EVOLUTIONARY MODELS OF MIGRATORY SYSTEMS. A BEHAVIOURAL APPROACH WITH APPLICATION ON INTRAURBAN MOBILITY IN TORONTO CMA, 1966-1976

by

© PAVLOS SAVA KANAROGLOU, B.Sc., M.A, M.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

September 1987

EVOLUTIONARY MODELS OF MIGRATORY SYSTEMS: A BEHAVIOURAL APPROACH

DOCTOR OF PHILOSOPHY (1987) (Geography) McMASTER UNIVERSITY Hamilton, Ontario

TITLE: Evolutionary Models of Migratory Systems. A Behavioural Approach with Application on Intraurban Mobility in Toronto CMA, 1966-1976.

AUTHOR: Pavlos Sava Kanaroglou, B.Sc. (Athens University)

M.A. (McMaster University)

M.Sc. (McMaster University)

SUPERVISORS: Prof. Y.Y. Papageorgiou, Prof. K.-L. Liaw

NUMBER OF PAGES: (xiv), 287

ABSTRACT

In this thesis a general and formal conceptualization, within which any type of migration can be studied is proposed. This framework results from the merging of two different traditions of research. The first has its roots in statistical mechanics and is expressed by what is known as evolutionary models. In this respect we follow the master equation formulation of Weidlich and Haag(1983).

The second tradition has its roots in probabilistic choice theory and is expressed by the nested logit model. Our particular formulation in this respect follows the derivation of Ben-Akiva and Lerman(1985).

We demonstrate that, as evolutionary, our framework is more general than the existing dynamic frameworks. In addition, due to its connection to choice theory it has behavioural characteristics. The connection between the master equation and the nested logit model is provided by the migration rates or equivalently by the set of probabilities of individuals moving between any two zones of the

system. We argue that to a set of observed migration rates and population distribution corresponds, through a set of linear equations derived from the master equation, a unique equilibrium population distribution which indicates the tendency of the system at the moment of observation. Thus, the dynamic equilibrium is dependent on the socioeconomic conditions prevailing in the system through the migration rates.

Due to the level of aggregation of the available intrametropolitan household head migration data for the Toronto CMA and for the two time periods; 1966-71 and 1971-76, only one part of the nested logit model, the destination choice model, can be meaningfully calibrated. The results show that distance, the percentage of houses built before 1946, the number of new houses built during the study period and the density of the population are statistically significant predictor variables. We also undertake a preliminary analysis of the impact of new house completions in Toront City in the 1971-76 time period on the tendency of the system expressed by the dynamic equilibrium discussed above.

ACKNOWLEDGMENTS

I find it imperative, before getting into the main part of this thesis, to express my gratitude for the support, cooperation and advice that I received during the years of my Ph.D. studies in the Geography department at McMaster University.

I am grateful to Canada Mortgage and Housing Corporation (CMHC) for financial support through the University Scholarship for Graduate Studies Program during part of my stay in McMaster. The rest of the time I was supported by a McMaster University graduate scholarship.

The cooperation, advice and friendship of Y.Y. Papageorgiou and K.-L. Liaw, my two supervisors, are greately appreciated. Many of the ideas presented in this thesis were formed through long discussions with them. My thanks are also extended to W. Anderson who as a member of my Ph.D. advisory committee provided several useful suggestions and comments.

I am also grateful to several other individuals that helped me in various direct or indirect ways. L.J. King provided me with his support in several occasions. J.R. Miron provided, free of charge, the data base used in this thesis and spent several hours of his time to explain the details in it. Faculty members, fellow graduate students and thersecreterial and technical staff in the Geography Depertment at McMaster created an interesting, pleasant and stimulating research environment.

Finally, I would like to thank my wife Vina for her encouragement and support throughout these years. Both her and my daughter Niki have exhibited a great deal of patience with my frequent absence from the dinner table during the preparation of this thesis.

TABLE OF CONTENTS

Page

CHAPTER

Page

CHAPTER

LIST OF FIGURES

LIST OF TABLES

TABLE		Page
V.15	1971-76 Female Net Migration	166
V.16	Entropy Measures	171
VI.1	Dependent and Independent Variables, Departure Model: 1966-71	191
VI.2	Dependent and Independent Variables, Departure Model: 1971-76	192
VI.3	1971-76 Destination Choice Results. Migrant Heads of All Ages	195
VI.4	1966-71 All Migrant Heads	207
VI.5	Inclusive Values	210
VI.6	Departure Model Beta Weights	211
AP1.VI.1	The Distance Matrix (Km)	234
AP2.VI.1	1966 Destination Specific Independent Variables	260
AP2.VI.2	1971 Destination Specific Independent Variables	261
AP3.VI.1	1971-76 Destination Choice Results. Migrant Heads of Age 15-34	263
AP3.VI.2	1971-76 Destination Choice Results. Migrant Heads of Age 35-54	264
AP3.VI.3	1971-76 Destination Choice Results. Migrant	

TABLE		Page

INTRODUCTION

The study of human populations trascents the boundaries of many disciplines. Over the last two centuries many prominent scholars have made the observation, identification and explanation of population phenomena their subject of interest. Examples of these phenomena, observed in many countries, are the demographic transition from high to low birth and death rates and the increasing concentration of the population in few major urban centres, followed recently, at least according to some observers, by a reversal of this trend.

In addition, population projection and forecasting has been an integral part in everyday urban and regional planning practices. The opening and closing of schools, water and sewage facilities, new transportation routes and police and fire protection are closely linked to the size and composition of the population to be supported. Typically, then, every planning exercise starts with some assumptions about the future population size and distribution in the study area.

Although much of the population research has been descriptive, the nature of population phenomena and the need for accurate projections has led to the development of models as formal tools of analysis. The possible sources of population change in an area are fertility, mortality and migration. Models of population change have traditionally indicated a bias depending on the discipline within which they were developed. Many demographic models assume away the effects of migration while in geography, economics and sociology the primary concern has been the relationship between population movement and a number of factors, such as distance, and social and economic conditions in the study area.

More recently A. Rogers and others have advocated that it is not appropriate, as in formal demography, to study population changes in a single area, where migration is viewed as a balance or a residual in the basic demographic equation after encountering changes due to natural growth. It is more appropriate to consider a region divided into a number of mutually exclusive and collectively exhaustive areas or zones, where the out-migrant of one zone is the in-migrant of another. This way, the view of formal demography is merged with that of population geography and the other disciplines. Efforts in this direction resulted in the development of multiregional models of population change, known as Marcov chain models.

The importance of migration, specificaly, has been emphasized recently not only because it provides a link between the zones of an area but also because of declining birth and death rates it usually constitutes the major factor of population change. Particularly since, over and beyond its direct impact, it has indirect ones, such as that of altering the age profile of the population of both the receiving and the sending zone due to the age selectivity in the migration process. Thus, through migration both the fertility and mortality of the zones are affected.

Studies of migration in geography, sociology and economics have emphasized the important view that an observed migration pattern is the result of individual decisions to migrate. These decisions are in turn dependent on the social, economic and other characteristics of the zones. This view of migration is appropriately termed behavioural.

We call an area divided into zones, with migrants moving between the zones, a migratory system. A particular population distribution in the zones of the system constitutes a state of the system. The typical Marcov chain model is expressed by a set of time difference or differential equations that describe a trajectory of the migratory system.

The major contribution of this thesis is that it unifies the ideas developed separately in the literature into a single, general framework, within which migration can be studied. We term this framework evolutionary as opposed to dynamic. An evolutionary model considers, at any point in time, a probability distribution over the states of the system rather than a single state, as in dynamic models. In this sense, an evolutionary model is more general than a dynamic one. Given an evolutionary model, specific dynamic models of migration, icluding the usual Markov type models, can be derived under certain assumptions. The evolutionary model we propose is expressed by the master equation.

We also maintain the behavioural view that a particular trajectory of the migratory system, or the macroevolution of the system, is due to the migratory behaviour of the participating individuals, or the microbehaviour in the system. In the master equation, individual behaviour is embodied in the migration rates between the zones in the system, or the probability of an individual moving from one zone to another. Using discrete choice theory we express the probability of moving between zones as a function of the characteristics of individual migrants and the social and economic characteristics of the zones. The specific functional form we end up with is the nested logit model.

The general migratory framework we propose then is made up of two basic, interlinked components; the first is the master equation which models the macroevolution of the system, the second is the nested logit model which deals with the microbehaviour of individuals in the system. Although migration can be classified in many different ways, the most usual classification is according to the type of the spatial framework used. In this type of classification we have international, interregional, interurban, rural-urban and intraurban migration. We argue that within the general framework proposed any type of migration can be handled.

As a demonstration, we apply these ideas in the case of intraurban migration in the Toronto Census Metropolitan Area (CMA) for the time period 1966-76. The aggregate nature of our data base does not allow us to fully exploit the properties of the nested logit model. The full model consists of two logically connected models; the departure model and the destination choice model corresponding to the decision to move and the decision to select a particular destination, respectively. The number of degrees of freedom available allow us to calibrate meaningfully only the destination choice part.

The final idea that we incorporate in this thesis is that the observed migration rates at one point in time embody the tendencies of the system at this time. A particular dynamic equation, the linear, derived from the master equation, is used with an observed population distribution and a set of migration rates to arrive at a unique equilibrium population distribution. This particular distribution is the state of the system that would eventually prevail if the observed migration rates persisted indefinately. As such, the equilibrium state indicates the tendency of the system at the time the migration rates were observed.

The ideas described above in condenced form are explained in detail in the six chapters following this introduction. Chapter II is a review of the migration literature. The purpose of this review is threefold. Firstly, it is aimed to expose the main ideas and theoretical frameworks upon which migration research is based. Secondly, to reveal the empirical findings in the intraurban migration research. These findings will serve as a guide in selecting predictor variables for the statistical analysis. And thirdly, to reveal the need for a general migratory framework that encompases the desirable properties of previously used frameworks.

Chapter III presents a detailed derivation of the master equation and the associated deterministic system of equations. The first is an evolutionary model whereas the second is a dynamic approximation to the master equation. The dependency of the system dynamics on the probability of an individual to migrate between any two zones in the system is emphasized.

In chapter IV we use discrete choice theory to express the probability of migration between two zones in the system as a function of the characteristics of individual decision makers and the socioeconomic conditions prevailing in the zones of the migratory system.

The following two chapters could be characterized as empirical. Chapter V presents the data and a preliminary descriptive analysis of the observed migration patterns in the Toronto CMA. Chapter VI concentrates on the statistical analysis, presenting the predictor variables used and the results obtained.

In chapter VII we argue that the state representing the equilibrium population distribution, obtained by using the associated deterministic system, can be used as a short run predictive tool of the trajectory of the system. A attempt is also made to connect this equilibrium distribution to social and economic characteristics of the zones through the probabilities to migrate. Through a computer program we perform a preliminary analysis of the impact of variations in new house completions in Toronto City on the equilibrium population distribution.

Finally, chapter VIII is a brief summary of the findings, some concluding comments and suggestions for further research on the topics discussed.

Before we close this introduction we make some comments on the conventions used in the presentation of the material. Equations, tables and figures within the chapters are referenced by simple arabic numerals. Whenever reference is made to an equation table or figure of another chapter, its number is preceded by the number of the chapter which always appears in capital roman numerals. Finally, when reference is made to an equation, table or figure of an appendix to a chapter, then its number is preceded by the string APn, where n is the number of the appendix. Thus, figure AP3.VI.1, for example, means figure 1 in appendix 3 to chapter VI.

CHAPTER II

MIGRATION IN PERSPECTIVE: A REVIEW

Over the last two centuries, as societies moved from the Agrarian to the Industrial age, new technologies were introduced at an ever increasing pace. More specifically the phenomenal improvement in transportation technology has progressively facilitated the movement of people over larger distances. In addition, we have witnessed a parallel increase in human migration at the international, interregional, interurban or intraurban level (Zelinsky, 1971).

Human migration, whether a cause or an effect, is recognized today as one of the very important processes of change in advanced or developing societies. As a result it has been the focus of many researchers in Geography, Economics, Demography, and Sociology. The volume of literature that has been generated is immense. To give a complete account of this literature is impossible. It is also unnecessary since for the purposes of this thesis our interest is to expose the theoretical or conceptual frameworks in this literature and how these conceptual frameworks have been translated into statistical models suitable for hypothesis testing.

Only representative studies in each case are mentioned. A further limitation of this review is that it concentrates on applications and findings in relatively advanced countries such as Canada and the United States.

It is also useful to mention at the outset that this chapter is structured to show the need for an integrated evolutionary conceptual framework from which dynamic models of migration can be derived. These models should be consistent with a behavioural theory of migration and should be flexible enough to allow the testing of hypotheses.

It is deemed useful therefore to examine the existing conceptual frameworks. For our purposes a conceptual framework is the basic set of ideas through which migration is viewed. The conceptual framework adopted influences to a large extent the approach taken in studying migration. An approach may be totally descriptive or it may materialize in formal models of migration. It is thus conceivable, and there are numerous examples in the literature, of descriptive or formal models that have their roots in the same conceptual framework. Equally numerous are the examples of models that borrow ideas from more than one framework.

We may distinguish at least three frameworks. The first derives from the works of Ravenstein (1885, 1889) and it is schematic and mechanistic in nature. The second has its roots in the economic theory of labor markets and it will be therefore termed economic. The third developed originally within the sociological literature although many of the ideas were borrowed from psychology with heavy concentration on individual behaviour. Thus, this third framework is termed behavioural.

Figure 1 illustrates the three frameworks that developed within migration research. The schematic framework materialized in gravity and entropy maximizing models, with applications in any type of migration. The literature in the economic framework has been empirical with heavy concentration on interregional and interurban migration. Traditionally, a linear model is fitted to macro-data, although since the mid-seventies many writers have advocated the use of non-linear type of models and the advantages of micro type of data. The classification in macro-adjustment, human capital investment and microeconomic is mainly due to the background ideas that influence the use of predictor variables in the empirical analysis. Finally, the literature in the behavioural framework has concentrated on intraurban migration. The terms economic and sociological in the classification of the behavioural framework are mainly due to the disciplines within which the branches developed. Within the sociological literature we can distinguish two

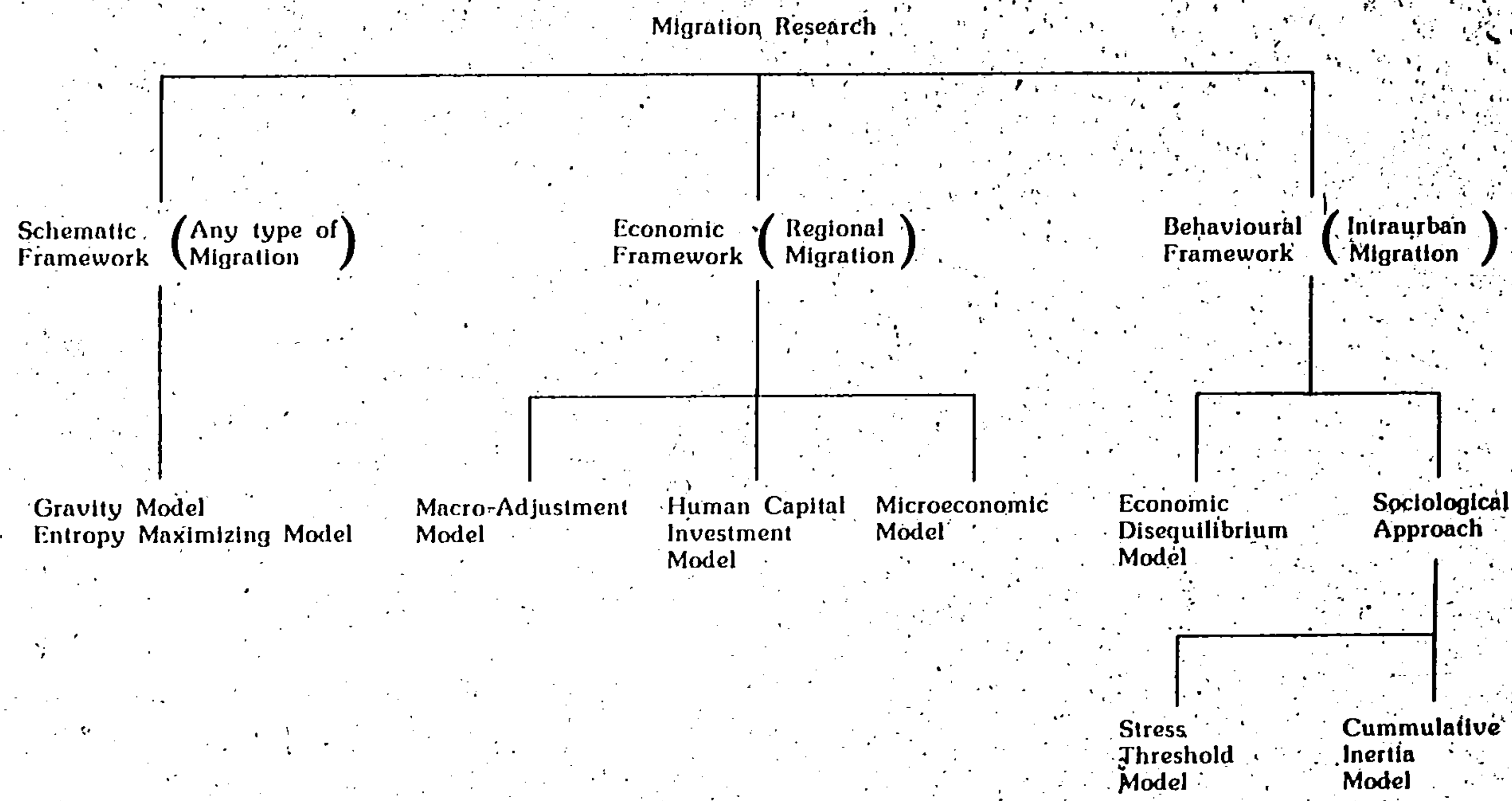

FIGURE 1. MIGRATION RESEARCH CLASSIFICATION

types of models depending on whether the emphasis is on the factors that encourage a move (Stress Threshold) or on those that inhibit a move (Cumulative Inertia). It should be emphasized, however, that any classification is not strict. Especially in the behavioural framework, much of the literature in the late seventies and eighties has attempted to integrate the various branches into a unifying behavioural framework.

The remainder of this chapter is divided into five sections. Sections 1, 2 and 3 provide descriptions of the three frameworks mentioned above. The basic ideas and models in intraurban migration are described as part of section 3, since most of the behavioural work has concentrated in this type of migration. We deem necessary however to include one more section (section 4) on intraurban migration where the accumulated empirical evidence in this area is described. This is because the particular application we are concerned with in this thesis deals with intraurban migration in the Toronto CMA. Finally, section 5 provides a brief overview on the work on dynamic and evolutionary models of migration and their link to the nested logit model. This section practically sets the stage of the introduction of the evolutionary model described in chapter III. We argue that there is no need to have different conceptual frameworks for different types of migration. The same type of model can be used to study intraurban, interurban or interregional migration by using appropriate independent variables.

1. The Schematic Framework

This framework is schematic or mechanistic in nature and derives from the works of Ravenstein (1885, 1889) whose main preoccupation was to uncover any existing laws in the human migration process at the macrolevel. Ravenstein by examining census data from Britain and twenty other countries uncovered seven regularities in the migration process which he called laws although he recognized that his "laws" did not have the rigidity of physical laws. The influence of Newtonian mechanics and the law of gravitation is profound in his work. Later on, his ideas were operationalized with a gravity type of model according to which a population centre attracts migrants from other centres in proportion to its population size and inversely in proportion to its distance from those centres. Migrants leave a centre according to the same principle. We mention some of the most important contributions in this direction.

Reilly (1929, 1931) used the concept of social gravitation to study the retail activity attracted by a city. His main difference with Ravenstein is that activity varied inversely with the square of distance rather than the distance itself. However, what gave way to a gravity model, as it appeared in several studies later on, were the works of Stewart (1941, 1948) and Zipf (1946). They suggested that the number of persons who move between any two communities whose respective populations are P_i and P_j and which are separated by a distance d_{ij}, will be proportional to the ratio $P_i P_j / d_{ij}$. Stewart (1948) extended the physical analogy to include the concept of "potential of population" which is a measure indicating the intensity of the possibility of interaction.

Further extensions in the model were introduced by Dodd (1950), Carrothers (1955) and Anderson (1956). In these extensions P_i, P_j and d_{ij} are raised in one way or another to a power which is to be estimated empirically.

By 1955, however, advocates of the approach appreciated its theoretical limitations. Carrothers (1955) summed up the then current feeling when he wrote:

> "As has been pointed out, the gravity and potential concepts of human interaction were developed originally from analogy to Newtonian Physics of matter. The behaviour of molecules, individually, is not normally predictable, but in large numbers their behaviour is predictable on the basis of mathematical probability. Similarly, while it may not be possible to describe the actions and reactions of the individual human in mathematical terms, it is quite conceivable that interactions of groups of people may be described this way. This possibility is suggested by the phenomenon, observable in all the social sciences and in city planning, that people behave

> differently in groups than they do as individuals. A fundamental difficulty arised from the different units of measure involved; the individual human being can make decisions with respect to his actions, while the individual molecule (presumably) cannot. This does not imply that interactions of humans in large numbers cannot be described mathematically, but it does mean that the threshold where the power of individual decision-making critically affects the results must be determined before the concepts can be broadly applied in practice."

The use of a gravity type-model, however, continued in various extended forms that allowed to include factors that affect migration, other than population and distance. Perhaps the most well known of those is the work of Lowry (1966) mainly because it was one of the first studies to suggest that the destination economic characteristics are more important than those of the origin; a finding supported by many other studies for different countries and using different types of models (Ledent and Liaw, 1985).

The work of Wilson (1967) constitutes the beginning of a new era for this type of modelling human migration and spatial interaction in general. Wilson by using ideas from statistical mechanics was able to derive the interaction between two places as a statistical average. The end result was a family of gravity type models.

His main argument goes as follows. Given an area divided into L subareas or zones and N individuals, a particular migration matrix

(M_{ij}) constitutes a state of the migratory system for a given time period. For any realistic system there is a very large number of possible states. The question is which one of all these states is the most likely. To answer this question one has to consider the concept of a microstate. For a given state (M_{ij}^{o}) any (i,j) element represents merely the number of individuals that moved from i to j within the given time period. This number tells us nothing about the particular individuals, like Mrs. Smith or Mr. Brown, that moved from i to j. If we wish to be that specific then certainly there are many different ways that the distinct individuals can be assigned to (i,j) pairs so that (M_{ij}^{o}) is obtained. Each one of these ways defines a microstate. Now if one is willing to accept that all the possible microstates have an equal probability of occurring then a likely state is the one associated with a high number of microstates. It follows that the most likely state is the one associated with the highest number of microstates. Thus, determining the most likely state becomes a relatively simple problem of calculus where the elements of the most likely state matrix are obtained as functions of the origin and destination populations and distance by maximizing the function that expresses the number of microstates associated with a state when population balance and cost constraints are satisfied. Because the function that expresses the number of microstates associated with a state is equivalent to the entropy function the resulting gravity type models are also called entropy maximizing models.

Despite the theoretical justification provided by Wilson for the use of the gravity type model, uncertainty surrounding the parameters of the model has persisted mainly because the underlying theory does not incorporate the individual decision process. A particular example of this uncertainty is the well known controversy and debate around the meaning of the distance exponent in a gravity model (see Curry, 1972; Curry et al., 1975; Sheppard et al., 1976; Cliff et al., 1974, 1975, 1976; and Johnston, 1975).

Within this particular framework one should also mention the work of Lee (1966) which represents an attempt to extend the ideas of Ravenstein (1885) by incorporating the intervening opportunities concept elaborated by Stouffer (1940, 1960). Thus, the decision to migrate and the process of migration is conceptualized to be affected by four types of factors:

(1) Factors associated with the origin

(2) Factors associated with the destination

(3) Intervening obstacles

(4) Personal factors.

Factors (1) and (2) above can be positive or negative reflecting attracting and repelling factors, respectively. The ideas of

Lee(1966) have led to a variation of the gravity model, the well known intervening opportunities model.

2. The Economic Framework

The second type of conceptual framework has its roots in the economic theory of labor. The origins can be traced to Hicks (1932). In this theory the human element is seen as labor which is an input in the production process. Spatial differences in the price of labor, or wage differentials, will cause labor to move from low to high wage areas. This move will increase the supply of labor in the receiving and it will decrease it in the sending area. This will result in lowering wages in the receiving area and increasing them in the sending area. Thus migration is seen as a mechanism that will equalize wages over the study area, or, to put it in a different way, migration is the process that will keep the system in equilibrium with respect to wages. A by-product of this line of reasoning is the link of migration to unemployment rate. Presumably, areas of high wages are those where labor is in short supply and therefore low unemployment rate. A migration stream will therefore be towards areas of low unemployment.

This theory gave rise to the well known "macro-adjustment model[4] of migration. It generated a voluminous empirical literature where a migration matrix (M_{ij}) is typically the dependent variable which is regressed against a set of independent variables. Most of the time (M_{ij}) represents gross out-migration. Net migration has also been used as a dependent variable and continues to be used (Graves, 1983); although many writers such as Vanderkamp (1971) and Liaw et al. (1985) have argued against its usage since net migration mixes two-way flows and as such it obscures the effects of certain independent variables. This is particularly so, since the influence of independent variables on in-migration and out-migration is not symmetric. In more recent studies as is the case of Shaw (1985) (M_{ij}) is normalized by dividing by the population of origin before it is used as a dependent variable.

Traditionally, migration has been considered to be a linear function of independent variables. Since the publication of Muth's (1971) paper considerable research effort has been allocated to the estimation of simultaneous linear equations (see Mueller, 1982, for a survey of this literature).

Given the theoretical foundation of this type of modelling one would expect that wage and unemployment levels would come out as the most significant of the independent variables used. In addition, the theory does not suggest that the impact on migration of these variables when measured for the origin should be different than that when measured at the destination. Empirical studies have consistently shown that wage or income level at the destination is significant whereas the same variable for the origin is not (Courchene, 1970). With respect to unemployment there is a significant absence of concensus among the researchers. Fields (1976), for example, finds that employment variables are more significant than income variables. Courchene (1970) reports that unemployment at the origin is more important than that at the destination. Finally, Greenwood (1975) in a literature review of empirical studies confronts migration scholars with the problem of explaining why unemployment at the place of origin appears to be insignificant in explaining migration. While plausible explanations have been offered, Shaw (1985) maintains that the basic problem is in the theoretical foundation of the model. According to him there are three basic limitations in this theory:

(1) Labor is not a commodity that can be traded freely with its price reacting quickly in situations of disequilibrium as a market clearing mechanism. On the contrary, labor as a commodity possesses characteristics that do not allow it to be traded freely. These are that:

(i) It is not homogeneous;

(ii) Information about its terms of exchange is not flowing freely;

(iii) The price (wage) is determined between a buyer and seller through a contract.

(2) The theory does not take into account migration for reasons unrelated to employment. Such reasons may be those related to stages in the life-cycle. These are moves, for example, for education, marriage or retirement reasons.

(3) The theory does not account for household moves. In a one income household even if the wage earner moves according to the theory, the rest of the household members will simply follow. Thus, a large percentage of the migration moves are of this type.

In search for a theoretical framework that allows a large number of degrees of freedom, many researchers turned to the human capital investment model proposed by Sjaastad (1962), an early critique of Hicks' model. He proposed to look at migration as "an investment which increases the productivity of human resources" rather than a labor price equilibrating mechanism. If migration is an investment then it must have some cost and some return. From the individual's point of view the cost is the lost present earnings and the monetary and non-monetary costs associated with the move. The returns, on the other hand, are the net discounted present value of the future earnings an individual expects to have due to migration. The model places more emphasis on the individual decision making process and as such it is more realistic. It allows for non-monetary costs and returns such as climate and culture to be taken into consideration. A lot of empirical evidence exists today that the non-monetary costs are more important in migration than the monetary ones. In an intermetropolitan migration study for the labor force entrants in Canada for the time period 1971-76, Liaw et al. (1986) reported that cultural dissimilarity between metropolitan areas is more important than employment growth at the destination both in terms of t-ratios and beta weights.

In addition, the human capital investment theory allows us to explain why the volume of migration decreases with age. This is, according to the model, because as individuals grow older, the time remaining to live is shorter and therefore any expected returns accumulated by the individual in the remaining life time will be less. The costs therefore tend to overweigh the returns as one grows older. Furthermore, the theory does not restrict the mobile unit to be an individual. It might, as well, be a household. One will have then to take into account the costs and expected returns for the whole household.

The human capital investment model has been criticized on theoretical grounds. Da Vanzo (1981), for example, argues that migrants in reality do not go about assessing their life time expected returns for every alternative destination before they decide to invest in migration. In addition on empirical grounds the expected life time returns discounted at the time of migration appear hard to implement in practice. If one is to interpret the returns of migration as income expectations then the existing empirical studies would not be supportive of the theory. More particularly studies in the U.S. by Yezer and Thurston (1976) and Martin and Lichter (1983) find that returns to migration five years after the move are either insignificant or negative. Moreover, in Canada Grant and Vanderkamp (1976) find that income returns in a short period after the move increases for interprovincial migrants. At a later and more detailed study, however, Grant and Vanderkamp (1980) using disaggregated data managed to contradict their previous result. In particular, they found that two to five years after a migration decision the monetary pay-off was still small and that it was smaller the higher the pre-migration income. A possible explanation offered by Shaw (1985) for this type of result is that in relatively well-off countries, such as Canada and the United States, the non-monetary motives of migration are more important. What becomes apparent, however, is that more empirical research is needed with microdata if one is to accept the hypotheses posed by the human capital investment model.

Recently, yet another way of thinking about migration started developing within the economic literature. This line of thinking emanates from the work of Graves and Linneman (1979) and it is strongly influenced by the approach taken in Urban Economics rather than that in Labor Economics. Individual migrants are assumed to possess a utility function which they are trying to optimize. The utility function is dependent on two types of goods: traded and non-traded. The non-traded goods are typically thought to be location-specific amenities. Change in demand for traded goods will not cause migration but a change in demand for non-traded goods will. An uneven change in prices over space for non-traded goods will cause a change in demand for these goods which will induce migration. Also, an uneven change in wages over space may cause a change in demand for non-traded goods and this will cause migratory adjustments.

The main difference between this view and the view of the labor economists is in the way equilibrium is perceived. For those that adhere to labor economics, equilibrium implies a fixed level of wages over space. By contrast, for the urban economist equilibrium implies the equalization of individual utility over space. The latter view appears to be broader and more general than the former for only if the location-specific amenities are assumed to be evenly distributed over space a utility equilibrium implies a wage equilibrium.

Later empirical work by Graves (1979, 1983) confirmed the significance of climatological amenities in net intermetropolitan migration in the United States. On this line is also the work of Krumm (1983) who contrasts locational amenity and economic variables in interregional migration in the United States. His conclusion is that regional wage differentials alone do not account for the observed migration patterns.

We may now conclude our discussion on the economic conceptual framework of migration by making these final observations. Firstly, most of the work that utilizes the economic conceptual framework has been oriented towards explaining interregional or intermetropolitan migration. Secondly, what we termed "the economic conceptual framework" encompasses into it three different ways of thinking which in the chronological order introduced are represented by:

(a) The labor economics model

(b) The human capital investment model

(c) The microeconomic model.

The line of thinking has progressively moved towards examining more closely the individual decision to migrate. More particularly the microeconomic model appears to be, at present, a more flexible framework within which one may conceive and study migration.

3. The Behavioural Framework

The last conceptual framework we review may be termed behavioural. This is because the conceptual structure of the model is based on assumptions about individual behaviour. Most of the applications of this framework relate to intraurban migration. Recently, however, some of the models developed within the behavioural theory, for example the increasingly popular logit model, find their way into any type of migration research. We may also say that the economic framework we examined in section 2 is becoming increasingly behavioural.

One of the main reasons why the behavioural framework was developed within the intraurban migration literature is because common sense dictates that intraurban moves are not related to employment. Researchers therefore in this area were forced to examine individual behaviour and the decision to migrate more closely since no dominant or obvious reason exists as to why households change their place of residence.

The basic ideas today have their roots in choice theory as it has been adapted in the field of economics although there is considerable influence from sociology. Although initial conceptions of the intraurban mobility process are attributed to Burgess (see

Park, Burgess and McKenzie, 1925) and Hoyt (1932), reference to traditional theoretical formulations of intraurban mobility and locational choice behaviour has usually alluded to the structural economic framework of Alonso (1964) and Muth (1969).

