol 6, pp 253-374.

Nyhus, D. (1980) "The INFORUM System of Input-Output Models and its Relationship to the Forestry Industry", Paper presented at the Nordic Workshop on Models for the Forest Sector, Hemavan, April.

Paelinck, J. and P. Nijkamp (1975) *Operational Theory and Method in Regional Economics*, Westmead: Saxon House, D.C. Heath Ltd.

Pattee, H.H. (1973) "The Physical Basis and Origin of Hierarchical Control", in H.H. Pattee (ed), Hierarchy Theory: The Challenge of Complex Systems, New York: Braziller.

Penrose, R. (1956) "On the Generalized Inverse of a Matrix", *Proc. Cambridge Phil. Society*, vol 51, pp 17-19.

Petri, P.A. (1972) "Convergence and Temporal Structure in the Leontief Dynamic Model", in A. Brōdy and A.P. Carter (eds), *Input-Output Techniques*, Amsterdam: North-Holland.

Polenske, K.R. (1970) "Empirical Implementation of a Multiregional Input-Output Gravity Trade Model", in A.P. Carter and A. Brōdy (eds), *Contributions to Input-Output Analysis*, Amsterdam: North-Holland.

Polenske, K.R. (1972) "The Implementation of a Multiregional Input-Output Model for the United States", in A. Brōdy and A.P. Carter (eds), *Input-Output Techniques*, Amsterdam: North-Holland.

Puu, T. (1979) *The Allocation of Road Capital in Two-Dimensional Space: A Continuous Approach*, Amsterdam: North-Holland.

Rēnyi, A. (1960), "Dimension, Entropy and Information", *Transactions, Second Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, Czechoslovak Academy of Sciences, Prague, pp 545-556.

Rēnyi, A. (1961) "On Measures of Entropy and Information", *Proceedings, 4th Berkley Symposium on Mathematical Statistics and Probability*, vol 1, pp 547-561.

Rēnyi, A. (1966) *Wahrscheinlichkeitsrechung VEB Deutscher*, Berlin: Verlag der Wissenschaften.

Reza, F.M. (1961) *An Introduction to Information Theory*, New York: Mc Graw Hill.

Richardson, H.W. (1972) *Input-Output and Regional Economics*, London: Weidenfeld and Nicolson.

Richardson, H.W. (1973) *Regional Growth Theory*, London: Macmillan.

Riefler, R.F. (1973) "Interregional Input-Output: A State of the Arts Survey", in G.G. Judge and T. Takayama (eds), *Studies in Economic Planning over Space and Time*, Amsterdam: North-Holland.

Rietveld, P. (1979) *Multi-objective Decision-making and Regional Planning*, Boston: Martinus Nijhoff.

Round, J.I. (1978 a) "On Estimating Trade Flows in Interregional Input-Output Models", *Regional Science and Urban Economics*, vol 8, pp 289-302.

Round, J.I. (1978 b) "An Interregional Input-Output Approach to the Evaluation of Nonsurvey Methods", *Journal of Regional Science*, vol 18, pp 179-194.

Roy, J.R., D.F. Batten and P.F. Lesse (1981) "Minimizing Information Loss in Simple Aggregation", *Environment and Planning*, in press.

Roy, J.R. and P.F. Lesse (1981) "On Appropriate Microstate Descriptions in Entropy Modelling", *Transportations Research B*, vol 15 B.

Schaffer, W.A. (ed) (1976) *On the Use of Input-Output Models for Regional Planning*, Leiden: Martinus Nijhoff.

Schaffer, W.A. and K. Chu (1969) "Nonsurvey Techniques for Constructing Regional Interindustry Models", *Papers, Regional Science Association*, vol 23, pp 83-101.

Scott, A.J. (1970) "Transportation and the Distribution of Population: Some Entropy Maximizing Models", Discussion Paper 32, Centre for Urban and Community Studies, University of Toronto.

Semple, R.K. and R.G. Golledge (1970) "An Analysis of Entropy Changes in a Settlement Pattern over Time", *Economic Geography*, vol 46, pp 157-169.

Sevaldson, P. (1970) "The Stability of Input-Output Coefficients", in A.P. Carter and A. Brōdy (eds), *Contributions to Input-Output Analysis*, Amsterdam: North-Holland.

Sevaldson, P. (1972) "Studies in the Stability of Input-Output Relationships: Effects of Aggregations and Changes in Coefficients on the Results of Input-Output Analyses", Working Paper No 72/6, Statistisk Sentralbyrå, Oslo.