In the simplest form of the theory households are assumed to be able to associate a particular consumption of land and a composite good (any service and good other than land) with a utility index. A decision on a particular consumption is made by maximizing utility while obeying a budget constraint imposed by a given income level. The income of the household is spent on three items: land, the composite good and transportation cost. In every respect this is the standard microeconomic model.

Consider now a monocentric city where locations in it are differentiated only by their distance from the centre of the city, the CBD. Locations closer to the CBD are associated with lower transportation cost. Consider also a number of households in the city with identical preferences and incomes. The households are going to compete for land in the city. Out of this competitive process an equilibrium is going to emerge according to which all households are located in the city and they have no incentive to move since they have all adjusted their consumption of the composite good and land for the location they occupy so that the same equilibrium level of utility prevails over the city. For a given equilibrium utility level one can show that the bid rent function, which is actually the price of land at equilibrium, declines monotonically and at a decreasing rate, with distance. In this sense there is trade-off between accessibility to the CBD (or transportation cost) and the price of land. Different bid rent functions will be associated with different equilibrium utility levels but the basic form of the function will remain the same. Wheaton (1974) clarified how an equilibrium utility level for a city is deter~~mined~~ by introducing the concepts of an open and closed city.

The basic model described above has been extended in several directions in what constitutes today the field of New Urban Economics with a considerable body of literature. For the interested reader a non-mathematical treatment of the theory and its extensions is given by Miron (1982). There are also some excellent literature reviews, among them Anas and Dendrinos (1976) and Wheaton (1979).

The land market theory has been criticized extensively. Miron (1982), p. 135) summarizes the criticisms into four types: (1) the heavy reliance on a competitive market; (2) the static nature of the model; (3) the diminishing returns in the insights gained by recent analyses; and (4) the unrealistic nature of individual behaviour assumed in the model. Answers to these criticisms are provided by

Miron himself and others. In addition, it is claimed that for the phenomena studied a lot of insight has been gained and continues to be gained. For our purposes, the theory, although insightful, is not directly operational. It provided however a set of ideas that had a direct influence on more operational frameworks that have evolved in the meantime. The reader should recall that the microeconomic ideas have been recently adopted by researchers adhering to the economic framework discussed in the previous section.

Most of the recent theoretical conceptualizations in intraurban or residential mobility literature derive from the work of Wolpert (1965). Emphasis is placed on individual behaviour rather than a general equilibrium theory. Wolpert (1965), influenced by Rossi (1955), developed three central concepts in mobility behaviour: place utility, search strategies and life-cycle effects.

Place utility is defined to be the net composite of utilities from the household's integration in space (Wolpert, 1965, p. 162) and is evaluated through an aspiration level that adjusts itself on the basis of experience and stage in the life-cycle. Dissatisfaction, which is identified with a negative place utility, serves as a stimulus to search behaviour. The household's perception of opportunities is limited due to imperfect information concerning the urban environment, but it is still capable of differentiating choices according to relative or expected utilities. In this sense, the household is "intendedly" rational in its mobility decision. The domain over which households search is limited, or spatially biased, due to the limited perception of the surrounding environment. This bias leads to clustered sampling of residences close to the household's present location. The search domains of a household is a function of its socioeconomic characteristics and its present residential location."

Two main lines of inquiry have evolved from Wolpert's place utility approach to residential mobility. The first has been extensively developed within the economic literature, and is usually referred to as economic. This is not be confused with the economic framework discussed in section 2 of this chapter. The second line of inquiry in intraurban mobility has been developed within the sociological literature. In the remainder of this section we present the main ideas in the two lines of research starting with the economic.

Mobility in the economic literature is considered to be a mechanism for adjusting housing consumption. The stimulus to move is related directly to changes in housing demand and disequilibrium in consumption. Given a certain income and a preference structure a household has a certain optimal housing consumption. This optimal level of consumption changes over time. Sub-optimal level of consumption arises when the household's demand for housing changes to a new desired level. The presence of transaction and moving costs prevent the household from moving. A move will take place if the expected utility gain is higher than the disutility associated with transaction, search and moving costs.

Several variations of the disequilibrium theory have been developed into models. Goodman (1976) has taken Wolpert's conception of place utility to articulate a probability model of mobility which is a function of the discrepancies between actual and optimal housing consumption, and moving costs. Housing consumption is factored into several dimensions of housing quality and locational attributes. The household is assumed to possess a utility function composed of consumption levels of housing attributes per time period and a composite commodity, with an income constraint binding on the allocation of resources.

A more precise characterization of the notion of disequilibrium in housing consumption has been developed by Hanushek and Quigley (1979). In this framework, mobility is considered as an adjustment process which allows households to attempt to minimize the gaps which develop over time between actual and equilibrium consumption levels. Time is explicitly recognized in this model. Disequilibrium in consumption is hypothesized to be composed of two distinct components. The first, termed a structural disequilibrium component, is the gap between actual and desired consumption at a given point in time. The second is the change in the equilibrium level of housing consumption between two successive points in time. It is suggested that households are likely to react over time more rapidly to exogenous changes in their equilibrium consumption levels rather than to their structural disequilibrium gaps. As in Goodman's model a household's utility function is dependent on the consumption levels of housing attributes. A disequilibrium is characterized by an overconsumption of certain attributes and an underconsumption of others. Hanuchek and Quigley (1979) postulate that underconsumption exerts a stronger impact on housing demand than overconsumption.

Although the disequilibrium theory of housing demand is intuitively appealing, it results in models with properties that are difficult to confirm empirically. This difficulty stems mainly from the notion of the equilibrium consumption level which is not directly observable. In addition, aggregate market conditions tend to obfuscate individual deviations from equilibrium that the theory seeks to address. Furthermore, the theory addressed only one aspect of the intraurban household mobility process, the decision to change residences.

We come now to examine the sociological perspective on residential mobility. While not independent of economic arguments, the sociological approach to mobility theory has tended to emphasize the socio-psychological forces which give rise to the mobility decision. These forces up until recently have not been given recognition in the economic literature.

Two major themes have developed independently within the sociological literature. The first revolves around the concept of residential dissatisfaction or locational stress and the factors that give rise to it. Most of the arguments are primarily related to the characteristics of the household and their changes over time. A typical example is the life-cycle model which states that discrete events in the family's life-cycle have major impacts on the probability of moving. The second theme in the literature is related to the household's resistance to moving. According to this, psychological attachment to the dwelling and neighborhood, through interpersonal relationships and community ties, intensify over time, resulting in, what is called, cummulative intertia which is expected to lower the probability of moving. Although the two themes have evolved separately, recent literature has begun to integrate them into a single consistent framework. In the remainder of this section we discuss the developments in the two sociological themes.

As an underlying factor of residential mobility, locational stress relates changes in socioeconomic traits and characteristics in the residential environment to the degree of satisfaction experienced by the household at a given time and location, as this compares to the satisfaction that might be attainable elsewhere. The stress-threshold model, originally suggested by Rossi (1955), was developed by Brown and Moore (1970) as a continuous process of evaluation by the household as its characteristics and the residential environment change. Locational stress, when inflated beyond a given threshold value, prompts the household to act. A reduction in stress level is achieved either by improving the current residence or by seeking an alternative residence.

Much of the analysis in the literature has remained at the intuitive level. Relatively few attempts have been made to place this framework in an operational mode. Speare, Goldstein and Frey (1975) have considered that the mobility process consists of three interrelated stages; the development of the desire to consider moving, the selection of an alternative location and the decision to stay or to move. The first stage of the mobility decision is characterized as a transition process from satisfaction to dissatisfaction and is expressed as the probability of considering to move. At the second stage the probability of moving is calculated, for those who consider moving. In order to achieve that, the probabilities of moving to all the alternative destinations are calculated and compared in a cost/benefit analysis way. The probability of moving to any destination is expressed as a function of two variables; the difference in the residential satisfaction at the origin and the expected residential satisfaction at the destination and a threshold value related to mobility costs. The two submodels are integrated into a functional form of the joint probability of desiring to move and then actually moving.

Compared to the conceptual framework, the formalization of the mobility process in the Speare et al. model is an elementary one. Only few of the concepts suggested by Wolpert (1965), and subsequently developed by Brown and Moore (1970), were actually specified in the model. In addition, Brummell (1977) has argued that sociological approaches such as the one in Speare et al. suffer from a number of fundamental problems. Most of these problems, however, are the result of the less than rigorous fashion by which the concept of residential satisfaction has been defined. To overcome these difficulties, Brummell (1977) proposed a model with elements from both the sociological and the economic approach. Households are assumed to possess a utility which is a function of the attributes of the residential environment and all other goods consumed. The household's general consumption is subject to three types of constraints; the usual budget constraint, a minimum consumption level characterizing subsistence consumption and a maximum consumption level above which substitutability among goods is meaningless. The household, within these constraints, attempts to maximize its utility.

At any point in time a household experiences two types of utility. The first is a suboptimal one attained by a household, fixed at some point is space, that allocates its income to all other desired goods. The second type of utility is the one that the household would achieve if it were possible to move, to one of the locations about which it has gathered information over time from social and other contacts. The first represents the actual or experienced utility and the second is the aspiration utility or the maximum utility that can be achieved by moving to one of the alternative locations. Residential stress then is defined as the difference between the maximum of the attainable place utility the household currently believes it could be receiving and the experienced place utility from its present locational situation. Mobility results if stress exceeds a threshold level which quantifies the household's resistence to moving.

Brummell's framework, although similar to the economic one, possesses some characteristics that make it distinct. Economic theorists like Goodman (1976) and Hanushek and Quigley (1979) have tended to define disequilibrium for a household as an absolute measure of deviance from the optimal consumption pattern which is endogenously defined by the household's socio-economic characteristics. By contrast, Brummell defines locational stress as a relative concept which relates to the viable alternatives of which the household is aware. Households with the same preferences and socioeconomic conditions may be subject to different stress levels if their information about alternative housing opportunities is qualitatively and quantitatively different. Precisely because of this, the concept of a deterministic equilibrium does not exist in this framework.

Concurrent with the concern for understanding the factors which encourage mobility has been a similar one for understanding those which inhibit it. The major inhibiting factor of mobility is thought to be what has been termed by McGinnis (1968) cummulative inertia. The underlying behavioural assumption is that as time progresses, the household develops stronger emotional ties to the current residence because of the establishment of interpersonal relationships within the community and the progressive biasing of the household's cognitive image of the urban environment around home. This theoretical idea was originally suggested by Goldstain (1954). McGinnis (1968), after a series of empirical observations indicating that duration of stay effects in mobility situations appear to influence the probability of further movement by the household, transformed the idea into the theoretical axiom of cumulative inertia. He further developed a Markov chain type of model. In its simplest form, the distribution of the population among alternative locations is represented by a row vector, in which each element is the number of households residing in a zone of the city at an arbitrary reference time. The mobility process is then characterized as a sequence of probabilistic events over a series of discrete, equal time intervals. During each event, a proportion of the population in each zone migrates to another zone according to a transition matrix whose elements are the probabilities of migrating from one zone to another over the specified time interval.

At the end of each mobility event the household location vector is partitioned into a set of classes according to the residential histories of the households involved. The axiom of cumulative inertia is invoked by assuming that the probability of moving from one location to another declines as the number of time periods the population remains immobile increases. McGinnis (1968, p. 718) also assumes that any household that moves in any given interval erases all memory of previous locational stability and is transferred to the class of households that has remained in the current residence for no previous time periods. Under these assumptions, it is shown that the aggregate household location vector converges to a distribution which is different from the limit distribution of an equivalent Markov chain model.

This model, although it embodies only one behavioural aspect of the mobility decision, is extremely complex and its properties have never been analyzed exhaustively. Furthermore, the model has been criticized for the primitive assumption regarding the total cancellation of duration of stay effects following mobility and the assumption regarding the rate of immobilization of the population as duration of stay increases. As a result most of subsequent research has attempted to address these issues. The McGinnis model, however, often referred to as the Cornell Mobility Model, has served as the major reference point for this type of research.

An alternative approach to modelling the mobility process has been taken by Ginsberg (1971). He argued that the treatment of time at discrete intervals in Markov mobility models ~~severely~~ limits their utility. Thus, mobility is formulated as a semi-Markov process so that the timing of moves and destination selection are explicitly recognized in the framework. The waiting time from one move to the next is a random variable with a probability distribution that depends on the origin and destination states, and the length of time since the last move. The model has all the standard properties of McGinnis' earlier formulation plus some additional properties related to duration of stay effects. The assumption, however, that the moving probabilities are independent of the household's mobility history prior to the last move is still part of the model. Relaxation of this assumption requires the introduction of a higher-order Markov process. Due to the expected structural complexity such a model has not been proposed.

From what we have seen then, it appears that sociological investigations of residential mobility behaviour have been diverse and fragmented. Although suggested as early as Ginsberg (1972), the mutual study of the affirmative and inertial effects on mobility has been largely ignored. One major attempt which integrates the locational stress conception of mobility with duration of stay effects is by Huff and Clark (1978).

The Huff and Clark (1978) model extends the Speare et al. (1975) concept of residential satisfaction to a composite one which recognizes both stress formulation and resistance-to-move forces which influence the mobility decision. The probability of moving is assumed to be a function of the difference between residential stress and resistence to mobility experienced by the household at time t. In addition, both stress and resistance factors are explicitly considered as functions of time. Qualitative analysis with this model revealed four general classes of temporal behaviour which depends on the relative discount rates applied to resistance and stress measures. Simulation experiments with different parameterizations of the model have produced complex, but realistic mobility patterns.

4. Intraurban Migration Empirical Evidence

In this section we review the most important of the empirical evidence that has accumulated in the migration literature. Although findings are not always strictly comparable we make an attempt to present what has received considerable substantiation.

Consistent with the theoretical literature, presented in section 3, mobility is considered as a response to three types of stimuli:

- needs generated by changing socio-economic circumstances
- social and physical attributes offered by locations
- changing standards a household uses to evaluate its situation.

Before we proceed to examine what type of predictor variables were used in each of the above categories we would like to clarify what is meant by mobility. Three distinct conceptions of mobility have been used most often: desired, expected and actual mobility. Although the exact interpretation of each term has varied with the specific wording on survey questionaires, a concensus has emerged whereby desired and expected mobility are interpreted as different degrees of moving potential. Thus desired mobility reflects the household's mobility preferences whereas expected mobility reflects precisely articulated plans for moving.

With these mobility conceptualizations in mind we now turn to review the impact of changes in household socio-economic characteristics, to mobility. The most frequently used household socioeconomic characteristics are: family life cycle, age, tenure and income.

The family life-cycle, defined as the stages through which a family ages from formation to dissolution, has consistently been found to be strongly related to any type of mobility. Thus, expected mobility has been positively associated with younger families with the oldest child in the pre-school age (Butler et al., 1969), and with small families and very large families (Roistacher, 1974). Although not as strongly related, empirical evidence suggests a positive relationship between actual mobility and family life-cycle (Roistacher, 1974).

Age is normally perceived as the age of the household head.

The hypothesis that is usually tested is that age is a deterrent to residential mobility. This hypothesis has been supported when using as dependent variable the desired mobility (Rossi, 1955), when using expected mobility (Speare, 1974) and when using actual mobility (Duncan and Newman, 1976).

With respect to tenure it has consistently been found that those who own homes are far less likely to move than those who rent (Rossi, 1955; Roistacher, 1974; Goodman, 1976). The role of tenure and its relationship to other variables such as family life-cycle and age has been investigated by Pickvance (1974). His major finding was that family life-cycle and age affect mobility but only indirectly through tenure which acts as an intervening variable.

Relationships of income to mobility have not been consistently demonstrated in the literature. Its relationship to desired mobility has ranged from slightly positive and significant (Speare, 1974) to insignificant (Pickvance, 1974). A small positive relationship was found by Speare (1974) between income and expected mobility, but income has been found to have little to do in translating moving expectations into actual mobility (Duncan and Newman, 1976). Goodman (1976), however, has found the probability of moving to be significantly related to a change in income.

We turn now to examine the social and physical attributes offered by locations, or for short the locational attributes, and their impact on mobility. If locational attributes, as perceived by households, deviate from the expected levels or are perceived to be attainable at another location at an acceptable cost, they may stimulate the household into considering relocation. Five dimensions of location have been discussed in the literature. These are: neighbourhood characteristics, housing characteristics, accessibility, neighborhood services and housing cost. We discuss them in this order below.

Neighborhood characteristics or neighborhood quality is cited as one of the most important of the above dimensions (Butler et al., 1969; Speare, 1974). Neighborhood quality is not only important for mobility but also for residential choice. Butler et al. (1969), for example, found that an overwhelming majority of metropolitan households preferred better neighborhood quality with a less desirable housing unit or a less accessible location, to a less desirable neighborhood with either a better housing unit or better accessibility.

Although accepted in the literature to be important, few studies have attempted to define the components of neighborhood quality which are actually evaluated by households. Studies which have attempted to identify these components have usually done so by interpreting the loadings of independent variables on a reduced set of factors related to neighborhood quality (Butler et al., 1969; Atkinson and Phipps, 1977; Cadwallander, 1979).

Most of these studies revealed that neighborhood quality is evaluated in terms of two major cognitive categories, comprising the physical and social attributes of the neighborhood. Significant physical attributes include the exterior condition of dwellings and lots, conditions of public facilities such as streets and sidewalks and characteristics related to lot size, privacy and traffic noise (Butler et al., 1969, Cadwallander, 1979). Social attributes of importance include measures of friendliness and interaction with neighbors (Butler et al., 1969), crime and the demographic composition of the neighborhood (Atkinson and Phipps, 1977; Cadwallander, 1979).

Housing characteristics, and their relationship to household satisfaction and mobility, have been studied by several authors and have been found to be of major importance (Rossi, 1955; Michelson, 1977). Although characterized by several dimensions, usually relating to quality, distinctiveness and conveniences, the primary housing-related determinant has been interior space. Rossi (1955) found space and design to be the most frequently cited set of criteria in the residential choice decision, each attribute being mentioned by half the movers surveyed. Also Butler et al. (1969) in their national survey, found the number of rooms to produce the highest loading on satisfaction. They also found households to prefer modern architecture over traditional.

Although traditionally hypothesized to be a major determinant of mobility and choice, accessibility is perhaps the most ambiguous of the major dimensions considered. Inspired by the microeconomic framework, described in section 3, almost every empirical study has examined the impact of accessibility upon moving behaviour and the locational decision. In most of the studies accessibility is measured as simply the distance to workplace or as a composite distance to workplace and other amenities.

Barrett (1973) notes that only 2.7 percent of the respondents in his sample stated that living closer to workplace was one of the deciding factors in the decision to move and households in Michelson's (1977) survey cited general access as one of their reasons for moving only about 7 percent of the time. But other studies have found accessibility's role in the location decision to be more important. Weinberg (1975), for example, by analyzing the mobility histories or residents in the San Francisco Bay area, identified a significant association between journey-to-work and actual mobility. In this study hypotheses that workplace change affects residential mobility and that residential change affects workplace mobility were both supported by the data. Furthermore, Friedman (1975) found journey-towork costs significantly related to the locational decisions of households. One reason for the ambiguous findings was thought to be the fact that accessibility benefits are obfuscated by proximity costs. The costs of locating near major destinations such as workplaces or shopping areas often overweigh the benefits because of the externalities these places generate.

This hypothesis has received partial substantiation in the literature. Empirical studies that have isolated several distinct components of accessibility, related to the types of destinations the household is accessible to, have generally found positive associations between residential satisfaction and accessibility to destinations that generate fewer externalities. Examples of this kind of studies are Butler et al. (1969) and Weinberg et al. (1981).

The impact of neighborhood services on mobility and choice behaviour have been investigated by a number of authors. These services are the local public goods provided to constituents of local jurisdictions by government and quasi-government agencies. Typically, public services of this kind have included fire, garbage and police services (Butler et al. 1969; Atkinson and Phipps, 1977), recreational facilities (Cadwallader, 1979) and schools (Friedman, 1975). Hypotheses pertaining to their effects on mobility are those originating from the Tiebout hypothesis (Tiebout, 1956), which conjectured that households tend to settle in communities in accordance with their incomes and preferences for outputs of local public services (Friedman, 1975, p. 18).

Moore (1972) has suggested that public goods are likely less responsible for stimulating mobility than they are for attracting households to locations. This contention has received some support in the empirical literature. On the other hand, with regard to residential choice, Atkinson and Phipps (1977) have shown that movers tend to relocate to areas providing more or better public services.

Not all studies have concurred with these findings, however. Friedman (1975, p. 84), for example, concluded that most public goods and services play a minor role in residential choice.

The impacts of local public goods and services become less ambiguous when income is taken into consideration. Studies of residential location decisions have uncovered that lower income households prefer greater quantities of public services, tending to locate in areas characterized by high effective tax rates (Friedman, 1975, p. 81), middle class areas and areas of high per capita educational expenditures. Friedman has suggested that low income households pursue this locational strategy in order to increase their real incomes, avoiding full payment for these services by consuming smaller quantities of housing in this community.

A final locational attribute investigated in this section is housing cost. Moore (1972) has asserted that the overall effect of housing cost on the mobility decision is weaker than might be expected. Furthermore when cost is cited, it is more often articulated by renters than by owners. Cost considerations have been found to be of minor importance both for the decision to move and the locational decision (Michelson, 1977).

The influence of housing cost on mobility is difficult to determine since, as discussed previously, it is dependent to a great extent upon household income. Few studies have attempted to address these interdependencies and how they influence subsequent mobility behaviour. Goodman (1976) has shown that households whose housing expenditures are less than average for its economic status are more likely to move than families whose expenditures are greater, but his analysis in only suggestive since income consideration in relation to housing costs were not explicitly considered.

Finally, we are going to conclude this section by examining the standards or stated otherwise the attitudes and values held by the households and which in some way influence its mobility or residential choice. Moore (1972, pp. 8-9) has discussed standards in terms of "push" and "pull" factors acting upon the household. He has argued that push factors relate to how a household feels towards its present locational environment. These feelings which are translated into mobility are typically manifested in the attitudes the household possesses towards the social composition of the neighborhood. Thus, complaints about the neighborhood composition observed in most mobility studies (Atkinson and Phipps, 1977, p. 105) are thought to arise from the neighborhood undergoing a social change which the household deems undesirable or the household's changing expectations for the social environment in which the household operates. Although several authors have found these attitudinal factors to be important in neighborhood satisfaction ratings (Atkinson and Phipps, 1977) and actual mobility (Rossi, 1955), they have typically been shown to be less important than objectively defined attributes of the residence such as public services provided to the neighborhood, specific characteristics of the dwelling and changes in the household's socioeconomic circumstances.

Pull factors relate to household's perceptions of alternative residential environments and are usually expressed through its lifestyle aspirations. Four kinds of aspirations have been identified that motivate households into changing residence. These are:

consumption oriented aspirations, social prestige oriented aspirations, family oriented aspirations and community oriented aspirations. However, little in the way of empirical evidence has managed to quantify the relative importance of these aspirations or the context and conditions under which they become significant. Butler et al. (1969), for example, found no significant differences between the aspirations of movers and non-movers. In addition, suggestions that lifestyle aspirations are only manifested in the mobility behavior of homeowners have been refuted by Michelson (1977) who asserts that renters rather than lacking long term aspirations, are actually adopting incremental strategies in achieving their goals, since the attainment of an ideal is currently constrained by the stage in the family-mobility cycle in which they find themselves.

A final factor which has been indirectly related to standards is the previous mobility of the household. Behavioral explanations for this phenomenon have usually entered around the household's establishment of interpersonal relationships near the home and the progressive spatial biasing of the household's cognitive image of the urban environment. Clark and Huff (1977) maintain that although tests have supported the existence of duration-of-stay effects, they have not necessarily proven that they result from cumulative inertia.

Quigley and Weinberg (1977, p. 53) have gone one step further by suggesting that the relationship between previous mobility and moving probabilities is entirely due to spurious correlation.

5. Behavioural Models and Evolution: A Synthesis

In this section we would like to put together the basic ideas of a desirable framework that we are going to expand on in the following chapters. All, or almost all, of the ideas we use have been proposed elsewhere and many have been explained in the first three sections of this chapter. In what follows, therefore, we are trying to build a general framework within which to study migration by pulling together ideas that have developed in independent research streams and traditions. By general framework we mean one within which we can study any type of migration be it intraurban, interurban or interregional.

Central to the idea of a general migratory framework is the notion of a migratory system. Loosely speaking, this is the study area divided into a number of zones along with the people that move between zones and the social, economic and demographic characteristics of individuals and zones. The advantage of this approach then is that any insight gained through theoretical explorations for the general migratory system applies immediately to any type of migration. In addition a migratory system should be defined in a way that would allow easy adaptation to any specific circumstances for empirical explorations.

A state of the system is a variable vector $\underset{\sim}{S}$ of observable and measurable attributes of the system. The system typically is in one state $\underset{\sim}{S}_i|t$ at a given point in time t and in a different state $\underset{\sim}{S}_j|t+1$ at the next observational point in time $t+1$. With the idea of a migratory system in mind we may say that migration research in general has attempted to deal with three types of issues.

(1) To identify the forces that made the system move from state $\underset{\sim}{S}_i|t$ to state $\underset{\sim}{S}_j|t+1$.

(2) To project with certain bounds the direction the system is going to move in some future time.

(3) To control or direct the system towards a desired direction.

Issues (2) and (3) are particularly relevant to planning and are clearly dependent on a satisfactory answer to question (1). In subsequent chapters we restrict ourselves to issue (1) and peripherally to issue (2).

For our purposes a state of the system is a particular distribution of the population in the zones of the area of interest. This idea was also used by Wilson (1967, 1970) as explained in section 1. The migratory system moves over time from one state to another. If it were possible to constantly observe the migratory system over an observational time period then the trace of the state of the system over this period would form the trajectory of the system. To provide answers to question (1) above then it is desirable to build explicitly dynamic models that will provide explanations for particular trajectories of the system.

Some attempts made by Feeney (1973), Batty (1978) and others to extend Wilson's model into a dynamic one resulted in quasi-dynamic models of moving equilibria. Our focus therefore is more into a considerable body of literature that developed around Andrei Rogers at the International Institute for Applied Systems Analysis (IIASA) in Vienna, Austria as part of the Migration and Settlement Study. Most of the work in this area is based on the multiregional demographic model explained in detail in Rogers (1968, 1971, 1975). This is a system of simultaneous linear equations expressed in matrix notation. Time is treated as a discrete or a continuous variable with the equations being time difference or differential equations, respectively. The model is a direct generalization of the uniregional demographic model, which in its continuous form is known as the Lotka model (Lotka, 1907) and in discrete form as the Leslie model (Leslie, 1945).

The uniregional demographic model is concerned with the evolution of the population of one region, taking into account mortality, fertility and net migration. By contrast, the multiregional demographic model is concerned with a study area divided into regions. An outmigrant of one region in the system is viewed as an inmigrant of another. Thus, the population of the study area changes due to natural causes and also continuously redistributes itself over the regions of the area due to migration. In this sense migration provides the process that links the regions of the study area. At this point, we should note that the system as perceived in the multiregional demographic model is no different than the migratory system described above.

The multiregional model emphasizes the importance of the age composition of the population under study. The majority of the studies based on it are concerned with the short and long run implications of the recent past and current demographic characteristics of the population. Thus although migration is viewed as an important factor of population change in a region, no attempt is made to associate it with a social or decision making process as in the economic or the behavioural frameworks discussed in sections 2 and 3.

Since many nations are experiencing or are expected to experience zero population growth, the multiregional model has been used by Rogers and Willekens (1976, 1978) to study the effects of the components of population change on the spatial population distribution when the national population remains unchanged. Liaw (1980, 1981) has also used the model to study the evolution of size and spatial distribution of a population when the rates of birth, death and interregional migration remain constant. Finally, in connection with population dynamics Ledent (1981) has used the model in order to construct multiregional life tables and population projections by introducing heterogeneity in migration rates by region of birth.

The multiregional model belongs to a general class of models built with the help of stochastic processes and known as stationary Markov models. The reader may recall that stationary Markov models have also been used in connection with the cumulative inertia research, as explained in section 3 and Simmons (1974) and Porell (1981) have used the model to study intraurban migration.

One aspect in these models, the time stationarity assumption has been criticized by Pickles (1980) and Huff and Clark (1978). Joseph (1974, 1975) also commented on the unrealistic nature of the assumption but he suggested that the model is useful in the analysis of the distributional consequences of certain migration patterns.

Although several authors, such as Rogerson (1979), have attempted to develop non-stationary Markov models, more recently, a more general class of models has been introduced. In these models, termed evolutionary, each state of the system is associated with a certain probability of occurrence. Some states are highly likely but others are very unlikely. Thus, we may imagine a probability distribution over the states of the system. Evolutionary models allow to study the change of the state probability distribution over time. In this sense evolutionary models are more general than the conventional dynamic models, just like the dynamic models are more general than the static ones.

The roots of evolutionary models are in statistical mechanics and have been introduced in social sciences by Haken (1977) and Nicolis and Prigogine (1977). Our formulation represents an evolutionary model and follows the works of Weidlich and Haag (1983), de Palma and Lefevre (1983) and Haag and Weidlich (1984). From the general evolutionary model, specific dynamic models can be derived under certain assumptions. Although a general evolutionary model can rarely be used in empirical applications, its generality allows theoretical investigations that can yield insights about any migratory system. In addition the circumstances under which dynamic models are derived are precisely known and in some cases empirically tested.

As mentioned earlier on, what forces the migratory system to move from one state probability distribution to another is the change in the probabilities to move from any zone i to any zone j. These probabilities, in turn, are affected by the social, economic and demographic characteristics of the zones and the individuals that move between them. Given that the probabilities (p_{ij}) appear in the evolutionary equations and the resulting dynamic ones, our next step is to express p_{ij} in terms of the social, economic and demographic characteristics of the system.

In the behavioural framework most of the papers that followed Wolpert's (1965) work, both in the economic or the sociological literature expressed the probability of moving in terms of socioeconomic characteristics of the system. It is important to note that most of these works dealt with the probability of moving and not with the probability of moving from one zone to another. Thus, representing the economic literature Goodman (1976) and Hanushek and Quigley (1979) express the probability of moving as a function of disequilibrium housing consumption. The Speare et al. (1975) model representing the stress models in the sociological literature expressed the probability of moving as a function of the difference of residential satisfaction and the cost of moving. Although the resistance to move models dealt with the probability of moving from a zone i to a zone j, they expressed it only as a function of previous mobility (McGinnis, 1968; Ginsberg, 1971).

The schematic framework, reviewed in section 1, has been criticized for the lack of an underlying behavioural theory. The economic framework, on the other hand, has become over the years increasingly behavioural. We would like, therefore, to adopt a model which is based on a behavioural theory. In addition, the model should express the probability of moving from any zone i to any zone j of the system, thus allowing to study simultaneously the decision to move and the decision to select a destination. As Onaka and Clark (1983) point out the nested logit model fulfills these requirements.

The nested logit model was introduced by McFadden (1978) in the residential relocation context. In our case we follow the derivation of the model by Ben-Akiva and Lerman (1985), since, to our knowledge, it is the one that puts forward all the underlying assumptions. The detailed derivation of the model in the migration context is the subject of chapter IV.

CHAPTER III

MACROEVOLUTION AND MICROBEHAVIOUR

A set of geographical areas exchanging migrants is conceived as a <u>system</u>. For simplicity the system is considered closed, which implies that the total population in the system remains constant over time. The geographic areas, hereafter referred to as areas, could be cities, zones within a city or provinces (states) in a country. A particular population distribution over the areas constitutes a <u>state</u> of the system. At a given point in time each possible state is associated with a probability of occurrence. One, then, may think of a probability distribution over all the possible states of the system at a particular point in time. The change of this probability distribution over time is termed the <u>evolution</u> of the system. This evolution is modelled with a set of time differential equations, which, in vector form, constitute the <u>master equation</u>.

The large number of equations involved (one per state) makes the master equation suitable only for theoretical explorations in systems of a very small number of areas. Under certain conditions, however, particularly suitable for migratory systems, the master equation can be approximated by a deterministic set of equations that models the change over time of the most likely state of the system. Each equation in this set is a time differential equation describing the change over time of the most likely population size of one area. This change depends on the prevailing set of <u>migration rates</u>.

The migration rate, from area A to area B, can be thought of as the number of individuals who migrate from area A to area B within a unit time interval. In a real system migration rates change continuously over time. In practical terms, however, since observations pertain to one-year or five-year intervals, migration rates are usually considered constant within the observation period, but in general they are different from one observation period to the next.

For an observed set of migration rates in a given time period, the deterministic set of equations will provide a unique path of the system to a unique <u>equilibrium</u> population distribution, or a steady-state. This equilibrium population distribution indicates the tendency of the system at the time the migration rates were observed.

It is further argued that an observed set of migration rates is a manifestation of the socioeconomic conditions in the system at the time of the observation. Alternatively, the migration rates exhibit the reaction of individuals to prevailing socioeconomic conditions and they thus provide a link between the microbehaviour of individuals and the macroevolution of the system. To put it simply, the socioeconomic conditions through individual behaviour and migration rates, determine the tendency of the system which is represented by the equilibrium population distribution.

The remainder of this chapter is divided into seven sections. Section 1 describes the basic ingredients of a migratory system and its states, as they are perceived in this thesis. Section 2 establishes a probability distribution over the states of the system and provides a Markovian type equation that describes the dependency of this probability distribution at time $t + \Delta t$ on the probability distribution at time t. This equation is used in section 3 as a stepping stone in order to derive the master equation which models the evolution of the probability distribution over the states of the system through time. By assuming that the individual decisions to migrate are independent, the master equation is refined and linked to individual behaviour in sections 4 and 5. The following two sections represent attempts to derive dynamic equations from the general master equation derived in section 5. These dynamic equations typically describe the path over time of a state that is presumed to adequately represent the probability distribution over the states of the system. Such states are the mean value state, investigated in section 6, and the most likely state, investigated in section 7.

The material presented in this chapter was published in a more condensed form in Kanaroglou, Liaw and Papageorgiou (1986a).

1. The System

This section introduces some terms that are used in later sections of this chapter. It also establishes an abstract world by retaining only those ingredients of the real world that are necessary, and sufficient at least to the understanding of the modeller, to study the phenomena of interest.

Consider a system having a total number N of <u>individuals</u> distributed among L <u>areas</u>. The geometry and size of the L areas is given exogenously. Both N and L are fixed. The distribution of individuals among the areas of the system at time t is a random vector, $R_n[t]$, with a realization $\underset{\sim}{n}$ being a vector of reals such that

$$\underset{\sim}{n} \equiv (n_1, \ldots, n_L) \ \underline{\text{with}}\ 0 \leq n_i \leq N \ \underline{\text{for}}\ i = 1, \ldots, L$$
$$\underline{\text{and}} \sum_i n_i = N. \tag{1}$$

A particular $\underset{\sim}{n}$ is called a <u>state of the system</u>. The set of all states of the system is denoted $\mathcal{N}$.

Since N is fixed, change in the system is limited to migrations of individuals between the areas of the system. Change in human systems, on the other hand, further involves growth due to natural causes and migrations of individuals between the system and the rest of the world. Therefore N fixed appears restrictive. In practice, however, time is discrete. Given sufficient information, the population of the areas of the system can be adjusted at the beginning of each observation time period. In the absence of sufficient information, which is often the case, one will have to resort to models of mortality and fertility from formal mathematical demography to obtain estimates of the distribution of growth.

Given a particular state $\underset{\sim}{n}$, we may describe all states of the system as follows. Define a vector of integers

$$\underset{\sim}{m} \equiv (m_1, \ldots, m_L) \ \underline{\text{with}}\ 0 \le n_i + m_i \le N \ \underline{\text{and}}\ |m_i| \le N \ \underline{\text{for}}\ i = 1, \ldots, L, \ \underline{\text{and}}\ \sum_i m_i = 0. \qquad (2)$$

Since $\sum_i (n_i + m_i) = N$, it follows by definition that $\underset{\sim}{n} + \underset{\sim}{m} \in \mathcal{N}$. Clearly $\mathcal{N} = \{\underset{\sim}{n} + \underset{\sim}{m}\}$. The set of all vectors $\underset{\sim}{m}$ is denoted $\mathcal{M}$.

Such a system affords interpretations at the regional, inter-urban and intra-urban levels. More precisely, areas could represent the regions of a country, so that analysis pertains to interregional migration. On the other hand, if the object of our analysis is the evaluation of an urban system, some areas could represent cities while the rest could represent hinterlands. Finally, at an even more detailed level, areas could represent the neighbourhoods of a city.

2. State Probability Distribution

The evolution of human systems depends on past history. A full recognition of this point proves analytically intractable. For this reason, we introduce the Markov assumption, whereby only the very recent past is recognized to be relevant.

<u>Assumption 1:</u> <u>The present state of the system depends only upon the last previously observed state.</u>

Notice however that the last observed state is the outcome of past history and as such, at least to some extent, may serve to justify the assumption 1.

Using assumption 1, let $\text{Prob}\{R_n[t + \Delta t] = \underset{\sim}{n} + \underset{\sim}{m} \,|\, R_n[t] = \underset{\sim}{n}\} \equiv P[\underset{\sim}{n} + \underset{\sim}{m};\, t + \Delta t \,|\, \underset{\sim}{n};\, t]$ represent the probability of the system being in state $\underset{\sim}{n} + \underset{\sim}{m}$ at time $t + \Delta t$ given that it was in state $\underset{\sim}{n}$ at time t. Since the system can only be in one state at any given time, we have

$$
\begin{aligned}
P[\underset{\sim}{n} + \underset{\sim}{m};\, t \,|\, \underset{\sim}{n};\, t] &= 1 \ \underline{\text{for}}\ \underset{\sim}{m} = \underset{\sim}{0} \\
&= 0\ \underline{\text{otherwise}}.
\end{aligned}
\tag{3}
$$

Furthermore, since the system must be found in some state at any given time, we also have

$$\sum_{\underset{\sim}{m} \in \mathcal{M}} P[\underset{\sim}{n} + \underset{\sim}{m};\ t + \Delta t \,|\, \underset{\sim}{n};\ t] = 1. \tag{4}$$

Let $\text{Prob}\{R_n[t + \Delta t] = \underset{\sim}{n}\} \equiv P[\underset{\sim}{n};\ t + \Delta t]$ represent the probability of the system being in state $\underset{\sim}{n}$ at time $t + \Delta t$. Since the last previously observed state, upon which the present depends, can be any one of the states of the system, we have

$$P[\underset{\sim}{n};\ t + \Delta t] = \sum_{\underset{\sim}{m} \in \mathcal{M}} P[\underset{\sim}{n};\ t + \Delta t \,|\, \underset{\sim}{n} + \underset{\sim}{m};\ t]\ P[\underset{\sim}{n} + \underset{\sim}{m};\ t]. \tag{5}$$

Equation (5) provides the cornerstone for the ensuing analysis.

3. Master Equation

Since our concern is about the evolution of the system, it is convenient to investigate how the probability of a particular state in (5) changes over time. This gives rise to the <u>master equation</u> of the system defined by the following lemma.

Lemma 1:

$$\frac{d}{dt}P[\underset{\sim}{n};\, t] = \sum_{\substack{\underset{\sim}{m}\in\mathcal{M}\\ \underset{\sim}{m}\neq 0}} (w[\underset{\sim}{n};\, \underset{\sim}{n}+\underset{\sim}{m}]P[\underset{\sim}{n}+\underset{\sim}{m};\, t] - w[\underset{\sim}{n}+\underset{\sim}{m};\, \underset{\sim}{n}]\, P[\underset{\sim}{n};\, t]),$$

where

$$w[\underset{\sim}{n};\, \underset{\sim}{n}+\underset{\sim}{m}] \equiv \frac{\partial}{\partial \Delta t} P[\underset{\sim}{n};\, t+\Delta t\,|\,\underset{\sim}{n}+\underset{\sim}{m};\, t]\,\big|_{\Delta t=0}.$$

Proof: Using (4), we may write (5) as

$$P[\underset{\sim}{n};\, t+\Delta t] = \sum_{\substack{\underset{\sim}{m}\in\mathcal{M}\\ \underset{\sim}{m}\neq 0}} P[\underset{\sim}{n};\, t+\Delta t\,|\,\underset{\sim}{n}+\underset{\sim}{m};\, t]P[\underset{\sim}{n}+\underset{\sim}{m};\, t] +$$

$$(1 - \sum_{\substack{\underset{\sim}{m}\in\mathcal{M}\\ \underset{\sim}{m}\neq 0}} P[\underset{\sim}{n}+\underset{\sim}{m};\, t+\Delta t\,|\,\underset{\sim}{n};\, t])\, P[\underset{\sim}{n};\, t]. \qquad (6)$$

Expanding the conditional probabilities in Taylor series about t leads to

$$P[\underset{\sim}{n};\, t+\Delta t] = \Delta t \sum_{\substack{\underset{\sim}{m}\in\mathcal{M}\\ \underset{\sim}{m}\neq 0}} \frac{\partial}{\partial \Delta t} (P[\underset{\sim}{n};\, t+\Delta t\,|\,\underset{\sim}{n}+\underset{\sim}{m};\, t])P[\underset{\sim}{n}+\underset{\sim}{m};\, t] +$$

$$(1 - \Delta t \sum_{\substack{\underset{\sim}{m} \in \mathcal{M} \\ \underset{\sim}{m} \neq 0}} \frac{\partial}{\partial \Delta t}(P[\underset{\sim}{n}+\underset{\sim}{m}; t + \Delta t | n; t])) P[\underset{\sim}{n}; t] + o[\Delta t^2]. \tag{7}$$

Upon re-arrangement, the limit of (7) as $\Delta t \to 0$ gives the result.||

The quantities $w[\underset{\sim}{n}; \underset{\sim}{n} + \underset{\sim}{m}]$ describe how the probability of transition from state $\underset{\sim}{n} + \underset{\sim}{m}$ to state $\underset{\sim}{n}$ changes as the length of the (infinitessimal) transition period increases. Using (3) and the definition of w in lemma 1, for $\underset{\sim}{m} \neq 0$,

$$w[n; n + m] = \lim_{h \to 0} \frac{P[\underset{\sim}{n}; t + h | \underset{\sim}{n} + \underset{\sim}{m}; t] - P[\underset{\sim}{n}; t | \underset{\sim}{n} + \underset{\sim}{m}; t]}{h}$$

$$= \lim_{\Delta t \to 0} \frac{P[\underset{\sim}{n}; t + \Delta t | \underset{\sim}{n} + \underset{\sim}{m}; t]}{\Delta t}. \tag{8}$$

That is, $w[\underset{\sim}{n}; \underset{\sim}{n} + \underset{\sim}{m}]$ is a <u>transition rate</u> of probability per unit of time from state $\underset{\sim}{n} + \underset{\sim}{m}$ to state $\underset{\sim}{n}$. Since $w[\underset{\sim}{n}; \underset{\sim}{n} + \underset{\sim}{m}] \geq 0$ for $\underset{\sim}{m} \neq 0$, the probability of a transition cannot decrease as the length of the transition period increases. On the other hand, using once more (3) and the definition of w in lemma 1, for $\underset{\sim}{m} = 0$,

$$w[\underset{\sim}{n}; \underset{\sim}{n}] = \lim_{\Delta t \to 0} \frac{P[\underset{\sim}{n}; t + \Delta t | \underset{\sim}{n}; t] - 1}{\Delta t} \leq 0. \tag{9}$$

In other words, the probability of no transition cannot increase as the length of the transition period increases.

Using these remarks, the master equation can be interpreted as a probability rate equation. In particular, the change over time of the probability of state $\underset{\sim}{n}$ can be decomposed into (1) the increase due to changes in the probability of transition from all other states $\underset{\sim}{n} + \underset{\sim}{m}$ to state $\underset{\sim}{n}$, and (2) the decrease due to changes in the probability of transition from state $\underset{\sim}{n}$ to all other states $\underset{\sim}{n} + \underset{\sim}{m}$.

The master equation consists of $\binom{N + L - 1}{L - 1}$ coupled linear differential equations, each one corresponding to a distinct state. In principle, it contains all information about how the system could possible evolve. However, the form of the master equation as given by lemma 1 is too general and too abstract. It is too general because it admits transitions from any state to any other state, a possibility that counters intuition. It is too abstract in the sense that it includes only macro-descriptors of the system: it seems difficult to propose reasonable assumptions directly on the behaviour of the transition rates. Following Haag and Weidlich (1984), we now address these problems by cutting drastically the number of admissible transitions between states and, at the same time, by replacing the aggregate transition rates with analogous entities at the micro-level -- where they can be explicitly linked with individual decision-making.

4. Adjacent States

A pair of states is said to be adjacent if and only if the transition from one to another involves the movement of a single individual. In order to study adjacent states, we introduce the translation operators

$$E_i^{\pm k} f[\underset{\sim}{n}] \equiv f[n_1, \ldots, n_i \pm k, \ldots, n_L] \tag{10}$$

where k is a positive integer.

Let $p_{ji}[t,\Delta t]$ denote the probability of migrating to j from i during the time interval $[t, t + \Delta t]$ for a particular individual in i. Then the corresponding migration rate to j from i, is given by

$$q_{ji}[t] \equiv \lim_{\Delta t \to 0} \frac{p_{ji}[t,\Delta t]}{\Delta t}. \tag{11}$$

For infinitessimal time intervals, (11) implies that

$$p_{ji}[t,\Delta t] = q_{ji}[t]\Delta t \tag{12}$$

is approximately true.

It is worth mentioning at this point that in practice migration data relate to a time period of some length, normally of one or five years. In empirical terms therefore, in order to obtain an estimate of q_{ji} we are forced to consider that the migration rate from one area to another is constant over the observation period. If, for example, the observation period is one year and we consider a day as the unit of time then p_{ji} is expressed as a dimensionless number with value approximately equal to the number of people that moved from i to j during the year divided by the population in area i. The estimate of the migration rate q_{ji}, on the other hand, will have the units of inverse of time (per day) and its value will be equal to $p_{ji}/365$. Thus since q_{ji} is not directly observable we obtain an estimate of it by considering that it is constant for every day of the year. Note however that if in this example we consider the year as the unit of time then the estimates of both p_{ji} and q_{ji} will attain the same value but they will be measured in different units.

On an intuitive basis, the probabilities $p_{ji}[t,\Delta t]$ and hence the migration rates $q_{ji}[t]$ depend on the relative attractiveness of areas, expressed in utility form, which, in turn, are affected by local and system characteristics. Hence these probabilities reflect the complex interaction of many factors leading to the decision of an individual to migrate. It is useful to write $q_{ji}[t] = f_{ji}[\underset{\sim}{n}[t]; \underset{\sim}{\Theta}[t]]$, where $\underset{\sim}{\Theta}[t]$ includes all factors, other than the state of the system, which influence the decision of an individual to migrate.

<u>Assumption 2</u>: <u>Individual decisions to migrate are independent</u>.

There are two fundamental implications of the assumption (2). Firstly, it reduces the enormous number of elements in the sum of the master equation by forcing the system to evolve only through adjacent states. Secondly, it allows transition rates to be expressed in terms of the migration rates.

Contrary to assumption 2, individual decisions to migrate often exhibit interdependence. A typical implication of such interdependence is the simultaneous movement of several individuals from one area to another. For example, the employment change of a household head may result in the migration of the entire household. In consequence, there are many ways in which an observed population change could happen: if migration toward area i involves h individuals during a unit time period, there are 2^{h-1} possible sequences of state transitions which can produce this result. Assuming no emigration during the unit time period, figure 1 represents the case h=3. Clearly, at any infinitessimal time period Δt between t and $t+1$, there can be a state transition involving zero, one, two, or three individuals in area i. In essence, assumption 2 breaks down any state transition involving more than a single individual into a corresponding sequence of single-individual transitions. For example, if the second case of figure 1 actually happened, assumption 2 could approximate it as in the fourth case.

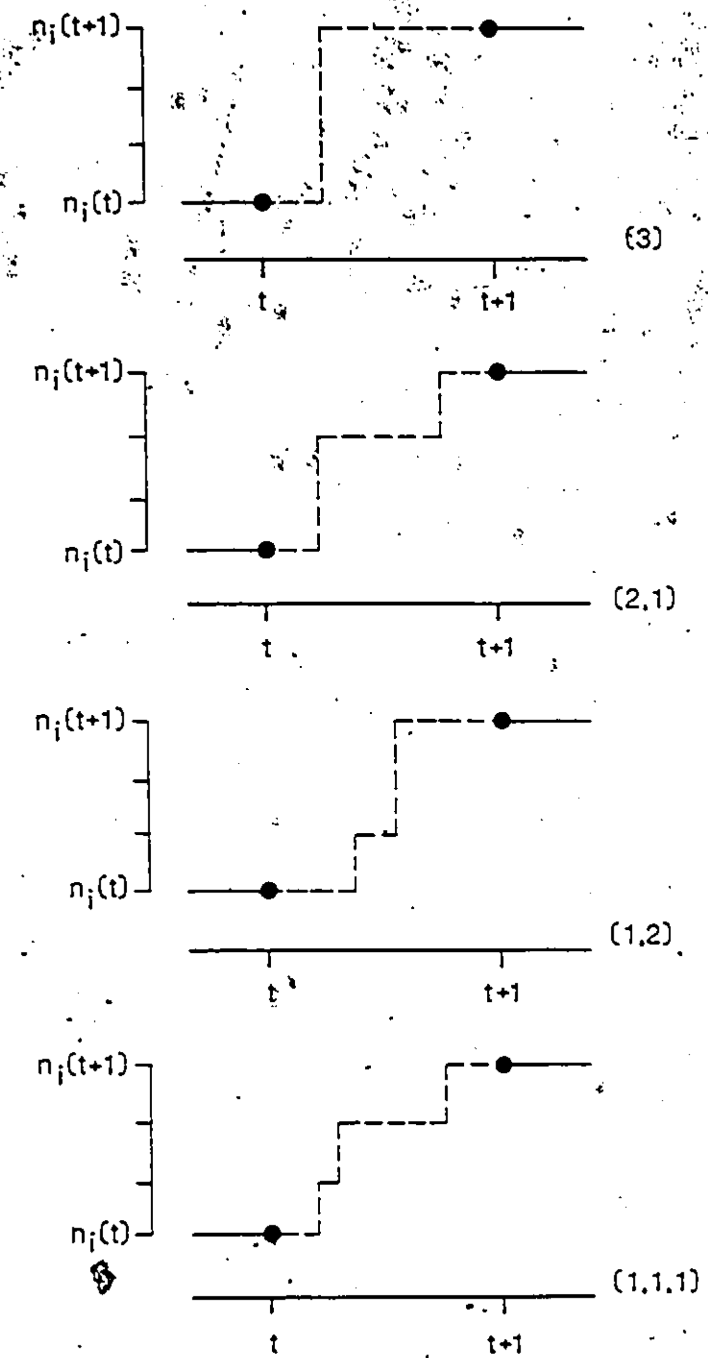

Figure 1: Alternative sequences of state transitions for three individuals.

This is established by the following two results.

<u>Lemma 2: The probability of h migrations from i to j during Δt is $o[\Delta t^h]$.</u>

<u>Proof</u>: Using assumption 2 the probability of h migrations from i to j during Δt is

$$P[E_j^{+h} E_i^{-h} \underset{\sim}{n};\, t + \Delta t | \underset{\sim}{n};\, t] = \binom{n_i}{h} (q_{ji}[t]\Delta t)^h (1-q_{ji}[t]\Delta t)^{n_i - h}. \| \quad (13)$$

In consequence, as $\Delta t \to 0$, it is sufficient to concentrate on the probability of a single migration from i to j during Δt.

<u>Lemma 3: The probability of h migrations during Δt is $o[\Delta t^h]$.</u>

<u>Proof</u>: Because of lemma 2, it is sufficient to examine migrations that originate in different places. Using once more assumption 2 together with (10), the joint probability of one migration from i to j and one migration from k to l during Δt is

$$P[E_l^{+1} E_k^{-1} (E_j^{+1} E_i^{-1} \underset{\sim}{n}); t + \Delta t | \underset{\sim}{n}; t] =$$

$$q_{ji}[t] n_i \Delta t (1 - q_{ji}[t]\Delta t)^{n_i - 1} \; q_{lk}[t] n_k \Delta t (1 - q_{lk}[t]\Delta t)^{n_k - 1} \qquad (14)$$

which is $o[\Delta t^2]$. The argument can be generalized for any pattern of migrations during Δt.||

In consequence, as $\Delta t \to 0$, it is sufficient to concentrate on the probability of a single migration during Δt. That is, change in the system is only represented by a sequence of adjacent states. This is the first implication of assumption 2. The second implication, which bridges the macro- and micro-levels of our system, is established by lemma 4.

<u>Lemma 4</u>: $\lambda[E_j^{+1} E_i^{-1} \underset{\sim}{n}; \underset{\sim}{n}] = q_{ji}[t] n_i$.

<u>Proof</u>: Consider the probability of a single migration from i to j during Δt. Using (13), this probability is written as

$$P[E_j^{+1} E_i^{-1} \underset{\sim}{n}; t + \Delta t | \underset{\sim}{n}; t] = q_{ji}[t] n_i \Delta t (1 - q_{ji}[t]\Delta t)^{n_i - 1}. \qquad (15)$$

Taking into account (8) and (15),

$$w[E_j^{+1} E_i^{-1} \underset{\sim}{n}; \underset{\sim}{n}] = \lim_{\Delta t \to 0} \frac{P[E_j \ E_i \ \underset{\sim}{n}; t + \Delta t | \underset{\sim}{n}; t]}{\Delta t}$$

$$= \lim_{\Delta t \to 0} q_{ji}[t] n_i (1 - q_{ji} \Delta t)^{n_i - 1}$$

$$= q_{ji}[t] n_i . \| \qquad (16)$$

5. Macroevolution and Microbehaviour

This section is based on Haag and Weidlich (1984).

We are now ready to express the master equation in terms of the migration rates, in other words to link decisions of individuals with the macro-evolution of our system.

Proposition 1: Under assumption 2,

$$\frac{d}{dt} P[\underset{\sim}{n}; t] = \sum_{i,j} (E_i^{+1} E_j^{-1} - 1) \, q_{ji}[t] n_i \, P[\underset{\sim}{n}; t].$$

Proof: Excluding all but adjacent states because of lemma 3, the master equation of lemma 1 becomes

$$\frac{d}{dt} P[\underset{\sim}{n}; t] = \sum_{i,j} w[\underset{\sim}{n}; E_i^{+1} E_j^{-1} \underset{\sim}{n}] \, P[E_i^{+1} E_j^{-1} \underset{\sim}{n}; t] -$$

$$\sum_{i,j} w[E_i^{+1} E_j^{-1} \underset{\sim}{n}; \underset{\sim}{n}] \, P[\underset{\sim}{n}; t]. \qquad (17)$$

Since the summations in (17) run over all i and j, exchanging indices in the second summation of the right hand-side simply changes the order of summation. Therefore,

$$\frac{d}{dt} P[\underset{\sim}{n}; t] = \sum_{i,j} w[\underset{\sim}{n}; E_i^{+1} E_j^{-1} \underset{\sim}{n}] \, P[E_i^{+1} E_j^{-1} \underset{\sim}{n}; t] -$$

$$\sum_{i,j} w[E_j^{+1} E_i^{-1} \underset{\sim}{n}; \underset{\sim}{n}] \, P[\underset{\sim}{n}; t]$$

$$= \sum_{i,j} E_i^{+1} E_j^{-1} \{w[E_j^{+1} E_i^{-1} \underset{\sim}{n}; \underset{\sim}{n}] \, P[\underset{\sim}{n}; t]\} -$$

$$\sum_{i,j} w[E_j^{+1} E_i^{-1} \underset{\sim}{n}; \underset{\sim}{n}] \, P[\underset{\sim}{n}; t]$$

$$= \sum_{i,j} (E_i^{+1} E_j^{-1} - 1) \, w[E_j^{+1} E_i^{-1} \underset{\sim}{n}; \underset{\sim}{n}] \, P[\underset{\sim}{n}; t]. \qquad (18)$$

We then use lemma 4 to establish the claim.||

At any given moment τ the number of possible states for a system of N individual and L regions is $\binom{N+L-1}{L-1}$. For any realistic system this is an extremely large number. Canada's 25 million, for example, can be distributed in its 10 provinces in $\binom{25000009}{9}$ ways. This is a number in the order of 10^{64}. Yet the system can be in only one state at any given time. The probability of any state $\underset{\sim}{n}$

at any moment τ is $P[\underset{\sim}{n}; \tau]$. Some of the states have a high probability of occurrence, others have a small one.

As defined in section 2 of this chapter, $P[\underset{\sim}{n}; \tau]$ represents a probability distribution of the random vector $R_n[\tau]$. The master equation is said to be an evolutionary model because the evolution of the probability distribution over the states of the system is modelled explicitly. By contrast, in dynamic models the concern is with a particular time path of the system. In this sense an evolutionary model is more general than a dynamic model, just as a dynamic model is more general than a static one.

It is important also to note that the master equation provides a link of the general evolution of the system with the migratory behaviour of individuals in the system which is embodied in the migration rates q_{ji}.

The master equation of proposition 1 represents a system of $\binom{N+L-1}{L-1}$ differential equations. A general analytical solution to this system of equations is not known. Neither is anything known about the general existence and stability characteristics of such a solution. What is then the use of this model?

Because of its generality this model is good for theoretical explorations. We can gain insight about the migratory system in hypothetical situations if we are prepared to make assumptions about the unknown probability distribution $P[\underset{\sim}{n}; t]$. In this thesis, however, this avenue is left unexplored. Instead we address the question of whether the evolutionary model represented by the master equation can be reduced to a dynamic model. In the following two sections we shall describe two ways of deriving dynamic models of the migratory system from the master equation.

6. Evolution of the Means

We know that a probability distribution can be adequately specified if we know a number of its moments. If the probability distribution is normal, then its first two moments, the mean and the variance, will completely specify the probability distribution.

The problem is that we do not know much about the probability distribution $P[\underset{\sim}{n}; t]$. As we mentioned in the previous section in any realistic system we have an enormous number of states. These states, however, are not equiprobable. A population distribution, for example, where one of the areas in the system has no population is an extremely unlikely one. Such highly unlikely states constitute a substantial proportion of the possible number of states, which implies that the probability distribution $P[\underset{\sim}{n}; t]$ must be associated with a rather low variance. This in turn implies that it is meaningful to consider the mean state as a representative of the probability distribution $P[\underset{\sim}{n}; t]$.

The mean of population size in j, $\bar{n}_j[t]$, is by definition

$$\bar{n}_j[t] = \sum_{\underset{\sim}{n}} n_j \, P[\underset{\sim}{n}; t]. \tag{19}$$

In order to determine its evolution, we must ~~first~~ establish the following lemma.

<u>Lemma 5:</u> <u>Let</u> $F_{ij}: \mathcal{N} \to \mathcal{R}$ <u>such that</u>

$E_j^{+1} E_i^{-1} F_{ij}[\underset{\sim}{n}] = 0$ <u>if and only if</u> $n_i = 0$

$F_{ij}[\underset{\sim}{n}] = 0$ <u>if and only if</u> $n_j = 0$.

<u>Then</u> $\sum_{\underset{\sim}{n}} E_j^{+1} E_i^{-1} F_{ij}[\underset{\sim}{n}] = \sum_{\underset{\sim}{n}} F_{ij}[\underset{\sim}{n}]$.

<u>Proof:</u> The number of terms in both summations is $\binom{N+L-1}{L-1}$, the total number of states. Of these, the number of states with $n_i = 0$ and the number of states with $n_j = 0$ are both equal to $\binom{N+L-2}{L-2}$. Hence the number of zero terms on the LHS equals the number of zero terms on the RHS. Every one of the remaining non-zero terms on the LHS equals a corresponding non-zero term on the RHS. To see this, consider a vector $\underset{\sim}{n}$ with $n_i > 0$. Then $E_j^{+1} E_i^{-1} F_{ij}[\underset{\sim}{n}] \neq 0$, and,

also, $F_{ij}[E_j^{+1} E_i^{-1} \underset{\sim}{n}] \neq 0$ because $n_j + 1 > 0$. In order to complete the argument, it suffices to notice that $E_j^{+1} E_i^{-1} F_{ij}[\underset{\sim}{n}] = F_{ij}[E_j^{+1} E_i^{-1} \underset{\sim}{n}]$ because of definition (10), and that $F_{ij}[E_j^{+1} E_i^{-1} \underset{\sim}{n}]$ is a non-zero term on the RHS.||

To illustrate lemma 5, consider the case L=3 and N=4. There are $\binom{N+L-1}{L-1} = \binom{6}{2} = 15$ possible states. Of these, the number of states with $n_i = 0$, $i = 1,2,3$, is $\binom{N+L-2}{L-2} = \binom{5}{1} \equiv 5$. The following table gives F_{ij} and $E_2^{+1} E_1^{-1} F_{ij}$ for all possible states. Using the first state (4,0,0) as an example, note that $F_{ij}[4,0,0] = 0$ because $n_2 = 0$ and $E_2^{+1} E_1^{-1} F_{ij}[4,0,0] = F_{ij}[E_2^{+1} E_1^{-1} (4,0,0)] - F_{ij}[3,1,0]$.

Following now Haag and Weidlich (1984), we are ready to express the evolution of the mean-value state independently of the unknown distribution $P[\underset{\sim}{n}; t]$.

<u>Proposition 2</u>: $\frac{d}{dt} \bar{n}_j[t] = \sum_i (\overline{q_{ji}[t] n_i[t]} - \overline{q_{ij}[t] n_j[t]})$ <u>for</u> $j = 1, \ldots, L$.

<u>Proof</u>: Differentiating (19) with respect to time and taking into account proposition 1,

STATE	F_{12}	$E_2^{+1}E_1^{-1}F_{12}$
(4,0,0)	0	$F_{12}[3,1,0]$
(3,1,0)	$F_{12}[3,1,0]$	$F_{12}[2,2,0]$
(3,0,1)	0	$F_{12}[2,1,1]$
(2,2,0)	$F_{12}[2,2,0]$	$F_{12}[1,3,0]$
(2,1,1)	$F_{12}[2,1,1]$	$F_{12}[1,2,1]$
(2,0,2)	0	$F_{12}[1,1,2]$
(1,3,0)	$F_{12}[1,3,0]$	$F_{12}[0,4,0]$
(1,2,1)	$F_{12}[1,2,1]$	$F_{12}[0,3,1]$
(1,1,2)	$F_{12}[1,1,2]$	$F_{12}[0,2,2]$
(1,0,3)	0	$F_{12}[0,1,3]$
(0,4,0)	$F_{12}[0,4,0]$	0
(0,3,1)	$F_{12}[0,3,1]$	0
(0,2,2)	$F_{12}[0,2,2]$	0
(0,1,3)	$F_{12}[0,1,3]$	0
(0,0,4)	0	0

$$\frac{d}{dt}\,\bar{n}_j[t] = \sum_{\underset{\sim}{n}} n_j \frac{d}{dt} P[\underset{\sim}{n};\, t]$$

$$= \sum_{\underset{\sim}{n}} n_j \sum_{i,k} (E_i^{+1} E_k^{-1} - 1)\, q_{ki}[t] n_i P[\underset{\sim}{n};\, t]$$

$$= \sum_{i,k} \sum_{\underset{\sim}{n}} E_i^{+1} E_k^{-1} (n_j - \delta_{ij} + \delta_{kj}) q_{ki}[t] n_i P[\underset{\sim}{n};\, t] -$$

$$\sum_{i,k} \sum_{\underset{\sim}{n}} n_j\, q_{ki}[t] n_i P[\underset{\sim}{n};\, t]. \qquad (20)$$

where δ_{ij} and δ_{kj} denote a Kroeneker delta. It can be established that the functions $p_{ki}\,[t] n_i P[\underset{\sim}{n};\, t]$ satisfy the conditions of lemma 5, hence that

$$\sum_{i,k} \sum_{\underset{\sim}{n}} E_i^{+1} E_k^{-1} n_j\, q_{ki}[t] n_i P[\underset{\sim}{n};\, t] = \sum_{i,k} \sum_{\underset{\sim}{n}} n_j\, q_{ki}[t] n_i P[\underset{\sim}{n};\, t]. \qquad (21)$$

Combining (20) and (21), we obtain

$$\frac{d}{dt}\,\bar{n}_j[t] = \sum_{i,k} \sum_{\underset{\sim}{n}} E_i^{+1} E_k^{-1} \delta_{kj}\, q_{ki}[t] n_i P[\underset{\sim}{n};\, t] - \sum_{i,k} \sum_{\underset{\sim}{n}} E_i^{+1} E_k^{-1} \delta_{ij}\, q_{ki}[t] n_i P[\underset{\sim}{n};\, t]$$

$$= \sum_{i} \sum_{\underset{\sim}{n}} (E_i^{+1} E_j^{-1} q_{ji}[t] n_i - E_j^{+1} E_i^{-1} q_{ij}[t] n_j) P[\underset{\sim}{n};\, t]$$

$$= \sum_{i} \sum_{n} (q_{ji}[t]\bar{n}_i - q_{ij}[t] n_j) P[\underset{\sim}{n};\, t], \qquad (22)$$

Using once more lemma 5.1 ||

When $P[\underset{\sim}{n}; t]$ is symmetric and unimodal, use of the mean-value state is appropriate for understanding something about the evolution of the system because it represents the most probable state. If however symmetry, or unimodality, or both are violated then using the mean-value state may not be appropriate. For example, in the case of bimodality, the mean-value state may be one of the least probable, rather than most probable, states. In such cases the evolution of the maximum, or maxima, of $P[\underset{\sim}{n}; t]$ should be directly investigated.