Sevaldson, P. (1974) "Studies in the Stability of Input-Output Relationships: Price Changes as Causes of Variations in Input-Output Coefficients", Working Paper No 74/14, Statistisk Sentralbyrå, Oslo.

Shannon, C.E. (1948) "A Mathematical Theory of Communication", *Bell System Technical Journal*, vol 27, pp 379-423 and 623-656.

Shannon, C.E. and W. Weaver (1949) *The Mathematical Theory of Communication*, Urbana: University of Illinois Press.

Sharpe, R. and D.F. Batten (1976) "A DREAM Model for Regional Economic Planning", *Papers, Regional Science Association (Australian - New Zealand Section)*, vol 1, pp 3-18.

Sharpe, R., D.F. Batten and M. Anderson (1981) "Dynamic Modelling of an Urban Economy", Paper presented at the 8th European Symposium on Urban Data Management, Oslo, June.

Sharpe, R. et al. (1977) "Modelling Regional Impacts of National Economic Policies", *Papers, Regional Science Association (Australian - New Zealand Section)*, vol 2, pp 105-117.

Sharpe, R., O. Ohlsson and D.F. Batten (1979) "Equity versus Efficiency: Regional Population and Building Industry Impacts", *Papers, Regional Science Association (Australian - New Zealand Section)*, vol 4, pp 165-178.

Sheppard, E.S. (1976) "Entropy Theory Construction and Spatial Analysis", *Environment and Planning* A, vol 8, pp 741-752.

Simon, H.A. (1957) *Models of Man: Social and Rational*, New York: Wiley.

Simon, H.A. (1973) "The Organization of Complex Systems" in H.H. Pattee (ed), *Hierarchy Theory: The Challenge of Complex Systems*, New York: Braziller.

Skagerstam, B.S. (1975) "On Notions of Entropy and Information", *Journal of Statistical Physics*, vol 12, pp 449-462.

Skolka, J.V. (1964) *The Aggregation Problem in Input-Output Analysis*, Czechoslovakian Academy of Sciences, Prague.

Smith, P. and W.I. Morrison (1974) *Simulating the Urban Economy: Experiments with Input-Output Techniques*, London: Pion.

Snickars, F. (1979) "Construction of Interregional Input-Output Tables by Efficient Information Adding", in C.P.A. Bartels and R.H. Ketellapper, (eds), *Exploratory and Explanatory Statistical Analysis of Spatial Data*, Boston: Martinus Nijhoff.

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

Snickars, F. and L. Lundqvist (1978) "Investments and Transport in Interdependent Regions - A Small Dynamic Model", in W. Buhr and P. Friedrich (eds), *Competition among Small Regions*, Baden-Baden: Nomos Verlag.

Sraffa, P. (1960) *Production of Commodities by Means of Commodities*, Cambridge University Press, Cambridge.

Statistisk Sentralbyrå (1978) *Standard Industrial Classification*, Oslo.

Stewart, J.Q. (1974) "Empirical Rules Concerning the Distribution and Equilibrium of Population", *Geographical Review*, vol 38, pp 461-485.

Stewart, J.Q. (1948) "Demographic Gravitation: Evidence and Applications", *Sociometry*, vol 11, pp 31-58.

Stone, R. (1962) "Multiple Classifications in Social Accounting", *Bulletin de l'Institut Internationale de Statistique*, vol 39, pp 215-233.

Su, T.T. (1970) "A Note on Regional Input-Output Models", *Southern Economic Journal*, vol 36, pp 325-327.

Taneja, I.J. (1974) "A Joint Characterization of Directed Divergence, Inaccuracy, and Their Generalizations", *Journal of Statistical Physics*, vol 11, pp 169-176.

Theil, H. (1957) "Linear Aggregation in Input-Output Analysis", *Econometrica*, vol 25, pp 111-122.

Theil, H. (1967) *Economics and Information Theory*, New York: American Elsevier; Amsterdam: North-Holland.

Theil, H. (1972) *Statistical Decomposition Analysis*, Amsterdam: North-Holland.

Theil, H. and P. Uribe (1965) The Information Approach to the Aggregation of Input-Output Tables, Report 6503, Centre for Mathematical Studies in Business and Economics, University of Chicago.

Tiebout, C.M. (1957) "Regional and Interregional Input-Output Models: An Approach", *Southern Economic Journal*, vol 24, pp 140-147.

Tilanus, C.B. and H. Thiel (1965) "The Information Approach to the Evaluation of Input-Output Forecasts", *Econometrica*, vol 32, pp 847-862.

Tokoyama, K. et al. (1976) "Structures of Trade, Production and Development", in K.R. Polenske and J.V. Skolka (eds), *Advances in Input-Output Analysis*, Cambridge, Massachusetts: Ballinger.