7. The Associated Deterministic System

It has already been mentioned that neither the solution to the master equation, nor the existence and stability characteristics of the stationary solution, are known in general. Consequently, as de Palma and Lefevre (1983) point out, several approximation methods have been proposed to overcome the impasse. All these are based on the property that, when the total population of the system becomes arbitrarily large, the variance of the corresponding probability distribution around the most probable state becomes arbitrarily small. Here, following de Palma and Lefevre (1983), we discuss an approximation method due to Kurtz (1978). His fundamental assumption concerns the structure of the migration rates q_{ji}. In general, these rates depend upon the state of the system $\underset{\sim}{n}$. The migration rates, however, may depend upon the state of the system through $\underset{\sim}{\rho} \equiv \underset{\sim}{n}/N$, rather than directly through $\underset{\sim}{n}$, where $\rho_i \equiv n_i/N$ is the proportion of individuals in region i. In this case we say that the migration rates are <u>extensive</u>, i.e. $q_{ji}[t] = f_{ji}[\underset{\sim}{\rho}[t]; \underset{\sim}{\Theta}[t]]$.

Let $R_\rho[t]$ be a random <u>vector</u> with realizations $\underset{\sim}{\rho}$. Since there is a one-to-one correspondence between $\underset{\sim}{n}$ and $\underset{\sim}{\rho}$, the random vectors $R_n[t]$ and $R_\rho[t]$ have the same probability distribution, in other words $\text{Prob}\{R_\rho[t] = \underset{\sim}{\rho}\} = P[\underset{\sim}{n}; t]$. Furthermore, using (1),

$$\underset{\sim}{\rho} \equiv (\rho_1, \ldots, \rho_L) \ \underline{\text{with}}\ 0 \le \rho_i \le 1 \ \underline{\text{for}}\ i = 1, \ldots, L$$

$$\underline{\text{and}} \sum_i \rho_i = 1. \qquad (23)$$

Notice that $\underset{\sim}{\rho}$ is a vector of reals. The set of all vectors $\underset{\sim}{\rho}$ is denoted $\mathcal{P}$.

It has been mentioned at the beginning of this section that the theorem of Kurtz hinges upon the property that $R_\rho[t]$ tends to a single state as N tends to infinity. In order to gain some intuition about this property, consider a migratory system with extensive migration rates at time zero, when the state of the system is $\underset{\sim}{\rho}^N[0]$. Perform now the following thought experiment. Gradually increase the total population of the system with initial condition $\underset{\sim}{\rho}^N[0]$, and examine how $R_\rho^N[0]$ behaves around this most probable state as the total population size tends to infinity. We denote by $p_i^N[0]$, $i = 1, \ldots, L$ the probability that a new individual will locate in area i. This probability will, in general, be a function of $\underset{\sim}{\rho}^N[0]$. As N increases, the impact of new additions on $p_i^N[0]$ will diminish -- provided that these probabilities are well-behaved. For N arbitrarily large, the change in these probability distributions will become arbitrarily small. However, due to the law of large numbers, the proportion of individuals allocated to any area will come arbitrarily close to the corresponding probability of locating there, that is,

$$\lim_{N \to \infty} R_\rho^N[0] = \lim_{N \to \infty} (p_1^N[0], \ldots, p_L^N[0]) \equiv \underset{\sim}{\rho}^\infty[0]. \qquad (24)$$

Therefore, if we add continuously individuals to an initial state

$\underset{\sim}{\rho}^N[0]$, $\lim_{N \to \infty} \text{Prob}\{R_\rho^N[0] = \underset{\sim}{\rho}^\infty[0]\} = 1$.

The same argument applies for the system of initial size N at another time arbitrarily close to zero; and so on... In this manner we obtain $\underset{\sim}{\rho}^\infty[\tau]$ over $\tau \geq 0$, which may be taken to represent the <u>deterministic</u> evolution of an associated limit-system with initial condition $\underset{\sim}{\rho}^{\infty}[0]$. The theorem of Kurtz provides the means to describe such evolution.

For $\underset{\sim}{x} \in \mathcal{P}$, consider the system of differential equations

$$\frac{d}{dt}x_j[t] = \sum_i (f_{ji}[\underset{\sim}{x}[t]; \underset{\sim}{\Theta}[t]]x_i[t] - f_{ij}[\underset{\sim}{x}[t]; \underset{\sim}{\Theta}[t]]x_j[t])$$

$$\equiv F[\underset{\sim}{x}[t]; \underset{\sim}{\Theta}[t]] \text{ for } j = 1, \ldots, L. \qquad (25)$$

Let $\underset{\sim}{x}[0]$ represent the initial condition for this system and $\underset{\sim}{x}^{\star}[t]$ a solution of the system at time t.

<u>Proposition 3:</u> <u>If F is Lipschitz, and if</u> $\underset{\sim}{\rho}^{\infty}[0] = \underset{\sim}{x}[0]$ <u>then, for every</u> $\tau < \infty$, $\limsup_{N \to \infty} \sup_{t \leq \tau} |\underset{\sim}{\rho}^{\infty}[t] - \underset{\sim}{x}^{\star}[t]| = 0$ <u>almost everywhere.</u>

<u>Proof:</u> See Kurtz (1978).

It is perhaps useful at this point to elaborate on some further details in the statement of the theorem of Kurtz. A continuous function $F[\underset{\sim}{x}[t]; \Theta[t]]$ is said to be Lipschitz if there exists a constant $c > 0$ such that $\|F[x^1[t]; \Theta[t]] - F[x^2[t]; \Theta[t]]\| < c\|\underset{\sim}{x}^1[t] - \underset{\sim}{x}^2[t]\|$ $\|\underset{\sim}{x}^2[t]\|$ for all $\underset{\sim}{x}^1[t]$, $\underset{\sim}{x}^2[t]$ and t in the domain of the function, where

$\|\cdot\|$ denotes the Euclidean norm. It can be shown that if F has continuous partial derivatives, it satisfies the Lipschitz condition. However, the converse is not always true. Lipschitz, therefore, is not as strong as requiring F to have continuous partial derivatives. Taking into account (23), it is sufficient that the migration rates obey the Lipschitz condition. On the other hand the requirement $\underset{\sim}{\rho}^{\infty}[0] = \underset{\sim}{x}[0]$ simply says that, in order to be comparable, the stochastic and the corresponding deterministic system must have the same initial conditions. Clearly, with N large but finite, the variance of the probability distribution will be non-zero. Thus the observed population distribution at the beginning of time, $\underset{\sim}{\rho}^{N}[0]$, might be close, but not identical to the corresponding most probable state and this, in turn, may differ from $\lim_{N \to \infty} R_{\rho}[0]$. Hence, in general, $\underset{\sim}{\rho}^{N}[0] \neq \underset{\sim}{\rho}^{\infty}[0]$. It follows that if the conditions of proposition 3 hold, the evolution of the system as implied by the master equation of proposition 1 can only be approximated by the system of differential equations (25) over a finite, but arbitrarily long, future. This, of course, requires perfect knowledge of $\underset{\sim}{\Theta}[t]$. Therefore, in reality, such approximations will be useful only over short time-horizons.

The deterministic system of equation (25) consists of L differential equations. This is a drastic reduction in the number of equations from the original system of proposition 1. Although a general solution is still not known, the associated deterministic system allows in many cases to study the behaviour of solutions, including their stability characteristics.

It has become clear in our discussions so far that the key to the evolution of the system is the dependency of the migration rates q_{ji} on the state of the system and $\underset{\sim}{\Theta}$. There are some special cases to this general dependency that are particularly interesting. If, for example, q_{ji} depends only on $\underset{\sim}{\Theta}$ and not on the state of the system then it is well known that the probability distribution $P[\underset{\sim}{n};t]$ is approximately a time dependent multivariate normal. In this case the mean value state is the most probable state and the equation of proposition 2 is equivalent to (25). In this case then we argue, that $q_{ji}[\tau]$ at a point in time τ is an expression of the socioeconomic characteristics of the system $\underset{\sim}{\Theta}[\tau]$. For the observed approximations of q_{ji} the system (25) attains a unique equilibrium or steady state which reflects the values of q_{ji} and thus the tendency of the system at time τ.

We will come back to examine these issues along with an empirical application in more detail in chapter VII. Before we are able to do this, however, we need to examine two issues. The first is the functional dependency of the migration rates q_{ji} on the socioeconomic characteristics of the migratory system $\underset{\sim}{\Theta}$. This is done in the following chapter. The second issue is the introduction of actual data and the calibration of the model we introduce in chapter IV. This is done in chapters V and VI.

CHAPTER IV

DISCRETE CHOICE THEORY AND MIGRATION

The intention of this chapter is to provide an explicit link of the migration rates q_{ji} $i,j = 1,2, \ldots L$ to the conditions $\underset{\sim}{\Theta}[t]$ that affect the decision of individuals to migrate. The migration rates are connected to the probabilities to migrate p_{ji} $[t,\Delta t]$ as described in section III.4. Therefore it is sufficient to concentrate on the probabilities to migrate. Much of the material presented in this chapter has been published elsewhere and especially in the literature related to modelling travel demand.

The basis for this kind of modelling is individual choice theory. Within this theory the aggregate behaviour of a population is considered to be the result of individual decisions. This idea is consistent with the dynamics of the master equation and its associated deterministic system described in chapter III.

Various implementations of choice theory have been proposed. There are however some common features to all of them. Namely,

A decision maker - This is an individual or a group of individuals such as a household or a family.

Alternatives - A choice is made by the decision maker from a non-empty set of alternatives. It is usually conceived that there exists a universal set of alternatives which is defined by the environment of a decision maker. Each decision maker considers only a subset of the universal set which is termed a choice set.

A choice set can be either continuous or discrete. It is continuous when it is noncountably infinite. Intuitively in this case the choice set contains a continuum of an infinite number of alternatives. In microeconomic consumer theory, for example, continuous choice sets are considered. A choice set on the other hand is discrete when it consists of a finite number of alternatives. A choice theory that deals with discrete choice sets is termed discrete choice theory. This is what we employ in this thesis.

Attributes - An alternative is characterized by a number of attributes. A vector of values of all the attributes defines the attractiveness of an alternative.

Decision Rule - This is the internal mechanism that allows a decision maker to evaluate the alternatives in the choice set in order to arrive at a unique choice. There is a variety of decision rules that have been proposed. We are interested in a specific category of decision rule whereby a utility function maps a vector of attributes to a scalar, referred to as utility. The decision maker by using explicit or implicit trade offs is assumed to select the alternative that will make his utility maximum. Because of the term utility, this branch of choice theory is known as utility theory.

From the observer's point of view, different decision makers in the system will assign different utility values to a particular alternative. The observer, therefore, may treat this utility value as a random variable. If this is the case, then we are dealing with an even more specific branch of choice theory referred to as random utility theory.

In this chapter we utilize random utility theory to express the probabilities to migrate p_{ij}, $i, j = 1, 2, \ldots L$ within the system as a function of the conditions Θ that influence the decision of individuals to migrate. The functional form we arrive at is the nested logit model. It is worth noting here that there are two basic ways of deriving the nested logit model.

The first is to use the Generalized Extreme Value model, introduced by McFadden (1979), and show that the nested logit model is a special case of it. The second is to derive the nested logit step by step. This approach was followed recently by Ben-Akiva and

Lerman (1985) who derived the nested logit in the case of travel demand. In this chapter we tailor the latter approach to migration.

The remainder of this chapter is divided into seven sections. In section 1 the probabilities to migrate are conceptualized within the random utility framework. Since the Gumbel distribution is central in the derivation of any logit model, section 2 is devoted to this distribution and its properties. Section 3 describes the basic assumptions of the utilities. Each of the two levels of the nested logit are discussed in the following two sections 4 and 5. The connection between the two levels is discussed in section 6. Finally, estimation issues are discussed in section 7.

1. Conceptual Framework

The decision of a particular individual in j to migrate to i during Δt can be partitioned in two: (1) the decision to leave j during Δt and (2) once the decision of leaving j has been taken, the decision to select i among the L-1 remaining alternatives. Lowry's (1966) original finding that destination characteristics are more important than origin characteristics in the decision to move was confirmed by several other migration studies (Ledent and Liaw, 1985). Using the two-level structure that distinguishes clearly between origin and destination, accounts well for this persistent asymetry.

At the intraurban level the need to consider mobility as a multilevel process has been recognized explicitly. Speare et al. (1975), for example, have introduced a three-stage process consisting of the development of the desire to consider moving, the decision to stay or to move and the selection of an alternative location.

As we shall demonstrate in section 6 of this chapter, considering the two level as opposed to the one level framework results in conceptually more general and empirically more appealing models.

In order to apply the two-level choice framework, we factorize $p_{ij}\,[t,\Delta t]$ as in

<u>Assumption 1</u>: $\overline{p}_{ij}[t,\Delta t] = p_{i|j}[t]\; p._{j}[t,\Delta t]$ <u>for</u> $i=j$, $j = 1, \ldots L$;

where $p._{j}[t,\Delta t]$ is an individual's probability of leaving area j during the time interval $[t,t+\Delta t]$, and $p_{i|j}[t]$ is the conditional probability of choosing area i given that the individual has moved out of area j. For convenience, we call $p._{j}$ the <u>departure probability</u> and $p_{i|j}$ the <u>destination choice probability</u>. According to assumption 1, only the departure probability depends on the length of the time interval Δt.

The two level decision of an individual in j can now be modelled as in figure 1.

The first level corresponds to the question of whether to stay or leave j during Δt with probabilities $p_{jj}[t,\Delta t]$ and $p_{\cdot j}[t,\Delta t]$ respectively. Thus, the choice set of the individual in j at this level contains two alternatives. The second level corresponds to the actual choice of a place other than j, once the decision to leave j has been taken in the first level. The choice set $_j$ at this level contains the L-1 possible destinations from j.

The departure probability $p_{\cdot j}$ depends on factors that affect the decision of an individual to migrate and on the time interval Δt. We would like to separate the two effects mainly for practical reasons because in empirical applications the time interval Δt varies depending on the available data.

Assumption 2:

$$p_{\cdot j}[t,\Delta t] \;/\; p_{jj}[t,\Delta t] = h[t,\Delta t]\; p_{\cdot j}[t,l[t]] \;/\; p_{jj}[t,l[t]]$$

with h a continuous, increasing function of Δt such that

$$h[t,0] = 0,\; h[t,l[t]] = 1$$

and $l[t]$ an intrinsic time interval, explained in more detail below.

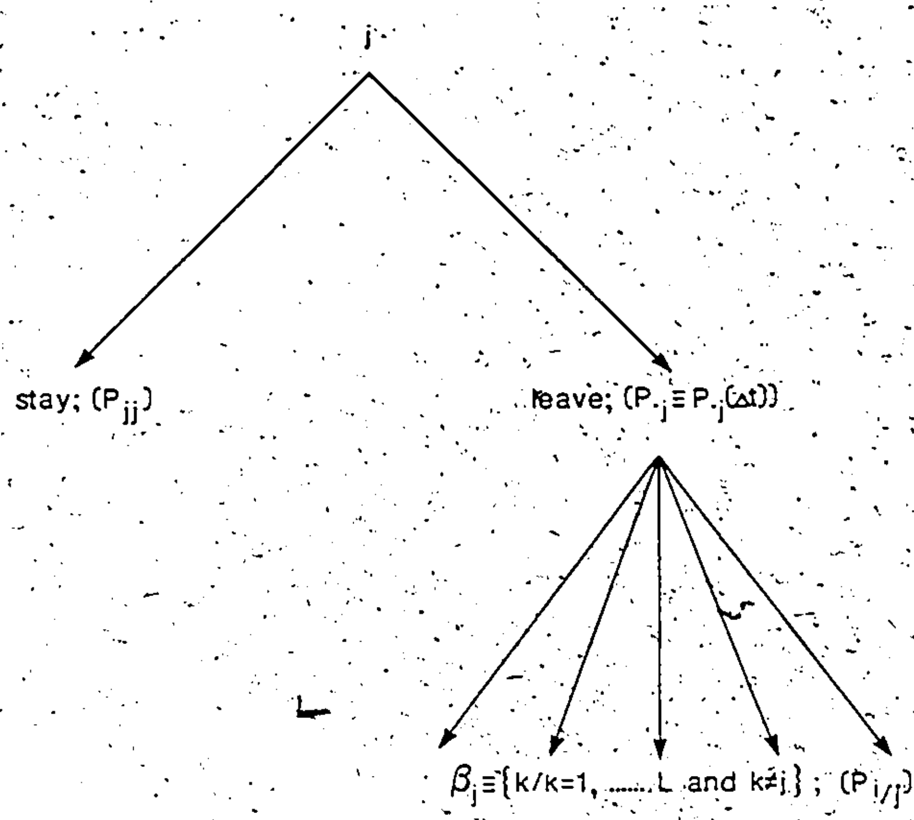

Figure 1: The Decision to Migrate

The ratio $p_{\cdot j}[t,\Delta t] \,/\, p_{jj}[t,\Delta t]$ represents the odds of outmigration from j within the time interval Δt. Because $p_{\cdot j}[t,\Delta t] + p_{jj}[t,\Delta t] = 1$, for any t and Δt, this assumption implies that:

$$p_{\cdot j}[t,0] = 0 \text{ and } p_{jj}[t,0] = 1$$

Thus, as the time interval Δt increases starting from zero, the probability of migrating out of j ($p_{\cdot j}$) increases and the probability of staying in j (p_{jj}) decreases. This means that the odds of outmigrating from j increase. This increase in the odds is captured by the function h.

From the way function h is defined it is geometrically evident that the value l exists for any time $t=\tau$. Consider, for example, figure 2. Function h will pass through the origin because $h[\tau,0] = 0$. Now, since h is continuous and increasing it is going to cut the line $h[\tau,\Delta t] = 1$ at some point, say A. This point then defines precisely the value l.

Because we wish to investigate the association of the probabilities to move with factors other than the time interval Δt, it is sufficient, at least for the time being, to concentrate on $p_{\cdot j}[t,l[t]]$. From now on, for notational simplicity, we shall omit the time index t. Thus we let $p_{\cdot j}[t,l[t]] \equiv p_{\cdot j}$, $p_{jj}[t,l[t]] \equiv \bar{p}_{jj}$, $p_{\cdot j}[t,\Delta t] \equiv p_{\cdot j}[\Delta t]$ and $p_{i|j}[t] \equiv p_{i|j}$. We consider every individual as a potential

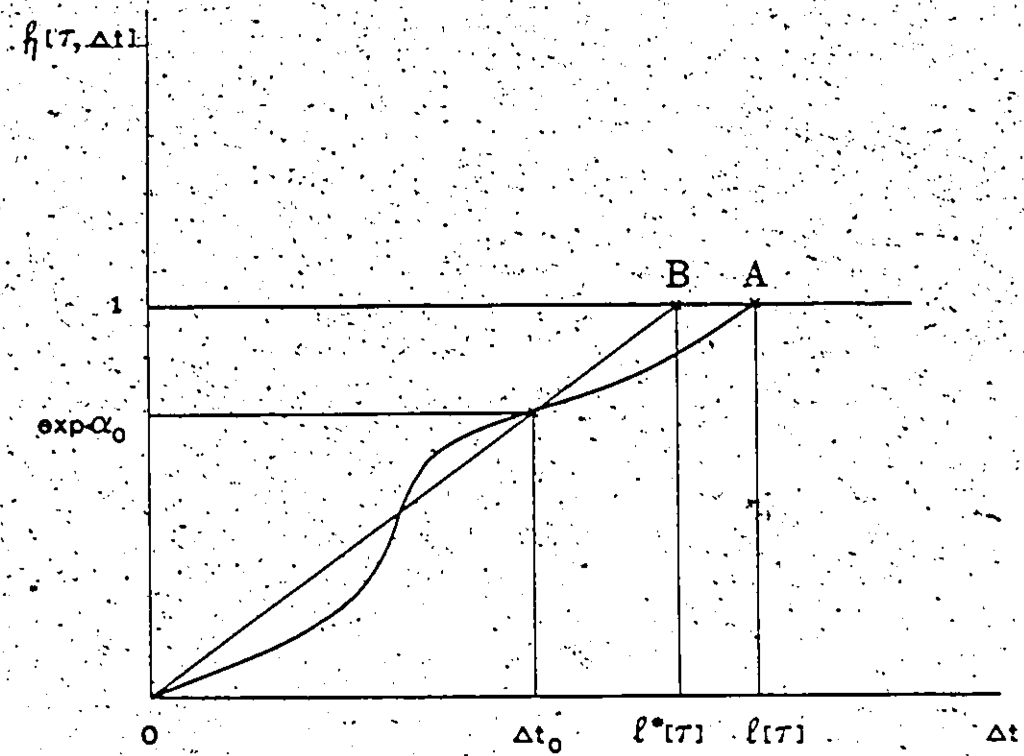

Figure 2: The Existence of $\ell[t]$

migrant. At every level of the decision making process an individual evaluates alternatives with respect to a vector of attributes. By mapping the vector of attributes to a scalar each alternative is associated with a utility level.

Assumption 3: Individuals aim to maximize utility

This assumption provides the behavioural mechanism that drives the evolution of the migratory system. Let u_{jj} be the utility of staying in j, and u_{kj} be the utility of moving from j to k.

Since the individual aims to maximize his utility, the probability of staying in j during the time interval l is the probability that the level of utility attained by staying in j is not lower than the maximum level of utility attainable by moving elsewhere:

$$p_{jj} = \text{Prob}\{u_{jj} \geq \max\{u_{kj}\};\ k \in \beta_j\} \tag{1}$$

Although these utilities vary through time, they are presumed fixed during the time interval l. On the other hand, the probability of leaving j in the same time interval is the probability that the maximum level of utility attained by leaving j is not lower than the level of utility attained by staying in j:

$$p_{\cdot j} = \text{Prob}\{\max\{u_{kj}\} \geq u_{jj};\ k \in \beta_j\} \tag{2}$$

If the decision taken at the first level is to leave, since the individual aims to maximize his utility, the probability of moving to i is the probability that the level of utility attained by moving to i from j is not lower than the maximum level of utility attainable by leaving j for elsewhere:

$$p_{i|j} = \text{Prob}\{u_{ij} \geq \max\{u_{kj}\}; \; i, k \in \beta_j\} \tag{3}$$

Since β_j does not include j,

$$\sum_{i \neq j} p_{i|j} = 1. \tag{4}$$

2. The Gumbel Distribution

Let utility u be composed by a deterministic part v and a stochastic part ϵ associated with the extent to which u can be predicted, i.e. let $u = v + \epsilon$. In addition to stochasticity arising from the structure of the model itself, the random component of the utility function can be justified in the following ways. Randomness may arise because individuals have different preferences expressed by different, deterministic utility functions.. Under these circumstances, different individuals confronted with the same situation and aiming to maximize utility may take different decisions. Consider an observer of the system. He knows perfectly well the current state of the system and the current value of all other factors in the system which determine utility. But, since he does not know the specific utility function of any individual, he cannot predict with certainty the behaviour of any individual sampled. If, however, he knows the distribution of ϵ, he can attach a probability to every alternative choice open to the individual sampled. This situation may be characterized as observer's uncertainty. On the other hand, individual's uncertainty arises whenever the information available to the individuals is imperfect. Then the random component of utility captures the difference between the actual utility v and the corresponding utility u as perceived by the individual sampled. Under these circumstances, the same individual may take different decisions. Clearly, in human systems, both sources of uncertainty compound each other to produce larger variance. The significant point is that, in any case, utility levels generated by the same situation are distributed around some "representative" utility level specific to that situation, where larger variance implies higher uncertainty and vice-versa. Under observer's uncertainty, "representative" utility levels may be generated by a "representative" individual. On the other hand, under individual's uncertainty, "representative" utility levels may be generated by the "true" utility function of an individual, which remains blurred due to missing information.

Because of the central limit theorem, it seems natural to associate the random component of utility with the normal distribution. This would lead us to the probit model. Since, however, the probit model is analytically and computationally difficult to handle, following McFadden (1974) and others, we shall assume that the random component of utilities associated with the decision to migrate has a Gumbel type II extreme value distribution. The Gumbel distribution serves as an approximation to the normal. Domencich and McFadden (1975, pp. 57-58) point out that the logit and probit models are, for practical purposes, statistically identical as far as their predictive capabilities are concerned. The same result was reported by Horowitz (1980).

A random variable ϵ with distribution function

$$F[\epsilon;x] \equiv \text{Prob}\{\epsilon \le x\} = \exp(-\exp(-\zeta(x-\xi))) \tag{5}$$

is Gumbel-distributed with parameters (ξ,ζ). Its density function is given by

$$f[\epsilon,x] = \zeta \exp(-\exp(-\zeta(x-\xi))) \exp(-\zeta(x-\xi)). \tag{6}$$

Since

$$\frac{df}{dx} = F\,\zeta^2 \exp(-\zeta(x-\xi))\,(\exp(-\zeta(x-\xi))-1), \tag{7}$$

the maximum of the density function occurs at $x=\xi$ (i.e. ξ is the mode). The density function is positively skewed. Hence the mean of ϵ is to be found on the right of the maximum of the density function. Indeed, the mean of ϵ is $\xi + .576/\zeta$ and its variance is $\pi^2/(6\zeta^2)$. We name ζ the scale parameter. The next two results are essential for modelling the probabilities to migrate.

Lemma 1: If ϵ is Gumbel-distributed with parameters (ξ,ζ) and α, β are scalar constants with $\beta > 0$ then $\alpha + \beta\epsilon$ is Gumbel-distributed with parameters $(\alpha + \beta\xi, \zeta/\beta)$.

Proof: From the definition of a Gumbel distribution (5),

$$F[\alpha + \beta\epsilon; x] \equiv \mathrm{Prob}\{\alpha + \beta\epsilon \leq x\}$$

$$= \mathrm{Prob}\{\epsilon \leq \frac{x - \alpha}{\beta}\}$$

$$= \exp(-\exp(-\zeta(\frac{x - \alpha}{\beta} - \xi)))$$

$$= \exp(-\exp(-\frac{\zeta}{\beta}(x - (\alpha + \beta\xi)))).|| \tag{8}$$

Lemma 2: If $\epsilon_1, \epsilon_2, \ldots, \epsilon_m$ are independent, Gumbel-distributed with parameters (ξ_1,ζ) and $(\xi_2,\zeta), \ldots, (\xi_M,\zeta)$ then $\max[\epsilon_1, \epsilon_2, \ldots, \epsilon_M]$ is Gumbel-distributed with parameters $(\frac{1}{\zeta} \ln \sum_{j=1}^{M} \exp(\zeta\xi_j), \zeta)$.

Proof: Since the random variables are independent,

$$\mathrm{Prob}\{\max\{\epsilon_1, \epsilon_2, \ldots, \epsilon_M\} \leq x\} = \prod_{j=1}^{M} \mathrm{Prob}\{\epsilon_j \leq x\}$$

$$= \exp(-\sum_{j=1}^{M} \exp(-\zeta(x - \xi_j)))$$

$$= \exp(-\exp(-\zeta x) \sum_{j=1}^{M} \exp(\zeta\xi_j)))$$

$$= \exp(-\exp(-\zeta x) \exp(\ln \sum_{j=1}^{M} \exp(\zeta\xi_j)))$$

$$= \exp(-\exp(-\zeta(x - \frac{1}{\zeta} \ln \sum_{j=1}^{M} \exp(\zeta\xi_j)))). \| \qquad (9)$$

Lemma 3: If ϵ_1 and ϵ_2 are independent, Gumbel-distributed with parameters (ξ_1,ζ) and (ξ_2,ζ) respectively, then

$$\mathrm{Prob}(\epsilon_1 > \epsilon_2) = \exp(\zeta\xi_1) / (\exp(\zeta\xi_1) + \exp(\zeta\xi_2)).$$

Proof: Let the joint density function of ϵ_1 and ϵ_2 be $f[\epsilon_1,\epsilon_2]$, and let the partial derivative of the joint distribution function with respect to ϵ_1 be $F_1[\epsilon_1,\epsilon_2]$. Then

$$\text{Prob}(\epsilon_1 > \epsilon_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\epsilon_1} f[\epsilon_1, \epsilon_2] d\epsilon_2 d\epsilon_1 = \int_{-\infty}^{\infty} F_1[\epsilon_1, \epsilon_1] d\epsilon_1$$

$$= \int_{-\infty}^{\infty} \zeta \exp(-\exp(-\zeta(\epsilon_1 - \xi_1)))\exp(-\zeta(\epsilon_1 - \xi_1))$$

$$\exp(-\exp(-\zeta(\epsilon_1 - \xi_2)))d\epsilon_1. \quad (10)$$

Let $\lambda \equiv \exp(-\zeta(\epsilon_1 - \xi_1))$ and $\beta \equiv \exp(\zeta(\xi_2 - \xi_1))$. Since $\ln\lambda = -\zeta(\epsilon_1 - \xi_1)$, it follows that $d\lambda/\lambda = -\zeta d\epsilon_1$. Upon substitution in (10) we obtain

$$\text{Prob}(\epsilon_1 > \epsilon_2) = \int_{\infty}^{0} -\exp(-\lambda)\lambda \exp(-\lambda\beta)(d\lambda/d\lambda)$$

$$= \int_{\infty}^{0} -\exp(-\lambda(1 + \beta))d\lambda$$

$$= (1/(1 + \beta))\exp(-\lambda(1 + \beta))\Big|_{\infty}^{0} = 1/(1 + \beta)$$

$$= 1/(1 + \exp(\zeta(\xi_2 - \xi_1))$$

$$= \exp(\zeta\xi_1)/(\exp(\zeta\xi_1) + \exp(\zeta\xi_2)) \| \quad (11)$$

3. Utilities as Random Variables

This section is based on Ben-Akiva and Lerman (1985).

We focus on utilities associated with the two-level decisions of individuals to migrate. The probability of staying in j, or of selecting i given that the decision to leave j has been taken, depends on the corresponding utilities u_{jj} and u_{ij}, $i \neq j$. In order to determine the structure of these utilities and, through them, the structure of the propensities to migrate, we adopt the following assumptions.

<u>Assumption 4</u>: $u_{ij} = v_{ij} + \epsilon_j^m + \epsilon_{ij}$; <u>either</u> $m = 0$ <u>and</u> $i = j$, <u>or</u> $m = 1$ <u>and</u> $i \neq j$; $i = 1, \ldots, L$,

where v_{ij} is the deterministic part of utility $\epsilon_j^m + \epsilon_{ij}$ is a random variable partitioned in two levels: firstly, the stochastic part of utility associated with staying (ϵ_j^0) or leaving (ϵ_j^1) during the time interval l; secondly, the stochastic part of utility associated with the choice of a place (once the decision at the first level has been taken) which can be either j (ϵ_{jj}) or other than j (ϵ_{ij}).

<u>Assumption 5</u>: <u>The random variable</u> ϵ_j^m <u>and</u> ϵ_{ij}, $m = 0, 1$ <u>and</u> $i, j = 1, \ldots, L$, <u>are independent</u>.

<u>Assumption 6</u>: <u>The random variables</u> ϵ_{ij} <u>are Gumbel-distributed with parameters</u> $(0, 1/\mu)$.

The last assumption to be made concerns the random variables ϵ_j^m.

<u>Assumption 7</u>: <u>For any Gumbel-distributed random variable</u> ϵ <u>with parameters</u> (ξ,ζ), <u>the random variable</u> $\epsilon_j^m + \epsilon$ <u>is Gumbel-distributed with parameters</u> $(\xi,\zeta\mu)$.

4. The Destination Choice Probabilities

<u>Proposition 1</u>: $p_{i|j} = \exp(v_{ij}/\mu) \,/ \sum_{k \neq j} \exp(v_{kj}/\mu)$.

<u>Proof</u>: By the definition (3) and assumption 4,

$$p_{i|j} = \mathrm{Prob}\{u_{ij} \geq \max\{u_{kj}\},\ i, k \in \mathcal{B}_j\}$$

$$= \mathrm{Prob}\{v_{ij} + \epsilon_j^m + \epsilon_{ij} \geq \max\{v_{kj} + \epsilon_j^m + \epsilon_{kj}\},\ i,k \in \mathcal{B}_j\}$$

$$= \mathrm{Prob}\{v_{ij} + \epsilon_{ij} \geq \max\{v_{kj} + \epsilon_{kj}\};\ i, k \in \mathcal{B}_j\}. \tag{12}$$

From Lemma 1, and assumptions 5 and 6, the random variables $v_{ij} + \epsilon_{ij}$, $i, j = 1, ..., L$, are independent and Gumbel-distributed with parameters $(v_{ij}, 1/\mu)$. On the other hand, from lemma 2, the random variables $\max\{v_{kj} + \epsilon_{kj}\}$, $k \in \Gamma_j$ are Gumbel-distributed with parameters $(\mu \ln \sum_{k \neq i,j} \exp(v_{kj}/\mu), 1/\mu)$. Application of lemma 3 to (12) yields

$$p_{i|j} = \exp(v_{ij}/\mu) / (\exp(v_{ij}/\mu) + \sum_{k \neq i,j} \exp(v_{kj}/\mu))$$

$$= \exp(v_{ij}/\mu) / \sum_{k \neq j} \exp(v_{kj}/\mu). \| \qquad (13)$$

We call (13) the destination choice model.

5. The Departure Probabilities

Proposition 2: $p_{\cdot j} = \exp(\mu I_j - v_{jj}) / (1 + \exp(\mu I_j - v_{jj}))$ where $I_j \equiv \ln(\sum_{k \neq j} \exp(v_{kj}/\mu)$.

Proof: By the definition 2 and assumption 4,

$$p_{\cdot j} = \mathrm{Prob}\{\max\{u_{kj}\} > u_{jj};\ k \in \mathcal{B}_j\}$$

$$= \mathrm{Prob}\{\max\{v_{kj} + \epsilon_j^1 + \epsilon_{kj}\} > v_{jj} + \epsilon_j^0 + \epsilon_{jj};\ k \in \mathcal{B}_j\}$$

$$= \text{Prob}\{\epsilon_j^1 + \max\{v_{kj} + \epsilon_{kj}\} > \epsilon_j^0 + (v_{jj} + \epsilon_{jj}); \quad k \in \beta_j\} \tag{14}$$

From lemmata 1 and 2, and assumption 6 $\max\{v_{kj} + \epsilon_{kj}\}$, $k \in \beta_j$ is Gumbel-distributed with parameters $(\mu ln \sum_{k \neq j} \exp(v_{kj}/\mu), \frac{1}{\mu}) \equiv (\mu I_j, \frac{1}{\mu})$.

Because of assumption 7 the random variable $\epsilon_j^1 + \max\{v_{kj} + \epsilon_{kj}\}$, $k \in \beta_j$ is Gumbel-distributed with parameters $(\mu I j, 1)$. Similarly $\epsilon_j^0 + (v_{jj} + \epsilon_{jj})$ is Gumbel-distributed with parameters $(v_{jj}, 1)$. Using lemma 3 on (14) we have:

$$p_{\cdot j} = \exp(\mu I_j) \,/\, (\exp(\mu I_j) + \exp(v_{jj}))$$

$$= \exp(\mu I_j - v_{jj}) \,/\, (1 + \exp(\mu I_j - v_{jj})). \| \tag{15}$$

Proposition 2 provides an expression for the probability of departure during the unit time interval. Combining this result with assumption 2 we can obtain an expression for the probability of departure during an interval Δt.

$$\frac{p_{\cdot j}[\Delta t]}{1 - p_{\cdot j}[\Delta t]} = h\,[\Delta t]\,\frac{p_{\cdot j}}{1 - p_{\cdot j}}$$

$$= h[\Delta t] \cdot \exp(\mu I_j - v_{jj})$$

$$= \exp(\alpha_0 + \mu I_j - v_{jj}) \tag{16}$$

where $\alpha_0 \equiv \ln h[\Delta t]$. Solving (16) for $p_{\cdot j}[\Delta t]$, we have:

$$p_{\cdot j}[\Delta t] = \exp(\alpha_0 + \mu I_j - v_{jj}) \,/\, (1 + \exp(\alpha_0 + \mu I_j - v_{jj})). \tag{17}$$

We call I_j the <u>inclusive value</u> and (17) the <u>departure model</u>.

Referring back to the motivation for introducing assumption 2, it now becomes clear that the effect of the time interval on the departure probability is captured by the parameter α_0.

In the case of $\Delta t = l$:

$$\alpha_0 = \ln h[\Delta t] = \ln h[l] = \ln 1 = 0$$

If the time interval is smaller (larger) than l, then α_0 will be smaller (larger) than zero.

The exact value of l is not known because the exact functional form of h is unknown. As we shall see in section 7, given proper data, the value of α_0 can be estimated statistically from model (17), for a value of $\Delta t = \Delta t_0$. Referring now to figure 2 we can see that if h is considered linear an estimate l^* of l may be obtained.

6. The Parameter μ

Consider an individual in area j and two possible destinations i and k. The utilities u_{ij} and u_{kj} of leaving j for i and k respectively are not independent random variables. According to assumption 4 these utilities can be decomposed as follows:

$$u_{ij} = v_{ij} + \epsilon_j^1 + \epsilon_{ij}$$

$$u_{kj} = v_{kj} + \epsilon_j^1 + \epsilon_{kj}$$

Both the utilities U_{ij} and U_{kj} share ϵ_j^1, which is the uncertainty involved in the decision to leave j.

The dependence of u_{ij} and u_{kj} is given by their covariance.

$$\begin{aligned}\mathrm{Cov}(u_{ij}, u_{kj}) &= \mathrm{Cov}(\epsilon_j^1 + \epsilon_{ij}, \epsilon_j^1 + \epsilon_{kj}) \\ &= E((\epsilon_j^1 + \epsilon_{ij} - E(\epsilon_j^1) - E(\epsilon_{ij}))\ (\epsilon_j^1 + \epsilon_{kj} - E(\epsilon_j^1) - E(\epsilon_{kj})))\end{aligned}$$

$$= E((\epsilon_j^1 - E(\epsilon_j^1))^2) = \mathrm{Var}(\epsilon_j^1). \tag{18}$$

Thus the dependence of u_{ij} and u_{kj} is entirely due to the uncertainty in the decision to leave j. When there is no uncertainty in the first level then the variance of ϵ_j^1 is zero which implies that u_{ij} and u_{kj} are independent. This is not surprising since the variability in this case is due entirely to ϵ_{ij} and ϵ_{kj} which by assumption 5 are independent.

Taking now into account assumptions 6 and 7 we can express the dependence of u_{ij} and u_{kj} in terms of the parameter μ.

<u>Proposition 3:</u> $\mathrm{Corr}(u_{ij}, u_{kj}) = 1 - \mu^2$.

<u>Proof:</u> Since the variance of a Gumbel-distributed random variable with parameters (ξ, ζ) is given by $\pi^2 / (6\zeta^2)$, assumptions 6 and 7 imply

$$\mu^2 = \frac{\mathrm{Var}(\epsilon_{ij})}{\mathrm{Var}(\epsilon_j^1 + \epsilon_{ij})}. \tag{19}$$

Moreover, since ϵ_j^1 and ϵ_{ij} are independent, we may write

$$\mu^2 = \frac{\mathrm{Var}(\epsilon_{ij})}{\mathrm{Var}(\epsilon_j^1) + \mathrm{Var}(\epsilon_{ij})}$$

$$= 1 - \frac{\mathrm{Var}(\epsilon_j^1)}{\mathrm{Var}(\epsilon_j^1) + \mathrm{Var}(\epsilon_{ij})}$$

$$= 1 - \frac{\mathrm{Var}(\epsilon_j^1)}{\mathrm{Var}(\epsilon_j^1 + \epsilon_{ij})} \tag{20}$$

Notice that

$$\mathrm{Var}(\epsilon_j^1 + \epsilon_{ij}) = \mathrm{Var}(\epsilon_j^1 + \epsilon_{ij})^{1/2}\ \mathrm{Var}(\epsilon_j^1 + \epsilon_{ij})^{1/2}$$

$$= \mathrm{Var}(\epsilon_j^1 + \epsilon_{ij})^{1/2}\ \mathrm{Var}(\epsilon_j^1 + \epsilon_{kj})^{1/2}$$

$$= \mathrm{Var}(u_{ij})^{1/2}\ \mathrm{Var}(u_{kj})^{1/2}, \tag{21}$$

using assumptions 4 and 7. It remains to substitute (18) and (21) in (20).||

Since $\mathrm{Var}(\epsilon_{ij}) \le \mathrm{Var}(\epsilon_j^1 + \epsilon_{ij})$, from (19) we have:

$$0 \le \mu^2 \le 1$$

The case $\mu^2 = 0$ occurs when $Var(\epsilon_{ij}) = 0$. In this case the entire variability in u_{ij} is due to ϵ_j^1. The same arguments apply for u_{kj}. Because of proposition 3 for any two destinations i and k the utilities u_{ij} and u_{kj} are perfectly correlated. This is because variability in u_{ij} and u_{kj} is due to ϵ_j^1 only, which is common in both of them. When $o < \mu^2 < 1$, stochasticity enters in the decision to stay or leave j and in the decision to choose a place (once the decision to stay or leave has been taken).

Finally, when $\mu^2 = 1$ then $Var(\epsilon_j^1) = 0$. In this case u_{ij} and u_{kj} are independent and the nested logit model becomes equivalent to the multinomial logit model with all t alternatives in the choice set.

On conceptual grounds therefore the nested logit model is more general than the multinomial logit model. In practical terms the nested logit model provides the means of avoiding the <u>independence from irrelevant alternatives</u> property for which the unilevel logit model has been extensively criticized. In the multinomial logit all the alternatives are assumed to belong to one choice set. The independence from irrelevant alternatives property stems from the requirement that the random parts of the utilities associated with the alternatives be independent. Usage of the model in practice may result in conflict whenever the utilities are not independent due to the nature of the problem or due to missing variables in the specification of utilities.

Several examples have been used in the literature to illustrate the conflict. The example used by Liaw, Kanaroglou and Moffett (1986) demonstrates the point in the migration context:

Suppose that an observer wishes to study migrants facing three possible destinations: two metropolitan areas (MA1 and MA2) and a non-metropolitan area (NA). Individual migrants in evaluating the alternatives will use a vector of attributes. From the observer's point of view some of these attributes are measurable, others are not. Those that are not measurable constitute the stochastic part of the utility. But even from those that are measurable some are missing due to lack of proper data. The more the missing variables the more difficult it is to distinguish between alternatives and the more correlated the random parts of the utilities become.

Suppose an important variable, say congestion, is missing in our example. This variable, if present tends to make the two metropolitan areas similarly less attractive than the non-metropolitan area. Thus the odds of the conditional probabilities Prob(MA1/MA1,MA2,NA) / Prob(NA/MA1,MA2,NA) and Prob(MA1/MA1,NA) / Prob(NA/MA1,NA) need not be equal. However, the multinomial logit model with all three alternatives at the same level implies the equality of the two odds. Precisely this constitutes the independence from irrelevant alternatives property of the multinomial logit model.

The problem is due to the fact that the unilevel multinomial logit model is derived from the assumption that the stochastic parts of the utilities are independent. But, as explained above, the missing variable forces the stochastic parts to be correlated. In this example, a two level choice where the migrant first chooses between a metropolitan and a non-metropolitan area and then having made this choice, decides on a particular metropolitan area would alleviate the above problem.

There are some final points to be made about the value of μ and its role in the departure probability $p._j$. Notice in the expression of proposition 2 that the probability of leaving j is determined both by characteristics associated with the destinations (I_j) and characteristics associated with the origin (v_{jj}). Observe that μ can not be negative, because $1/\mu$ is the scale parameter of ϵ_{ij}. When μ=0, the utilities u_{ij}, i = 1, 2, ..., L and $i \neq j$ are perfectly correlated. In this case μI_j is equivalent to $\mathrm{Max}\{v_{kj}\}$; $k \epsilon \mathcal{B}_j$ in the expression of $p._j$. At the other extreme, when μ=1 the utilities u_{ij}, i = 1, 2, ..., L, i=j are perfectly uncorrelated and the inclusive value fully enters the specification of

$p_{\cdot j}$. Normally, however, the utilities of the alternatives in the lower level are neither distinct nor perfectly correlated. This case corresponds to $0 < \mu < 1$.

Finally observing that μI_j corresponds to the deterministic part of $\max\{v_{kj} + \epsilon_{kj}\}$, $k \in \mathcal{B}_j$ we can give the following interpretation to the nested logit model. Individuals living in a certain area j are constantly evaluating alternative areas. Out of this evaluation they come up with an alternative as the one that appeals best to them. This alternative is compared to the current location j. If the alternative is sufficiently better than the current location j then a decision is taken to move. In the nested logit formulation, evaluation of areas is done by associating a utility level with each area which is a function of the attributes of the area. For an individual in j then μI_j corresponds to the utility of the best alternative area. This is compared to the utility v_{jj} of staying in j. The probability of moving then out of j is a function of the difference of the two utilities.

As a behavioural framework the nested logit model allows us to deal with the decision to move and the locational decision simultaneously. It also includes some of the concepts discussed in section II.3. Thus the difference between μI_j and v_{jj} can be thought of as a stress level on which the probability to move is dependent.

7. Calibration of the Nested Logit

The nested logit model represented by the equations (13) and (17) links the migration probabilities $p_{\cdot j}[\Delta t]$ and $p_{i|j}$ to the deterministic parts of the utilities expressed by v_{jj} and v_{ij}. The question we address in this section is how this model can be operationalized. In other words, we seek to express the utilities v_{jj} and v_{ij} in terms of independent variables and parameters to be estimated.

As observers we assume that potential migrants by using explicit or implicit trade offs and compensatory offsets select the alternative that maximizes their utility. But what are the factors that the average potential migrant is considering and what is the functional form of the utility that will reflect adequately the way he blends these factors in order to choose an alternative?

Although the independent variables to be used would differ from one type of migration to another, one may, in general, classify them in two broad categories:

(1) Relational variables - They refer to both the origin and the destination. These are typically measures of physical distance or measures of social distance such as the ethnic dissimilarity between two areas.

(2) Destination specific variables - They reflect the destination characteristics. We can think of two types of variables of this kind. Firstly, those that measure the physical attributes of the destination such as climate, type of soil, level of pollution, and so on. Secondly, those that measure the social and economic attributes of the destination such as employment opportunities, population density, etc.

We would like also to point out at this stage, that since the utilities are viewed by the modeller as random variables over the decision makers, independent variables related to the characteristics of the decision makers are often included in the deterministic part of the utility function.

To define the functional form of the utility we must meet two criteria that are usually conflicting:

(i) The attributes must affect the utility in a realistic way;

(ii) Estimation of the parameters should be computationally tractable.

The most widely accepted functional form is the linear in parameters. One should note that linearity in parameters does not necessarily imply linearity in the independent variables used. Real transformations of the variables can be used as long as the linearity in parameters is preserved.

The following two assumptions express v_{ij} and v_{jj} in terms of independent variables in a linear-in-parameters form.

Assumption 4: For $i \neq j$, $(1/\mu)\ v_{ij} = \underset{\sim}{b}'\underset{\sim}{Y}_{ij}$ where $\underset{\sim}{Y}_{ij}$ is a vector of s explanatory variables, and $\underset{\sim}{b}'$ is the corresponding vector of unknown parameters divided by μ.

Assumption 9: $v_{jj} = \underset{\sim}{\alpha}'\underset{\sim}{W}_j$ where $\underset{\sim}{W}_j$ is a vector (r-2) explanatory variables and $\underset{\sim}{\alpha}'$ is the corresponding vector of unknown parameters.

Substituting v_{ij} and v_{jj} given by assumptions 8 and 9 into the destination choice and departure models of propositions 1 and 2 yields

$$p_{i|j} = \exp(\underset{\sim}{b}'\underset{\sim}{Y}_{ij}) \;/\; (\sum_{k \neq j} \exp(\underset{\sim}{b}'\underset{\sim}{Y}_{kj}) \tag{22}$$

and

$$p_{\cdot j}[\Delta t] = \exp(\alpha_0 - \underset{\sim}{\alpha}'\underset{\sim}{W}_j + \mu I_j) \;/\; (1 + \exp(\alpha_0 - \underset{\sim}{\alpha}'\underset{\sim}{W}_j + \mu I_j))$$

$$= \exp(\underset{\sim}{a}'\underset{\sim}{X}_j) \;/\; (1 + \exp(\underset{\sim}{a}'\underset{\sim}{X}_j)) \tag{23}$$

where

$$I_j \equiv \ln \sum_{k \neq j} \exp(\underset{\sim}{b}' \underset{\sim}{Y}_{kj}) \tag{24}$$

$$\underset{\sim}{a}' \equiv (\alpha_0, -\alpha_1, -\alpha_2, \ldots, \mu) \tag{25}$$

$$\underset{\sim}{X}_j' \equiv (1, W_{1j}, W_{2j}, \ldots, I_j). \tag{26}$$

Note that the dimensions of the vectors of independent variables $\underset{\sim}{Y}_{ij}$ and $\underset{\sim}{X}_j$ are s and r, respectively.

The model in the form presented by equations (22) to (26) provide the basis for the empirical estimation of the parameters in $\underset{\sim}{a}$ and $\underset{\sim}{b}$. The usual methodology used for such a statistical estimation is <u>maximum likelihood</u>. The relevance of the maximum likelihood methodology in the case of the logit model, together with the related asymptotic statistical theory, was established by McFadden (1974). Jennrich and Moore (1975) have shown that a maximum likelihood estimation problem for the logit can be translated into a weighted non-linear least squares problem which can be handled by several commercially available software packages. The particular methodology used here is based on Liaw and Bartels (1982) who applied the previous ideas in the case of the BMDP package.

Both macro- and micro-data bases are relevant to the analysis of migratory systems. In the former case the unit of observation is a geographical area such as a census tract, a metropolitan region or a province. The information usually available is the number of mobile units (individuals, households) having moved between geographical areas of the system within a specified time-period, together with socio-economic characteristics of the geographical areas. The problem with macro-data could be inadequate degrees of freedom affecting the reliability of parameter estimates. On the other hand, with micro-data, observations concerning movement and socio-economic characteristics pertain to the mobile unit itself. Since the sample is usually large, degrees of freedom do not present a problem. However, numerical problems due to the large dimension of the resulting matrices, together with problems of extensive data noise, might appear in the case of micro-data. This case has been treated by McFadden (1976).

Although the same methodology can be used with micro-data, here we emphasize the case of macro-data bases which is more often encountered in migration research. Let $\underset{\sim}{n} = (n_1, n_2, \ldots, n_L)$ be the observed state of the system at a particular moment τ. Also, let n_j^1 represent the observed number of migrants leaving area j during Δt, and n_{ij} be the observed number of migrants who having left j have chosen i as their destination. It follows that:

$$\sum_{\substack{i=1 \\ i \neq j}}^{L} n_{ij} = n_j^1 ; \; j = 1, 2, \ldots L$$

The ratio (n_{ij}/n_j^1) represents the proportion of outmigrations from j that select i as their destination and constitutes a particular observation of a random variable whose asymptotic expected value is $p_{i|j}$ in the destination choice model (22). The number of observations in this case is equal to the number of elements in the matrix (n_{ij}/n_j^1) minus the diagonal elements, which is $(L^2 - L)$.

Similarly, the ratio (n_j^1/n_j) represents the proportion of the population of area j that outmigrates during the time interval Δt. This ratio constitutes an observation of a random variable whose asymptocic expected value is $p_{\cdot j}[\Delta t]$ of the departure model (23). Obviously the number of observations is equal to the number of areas L in the system.

The full mathematical justification for the validity of the estimation methodology and the algorithms of the BMDP package used is given in Kanaroglou, Liaw and Papageorgiou (1986b). Here we just give an outline of the general procedure on a conceptual level.

The estimation proceeds in two steps. Firstly, an estimate $\underset{\sim}{b}^*$

of $\underset{\sim}{b}$ in equation (22) is obtained using the destination choice model. Once this is available, an estimate $I_j^{\star}$ of I_j is determined through (24). Secondly, $I_j^{\star}$ is used in the departure model to provide $\underset{\sim}{\alpha}^{\star}$ of $\underset{\sim}{\alpha}$ in (23). The last element of this vector is $\mu^{\star}$, which is then multiplied by $\underset{\sim}{b}^{\star}$ to yield the estimated coefficients of $\underset{\sim}{Y}_{ij}$.

Thus we obtain estimates for an important quantity I_j and two equally important parameters α_o and μ. As we argued in section 6 of this chapter μI_j represents a measure of the utility of the best alternative (alternative associated with maximum utility) other than j, which is compared in the higher level with v_{jj}, the utility of staying. Also, μ is an index of intercorrelation of the utilities in the lower level. Finally, α_o absorbes the effects of variations in the time interval Δt on the probability of moving.

CHAPTER V

INTRAMETROPOLITAN MOBILITY IN TORONTO: 1966-1976 — A DESCRIPTIVE ANALYSIS

Chapters III and IV presented us with a logical theory that materialized in a set of mathematical models. We argued that these models can be used to analyse migration patterns at the interregional, interurban or intraurban level. We would like now to apply these ideas to analyse intrametropolitan migration patterns in the Toronto Census Metropolitan Area for the decade 1966-1976.

This is the first of a series of three chapters that could be classified as empirical. Its main purpose is to introduce the data set and perform some preliminary analysis of descriptive nature. It, therefore, represents an attempt to get a "feel" of what the data are telling us before we engage into more elaborate statistical analysis.

Another goal of this chapter is to provide empirical support to the argument made in section 1 of chapter III, that given sufficient data we can account for all the sources of change in a migratory system. This way we can practically consider the migratory system closed.

There are seven sections in this chapter. Section 1 provides a basic familiarization with the data set and the partitioning of the Toronto CMA into zones. The basic mobile unit in this data set is the household. Gross changes in the population and the number of households by zone in the CMA throughout the study period are discussed in section 2. From this discussion the need for a systematic accounting scheme of change in the number of households from one point in time to the next is identified. Such an accounting scheme is introduced in the form of two identities, explained in detail in section 3.

One of the components of change, household formation from one time period to the next, appears impossible to quantify through the use of data, since the relevant information is not collected by the Canadian Census. Following Miron (1983) then in section 4 we make use of a model of household formation that at least in some way compensates for the deficiency in our data base.

After the mobility figures have been cleared from their household formation components, the mobility trends for various demographic groups are examined in sections 5 and 6. Finally, in section 7, entropy as a measure of dispersion to various destinations of migrants from a specific origin is introduced and used in order to determine strong migration exchanges between zones in the system.

1. The Data Set

The data set used in this study was kindly provided by Dr. J. Miron of the University of Toronto. It is a Statistics Canada special tabulation of household head migrations in the Toronto CMA for the intercensal time periods 1966-71 and 1971-76.

Household head moves in the data set are disagregated by household size, and by the age, sex and marital status of the household head. Information also is provided about the migrant heads entering the areas from places in Ontario other than the CMA, from Canada outside Ontario and from outside Canada.

At this point it is useful to define what is meant by a household and the way the CMA is partitioned into areas. For census purposes a household is defined as:

> "... a person or a group of persons occupying one dwelling. It usually consists of a family group with or without lodgers, employees, etc. However, it may consist of two or more families sharing a dwelling, of a group of unrelated persons or of one person living alone. Every person is a member of some household and there is a one-to-one relationship between households and occupied dwellings except in the case of certain special households,

> such as those of military and diplomatic personnel stationed overseas, from which no household information was collected." (Statistics Canada 1973b)

Thus, the number of households corresponds approximately to the number of occupied dwelling units. For census purposes again a dwelling unit is considered to be a set of living quarters with an entrance not through someone else's living quarters. The number of occupied dwellings, therefore, tends to understate the real number of households.

A household is usually identified by a household head. According to a definition given by Statistics Canada one member of a household is designated to be the head of the household, so that there is a one-to-one correspondence between households and household heads. Unfortunately, the Statistics Canada definition has changed over time. The 1971 definition designates the husband as the household head if both husband and wife are present, the parent if living with unmarried children, or any member of a group sharing a dwelling equally (Statistics Canada, 1972). In 1976 this definition was modified so that either the husband or the wife in a household containing both can be designated as head. For consistency, in the data set used in this study the 1971 definition is followed.

Within the data set the Toronto CMA is conceptualized as being divided into thirteen areas. Each area is a municipality or an aggregation of less populous municipalities at the outskirts of the CMA. The exact zonal system adopted is illustrated in figure 1 and described in detail in table 1. Within the boundaries of the Toronto CMA there were thirty-two municipalities in 1966, twenty-nine in 1971 and twenty in 1976. Note that only thirty municipalities appear in table 1 for 1966. The remaining two, Milton and Stouffville, are aggregated with other municipalities introduced into the CMA in 1971, to form the zones of Halton and Outer York. Note also that in 1976 we have only 12 areas because Halton was not considered part of the CMA at this time.

Table 2 shows the population and household distribution in the perceived areas at three points in time relevant to the study period. It is important to note that beyond the gross CMA adjustments shown in table 1 there are minor adjustments in the boundaries of the areas, between censuses. Such adjustments certainly affect the population and the number of households in an area. To give an idea about the magnitude of such an effect, the 1971 population with the 1976 area boundary is shown in parentheses in table 2. Thus, six zones have changed in areal size. The most substantial change shown is that of Outer York. This is because a whole municipality, East

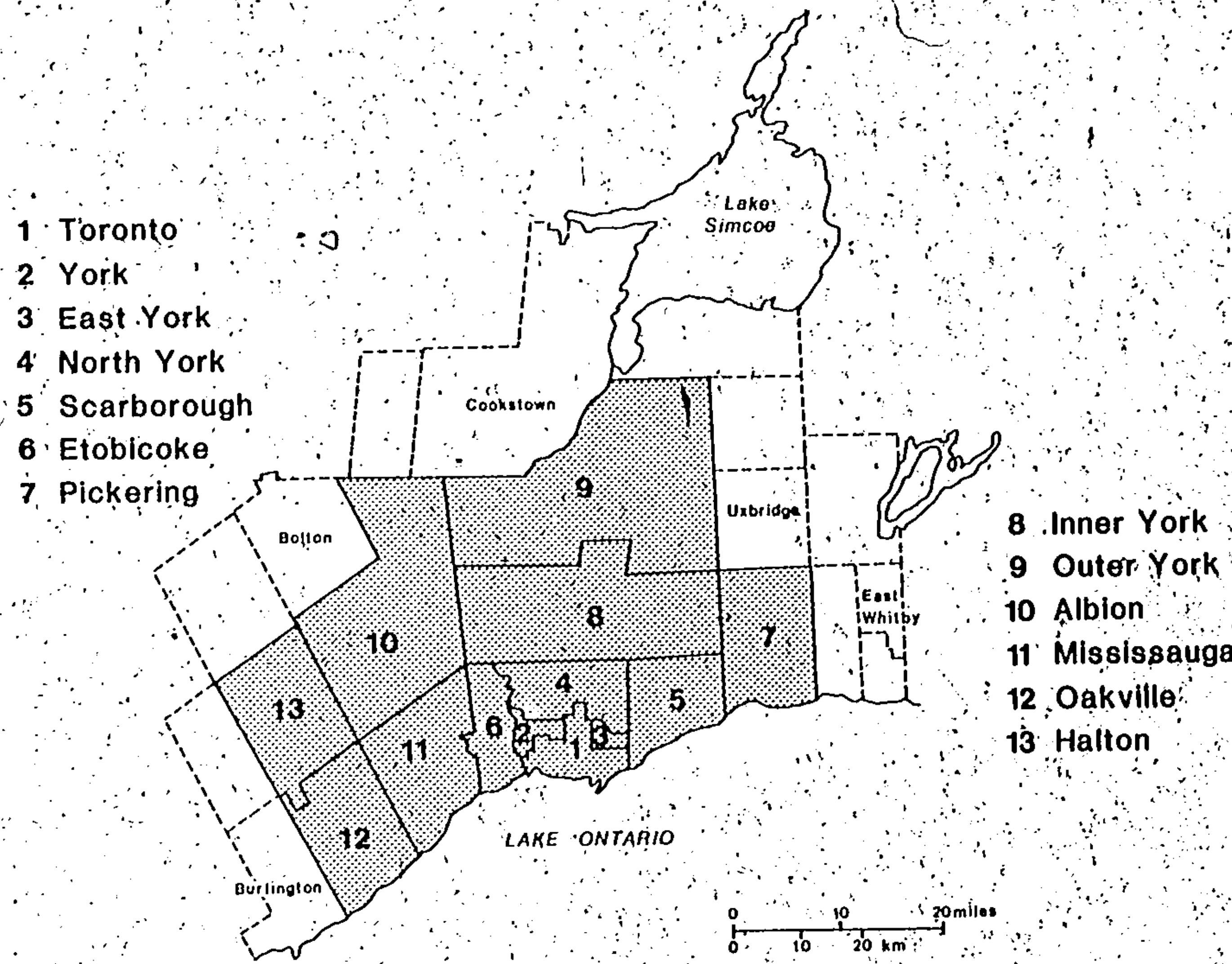

Figure 1. Toronto CMA and Study Zones

AREA	1966	1971	1976
1	City of Toronto Swansea Vl. Forest Hill Vl.	City of Toronto	City of Toronto
2	York Township Weston	York Borough	York Borough
3	York East Twp. Leaside	East York	East York
4	York North Twp.	North York	North York
5	Scarborough Twp.	Scarborough	Scarborough
6	Etobicoke Twp. Mimico New Toronto Long Branch Vl.	Etobicoke	Etobicoke
7	Pickering Twp. Ajax Pickering Vl.	Pickering Twp. Ajax Pickering Vl.	Pickering Twp. Ajax
8	Markham Twp. Richmond Hill Twp. Vaughan Twp. Woodbridge Vl. Markham Vl.	Markham Richmond Hill Vaughan	Markham Richmond Hill Vaughan
9	Not Part of CMA	Aurora King Twp. Newmarket Whitchurch- Stouffville	Aurora King Twp. Newmarket Whitchurch- Stouffville E. Gwillinbury
10	Albion Twp. Chingacousy Twp. Toronto Gore Twp. Brampton Bolton Vl.	Albion Twp. Chingacousy Twp. Toronto Gore Twp. Brampton Bolton Vl.	Caledon Brampton
11	Toronto Twp. Port Credit Streetsville	Toronto Twp. Port Credit Streetsville	Mississauga
12	Oakville	Oakville	Oakville
13	Not Part of CMA	Esquesing Twp. Acton Georgetown Milton	Not Part of CMA

1 - Toronto
2 - York
3 - E. York
4 - North York
5 - Scarborough
6 - Etobicoke
7 - Pickering
8 - Inner York
9 - Outer York
10 - Albion
11 - Mississauga
12 - Oakville
13 - Halton

Table 1: Municipalities Included in the Area

	1966		1971		1976	
Area	Population	Households	Population	Households	Population	Households
Toronto	697422	190137	712785	224695	633318	230395
York	145721	42264	147305	46525	141367	49055
East York	95450	32214	104785	38270	106950	42430
North York	399534	107759	504150	146155	558398	176090
Scarborough	278377	70630	334310 (338022)	91535	387149	118060
Etobicoke	265167	73732	282690	82290	297109	96935
Pickering	39254	9362	46785 (43074)	11750	48674	13570
Inner York	66423	16827	84942	22510	108704	31010
Outer York	47611	—	56615 (56040)	15170	76593	21960
Albion	59729	14900	81260 (87832)	21250	125893	35445
Mississauga	107851	27431	172355 (172042)	46930	250017	74155
Oakville	52678	13452	61480 (58440)	16655	68950	20425
Halton	30559	—	38515	10365	—	—
Total	2285796	598708	2627977	774100	2803122	909530

Source: Statistics Canada (1973a)
Statistics Canada (1968)
Statistics Canada (1978)

Table 2: Population and Households by Area

Gwillinbury, with a population of nearly 10,000 in 1971 was included in the CMA in 1976 and in this data set it is considered as part of Outer York. For the rest of the areas, the boundary changes induce population differences ranging, on a percentage basis, from minor (Mississauga, Scarborough) to substantial (Albion, Pickering).

As mentioned before, the mobile unit in this study is the household. The distribution of households shown in table 2 has been checked against the distribution in the special tabulation data used in this study. With these problems in mind we now turn to an analysis of the data in table 2. This analysis will give us an idea about the overall changes that occured in the CMA during the study period in terms of the household and population distributions.

2. Population and households: Gross changes

The population of an area changes over time due to natural growth (births minus deaths) and net migration (in-migration minus out-migration). Similarly, the number of households in an area changes because of household formation and dissolution and because of in- and out-migration of households. The process of household formation, however, is a significantly more complicated process than a simple birth. A household may be formed because of a marriage, a divorce, or simply because a youngster decided to leave the parent home and live alone. Also, as the household definition in the previous section implies and as Miron (1979) points out, a new household may appear because of minor renovations in a two-family dwelling so that each family gets its own entrance to its living quarters.