Tolman, R.C. (1938) *The Principles of Statistical Mechanics*, New York: Oxford University Press.

Tribus, M. (1969) *Rational Descriptions, Decisions and Design*, New York, Pergamon Press.

Tribus, M. and R. Rossi (1973) "On the Kullback Information Measure as a Basis for Information Theory: Comments on a Proposal by Hobson and Cheng", *Journal of Statistical Physics*, vol 9, pp 331-338.

Tsukui, J. (1961) "On a Theorem of Relative Stability", *International Economic Review*, vol 2, pp 229-230.

Tsukui, J. (1966) "Turnpike Theorem in a Generalized Dynamic Input-Output System", *Econometrica*, vol 34, pp 396-407.

Tsukui, J. (1968) "Applications of a Turnpike Theorem to Planning for Efficient Accumulation: An Example for Japan, *Econometrica*, vol 36, pp 172-186.

Tsukui, J. (1979) *Turnpike Optimality in Input-Output Systems: Theory and Application for Planning*, Amsterdam: North-Holland.

United Nations (1968) "A System of National Accounts", *Studies in Methods* (Series F, No 2), New York: United Nations.

Uribe, P., C.G. de Leeuw and H. Theil (1966) "The Information Approach to the Prediction of Interregional Trade Flows", *Review of Economic Studies*, vol 33, pp 209-220.

Vanwynsberghe, D (1976) "An Operational Nonsurvey Technique for Estimating a Coherent Set of Interregional Input-Output Tables", in K.R. Polenske and J.V. Skolka (eds), *Advances in Input-Output Analysis*, Cambridge, Massachusetts: Ballinger.

Wallenius, J. (1975) *Interactive Multiple Criteria Decision Methods*, Helsinki: Helsinki School of Economics.

Walsh, J.A. and M.J. Webber (1977) "Information Theory: Some Concepts and Measures", *Environment and Planning* A, vol 9, pp 395-417.

Webber, M.J. (1975) "Entropy Maximizing Location Models for Nonindependent Events", *Environment and Planning* A, vol 7, pp 99-108.

Webber, M.J. (1976) "Entropy Maximizing Models for the Distribution of Expenditures", *Papers, Regional Science Association*, vol 37, pp 185-198.

Webber, M.J. (1977 a) "An Elementary Entropy-Maximizing Model of Urban Consumers", *Papers, Regional Science Association*, vol 39, pp251-271.

Webber, M.J. (1977 b) "Pedagogy Again: What is Entropy?", *Annals, Association of American Geographers*, vol 67, pp 254-266.

Webber, M.J. (1979) *Information Theory and Urban Spatial Structure*, London: Croom Helm.

Wiener, N. (1948) *Cybernetics*, New York: Wiley.

Wilkinson, J.H. (1965) *The Algebraic Eigenvalue Problem*, Oxford: Clarendon Press.

Williams, H.C.W.L. and A.G. Wilson (1980) "Some Comments on the Theoretical and Analytical Structure of Urban and Regional Models", *Papers, Regional Science Association*, vol 45.

Wilson, A.G. (1967) "A Statistical Theory of Spatial Distribution Models", *Transportation Research*, vol 1, pp 253-269.

Wilson, A.G. (1969 a) "Entropy Maximizing Models in the Theory of Trip Distribution, Mode Split and Route Split", *Journal of Transportation and Economic Policy*, vol 3, pp 108-126.

Wilson, A.G. (1969 b) "Developments of Some Elementary Residential Location Models", *Journal of Regional Science*, vol 9, pp 377-385.

Wilson, A.G. (1970 a) "Interregional Commodity Flows: Entropy Maximizing Approaches", *Geographical Analysis*, vol 2, pp 255-282.

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

Wilson, A.G. (1973) "Further Developments of Entropy Maximizing Transport Models", *Transportation Planning Technology*, vol 1, p 183.

Wurtele, Z.S. (1960) "Equilibrium in a Uniformly Expanding Closed Leontief-Type System", *Review of Economic Studies*, vol 28, pp 23-28.

Ya Nutenko, L. (1970) "An Information Theory Approach to the Partitioning of an Area", *Soviet Geography: Reviews and Translations*, vol 11, pp 540-544.

Zipf, G.K. (1949) *Human Behaviour and the Principle of Least Effort*, Cambridge, Massachusetts: Addison-Wesley Press.

Åberg, M. and H. Persson (1981) "A Note on a Closed Input-Output Model with Finite Lifetimes and Gestation Lags", *Journal of Economic Theory*, in press.