It is therefore expected that the pattern of population change will most likely be different from that of household change. From table 3 we see that the CMA population increased by 15% in the 1966-71 period and by only 8.2% in the period 1971-76 (excluding Halton in the calculations of the latter case). The growth of households, however, is significantly higher (25% and 19%) for both the time periods. Four reasons have been forwarded by Miron (1979) that help to explain why the number of households grew relatively faster than the population in the Toronto CMA. The first is related to the natural process of the maturing baby boom cohort which produces a growing number of families. By contrast, the other three reasons are the result of changing attitudes and preferences. The increasing number of one-family households reflects a decreasing preference of families to share accommodation with another family. In addition an increase in the one-person households has been observed which was made possible by the availability of bachelor and one-

	Increase (%) 1966-71		Increase (%) 1971-76	
AREA	Population	Households	Population	Households
Toronto	2.20	18.10	-11.15	2.54
York	1.09	10.08	-4.03	5.44
East York	9.78	18.80	2.07	10.87
North York	26.18	35.63	10.76	20.48
Scarborough	20.09	29.60	15.81	28.98
Etobicoke	6.60	11.61	5.10	17.80
Pickering	19.19	25.51	4.03	15.49
Inner York	27.88	33.77	27.97	37.76
Outer York	18.91	--	35.29	44.76
Albion	36.05	42.62	54.93	66.80
Mississauga	59.81	71.08	45.06	58.01
Oakville	16.71	23.81	12.15	22.64
Halton	26.03	--	--	--
Overall Increase	14.97	25.03	8.25	19.09

Source: see text

Table 3: Population and Households Change by Area

bedroom apartments. Finally, the study period is characterized by a decrease in the popularity of marriage and an increase in the divorce rate.

The last three points imply that the size of the households became smaller over time. This point is indeed true for the Toronto CMA and it is substantiated in table 4. The figures in this table are calculated from table 2 by dividing the population by the number of households. The average size of the household for the CMA has decreased from 3.69 persons per household in 1966 to 3.08 persons per household in 1976.

In terms of the household size distribution in the CMA, two trends became clear from table 4. Firstly, the average household size decreased over time in <u>every</u> area of the CMA. The largest decrease was in Toronto City from 3.67 to 2.75, a drop of .92 persons per household in one decade. Secondly, the average household size is smaller in the inner city than in the suburbs. Throughout the decade the smallest average household size was observed in East York and the largest in Pickering. The range has changed marginally over time from 1.23 to 1.24 to 1.07 persons per household.

Going back to table 3 we can see how the population change and the change in the number of households is distributed in the areas of the CMA. The rate of increase in the number of households exceeds

AREA	1966	1971	1976
Toronto	3.67	3.17	2.75
York	3.45	3.17	2.88
East York	2.96	2.74	2.52
North York	3.71	3.45	3.17
Scarborough	3.94	3.65	3.28
Etobicoke	3.60	3.44	3.07
Pickering	4.19	3.90	3.59
Inner York	3.95	3.77	3.51
Outer York	--	3.73	3.49
Albion	4.01	3.82	3.55
Mississauga	3.93	3.67	3.37
Oakville	3.92	3.69	3.38
Halton	--	3.72	--
CMA	3.69	3.39	3.08

Source: see text

Table 4: Average Household Size by Area

that of the population in every area over both the two time periods. The highest difference in the rate of increase is in Toronto city for 1966-71. The increase, in any case, is much more substantial in the suburbs as opposed to the central part of the city. In fact, in 1971-76 we see a decrease in Toronto City and York and a substantial increase in Outer York, Albion and Mississauga.

But to understand these patterns better we need to examine the distribution of in-migrants in the CMA. As indicated in table 5 the CMA received slightly more than 140,000 heads in each of the two time periods. These figures constitute a sizeable proportion (18.43 and 15.57 respectively) of the total number of households found in the CMA at the end of the two time periods. The total impact of the number of households that entered the CMA, however, is even higher if one takes into account the net number of households that were formed due to the households entering the CMA during the five-year period. Toronto City was the recipient of a large proportion of the households that entered the CMA. For 1966-71 Toronto City received 34.51% of the CMA's in-migrants and for 1971-76 the same figure was 29.25%. A very large portion of these in-migrants came from outside Canada. These data therefore support the argument that immigrants tend to settle in the inner-city immediately after their arrival.

AREA	Abroad	Total	Abroad	Total
Toronto	30695	49245	21665	41425
York	5740	8505	4225	6855
East York	3575	6220	3420	6190
North York	14005	27320	15630	28265
Scarborough	6575	13485	7950	15310
Etobicoke	4690	10905	4900	11545
Pickering	655	1930	540	1795
Inner York	775	2590	1100	3560
Outer York	510	2135	480	2355
Albion	1510	4500	2035	6055
Mississauga	3680	10625	5200	14170
Oakville	1255	3505	1110	4065
Halton	435	1715	--	--
CMA	74100	142680	68255	141590

Table 5: Household heads entered the CMA

We are faced now with the following question: since the city of Toronto received the highest share of the CMA in-migrants then why is the increase in the number of households so moderate in comparison to that of suburbs such as Albion, Mississauga and Inner and Outer York?

We would like to suggest that during the study period there was a net shift of households from the inner city to the suburbs. To see this clearly we need to account not only for the households that entered or left the study area during the intercencal time period but also for the net number of households that were formed during the same time period.

3. Some Organizing Identities

The migratory system we described in chapter 3 was considered to be <u>closed</u>. We argued there that given sufficient data this assumption may not be restrictive. The purpose of this section, therefore, is twofold. Firstly, to validate the argument given in chapter 3, that it is indeed possible to account for growth due to household formation and migration within each time period considered. Secondly, to provide a general accounting scheme that will incorporate all the components of growth from one period to the next.

Let us go back to the system of L areas as described in the first section of chapter III. The study area is now a CMA and the mobile units are households. It is assumed to be a one to one correspondence between households and household heads. Thus in the following discussion we are referring to household heads, whenever household formation or dissolution is assumed to have taken place.

The CMA at time t is considered to have $N[t]$ households in it. In particular, the spatial distribution of households at time t is given by the vector $\underset{\sim}{n}[t]$ and is expressed as:

$$\underset{\sim}{n}[t] = (n_1[t], n_2[t], \ldots, n_L[t])$$

where $n_i[t]$, $i = 1, 2, \ldots L$ is the number of households in area i at time t.

Similarly $n_i[t + \Delta t]$, $i = 1, 2, \ldots L$ will be the number of households in area i at time $t + \Delta t$.

We need now to define a few more terms.

$n_{ii}[t, \Delta t]$ - The number of household heads in area i at time $t + \Delta t$ that were also household heads in area i at time t.

$n_{ji}[t, \Delta t]$ - The number of household heads in area j at time $t + \Delta t$ that were household heads in area i at time t.

$n_{di}[t, \Delta t]$ - The number of household heads in area i at time t that became non-heads or died within the time interval Δt.

$n_{li}[t, \Delta t]$ - The number of households heads in area i at time t that left the CMA within the time interval Δt.

$n_{fi}[t, \Delta t]$ - The number of household heads in area i at time $t + \Delta t$ that were non-heads at time t.

$n_{ai}[t, \Delta t]$ - The number of household heads in area i at time $t + \Delta t$, that were outside the CMA at time t.

At this point we should mention that migration is a continuous process. We should imagine that migrants are moving in and out of area i throughout all the time period Δt, which may potentially be large. In our case, for example, Δt is five years. During this time period household formation and dissolution takes places not only among those that were in area i at time t but also among those that are moving into area i, during the time interval Δt. Thus, n_{fi} is composed of new households formed from the stayers and of the in-migrants, in area i.

We can now write the identity:

$$n_i[t] \equiv n_{ii}[t,\Delta t] + \sum_{\substack{j=1 \\ j \neq i}}^{L} n_{ji}[t,\Delta t] + n_{di}[t,\Delta t] + n_{li}[t,\Delta t] \qquad (1)$$

This identity is merely saying that the total number of household heads in area i at time t can be decomposed to:

- Those who stayed in i during Δt;
- Those who moved from i to all the other areas in the CMA within Δt;
- Those who became non-heads or died during Δt; and
- Those who left the CMA altogether.

Another identity we can write relates to the number of households in area i at time $t + \Delta t$.

$$n_i[t+\Delta t] \equiv n_{ii}[t,\Delta t] + \sum_{\substack{j=1 \\ j \neq i}}^{L} n_{ij}[t,\Delta t] + n_{fi}[t,\Delta t] + n_{ai}[t,\Delta t] \quad (2)$$

A household in area i at time $t + \Delta t$ either was in area i at time t, or it came in i from some other area in the CMA during Δt, or it was formed during Δt, or it came from outside the CMA. Subtracting (1) from (2) we get the change in the number of households in area i, from all possible sources, from time t to time $t + \Delta t$.

$$n_i[t+\Delta t] - n_i[t] \equiv \left(\sum_{\substack{j=1 \\ j \neq i}}^{L} n_{ij}[t,\Delta t] - \sum_{\substack{j=1 \\ j \neq i}}^{L} n_{ji}[t,\Delta t] \right) +$$

$$(n_{fi}[t,\Delta t] - n_{di}[t,\Delta t]) +$$

$$(n_{ai}[t,\Delta t] - n_{li}[t,\Delta t]) \quad (3)$$

There are three terms in the right hand side of this identity. The first represents the change due to intraurban moves, the second is the net household formation and the third is the net migration into area i from outside the CMA. Recall now from chapter III that $p_{ij}[t,\Delta t]$ was defined as the probability of a household moving from j to i in the time interval between t and $t+\Delta t$. This probability can be approximated by the ratio of the migrants from j to i in the time period $[t, t+\Delta t]$ to the number of households in i in the beginning of the time period.

$$p_{ij}[t,\Delta t] \simeq \frac{n_{ij}[t,\Delta t]}{n_i[t]} \qquad (4)$$

Using (4) and ignoring household formation and migration in and out of the CMA the identity in (3) can be written as an equation.

$$n_i[t+\Delta t] - n_i[t] = \sum_{\substack{j=1 \\ j \neq i}}^{L} n_i[t] \cdot p_{ij}[t,\Delta t] - \sum_{\substack{j=1 \\ j \neq i}}^{L} n_j[t] \cdot p_{ji}[t,\Delta t] \qquad (5)$$

Recall now equation (III.12) and the discussion following it in chapter III. In practical terms a constant migration rate q_{ij} over Δt is given by:

$$q_{ij}(t) = \frac{p_{ij}[t,\Delta t]}{\Delta t}$$

If we, therefore, divide both sides of equation (5) by Δt we get a discrete version of equation (III.27).

What we have accomplished in this section is to provide an accounting scheme that allows us to see all the components that contribute to the change in the number of households in an area. In addition we have shown that if household formation and net migration from outside the CMA are ignored then the balance equation (3) is equivalent to (III.12) thus providing a link with the dynamics discussed in chapter III.

In the context of our empirical analysis, however, we are interested to know the migration matrix (n_{ij}) for the two time periods of interest. With the data set available it is suitable to look at the balance equation (2) above. Although a migration matrix can be extracted from the data set, any element n_{ij} of this matrix will not represent a pure flow of migrants from j to i. It will contain also part of n_{fi}, the number of households that were formed in the intercensal time period, and ended up in area i.

To be able to assess the contribution of the new households to the migration flows we need to know the number of the household heads at $t + \Delta t$ that were not heads at time t. This type of information however is not available from our data set and from the census in general. To be able to circumvent the problem, we resort to a model, due to Miron (1983), which provides an estimate of the probability that a head at $t + \Delta t$ was also a head at time t. We discuss this model in the next section.

4. A Model of Household Formation

This section is drawn entirely from Miron (1983). Let us consider a five-year age-sex cohort c at time t and, five years later, at time $t + 5$. The population of the cohort at those two points in time is denoted by $N_c[t]$ and $N_c[t+5]$, respectively. Let also, $H_c[t]$ and $H_c[t+5]$ represent the household heads in the cohort, at the beginning and the end of the time period considered. Thus, the cohort contains at time t, $H_c[t]$ household heads and $(N_c[t] - H_c[t])$ non-heads.

Over a five year period the population of the cohort will change due to mortality and in and out migration. Let us denote by S_c the cohort specific rate of population change over the five years. Then:

$$N_c[t + 5] = S_c \cdot N_c[t]$$

We make now the assumption that the number of heads in the cohort will change due to mortality and migration at the same rate as the whole cohort population. Thus, in the absense of any other source of change the number of heads at the end of the five year time period would be $S_c H_c[t]$.

There are however changes in the headship status due to other reasons, such as marriages and divorces. Let s be the proportion of heads that become non-heads and r be the proportion of non-heads that become heads due to changes in the headship status. The total number of heads at time $t+5$ is:

$$H_c[t+5] = (1-s)\ S_c H_c[t] + r\ S_c\ (N_c[t] - H_c[t]) \qquad (6)$$

Because of the way s is defined $(1-s)$ represents the proportion of heads that remain heads. The first component in the right hand side of equation (6) then represents those heads in the cohort that are still heads at the end of the five-year time period. The second component represents those non-heads in the cohort that became heads within the five-year time period.

Similarly, the total number of non-heads in the cohort at time $t+5$ is:

$$N_c[t+5] - H_c[t+5] = sS_c H_c[t] + (1-r)\ S_c (N_c[t] - H_c[t]). \qquad (7)$$

The first component in the right hand side represents the number of heads that became non-heads, and the second component represents the non-heads that remained non-heads throughout the five year time period.

As we mentioned before the purpose of this exercise is to assess the probability that a household head at time $t + 5$ was <u>not</u> a household head at time t. Let us denote this probability by P_c, then:

$$P_c = \frac{r \cdot S_c(N_c[t] - H_c[t])}{(1-s)\ S_c H_c[t] + rS_c(N_c[t] - H_c[t])} \tag{8}$$

This is the number of those that were not household heads at time t and became heads in the five years after t, divided by the total number of household heads at time $t + 5$. Eliminating r from (8) by utilizing (6) and (7) we get:

$$P_c = [1 - \frac{H_c[t]/N_c[t]}{H_c[t+5]/N_c[t+5]}] + [\frac{H_c[t]/N_c[t]}{H_c[t+5]/N_c[t+5]}]\, s \tag{9}$$

We have now expressed P_c in terms of the number of households and the population of the five-year age-sex cohort and in terms of s. The number of households and the population in any cohort are readily obtained from the census. The value of s, however, still cannot be inferred from the census data.

As Miron (1983), however, points out, one can guess intuitively which cohorts will have a high or low value of s. In addition it is argued that the value of P_c is not expected to be sensitive to the value of s because the coefficient of s in the linear expression (9) is expected to be significantly less than 1. There are two main reasons for that:

Firstly, the headship status H_c/N_c increases over time. An idea about this we get from table 4 for the whole population. The average household size is the inverse of the headship status. Because the average household size decreases over time, the headship status increases for every region. The same empirical observation is true for specific cohorts. Secondly, the headship status tends to increase with age. Since from t to $t + 5$ every member in the cohort will be five years older, the headship status is expected to increase.

The s values adopted here are exactly the ones used by Miron (1983) and they are shown in table 6, along with the calculated values of the corresponding P_c. The weighted average of the P_c values in table 6 were calculated to arrive at the ten year age groups values shown in table 7. The values shown in table 7 can now be applied to the data set to determine the components in the three identities (1), (2) and (3). Miron (1983) presents an accounting table of this kind by area.

In this study, however, we are interested in intraurban

1976 Age Group	Assumed s	Females Pc	Assumed s	Males Pc
15-19	0.06	0.889	0.05	0.959
20-24	0.18	0.490	0.03	0.616
25-29	0.08	0.482	0.01	0.174
30-34	0.04	0.404	0.01	0.073
35-39	0.02	0.330	0.01	0.077
40-44	0.02	0.380	0.01	0.026
45-49	0.02	0.291	0.01	0.033
50-54	0.02	0.289	0.01	0.016
55-59	0.02	0.351	0.02	0.036
60-64	0.06	0.268	0.06	0.055
65 & older	0.06	0.170	0.06	0.032

Source: Miron (1983)

Table 6

Probability of a Household Formed by Five-Year Age-Sex Cohorts

1976 Age Group	Females Pc	Males Pc
15-24	0.525	0.831
25-34	0.446	0.125
35-44	0.359	0.052
45-54	0.290	0.025
55-64	0.309	0.026
65 and older	0.170	0.032

Source: Miron (1983)

Table 7

Probability of a Household Formed by Ten-Year Age-Sex Cohorts mobility. Therefore, we would like to go back to the question posed at the end of section 3 and examine the intraurban moves of those household heads within the CMA that either stayed in place or moved to another area within the CMA boundary. What we really need then is to end up with a migration matrix (n_{ij}), for each time period, with elements n_{ij} free from the new households that were either formed or entered in the CMA in the intercensal time period. In this way we achieve the closed system discussed in chapter III.

To this end, two steps are followed. The first is to extract from our original data set ten migration matrices, one per ten-year age-sex cohort, for each time period. The elements in these matrices result from the question posed to migrant heads at census time as to where they were located five years before. The answers to this question are aggregated and reported without any examination as to whether a head in the end of the time period was also a head in the beginning of the time period. The second step, therefore, is to apply the probabilities of table 7 to the matrices that resulted from the first step. The resulting matrices can then be aggregated to any age-sex cohort level desired.

5. Mobility Trends in the Toronto CMA

The discussion in this section refers to the mobility of those who were household heads in the beginning and the end of each one of the time periods in question. For each one of the time periods the ten-year age-sex cohort migration matrices have been aggregated to produce the destination-by-origin matrices shown in tables 8 and 9. The rows and the columns have been added up to produce the last row and the last column. The last row indicates the total number of household heads that left an area and the last column indicates the total number of households that arrived in an area. The number at the intersection of the last column and last row is the sum of all the outmigrants from all the areas or equivalently, it is the sum of all the inmigrants in all the areas. This number indicates the total migration activity in the CMA in a particular time period. Thus, the total migration activity in the CMA for 1966-71 was 40,714 and for 1971-76 is was almost double, 80,246.

On the outmigration side two-fifths of this activity for 1966-71 and almost half for 1971-76 is due to outmigration from Toronto City. On the inmigration side the major recipients of households seem to be Missisauga and North York. To get a clear view, the net migration for each area was computed by subtracting the last column

1	2	3	4	5	6	7	8	9	10	11	12	13	
0	106	59	471	445	246	86	241	171	122	412	138	52	2549
177	0	4	105	43	87	11	25	27	36	51	17	9	592
109	11	0	72	174	28	54	49	50	11	26	2	5	591
1891	323	156	0	679	359	151	620	243	134	200	80	52	4888
1698	114	178	591	0	160	476	218	83	62	87	44	20	3731
1115	269	25	374	155	0	33	134	53	183	523	97	37	2998
1051	64	48	263	772	85	0	39	9	5	46	9	5	2396
1492	103	46	1056	621	216	28	0	190	64	71	14	22	3923
1047	110	9	465	320	140	68	531	0	37	59	31	5	2822
1617	275	11	425	192	556	18	126	59	0	417	33	177	3906
3873	400	33	726	313	2285	54	116	63	358	0	305	114	8640
835	61	4	127	85	241	0	15	11	37	467	0	56	1939
499	93	9	99	90	152	5	42	18	355	271	106	0	1739
15404	1929	582	4774	3889	4555	984	2156	977	1404	2630	876	554	40714

1 – Toronto
2 – York
3 – East York
4 – North York
5 – Scarborough
6 – Etobicoke
7 – Pickering
8 – Inner York
9 – Outer York
10 – Albion
11 – Mississauga
12 – Oakville
13 – Halton

Source: See Text

Table 8: Toronto CMA Household Mobility 1966 – 71

1	2	3	4	5	6	7	8	9	10	11	12	
0	206	96	963	804	263	98	290	186	185	529	269	3911
3001	0	10	297	59	145	5	23	42	37	85	11	3715
2291	41	0	255	341	34	40	55	42	25	73	35	3232
8066	636	299	0	1368	424	112	742	324	221	306	71	12569
4866	135	455	1384	0	203	403	319	179	64	186	42	8236
4241	749	31	670	242	0	21	145	96	225	1059	95	7574
963	18	37	431	1251	56	0	56	23	10	59	23	2947
2151	221	82	2552	901	291	89	0	318	89	105	28	6827
1579	123	40	861	548	176	80	797	0	59	123	19	4405
2910	503	42	1238	386	1013	25	186	93	0	1069	52	7516
8309	739	110	1601	721	2946	57	118	71	667	0	410	15749
964	63	18	230	166	255	14	31	14	68	867	0	2687
39361	3434	1220	10502	6787	5806	942	2762	1390	1650	4451	1055	79370

1 – Toronto 2 – York 3 – East York
4 – North York 5 – Scarborough 6 – Etobicoke
7 – Pickering 8 – Inner York 9 – Outer York
10 – Albion 11 – Mississauga 12 – Oakville

Source: See Text

Table 9: Toronto CMA Household Mobility 1971 – 76

from the last row in tables 8 and 9. In addition the inmigration was added to the outmigration to get the total migration per area. The results are shown in table 10.

Most of the migration activity for 1966-71 took place in Toronto and Mississauga. The first was a major net loser whereas the other was a major net gainer. The same seems to be the case in 1971-76. Overall, the "flight to the suburbs" which is evident in 1966-71 was intensified in 1971-76. The inmigrants that moved into the CMA located mainly in the inner city as it is evident from table 5. It looks then, as if the interzonal migrants left an inner city vacuum behind, which was filled quickly by inmigrants in the CMA. Toronto City specifically received almost the same number of CMA inmigrants from Canada outside the CMA in the two time periods. In 1966-71 Toronto City received approximately 10,000 CMA inmigrants more than in 1971-76. This difference is due to the larger number of inmigrants Toronto City received from abroad.

A likely explanation for this pattern is that new arriving inmigrants tend to share accommodation with friends and relatives and after an initial settling period seek independent accommodation. Thus most of the immigrants arriving in Toronto City in 1966-71 will seek their own accommodation in 1971-76. This event coupled with the large number of inmigrants in Toronto City from Canada

Area	1966 - 71 Net Migration	1966 - 71 Total Migration	1971 - 76 Net Migration	1971 - 76 Total Migration
Toronto	-12855	17952	-35449	43271
York	-1336	2521	280	7147
East York	9	1173	2013	4451
North York	117	9666	2068	23071
Scarborough	-154	7621	1650	15226
Etobicoke	-1564	7550	1764	13381
Pickering	1414	3380	1998	3885
Inner York	1767	6060	3865	9389
Outer York	1843	3600	3017	5000
Albion	2503	5314	5871	9171
Mississauga	6010	11272	11286	20208
Oakville	1064	2817	1636	3745
Halton	1186	2292	---	---

Source: See Text

Table 10: Net and Total Migration by Area

outside the CMA in 1971-76 and the tendency in the number of the Canadian households to increase would generate enough pressure to justify the large negative number of net migration in 1971-76. The large negative net migration for Toronto City, therefore, is not expected to create any vacant dwellings. This is already seen in table 3 where the number of household heads (occupied dwellings) actually increased by 2.54% in the 1971-76 period.

We would like now to examine whether the male and female household heads, at different ages, exhibited different movement patterns. This is the subject of the next section.

6. Mobility by Age and Sex

The number of households headed by females in the CMA has increased relative to the number of households headed by males, from 1971 to 1976, as shown in table 11. Overall the male to female heads ratio has dropped from 4.6 to 3.5. This trend is evident for every single area in the CMA. Another observation from table 11 is that the households headed by females tend to gravitate to the inner city. The male to female heads ratio tends to increase as we move from the inner city to the suburbs, both for 1971 and 1976. Specifically, in 1971 for every one female head in Toronto City we

Area	1971 Male	1971 Female	M/F	1976 Male	1976 Female	M/F
Toronto	160775	63620	2.5	154310	76080	2.0
York	36710	9815	3.7	36705	13355	2.7
East York	29100	9170	3.2	29915	12515	2.4
North York	126425	19735	6.4	141375	34715	4.1
Scarborough	80160	11375	7.0	95390	22670	4.2
Etobicoke	70965	11325	6.3	78015	18915	4.1
Pickering	10970	780	14.1	12145	1425	8.5
Inner York	20225	2285	8.9	26875	4130	6.5
Outer York	13330	1850	7.2	18900	3065	6.2
Albion	19450	1815	10.7	31650	3795	8.3
Mississauga	42820	4110	10.4	64420	9735	6.6
Oakville	14975	1675	8.9	17625	2795	6.3
Halton	9355	1010	9.3	---	---	---
Total	635260	138565	4.6	707325	203195	3.5
	1966 - 71			1971 - 76		
Total movers within CMA	37515	3204	11.7	68742	10631	6.5

Source: See Text

Table 11: Female and Male Headship Comparisons

had 2.5 male heads and in Pickering we had 14.1 male heads.

In the 1966-71 time period households headed by females were much less mobile than households headed by males. For every one household with a female head that moved there were 11.7 mover households headed by males. Note that we are referring only to intraurban moves and to household heads that were heads in 1966.

The same trend but at a much smaller scale is observed in 1971-76 with a ratio of male to female household head movers of 6.5. These figures imply, however, that the intraurban mobility of the female heads increased dramatically (almost doubled) with respect to that of male heads, from 1966-71 to 1971-76. Tables 12, 13, 14 and 15 give the net migration moves by age of the household heads for both the time periods and for both sexes. The overall mobility, measured by the total number of household outmigrants from all the areas, is also shown in these tables.

From 1966-71 the households with male heads of age 25-44 appear to be the most mobile. By contrast, among the female headed households the most mobile appear to be those headed by elderly of age over 65. Clearly, for the male heads the trend is to move from the inner city to the suburbs with the major loser being the city of Toronto for all the ages. This trend is not obvious for the female heads although the city of Toronto appears to have lost households for

AGE / AREA	15-24	25-34	35-44	45-54	55-64	65+
Toronto	-159	-5421	-3777	-1838	-852	-474
York	-9	-521	-410	-293	-86	5
East York	18	-88	-28	-20	15	19
North York	124	-131	-367	39	151	60
Scarborough	-4	-92	-61	-107	-29	15
Etobicoke	-41	-814	-617	-107	-63	19
Pickering	-20	774	419	219	73	24
Inner York	-44	516	895	322	141	24
Outer York	-30	849	556	312	107	58
Albion	35	1168	725	346	80	92
Mississauga	159	2428	2039	853	209	145
Oakville	-17	578	339	137	54	0
Halton	-13	753	287	98	68	5
Overall Mobility	1256	14333	11200	6230	2941	1554

Source: see text

Table 12: 1966-71 Male Net Migration

AGE / AREA	15-24	25-34	35-44	45-54	55-64	65+
Toronto	147	-61	-42	-53	-73	-253
York	2	-8	-3	-14	-21	17
East York	19	17	10	0	3	4
North York	10	55	29	14	31	95
Scarborough	-14	-8	13	39	14	79
Etobicoke	-14	-30	-29	-28	-17	17
Pickering	-21	-3	-19	-14	-10	-8
Inner York	-38	-17	-16	0	-21	4
Outer York	-45	-6	0	21	24	-4
Albion	-12	14	-3	16	7	25
Mississauga	0	50	48	43	10	25
Oakville	-21	3	3	-21	7	4
Halton	-12	-6	3	-4	10	-4
Overall Mobility	371	521	558	557	442	775

Source: see text.

Table 13: 1966-71 Female Net Migration

AREA \ AGE	15-24	25-34	35-44	45-54	55-64	65+
Toronto	-633	-12464	-8652	-4787	-2445	-2110
York	33	-267	-151	-20	107	252
East York	85	324	254	239	336	203
North York	168	-976	-71	605	487	610
Scarborough	4	-44	268	210	88	310
Etobicoke	9	-223	14	527	390	344
Pickering	13	1435	447	117	24	-15
Inner York	-24	1518	1427	595	175	87
Outer York	13	1558	947	375	131	29
Albion	74	3071	1757	624	244	44
Mississauga	280	5206	3222	1336	424	198
Oakville	-22	862	537	180	39	48
Overall Mobility	2194	28354	18755	10535	5250	3654

Source: see text

Table 14: 1971-76 Male Net Migration

AGE / AREA	15-24	25-34	35-44	45-54	55-64	65+
Toronto	-29	-289	-669	-643	-618	-1490
York	31	50	19	46	38	141
East York	69	116	74	82	107	125
North York	114	188	160	217	135	432
Scarborough	-36	119	96	114	111	411
Etobicoke	19	100	131	124	114	216
Pickering	-29	33	10	4	-17	-25
Inner York	-64	28	38	7	48	29
Outer York	-52	11	0	4	-7	8
Albion	-45	44	26	-7	24	17
Mississauga	45	222	128	60	69	95
Oakville	-24	-22	6	-7	-3	42
Overall Mobility	1131	2512	1609	1452	1334	2594

Source: see text

Table 15: 1971-76 Female Net Migration

all the ages except of those with heads in the category of 15-24.

For 1971-76 the most mobile households appear again to be those with male heads in the age of 25-44. The same trend of a move from the inner city to the suburbs is obvious although in this time period the negative net migration is more focused to the city of Toronto especially for the households with heads of age over 45. Among the households with female heads the most mobile appear to be those headed by elderly. The difference from the 1966-71 period is that for all the households except for those headed by young females (15-24) there is a clear negative net migration out of Toronto City. Again the households with female heads of age over 65 constitute the most mobile group.

It is a well known empirical regularity that mobility falls with age. How can we explain the fact that households headed by female elderly are the most mobile among all the households headed by females?

Let us look first at the nature of those households. They consist typically of one, unmarried, female person. From our data base we can see that in 1971 there were a total of 44,920 households headed by elderly females. Out of those 29,565 were one-person households. Similar figures are observed for 1976; out of 58,550 households headed by female elderly 42,485 were one-person households. We suspect that the majority of the observed moves are newly widowed older ladies that move to nursing homes or a different type of dwelling. We note that major gainers from these migrations are North York and Scarborough.

7. Entropy and Mobility

In this section we would like to examine whether outmigrants from a particular area tend to distribute themselves over several destinations or whether the bulk of them is channelled towards a specific destination. Such a measure of dispersion has been used in migration studies by Liaw and Kanaroglou (1986) and Liaw, Kanaroglou and Moffet (1986). It is provided by the well known function of <u>entropy</u>.

We recall from sections 4 and IV.7 that $n_{ji}[t, \Delta t]$ represents the number of migrants from area i to area j in the time interval between t and $t + \Delta t$. Also, $n_j^1 [t, \Delta t]$ represents the total number of outmigrants from region i in the same time interval. For brevity in this section we refer to a specific time τ and drop the time symbols t and Δt.

The proportion of outmigrants from i that choose j as their destination, given by $\dot{P}_{j|i} = n_{ji}/n_i$, is an approximation to the conditional probability $P_{j|i}$. Because $\sum_{\substack{j=1 \\ j \neq i}} n_{ji} = n_i$ we have

$$\sum_{\substack{j=1 \\ j \neq i}}^{L} \dot{P}_{j|i} = 1.$$

The entropy for area i is then defined as:

$$E_i = -\sum_{\substack{j=1 \\ j \neq i}}^{L} \dot{P}_{j|i} \log_2 \dot{P}_{j|i} \qquad i = 1, 2, \ldots L$$

When the base 2 logarithm is used, as above, the entropy is measured in bits. Because the proportion $\dot{P}_{j|i}$ is always less than 1 the logarithm gives a negative value which implies that the entropy will always assume a positive value. In theory, however, entropy attains the value of zero, the minimum possible, when all the outmigrants from i are directed to a specific destination k. In this case we have:

$$\dot{P}_{j/i} = 0 \quad j = 1, 2, \ldots, L. \ \ j \neq i,k \ \text{ and } \ \dot{P}_{k/i} = 1$$

By contrast the maximum entropy is obtained when the outmigrants from j are evenly distributed to the L-1 destinations. In this case we have:

	1966-71			1971-76		
AREA	Male	Female	Total	Male	Female	Total
Toronto	3.20	3.18	3.21	3.14	2.92	3.14
York	3.20	2.55	3.21	2.87	2.67	2.88
East York	2.80	1.56	2.77	2.69	2.09	2.65
North York	3.20	2.93	3.22	3.10	2.98	3.13
Scarborough	3.18	2.66	3.19	3.11	2.77	3.12
Etobicoke	2.52	2.67	2.57	2.31	2.75	2.39
Pickering	2.53	1.61	2.48	2.65	2.45	2.67
Inner York	2.86	2.59	2.88	2.73	2.73	2.76
Outer York	3.06	2.48	3.03	2.96	2.83	2.97
Albion	2.89	2.52	2.91	2.68	2.57	2.69
Mississauga	3.07	2.65	3.07	2.79	2.71	2.83
Oakville	2.84	2.52	2.84	2.68	2.13	2.60
Halton	2.88	2.51	2.87	--	--	--
Maximum		3.58			3.46	

Source: see text

Table 16: Entropy Measures

$$P_{j/i} = \frac{1}{L-1} \qquad j = 1, 2, \ldots, L \qquad j \neq i$$

and the entropy is given by:

$$E_i = - \sum_{\substack{j=1 \\ j \neq i}}^{L} \frac{1}{L-1} \log_2 \frac{1}{L-1} = - \log_2 \frac{1}{L-1} = \log_2 (L-1)$$

The maximum entropy, therefore, in the case of 12 regions is 3.46 and in the case of 13 is 3.58.

Table 16 summarizes our calculations of entropy for males, females and total number of migrants both for 1966-71 and 1971-76. The values for the total number of moves are slightly smaller for 1971-76 by comparison to those of 1966-71. This is probably due to the fact that we have one destination less in 1971-76 and therefore the maximum entropy is smaller. For the total number of migrants, the minimum value obtained in 1966-71 was for Pickering since almost half of its outmigrants went to Scarborough. During the same time period Etobicoke and East York also obtained low values. The former because half of its outmigrants chose Mississauga as their destination and the latter because North York and Scarborough received over half of the outmigrants. In 1971-76 the lowest value was for Etobicoke because of its strong association with Mississauga.

Comparing males and females we see that the females almost always are associated with lower entropy values than the males. This is because the number of female outmigrants is in general smaller than that of males, and the small number of female outmigrants is distributed to a small number of destinations. In this sense we could say that the female migration pattern was more destination selective than the male. Some very low values were computed for females in 1966-71. The very low value for East York, for example, is attributed to the fact that there were only 57 outmigrants from East York distributed to only three destinations (Toronto, North York and Scarborough). The highest value was consistently obtained for Toronto city with the largest number of outmigrants spread all over the CMA.

CHAPTER VI

INTRAMETROPOLITAN MIGRATION IN TORONTO, 1966-76

FORMAL STATISTICAL ANALYSIS

In chapter V we identified the mobility patterns and trends in the Toronto CMA for the study period 1966-1976. We would like now to link the observed patterns to social, economic and demographic conditions by means of statistical analysis. A mobility pattern is seen as the result of individual migrations. The logical theory that links the probability to migrate with individual decisions to migrate was the subject of chapter IV. The end result of this logical theory, the nested logit model, is the basis of the statistical analysis presented in this chapter.

Maximum likelihood is the particular statistical methodology used. The nonlinear least squares program of BMDP3R of the BMDP package forms the basis of the particular implementation of the maximum likelihood methodology we use in this analysis. Details of this implementation are discussed extensively in Liaw and Bartels (1982) and Kanaroglou, Liaw and Papageorgiou (1986b). The statistical theory upon which this methodology is based is an asymptotic one and was introduced by McFadden (1974) and his associates.

Technical details of the statistical theory and methodology are not discussed in this chapter. Instead, we concentrate on the way model and data are brought together, the description of the independent variables used and the discussion of the results obtained. It should be pointed out at the outset that the number of degrees of freedom in our data set allows meaningful estimation of the destination choice model. Although the departure model is also estimated, a large proportion of this chapter concentrates on the destination choice model and the independent variables used to calibrate it.

The empirical literature on intraurban migration, reviewed in section II.4, strongly suggests that the vast majority of the studies concentrated on the decision to move, and therefore they are relevant to the departure model on which we pay little attention here as mentioned above. The lack of an appropriate guiding theory and the scarcity of previous empirical findings for the choice of a destination makes the selection of independent variables a difficult task. In this respect, we consider the findings from the calibration of the destination choice model a significant contribution to the almost non-existent empirical evidence.

Here is a guide to the rest of this chapter which is divided into six sections. Section 1 concentrates on the way the statistical models are interfaced with the data and discusses the overall strategy adopted in the statistical analysis. Sections 2 and 3 discuss the independent variables for the destination choice and departure models respectively. The results of the destination choice model are discussed in the following two sections. In particular, section 4 concentrates on the results of the analysis for the 1971-76 time period and section 5 on the results of the 1966-71 time period. Finally, the results for the departure model for both the time periods are discussed in section 6.

1. Model/Data Interface

In this section we bring together the model and the ideas described in section IV.7 with the data set described in section V.1.

Certain compromises had to be made, on the number of areas that we would use in the study. Because Halton was not part of the CMA in the 1971-76 period, it was excluded from the statistical analysis for both the time periods under consideration in order to maintain compatibility. Outer York was also excluded since its corresponding values for certain critical independent variables were missing. Thus, the analysis was restricted to 11 areas.

The time interval Δt that we referred to in earlier chapters is the intercensal time period which is five years. In our case the models in (IV.28) and (IV.32) are expressed by:

$$p_{i/j} = \exp(\underset{\sim}{b}'\underset{\sim}{Y}_{ij}) \Big/ \Big(\sum_{\substack{k=1 \\ k \neq j}}^{11} \exp(\underset{\sim}{b}'\underset{\sim}{Y}_{kj}) \Big) \tag{1}$$

$$p_{\cdot j} = \exp(\alpha_o + \mu I_j + \underset{\sim}{a}'\underset{\sim}{X}_j) \Big/ (1 + \exp(\alpha_o + \mu I_j + \underset{\sim}{a}'\underset{\sim}{X}_j)) \tag{2}$$

$$I_j = \ln\Big(\sum_{\substack{k=1 \\ k \neq j}}^{11} \exp(\underset{\sim}{b}'\underset{\sim}{Y}_{kj}) \Big) \tag{3}$$

As mentioned in chapter IV there are three basic steps in the estimation procedure. In the first step through the destination choice model (1) the parameters $\underset{\sim}{b}$ are estimated. The second step involves the calculation of the inclusive value I_j with the estimates of $\underset{\sim}{b}$ obtained in step 1. Finally, through the departure model (2) the parameters α_o, μ and $\underset{\sim}{a}$ are estimated. We may recall that α_o captures the effect of the time interval on the probability $p_{\cdot j}$ to outmigrate and I_j represents the attractiveness of the best alternative area in the eyes of a typical outmigrant from j.

The estimation consisting of the three steps mentioned above was done separately for the two time periods 1966-71 and 1971-76. For the destination choice model, the proportion that corresponds to $p_{i/j}$ is the dependent variable. As explained in detail in section IV.7 the observations for the dependent variable are the elements of the matrix (n_{ij}/n_j^1), excluding its diagonal elements. It is reminded that n_{ij} is the number of migrants from area j to area i and n_j^1 is the total number of outmigrants from j. In our case for each time period, we have an 11×11 matrix. Such matrices are obtained from tables V.8 and V.9 by excluding Outer York and Halton and dividing the last row (the total number of outmigrant household heads from an area) into each row. The total number of observations are $11^2 - 11 = 110$.

For the departure model (2) the dependent variable is the proportion that corresponds to $p_{\cdot j}$. A particular observation relates to an area, say j, and is the total number of outmigrants from j divided by the population at risk, which, in this case, is the number of households in area j at the beginning of the five-year period. The number of observations is equal to the total number of areas, which, in this case, is 11.

The statistical theory upon which this estimation is based is an asymptotic one. To justify any inference, therefore, we need a large number of degrees of freedom. The degrees of freedom are calculated by subtracting the number of independent variables from the number of independent observations for the dependent variable. For the departure model the number of degrees of freedom is going to be at best 10, which is inadequate. The calibration was, however, done for the sake of completeness. In addition, these runs serve to demonstrate the performance of the estimation procedure in the case of small samples. In general, the results obtained are characterized by lack of robustness. This is in direct contrast with the results for for the destination choice model. Since the underlying statistical theory does not provide any indication about an adequate sample size, the runs for the departure model support in an indirect way the claim that the sample size for the destination choice model is adequately large.

Reviewing the empirical literature in section II.4 we identified the age of the household head as one of the most consistent predictor variables of intraurban desired, expected and actual mobility (Rossi, 1955; Speare, 1974; Duncan and Newman, 1976). For our data base the various age groups of household heads were found in section V.6 to exhibit differences in their mobility patterns. In trying to explain these patterns, instead of considering the age of the household head as a predictor variable, we disaggregate our data base according to age and carry out the analysis for different age groups. We also carry out the analysis for all the household heads. The age groups we considered are the young household heads of age less than thirty-five, the mature household heads of age between thirty-five and fifty-five and finally the elderly household heads of age over fifty-five. Our data base would have allowed us to run the model for the six different age groups shown in tables V.12 to V.15. However, the resulting migration matrices, especially the ones for the household heads over sixty-five would become extremely sparse, which would render any statistical analysis meaningless.

We would like now to turn our attention to the independent variables. It is well known that there is no unique way or method that will lead to a specific set of independent variables. It depends a lot on theoretical expectations and the researcher's experience. Broadly speaking, the process involves two stages:

In the first stage the variables ~~to~~ be used are conceptualized. The list of variables that one ends up with in this stage is dependent upon intuition, experience and knowledge of the literature.

The second stage is the actual implementation or measurement. At this stage, many of the conceptualized variables will be dropped because of data unavailability. This same issue of data availability will affect the way the rest of the variables are measured. But even if plenty of data are available there is no unique way on how a conceptualized variable can be translated into a concrete set of numbers. Each way may give a slightly different meaning to the conceptualized variable. The issue of measurement becomes more complicated when one is trying to avoid problems of mutlicolinearity and serial autocorrelation.

With these problems in mind we now turn to examine the independent variables chosen for this study. Reviewing the empirical literature on intraurban migration in section II.4, we have seen that most of the effort has concentrated on the decision to move, rather than the locational decision. The findings in the literature, therefore, are more helpful in the search of independent variables for the departure model rather than the destination choice model.

2. Independent Variables-Destination Choice

In our discussion of section IV.7 we broadly categorized the independent variables to be used in the destination choice model into two groups. Those that refer to both the origin and the destination, which we called relational variables, and those that refer to the destination characteristics, which we called destination specific variables.

In this study we used a total of thirteen independent variables for the destination choice model. The data for these variables are from Statistics Canada publications for the 1966 and 1971 censuses. For 1966 we were able to find information only for six of the thirteen variables. Fortunately, as it will become obvious later on, these were the most critical ones.

Six of the thirteen variables are of the relational type. As Porell (1982) points out one can think of them as measures of physical or social distance. The first of these variables is physical distance. This variable was used by Miron (1983) within the same context and data and was found to be statistically significant and negatively correlated to the destination choice probability. It was measured there by the minimum number of boundaries one has to cross in order to go from one area to another. In our case the distance between two areas was measured as the centroid distance with the help of a map of the Toronto CMA, provided in the 1976 census publications. A point was taken as a representative of an area. For sparsely populated areas such as Albion or Outer York the point was taken closer to the populated areas whereas for densely populated areas, such as Toronto City, a point close to the centre of the area was taken to represent the area. The distance between any two points was measured on the map in centimeters. Using then the map scale these distances were translated in kilometers. The end result is shown in table 1 of appendix 1.

From the behavioural point of view shorter distances are associated with more familiar surroundings to the individual mover. Furthermore as distance increases familiarity decreases at a decreasing rate. This implies that although in the mind of a potential mover two locations 1km and 3km away from the present location can be clearly distinguished, two other locations 41km and 43km away cannot. It is expected, therefore, that the migration flows will correlate better with the logarithm of distance rather than the distance itself. Thus the natural logarithm of distance is actually used as a variable in this study. Porell (1982), thinking along the same lines, uses the first and the second power of distance as predictor variables. With the logarithm function we achieve the same effect and at the same time we avoid any unnecessary multicolinearity problems. The logarithm of distance is expected to have a negative relationship to the dependent variable.

Distance, as one of the strongest predictor variables of migration flows, has been advocated for a long time by those adhering to the schematic framework as we discussed extensively in section II.1. Here we would like to point out that distance captures only one aspect of the geometry of the study area. It says nothing about the relative centrality of the areas since it does not differentiate between a central city to a suburb move from a suburb to a suburb move of equal distance. Thus, preferences for suburban or central locations are not reflected by the distance variable.

Since most of the intraurban moves take place within a short distance, due to the large size of the areas used in this study many of the moves that take place in the CMA are not recorded in our migration flow matrices. Clearly, distance does not capture the effects of variations in the size of the zones used. In general the way the zonal size variation effects can be captured in a nested logit model of migration is an issue that requires further investigation, not undertaken within this thesis. The treatment of spatial effects with the nested logit model is discussed extensively in Boots and Kanaroglou (1987).

The remaining five of the relational variables are all dissimilarity indices or measures of social distance among the areas. A dissimilarity index between two areas i and j for a certain variable X, say education, is calculated as follows:

Let us assume that the population in the two areas can be partitioned in M mutually exclusive groups with respect to the nominal or ordinal variable X. Let X_{im} be the percentage of the population in area i that belongs to group m. We have then:

$$\sum_{m=1}^{M} X_{im} = 100 \quad \text{and} \quad \sum_{m=1}^{M} X_{jm} = 100$$

The dissimilarity index for this variable is defined as:

$$DI = \frac{1}{2} \sum_{m=1}^{M} |X_{im} - X_{jm}|$$

This index indicates the percentage of population in area j that needs to be rearranged in the M groups so that area j becomes exactly similar to area i with respect to variable X. It is expected that the more dissimilar the two areas are, the less the migration flow between the two areas. Alternatively, migrants prefer to move to areas with a high percentage of people similar to them. Thus, a negative relationship is expected between migration and any one of the dissimilarity indices used.

What remains still undefined, however, is what is meant by similarity among people. In general, people are similar when they belong to the same "social class", attain the same "social status", and have common attitudes, values and lifestyles. Although a considerable body of the sociological literature has attempted to give precise meaning to the terms "social class" and "social status", Hodges (1964) maintains that little agreement has been reached and these terms mean different things to different people. Yet according to Hodges (1964, p. 13) the majority of the sociologists would agree that social status is a "blended product of shared and analogous occupational orientations, educational backgrounds, economic wherewithal, and life experiences". For our purposes then we accept that social status can be defined as a vector of attributes, although it is somewhat arbitrary what attributes go into the vector and how they are measured. Occupation, education and income are relatively straightforward to conceptualize as indicators of social status. To reflect the term "life experiences" in Hodge's definition we use two more attributes. The first of them is "religion" as a surrogate for similarities in ethnic and cultural backgrounds of people. The use of this variable was also triggered by the highly visible concentration of ethnic groups in Toronto. The second variable is "household size" since to a certain extent household heads leading households of similar size have similar life experiences.

Although some of the variables mentioned above are expected to be highly intercorrelated no attempt was made here to combine them into one or two composite indices (for example, by means of a factor analysis). The results obtained for an index of this kind are difficult to disentangle. Also the weighting schemes used in such methods are largely ad hoc. Details about the sources of the relational variables along with their definitions and comments are in appendix 1 to chapter VI. Information for these variables was available only for 1971.

Let us now turn to examine the destination specific variables and the a priori hypothesis concerning their presence in this calibration. Details about their definition and measurement are in appendix 2 to chapter VI. It is reiterated here that there is no sound theory to back up the selection of these variables, neither is there any empirical literature in precisely this context with the exception of Miron (1983). Porell's (1982) study of Wichita, Kansas bears some resemblance to the way destination choice is perceived here.

Some of the findings of the traditional empirical literature, reviewed in section II.4, particularly those related to the social and physical attributes of locations, provide some pointers for the selection of destination specific variables. Most of these variables, however, refer to neighbourhoods, and the areas in our study can hardly be thought of as neighbourhoods due to their large size. We recall that five types of factors have been identified in the literature as affecting the choice of a destination. As discussed in detail in section II.4 these were the characteristics of the neighbourhood, neighbourhood services, housing characteristics, accessibility and housing cost. We used one or more gross surrogates to represent each one of these factors, in the analysis.

The characteristics of a neighbourhood were represented by two variables, the density of the population and the age of the housing in an area. We recall that two types of neighbourhood characteristics, social and physical were identified as important. Population density is mostly associated with physical characteristics. High population density implies a smaller lot size, lack of privacy and increased traffic noise. It may, however, also be associated with social characteristics such as crime rate. There are also positive characteristics that can be associated with high population densities such as increased (presumably positive) interaction with neighbours. We expect however, the negative aspects to prevail and thus the sign of the population density is expected to be negative.

Other neighbourhood characteristics that were found to be of importance in the literature include the physical characteristics of the housing stock, such as the exterior condition of dwellings and lots. Assuming that older dwellings will be in worse physical condition than newer ones we include the "housing age" variable in the analysis. This variable represents the percentage of the occupied dwellings in an area that were built before 1946. If the above assumption holds, then this variable is expected to appear with a negative sign.

The "population share" variable, measured as the percentage of the population of the total area, was included as a gross surrogate for the public services provided in the area. Services of this kind are, for example, fire, garbage, police and public transportation services. It is assumed that a larger population share is associated with better services. The sign of this variable is, therefore, expected to be positive.

Two additional variables, "New Housing" and "Housing Availability" ("Housing Stock"), are incorporated as surrogates of the vacancy rate in an area since direct vacancy rate information is, in general, unavailable. "New Housing" is measured by the new house completions during the time period of interest and "Housing Availability" is measured by the number of the occupied dwellings in the beginning of the time period of interest. The underlying hypothesis is that the more the number of the housing units in an area the higher the expected vacancy rate. Both of these variables are expected to be associated with a positive sign.

As mentioned already in section II.4 accessibility has been associated with ambiguous findings in the literature. In most studies accessibility is measured by the distance from the place of residence to workplace or the composite distance to workplace and other amenities. In our case, with a highly aggregate data base, we are representing this idea by the number of jobs at the destination. An area with a high level of employment will attract migrants, who will locate there in order to minimize their transportation bill to work. We, thus, expect a positive relationship between this variable and the dependent variable.

The final variable we use is "Housing Value", measured as the median price of dwelling units in the area of interest. The impact of such a variable on migration inflows is difficult to determine, since migration inflows affect the price of houses in the migrant receiving area. In addition, the prices relate to the beginning of the time period, which in our case is five years. It is possible for the relative value of houses to change within a five year time period. In using this variable, we are forced to make the assumption that individual migrants in selecting a destination are taking the spatial distribution of housing prices as given and that this spatial distribution is maintained for the duration of five years. Even with this assumption, however, it is difficult to determine a priori the expected sign of this variable since this depends on the income profile of the migrant households. The exact way the destination specific independent variables are measures is described in appendix 2 to chapter VI.

3. Independent Variables - Departure Model

As mentioned in section 1 of this chapter parameter estimation for the departure model was done only for completeness due to the lack of adequate degrees of freedom. In an attempt to keep the number of independent variables to a minimum a total of three were used for both the time periods.

The dependent, as well as, the independent variables are presented in tables 1 and 2, for the two time periods 1966-71 and 1971-76. From the second column on both tables we can see that the highest proportion of outmigration in 1966-71 was for Inner York which, however, was characterized by net in-migration. This means that the proportion of in-migration was even higher. Toronto City, despite its high volume of out-migration (Table V.10) appears to have a relatively low proportion of outmigration due to its high number of households. For the 1971-76 time period Inner York, Albion, Mississauga and Oakville remained in the 1966-71 level. For Toronto City, however, the outmigration proportion doubled its 1966-71 level. This was due to the increase in the outmigration level from one time period to the next. The same, at a slightly smaller scale, happened to York, East York and North York. Here is a description of the three independent variables used:

Area	Population of Out Migration	Population Share	Tenure
Toronto	0.087	32	[illegible]
York	0.045	7	42.49
East York	0.018	4	40.58
North York	0.044	[illegible]	37.74
Scarborough	0.055	[illegible]	26.29
Etobicoke	0.062	12	33.63
Pickering	0.105	2	17.33
Inner York	0.128	3	18.49
Albion	0.094	3	20.00
Mississauga	0.090	5	23.74
Oakville	0.065	2	19.58

Table 1
Dependent and Independent Variables
Departure Model: 1966-71

* Gross Estimate

Area	Population of Out Migration	Population Share %	Tenure %
Toronto	0.175	28	58.20
York	0.074	6	48.60
East York	0.032	4	50.50
North York	0.072	20	48.70
Scarborough	0.074	13	36.80
Etobicoke	0.071	11	37.60
Pickering	0.080	2	15.60
Inner York	0.123	3	22.00
Albion	0.078	2	27.00
Mississauga	0.095	7	36.50
Oakville	0.063	2	30.00

Table 2
Dependent and Independent Variables
Departure Model: 1971-76

Inclusive Value - The inclusive value calculated through equation (3) is treated as an independent variable in the departure model. The calculated inclusive values will be discussed in section 6 of this chapter after we present the results of the destination choice model.

Population Share - The population share of the zones has been presented and discussed for both the time periods in section 2 of this chapter. It was argued there that the population share represents public services. In connection with the outmigration rates, the sign of the estimated coefficient would be expected, then, to be negative.

Tenure - As pointed out in section II.4, one of the consistent findings in the empirical literature of intraurban migration is that renters are more likely to move than homeowners (Roistacher, 1974; Goodman, 1976). This variable is calculated as the percentage of the total occupied dwellings, in an area, that are rented. The source for this information is Statistics Canada (1968) and Statistics Canada (1973a) for the 1966 and 1971 variables, respectively.

In 1966, then, almost 50% of the dwellings in Toronto city were rented. The smallest percentage of rented dwellings was in Pickering, Inner York and Oakville. Because the data were unavailable for Albion, the value provided is only a rough estimate.

Observing the same information in table 5 we can see that with the exception of Pickering the percent of rented dwelling units has increased in ~~all~~ the zones. Because households that rent their dwelling are expected to be more mobile the coefficient of this variable is expected to appear with a positive sign.

4. Results, Destination Choice Model: 1971-76

This section presents the results of the formal statistical analysis for the destination choice model for the 1971-76 time period. The results for the 1971-76 time period are discussed first, since many more independent variables in comparison to the 1966-71 period, are available which allows a richer analysis through the use of different combinations of independent variables. The results for the 1966-71 time period are discussed in section 5.

Four different types of runs were made. One for all the migrant heads and one for each of the following three age groups:

(1) Young adult heads between 15-34 years of age

(2) Middle age heads between 35-54 years of age

(3) Elderly heads of over 55 years of age.

The results for all the migrant heads are presented in table 3 whereas the results for the three age groups are summarized in appendix 3 of this chapter. Ten different runs are presented for each

Variable	1	2	3	4	5	6	7	8	9	10
Distance	-12.40 (-9.40)	-12.90 (-8.90)	-8.20 (-6.90)	-4.70 (-3.60)	-12.70 (-8.90)	-9.00 (-7.80)	-8.20 (-6.90)	-10.50 (-8.20)	-4.00 (-2.80)	-12.40 (-6.50)
Education	-1.90 (-1.50)	—	—	—	—	—	—	—	—	—
Household Size	-0.40 (-.50)	—	—	—	—	—	—	—	—	—
Income	1.50 (1.80)	—	—	—	—	—	—	—	—	—
Occupation	0.50 (.50)	—	—	—	—	—	—	—	—	—
Religion	-1.70 (-1.30)	—	—	—	—	—	—	—	—	-2.40 (-1.40)
Population	—	—	—	—	10.20 ([illegible])	7.40 (5.20)	—	—	—	—
Housing Value	—	—	—	—	—	—	—	—	—	1.90 (.95)
Housing Age	7.40 (9.00)	8.10 (9.00)	—	-2.00 (-2.00)	7.90 (9.50)	—	—	—	—	8.20 (9.40)
New Housing	11.50 (8.00)	13.20 (8.60)	9.10 (4.20)	5.00 (3.70)	9.40 (8.70)	—	4.60 (4.00)	6.30 (3.00)	9.30 (4.90)	10.70 (7.90)
Employment	—	—	—	—	—	—	—	4.90 (4.60)	—	—
Pop. Density	-11.11 (-12.80)	-11.40 (-13.30)	-6.50 (-6.00)	—	-10.10 (-13.20)	-9.50 (-9.70)	-5.00 (-6.30)	-8.60 (-9.20)	—	-10.90 (-12.90)
Housing Avail.	—	—	—	—	—	—	0.40 (.40)	—	—	—
ρ^2	0.809	0.792	0.621	0.505	0.790	0.560	0.623	0.715	0.474	0.803
$\bar{R}^2$	0.783	0.779	0.702	0.461	0.779	0.660	0.670	0.751	0.425	0.754
D.F.	90	95	96	96	94	96	95	95	97	93

Table 3: 1971 – 76 Destination Choice Results Migrant Heads of all Ages

age group. This set of runs represents a selection from a much larger set of runs. They were selected as summarizing in the best possible way the relationship found between different combinations of independent variables and the dependent variable.

In run 1 nine variables were included. This run was done in order to obtain an initial idea about the relative merit of each one of the independent variables. The value shown beside each variable is the *t*-ratio of the coefficient, associated with the independent variable. The *t*-ratio represents a measure of the relative likelihood of a change in the dependent variable caused by a change in the independent variable. The normal way of calculating the *t*-ratio of an independent variable is by dividing the parameter estimate of the variable by the corresponding standard error. As it has been pointed out, however, in Liaw and Bartels (1982) and in Kanaroglou, Liaw and Papageorgiou (1986b), because the statistical theory upon which this estimation is based is asymptotic, the *t*-ratios calculated in this way will tend to be inflated. For finite samples the asymptotic standard error used in the calculation of the *t*-ratio will tend to understate the real value of the standard error. A suitable correction is to multiply the asymptotic standard error by the square root of the weighted residual mean square before dividing into the parameter estimate. The *t*-ratios shown in the tables have been corrected this way.

As a measure of the intensity of the relationship between the dependent and an independent variable the beta weight is calculated shown in parenthesis under the t-ratio.

For the overall goodness of fit we use the following measure proposed by McFadden (1974):

$$\rho^2 = 1 - \frac{\hat{S}^2}{S^2}$$

where: $\hat{S}^2$ - is the weighted residual mean square of the maximum likelihood solution.

S^2 - is the weighted residual mean square computed under the null hypothesis that all the parameters are zero.

In addition, we use the coefficient of determination R^2 where R is the simple correlation coefficient between the observed and the estimated migration proportions.

The number of degrees of freedom for L areas and s independent variables in the destination choice model is given by:

$$DF = L \cdot (L-2) - s$$

The overall goodness of fit measured by ρ^2 indicates that almost 81% in the variation of the dependent variable can be accounted for by the independent variables used.

There is practically no evidence that household size and occupational dissimilarity are related to the destination choice probabilities. The rest of the relational variables appear also to be relatively insignificant. More specifically the dissimilarity in income appears with a *t*-ratio of 1.5 (positive sign) which tends to suggest that high income migrants prefer low income destinations and vice-versa.

The dissimilarity in education is the most significant of the relational variables and appears with the expected sign. In other runs (not shown in the tables) the absolute value of the *t*-ratio for this variable never exceeded 1.9. By contrast the "Religion" variable appeared to be more significant in subsequent runs reflecting the clustering of ethnic groups in Toronto.

As mentioned in section 2, in a similar study for Wichita, Kansas, Porell (1982, p. 132) used income, occupation and education dissimilarities between areas. Treating these dissimilarities as social distances between areas, he also included the squares of the dissimilarity measures as independent variables. Furthermore, since the dissimilarity is measured as the absolute value of a difference between the two areas, a dummy variable was incorporated to indicate which one of any two areas had the higher value of the corresponding variable. He presents runs for various income and race (white/non-white) groups. Although the physical distance variable appeared to be significant for every group, this was not true in general for the social distance variables. The income difference, however, appeared to be significant for the low income, white group.

Since the calibration in the Porell (1982) study was done with only the social and physical distance independent variables with small areas and with data that related to a one year period, we tend to believe that disaggregated data will produce better results for the dissimilarity variables.

The remaining four variables, "Distance", "Housing Age", "New Housing" and "Population Density" are highly significant. The "Housing Age" variable appears with a positive sign which implies that our original hypothesis will have to be reevaluated.

In run 2 only the four most significant variables were included. In comparison to run 1, the overall goodness of fit measured by either ρ^2 or R^2, is virtually unchanged. The "New Housing" variable is associated with the highest t-ratio (13.2) which is strong evidence of a relationship of this variable to the dependent variable. The beta weight for the "New Housing" variable, however, is the lowest (8.6) in this run indicating that, relative to the other variables in this run, "New Housing" has the lowest impact on the dependent variable.

The "Housing Age" variable is still highly significant with the second highest beta weight. The positive sign of the coefficient, however, still persists. We would like to investigate further the exact nature of this variable.

In order to examine the incremental contribution of the "Housing Age" variable to ρ^2 we excluded it in run 3. The value of ρ^2 drops significantly to .621. In addition, the significance of the other three variables was reduced both in terms of the t-ratio and beta weight, with the population density variable affected the most. Thus, the "Housing Age" variable appears to be complementary to the other three variables and especially to the "Population Density" variable. As explained in table 1 of appendix 2 the variable assumes high values for the more densely populated areas such as Toronto City, York and East York and those that represent more established communities such as Pickering and Inner York. Since these areas are anticipated to have a better infrastructure and provision of services and ammenities we tend to change our original hypothesis and assume that the "Housing Age" variable represents the positive effects of urbanization. By contrast, the "Population Density" variable captures the negative effects of overcrowding. To check this claim further, in run 4 we included only the variables, "Distance", "Housing Age" and "New Housing". Comparing with run 2 we see that "Distance" and "New Housing" are substantially less significant, which indicates the complementarity of these variables with the "Density" variable. What is important, however, in this run is that the "Housing Age" variable becomes marginally significant and with a negative sign. From a large number of similar runs (not shown in table 3) we verified that the "Housing Age" t-ratio assumes a negative value only when the "Population Density" variable is not present. According to our earlier claim, in this run both the positive and negative aspects of overcrowding are captured by the "Housing Age" variable. These two effects tend to cancel each other out with the negative effects being stronger. To a certain extent, therefore, this run supports our claim.

"Population Share" has been used in several (mostly interregional) migration studies and most of the time it appears to be significant. In run 5 we add "Population Share" as a fifth variable to the basic four variables of run 2. The goodness of fit of the model is virtually unchanged ($\rho^2 = .790$). The new variable is insignificant and the t-ratio of the other variables, with the exception of the "New Housing" variable, are unaffected. In the presence of the "Population Density" variable, "Population Share" would express the positive aspects of a high level of urbanization. But the presence of the housing variables renders the "Population Share" variable insignificant in this respect.

When the two housing variables are eliminated in run 6 then the "Population Share" variable becomes significant. The performance of the model, however, is not as good as in run 2 since the coefficient of determination drops to .560.

In run 7 we attempt to examine whether the "Housing Availability" variable can adequately replace the "Housing Age" variable. Thus, run 7 is identical to run 2 except that the "Housing Age" variable is replaced with the "Housing Availability" variable. The results of run 7 do not provide any evidence that the "Housing Availability" variable is related to the dependent variable, at least in the context of this run where the "New Housing" variable is present.

One positive aspect of higher population densities is the availability of employment. Through several runs we established that the "Employment" variable is competitive to the "Housing Age" variable. Whenever the "Housing Age" variable is included the "Employment" variable appeared to be insignificant with no impact on the overall explanatory power of the model. We show run 8 which is similar to run 2 except that the "Housing Age" variable is replaced by the "Employment" variable. In this case the "Employment" variable is significant with a t-ratio of 4.9 and beta weight 4.6. This variable, however, does not perform as well as the "Housing Age" variable for which in run 2 the t-ratio was 8.1 and the beta weight 9.5. In addition, the "Employment" variable does not complement with the other three variables as well as the "Housing Age" variable. As a result the overall goodness of fit is .715 as opposed to .792 that we had in run 2.

The same set of data was used by Miron (1983) to calibrate the same model with two independent variables, "Distance" and "New Housing". "Distance" was measured as the minimum number of boundaries one has to cross to go from one area to another. "New Housing" was identical to the variable used in this study. The methodology used by Miron (1983) was first to linearize the model by considering as a dependent variable the logarithm of the odds of outmigration and then to apply Ordinary Least Square. He reports two runs with the same variables, one for 1966-71 and one including the data for both the time periods 1966-71 and 1971-76. The coefficient of determination for those two periods were 0.460 and 0.505. To compare our results with those of Miron we did the same run with our method. The results are shown in run 9 where ρ^2=.474 and R^2=.423 which are comparable to Miron's. Moreover, with the addition of two more variables, "Housing Age" and "Population Density", the overall explanatory powers of the model, measured by ρ^2, almost doubles.

Finally, in run 10 we use two more variables in addition to the variables used in run 2. The first variable is the value of housing at the destination. The t-ratio for this variable is rather low (1.9) which does not provide a very strong evidence of a relationship between the "Housing Value" variable and the dependent variable. In addition, the t-ratio value is positive and it has been positive in several other runs, which tends to suggest that migrants gravitate towards destinations with high priced housing.

The second variable included in run 10 was the dissimilarity in religion between the areas under consideration, as a measure of concentration of ethnic groups in the areas of the Toronto CMA. The variable appears to be significant and with the expected sign. Its incremental contribution to ρ^2, however, is very small. On the basis of the performance of similar variables in intermetropolitan studies by Liaw and Kanaroglou (1986) and Liaw, Kanaroglou and Moffet (1986) and in an interprovincial study by Ledend and Liaw (1986) we expected that the "Religion" variable would perform better. Despite out efforts, however, in various other runs, the absolute value of the t-ratio for this variable hardly ever rose above 2.4. We provide two possible explanations for that.

Firstly, the areas we consider are relatively large to exhibit substantial differences in ethnic concentrations. Perhaps, this variable <u>would</u> have been more significant if the areas we had were smaller.

Secondly, as explained earlier on in this chapter religion was used as the basis of differentiation between the ethnic groups. The same census publication also provides information on the ethnic or cultural group persons or their paternal ancestors belonged to upon coming to this continent. Language spoken by a person or by his ancestor on the male side upon arrival on this continent was a guide in the determination of the ethnic group. Perhaps a variable formed around this information would have yielded better results.

In closing this section it is appropriate to make two important comments related to the statistical analysis just presented. The first comment relates to the difficulty in providing clear cut explanations for our findings. As we mentioned on several occasions in our description, we believe that the problem is with the aggregate nature of our data. The areas we are dealing with are very large for our purposes and the independent variables we use are "too gross" as surrogates of what we intend to represent. This weakness should not be viewed as a weakness of the model, which, as presented in chapter IV, has a sound behavioural theory to back it up.

The second comment refers to the relevance of these results. As compared to Miron's (1983) runs, with two additional variables we have managed to almost double the overall goodness of fit, although we recognize that the two additional variables introduced

("Population Density" and "Housing Age") might be highly intercorrelated to the "New Housing" variable.

As mentioned before the same runs were made for three different age groups. The results are presented in appendix 3 to chapter VI.

5. Results - Destination Choice Model (1966-71)

In this section we discuss the results of the formal statistical analysis for the time period, 1966-71. The analysis done is not as extensive as that of the 1971-76 period due to the fact that a much smaller number of explanatory variables are available. Fortunately, we have the four variables: distance, housing age, new housing and population density that proved to be the most important ones for the time period 1971-76. We also have the "population share" variable. We should remember, however, that the values of the 1966 "housing age" variable represent only estimates derived under certain assumptions from the 1971 values.

The overall results are very similar to the 1971-76 results except that the coefficient of determination (ρ^2) for most of the corresponding runs is smaller.

Thus, run 1 of table 4 corresponding to run 2 of table 3 with

Variable	1	2	3	4
Distance	-7.2 (-6.5)	-4.7 (-5.1)	-7.7 (-7.1)	-5.8 (-5.6)
Population	--	--	2.9 (4.9)	5.9 (6.8)
Housing Age	6.0 (9.2)	--	5.8 (8.5)	--
New Housing	8.8 (7.4)	5.6 (3.6)	4.0 (4.8)	--
Pop. Density	-9.7 (-17.0)	-6.8 (-9.0)	-9.5 (-20.0)	-8.5 (-12.9)
# of Variables Used	4	3	5	3
ρ^2	0.614	0.395	0.626	0.450
$\bar{R}^2$	0.540	0.530	0.566	0.540
D.F.	96	97	95	97

Table 4: 1966 - 71 All Migrant Heads

the four variables, Distance, Housing Age, New Housing and Population Density gives an overall goodness of fit of .614 as opposed to .792. All the four variables are significant.

Run 2 is identical to run 1 except that the "Housing Age" variable is excluded. The overall-goodness of fit drops substantially to .396. In addition the t-ratio and the beta weights for the other variables are reduced.

In run 3 we include the "Population Share" variable with the basic four variables of run 1. The incremental contribution of "Population Share" is very small since ρ^2 rises from .614 to .626. The "Distance" and "Population Density" variable are almost unaffected by the presence of "Population Share". The "New Housing" variable, however, is mostly affected since its t-ratio drops from 8.8 to 4.0. Thus the "Population Share" variable is competitive to the "New Housing" variable. The t-ratio of "Population Share" increases substantially only when "New Housing" is excluded. But then, the performance of the other three variables, and especially that of "Housing Age" are affected substantially with a significant drop in the value of ρ^2. This is seen in run 4, where excluding the two housing variables raises the t-ratio for the population share to 5.9 but the coefficient of determination in this run is only .450.

The same four runs of table 4 are presented for the three age groups in appendix 4 to chapter VI.

6. Results - Departure Model

In this section we present the results of fitting the departure model, given by equation VI.2, to the dependent and independent variables presented in tables 1 and 2 of this chapter for the time periods 1966-71 and 1971-76, respectively. To proceed with this fitting, however, we need the set of the inclusive values I_j, $j = 1, 2, \ldots, 11$ which is treated as an independent variable in its own right. We used formula VI.2 to calculate the inclusive values. The information needed for the calculation is the independent variables used and the parameters estimated in the destination choice model. Thus, for each one of the runs we made for the destination choice model shown in tables 3 to 4 we can have a set of inclusive values calculated, and for each set a separate analysis can be done at the departure level. As we have pointed out, however, in section 1 of this chapter, the small number of degrees of freedom does not allow a serious statistical analysis at this level. We confine, therefore, ourselves to one representative run from each time period and for all the age groups. We therefore choose run 2 of table 3 for the 1971-

76 time period and run 1 of table 4 for the 1966-71 time period. In both these runs the same explanatory variables were used. These were distance, housing age, new housing and population density. Using these variables and the parameters estimated in the destination choice model we calculated the inclusive values shown in table 5.

As explained in chapter IV these values provide the link between the destination choice model and the departure model., In several runs of the departure model with different combinations of explanatory variables we were not able to observe a pattern in the values of the estimated coefficients. This lack of robustness is attributed to the extremely small number of the degrees of freedom. A sample of these runs for both the time periods is shown in table 6, where only the beta weights are provided.

For 1966-71 the fit of the model is very poor ($\rho^2 = .157$) and the value of μ appears to be outside the expected bounds. In addition, "tenure" attains the opposite expected sign which tends to suggest that for an area the higher the proportion of those that rent their dwelling the lower the percentage of those that move outside this area. For 1971-76, however, the overall results seem to be better.

With this limited discussion on the results of the departure model we conclude this section and with it we conclude chapter VI. In this chapter we have examined how the nested logit model,

Area	1966 - 71	1971 - 76
Toronto	0.16	-0.61
York	0.37	-0.29
East York	0.32	-0.25
North York	0.09	-0.60
Scarborough	-0.02	-1.02
Etobicoke	0.23	-0.49
Pickering	-0.67	-1.72
Inner York	-0.13	-1.02
Albion	-0.30	-1.24
Mississauga	-0.31	-1.28
Oakville	-0.45	-1.37

Table 5
Inclusive Values

Variable	1966 – 71	1971 – 76
Inclusive Value	-10.70	-9.36
Pop. Share	7.60	9.17
Tenure	-4.00	4.75
α_0	-2.60	-2.15
r^2	0.157	0.575
R^2	0.419	0.490
μ	-1.016	-0.718

Table 6
Departure Model Beta Weights

introduced in chapter IV, can be interfaced to the intraurban migration data, discussed in chapter V (section 1). Furthermore, we have introduced the independent variables we used in this study, for the destination choice model (section 2) and the departure model (section 3). The results for the destination choice model, discussed in sections 4 and 5 for 1971-76 and 1966-71 respectively, revealed that "Distance", "Housing Age", "New House Completions" and "Population Density" constitute the most statistically significant combination of variables for all the age groups and for both the time periods. Finally, in the present section we discussed the results of the departure model. We now turn to examine the impact of independent variables on the equilibrium population distribution which was introduced and extensively discussed in section III.8. This is the subject of the next chapter.

CHAPTER VII

EQUILIBRIUM HOUSEHOLD DISTRIBUTION VARIATIONS. THE CASE OF NEW HOUSE COMPLETIONS

At the end of section III.7 we have introduced that for a closed migratory system the equilibrium population distribution indicates the tendency of the system for the time to which the equilibrium corresponds. The first task of this chapter (section 1) is to explore the relationship between equilibrium and the tendency of the system in more detail. In particular it is argued that the equilibrium population of a zone indicates the short run tendency in the direction of population change in that zone. It is further argued that the equilibrium population distribution corresponding to time τ is dependent on the actual population distribution at time τ and the matrix (p_{ij}) of the probabilities to migrate between the zones in the system.

The subject of chapter IV was to express the probabilities to migrate between zones in terms of social and economic conditions in the system, as well as the characteristics of individual migrants, with the help of probabilistic choice theory. The models we ended up with in chapter IV were actually calibrated in chapter VI with intrametropolitan migration data for the Toronto CMA and for the two intercencal time periods 1966-71 and 1971-76.

An obvious question then is this: If socioeconomic conditions and the characteristics of migrants affect the probabilities to migrate, which in turn affect the equilibrium population distribution then how do socioeconomic conditions and migrant characteristics affect the equilibrium population distribution? An answer to this question will provide an immediate link between the characteristics of the system and its short run population tendencies.

In this chapter we argue that such a link is not analytically obvious. We can, however, investigate this relationship with the use of a computer program. Thus, the problem of this link is put in sharp focus in section 2. Appendix to chapter VII elaborates on the structure of the computer program that was developed for the purpose of investigating this link. In essense this program allows us to calculate the equilibrium population distribution when one or more of the probability to migrate predictor variables are allowed to vary. In our specific case the program is used to explore the dependency of the equilibrium household distribution in the Toronto CMA for the 1971-76 time period on the number of new house completions in the city of Toronto. The results are discussed in section 3.

It is important to note at this point that sections 2 and 3 should not be approached as a definitive statement on the relationship between equilibrium population distribution and the characteristics of the system. It is rather an attempt to synthesize the elements discussed in previous chapters and provide pointers to an avenue of research that, in this thesis, is left mostly unexplored.

1. Equilibrium

As we noted in section III.4 the migration rates can be written:

$$q_{ij}[t] = f_{ji}[\underset{\sim}{n}[t];\ \underset{\sim}{\Theta}[t]]$$

where $\underset{\sim}{\Theta}[t]$ includes all factors, other than the state of the system, which influence the decision of an individual to migrate.

A well known result is that when the rates q_{ji} do not depend on the state of the system then the probability distribution of the states of the system $P[\underset{\sim}{n};t]$ is approximately a time dependent multivariate normal. In this case the sytem is said to be linear and the mean value state becomes the most probable state. Then using (III.22) we have:

$$\frac{d}{dt}\bar{n}_j[t] = \sum_i \left(q_{ji}[t]\left(\sum_{\underset{\sim}{n}} n_i[t]P[\underset{\sim}{n};\, t]\right) - q_{ij}[t]\left(\sum_{\underset{\sim}{n}} n_j[t]P[\underset{\sim}{n};\, t]\right)\right)$$

$$= \sum_i \left(q_{ji}[t]\bar{n}_i[t] - q_{ij}[t]\bar{n}_j[t]\right). \tag{1}$$

If we divide both sides of this equation by N we see that this is precisely the structure of the associated deterministic system (III.25) which describes the evolution of the most probable state. That is, when the system is linear, the state in the associated deterministic system is the expected state in the stochastic system.

This result is not surprising since, in this case, $P[\underset{\sim}{n};\, t]$ is approximately a multivariate normal and therefore the expected state $\bar{\underset{\sim}{n}}[t]$ closely approximates the most probable state of this distribution.

An observed state $\underset{\sim}{n}[t]$ may be assumed to be the most likely state or at least close to the most likely state. Under this assumption then equation (1) can be written as:

$$\frac{d}{dt}n_j[t] = \sum_i (q_{ji}[t]n_i[t] - q_{ij}[t]n_j[t]) \tag{2}$$

At any moment τ, the set of migration rates $q_{ji}[\tau]$ reflects the prevailing conditions $\underset{\sim}{\Theta}[\tau]$. Thus, if we know the exact functional dependency of q_{ji} on $\underset{\sim}{\Theta}$ and an initial observed state $\underset{\sim}{n}_o$ at time $\tau=0$, we can approximate very closely the real path of the system for

$t > \tau$ by using equation (2). In doing so, however, we are going to face two difficulties.

The first problem is that the exact functional relationship of of q_{ji} on Θ is not known although a plausible model that links the migration probabilities to the socioeconomic conditions Θ was introduced in chapter IV. The second difficulty relates to the nature of the migration data: Migration observations normally relate to a time period Δt rather than a moment t. Recalling our discussion on section III.4, if the migration rates q_{ji} are constant over the observational time period $[t, t+\Delta t]$ then (2) can be written as:

$$\frac{n_j[t+\Delta t] - n_j[t]}{\Delta t} = \sum_i (q_{ji}[t]n_i[t] - q_{ij}[t]n_j[t])$$

or because of (III.12)

$$n_j[t+\Delta t] - n_j[t] = \sum_i (p_{ji}[t,\Delta t]\, n_i[t] - p_{ij}[t,\Delta t]n_j[t]) \qquad (3)$$

Equation (28) is the well known <u>basic demographic equation</u> when the natural population growth is ignored.

Equation (2), with q_{ij} constant, is no different than the basic stationary Markov Model that has extensively been used in the literature as a predictive model (see section II.5). The stationarity assumption, which is synonymous to q_{ij} constant, has been criticized by many authors including Huff and Clark (1978) and Pickles (1980). In the intraurban migration context, on the other hand, equation (2) has been used in connection with the resistence to move literature. In this case q_{ij} was considered a function of the duration a household has stayed in the same place. Or, to put it in a different way, vector $\underset{\sim}{\Theta}$ consisted of only one element which was the duration a household stayed in a house.

Vining (1975) has used equation (2) in a different way. It is well known that when q_{ij} does not change over time then equation (2) represents a homogeneous system of L simultaneous differential equations with constant coefficients (q_{ji}). Such a system has a unique solution $\underset{\sim}{n}^*$ which is dependent on initial population distribution and the values of the coefficients q_{ji}, $i,j = 1,2, \ldots L$. We call $\underset{\sim}{n}^*$ an equilibrium population distribution. As the coefficients q_{ji} change over time then, to any observed state $\underset{\sim}{n}[\tau]$ at any moment $t=\tau$ corresponds an equilibrium $\underset{\sim}{n}^*[\tau]$. This equilibrium population distribution represents the tendency of the migratory system at time τ or the population distribution that would prevail in the long-run if the observed coefficients q_{ji} at time τ remained unchanged indefinitely. In this sense Vining (1975, p. 162) argues that a population distribution $n[t]$ continually chases its equilibrium distribution $n^*[t]$. Tarver and Curley (1965) and Rogers (1967) have also used equilibrium population distributions to analyze migration patterns before Vining.

More recently, Kanaroglou and Papageorgiou (1986) have considered the distance between the observed and the equilibrium population distribution as a measure of disequilibrium in the Canadian regional system. Figure 1 is borrowed from this study. The broken lines indicate the trajectory of the equilibrium population whereas the solid line indicates the real time path of the population for Quebec, Ontario, Saskatchewan and Alberta for the time period 1960 to 1982. Although no adjustment was made for international migration and natural growth, from the figure it is obvious that the real path has the tendency to follow the equilibrium path. Very similar figures are presented by Vining (1975) for Japan.

In the face of all the empirical evidence concerning different migratory systems one is tempted to suggest that if for a region j the equilibrium population $n_j^*[\tau]$ at time τ is greater than the observed population $n_j[\tau]$ at time τ ($n_j^*[\tau] > n_j[\tau]$) then the population of region j will tend to increase immediately after τ, and vice-versa. This suggestion is also intuitively appealing. In mathematical terms it can be expressed as follows:

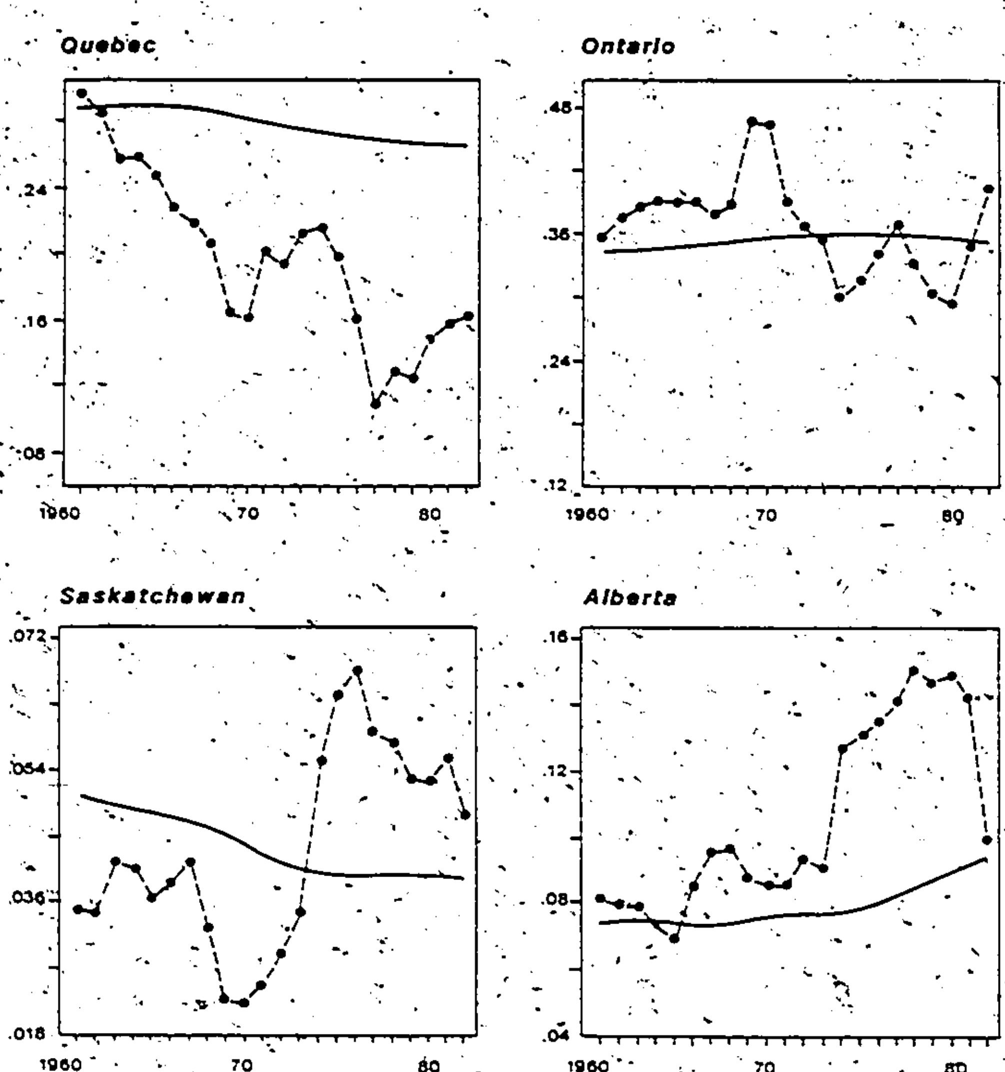

Figure 1: Observed and Steady-states (linear).

"In a closed migratory system let $\underset{\sim}{n}[\tau]$ be the observed population distribution at time τ and $\underset{\sim}{n}^{*}[\tau]$ be the corresponding equilibrium population distribution.
If for any area j in the system we have

$$n_j^{*}[\tau] > n_j[\tau] \qquad (n_j^{*}[\tau] < n_j[\tau])$$

then:

$$\frac{dn_j}{dt}\Big|_{\tau+} > 0 \qquad \left(\frac{dn_j}{dt}\Big|_{\tau+} < 0\right) \text{"}$$

The derivatives are evaluated at a point in time $\tau+$ which represents some time that immediately follows τ. Can we prove analytically a proposition of this kind? The answer unfortunately to this question is no. However, since the monotonicity of n_j as a function of time is dependent on the characteristics of the system; which are the relative magnitude of $\underset{\sim}{n}[\tau]$ and q_{ji} at time τ, perhaps a better way to put this question is: "Are there any conditions (necessary and/or sufficient) under which the above proposition holds?" Despite our efforts we were not able to derive such conditions. The remainder of this section explains the reasons. The general solution to system (2) is given by:

$$n_j[t] = \sum_{k=1}^{L} c_{kj} e^{r_k t} \quad j = 1, 2, \ldots, L \qquad (4)$$

where $r_k, k = 1, 2, \ldots, L$ are the characteristic roots or latent roots or eigenvalues of the matrix of the coefficients (q_{ij}) with i'th diagonal elements

$$s_i = -\sum_{\substack{j=1 \\ j \neq i}}^{L} q_{ji}$$

The dynamic theory of a linear and closed system like (2) is well known (Hearon, 1962, pp. 42-47). In particular, since the system is closed, the matrix (q_{ij}) is singular. The singularity of the (q_{ij}) matrix is easily verified since the sum of the elements in any one column vanishes. Thus, closedness of the system guarantees that one of the latent roots of the (q_{ij}) matrix is zero. Furthermore, if the rank of the matrix is L-1, then the zero root is a single, as opposed to a multiple root. From the rest of the roots none of them is pure imaginary, they are either negative real or complex with negative real parts.

Without any loss of generality let us consider that r_1 is the <u>dominant</u> zero root. Thus, r_1=0. Then, the solution (4) can be written as:

$$n_j[t] = c_{1j} + \sum_{k=2}^{L} c_{kj} e^{r_k t} \qquad j = 1, 2, \ldots, L.$$

When t tends to infinity each term of the summation in the right hand side tends to zero, since r_k, $k = 2, \ldots L$ are all real negative or complex with negative real parts. Even in the case of repeated roots when terms of the type $t \cdot e^{r_k t}$, appear, the summation will still tend to zero (Chiang, 1974, p. 539). This is because:

$$\lim_{t \to \infty} t^l \cdot e^{r_k t} = 0 \qquad \text{for real}[r_k] < 0 \text{ and any integer } l.$$

When an initial condition is given the equilibrium c_{1j} is to be determined in terms of the coefficients (q_{ij}) and the value of the state $\underset{\sim}{n}$ at the initial time τ. There are two cases to consider. Firstly, when the roots r_k, $k = 2, \ldots, L$ are all real then the solution n_j will not be necessarily a monotonic function of time. The form of the function depends on the magnitude and absolute value of the coefficients c_{kj}, $k = 2, \ldots, L$ which, in turn, are dependent on the relative magnitude of $n_j[\tau]$ and $q_{ji}[\tau]$. This is because,

$$\frac{dn_j[t]}{dt} = \sum_{k=2}^{L} c_{kj} r_k e^{r_k t} \qquad j = 1, 2, \ldots, L$$

Secondly, when complex roots are present, the path to equilibrium will be a damped oscillation. Whether $\frac{dn_i}{dt}$ will be positive or negative in the vicinity of τ depends on a number of factors such as the distance of the equilibrium from the initial condition, the amplitude of the oscillation and so on.

In this context, it is instructive to consider the solution of the system (2) in the special case of two regions keeping in mind that the case of three regions is much more complex. Then:

$$\frac{dn_1}{dt} = -q_{21}n_1 + q_{12}n_2$$

$$\frac{dn_2}{dt} = q_{21}n_1 - q_{12}n_2$$

The characteristic equation is:

$$\begin{vmatrix} -q_{21}-r & q_{12} \\ q_{21} & -q_{12}-r \end{vmatrix} = 0$$

The latent roots then are:

$$r_1 = 0 \text{ and } r_2 = -(q_{21} + q_{12})$$

As expected one of the roots is zero and the other is real and negative.

The general solution that will satisfy the given system of equations is then given by:

$$n_1[t] = A_1 + A_2 \cdot e^{-(q_{21}+q_{12})t}$$

$$n_2(t) = \frac{q_{21}}{q_{12}} A_1 - A_2 \cdot e^{-(q_{21}+q_{12})t}$$

Given two initial conditions, the constants A_1 and A_2 can be specified. Consider that in the beginning of the time period of interest τ, the populations $n_1(\tau)$ and $n_2(\tau)$ are known. The specific solution then to the system is given by:

$$n_1[t] = \frac{1}{q_{12}+q_{21}} (q_{12}(n_1[\tau] + n_2[\tau]) + (q_{21}n_1[\tau] - q_{12}n_2[\tau]) e^{-(q_{21}+q_{12}) \cdot t}) \quad (5)$$

$$n_2[t] = \frac{1}{q_{12}+q_{21}} (q_{21}(n_1[\tau] + n_2[\tau]) - (q_{21}n_1[\tau] - q_{12}n_2[\tau]) e^{-(q_{21}+q_{12}) \cdot t}) \quad (6)$$

We can make several interesting observations in equations (5) and (6). Firstly, by adding the two equations we see that the total population is preserved over time, since

$$n_1[t] + n_2[t] = n_1[\tau] + n_2[\tau] \text{ for any } t.$$

Secondly, the equilibrium population of an area is the total population weighted by the ratio of the inmigration rate to the total inmigration and outmigration rate:

$$n_1[t] = \frac{q_{12}}{q_{12}+q_{21}}$$

$$n_2[t] = \frac{q_{21}}{q_{12}+q_{21}}$$

Thirdly, $n_1[t]$ and $n_2[t]$ are both monotonic functions of time since:

$$\frac{dn_1[t]}{dt} = (q_{12}n_2[\tau] - q_{21}n_1[\tau])e^{-(q_{21}+q_{12})\ t}$$

$$\frac{dn_2[t]}{dt} = (-q_{12}n_2[\tau] + q_{21}n_1[\tau]) \cdot e^{-(q_{21}+q_{12})\ t}$$

Because the term $e^{-(q_{21}+q_{12})}$ is positive, n_1 is monotonically increasing if and only if $q_{12}n_2[\tau] > q_{21}n_1[\tau]$. In this case n_2 will be monotonically decreasing. Similarly n_1 will be monotonically decreasing and n_2 monotonically increasing if and only if

$$q_{12}n_2[\tau] < q_{21}n_1[\tau].$$

Unfortunately, the conditions derived in the case of two regions cannot be generalized for more than two regions. We are going to rely on the empirical evidence we have and our intuition and accept that $\underset{\sim}{n}^{*}[\tau]$ shows the tendency of the system at time τ, as Vining (1975) and others have done in the past.

2. Equilibrium Dependency on Predictor Variables

If we consider the observational time period Δt as the unit of time then the probabilities p_{ij} to migrate from a zone j to a zone i is equivalent to the corresponding migration rates q_{ij}. In particular, using equations (III.12) and (2), we have the system of equations:

$$\frac{d}{dt} n_j[t] = \sum_i (p_{ji}[t]\, n_i[t] - p_{ij}[t]\, n_j[t]) \qquad j = 1, 2, \ldots L \qquad (7)$$

Clearly the equilibrium population distribution $n^{*}[\tau]$ for this system is a function of the actual population distribution $\underset{\sim}{n}[\tau]$ at time τ and the values of the elements in the matrix (p_{ij}). Concerning the relationship of $\underset{\sim}{n}^{*}[\tau]$ to $\underset{\sim}{n}[\tau]$ we have argued in section 1 that if for a specific area j, $n_j^{*} > n_j$ then the tendency of n_j will be to increase in the short run and vice-versa. In this section we are more interested in the relationship of $\underset{\sim}{n}^{*}[\tau]$ to (p_{ji}). More specifically, we are interested in what affects (p_{ji}) which, in turn, will affect $\underset{\sim}{n}^{*}[\tau]$.

In general, at any moment τ, p_{ji} is a function of the state of the system $\underset{\sim}{n}[\tau]$ and a vector $\underset{\sim}{\Theta}[\tau]$ with elements corresponding to socioeconomic conditions and migrant characteristics. The empirical evidence in section VI.4 and VI.5 suggests that at least in the intrametropolitan case $n[\tau]$ affects p_{ji} insofar as it is perceived as a gross surrogate for many socioeconomic variables. Given a sufficient number of appropriate variables, the relationship of p_{ji} to $\underset{\sim}{n}[\tau]$ becomes statistically insignificant. We therefore concentrate on the relationship between p_{ij} and $\underset{\sim}{\Theta}$ given by assumption IV.1 and equations IV.22 to IV.26, which collectively constitute the nested logit model.

Estimation of p_{ji} can be done by separately estimating the departure probability $p._j$ and the destination choice probability $p_{j/i}$ and then multiply the two. If one is interested in a predictor variable affecting p_{ji} through $p_{j/i}$ only, then an estimate of $p_{j/i}$ can be obtained through the destination choice model whereas an estimate of $p._j$ can be obtained by simply dividing the population in region i in the beginning of the time period in question into the total number of outmigrants from i during the same time period. We are interested in this particular case since as discussed in section VI.6, estimates of $p_{.j}$ through the departure model are not particularly good.

In the remainder of this section we argue that although the effect of a predictor variable on the destination choice probabilities, as they are expressed through the destination choice model, can be analytically investigated, the effect of the probabilities of $\underset{\sim}{n}^{*}$ cannot. To see the first argument consider the destination choice model:

$$p_{ij} = \frac{\exp(r \cdot Y_i + S_{ij})}{\sum_{k \neq j} \exp(r \cdot Y_k + S_{kj})} \quad i,j = 1, 2, \ldots L; \quad i \neq j \tag{8}$$

where Y_i is the value the destination specific predictor variable of interest Y attains for destination i,

r is the estimated coefficient for Y, and

S_{ij} is the value for the rest of the utility function, excluding Y, for origin j and destination i.

A change in Y_i will affect some of the elements in the $(p_{i/j})$ matrix, while others will be left unaffected. In particular the elements in the i'th column $(p_{k/i}, k = 1, , \ldots L, k \neq i)$ of the matrix will not be affected since all these elements correspond to destinations other than i and therefore are not dependent on Y_i. All the other elements, except of the diagonal elements that are zero, are affected by a change in Y_i. More specifically we have:

$$\frac{\partial p_{i/j}}{\partial Y_i} = r\, p_{i/j}\,(1 - p_{i/j}) \qquad (9)$$

$$j = 1,2,\ldots L;\ j \neq i$$

and

$$\frac{\partial p_{l/j}}{\partial Y_i} = -\, r\, p_{l/j}\, p_{i/j} \qquad (10)$$

$$l = 1,2,\ldots L;\quad l \neq i$$
$$j = 1,2,\ldots L;\quad j \neq i$$
$$j \neq l$$

Equalities (9) and (10) imply that:

$$\frac{\partial p_{i/j}}{\partial Y_i} > 0, \quad \frac{\partial p_{l/j}}{\partial Y_i} < 0 \qquad \text{if } r > 0$$

and

$$\frac{\partial p_{i/j}}{\partial Y_i} < 0, \quad \frac{\partial p_{l/j}}{\partial Y_i} > 0 \qquad \text{if } r < 0$$

Thus, for $r > 0$ an increase in Y_i will cause the values of the conditional probabilities for the i'th row to increase and for any other row to decrease.

The opposite happens for $r < 0$. However, column summations will always add up to one,

$$\sum_{\substack{i=1 \\ i \neq j}}^{L} p_{i/j} = 1 \qquad (11)$$

The discussion of section 1 suggests that the dependency of n^* on (p_{ij}) is not as straightforward as the dependency of (p_{ij}) on Y_i. To investigate the impact of changes in Y_i on n^* we make use of a computer program, the algorithm of which is described in the Appendix to chapter VII.

3. Equilibrium Variations with New House Completions

The algorithm described in the appendix to this chapter was implemented into a computer program, using FORTRAN 77, and applied in the case of the Toronto CMA for the 1971-76 time period.

As pointed out in section 2, to calculate the equilibrium distribution iteration we need first an estimate of the probabilities in $(p_{i/j})$. This is obtained by making use of the statistical run that appears as run 2 on table VI.6 which was considered to be the best run when the migration household heads of all ages were used.

Four predictor variables were used in this run. These were; distance, the percentage of houses built before 1946 (Housing Age), new house completions (New Housing) and population density. The equilibrium household distributions were calculated when the "New Housing" variable was allowed to vary for one of the zones, the city of Toronto. The observed value of the "New Housing" variable for the city of Toronto was 16362 (see table VI.3). This variable was allowed to vary between 10362 and 22362 with increments of 200 housing units. The results are shown in figure 2. The real number of households in 1971 in the city of Toronto was 224695. The equilibrium number of households found in "Equilibrium Distribution (1)" (explained in Appendix to this chapter) was 98040, which is substantially lower than 224695, the real number of households. This result is consistent with the discussion in section 1, since excluding migration in and out of the CMA and household formation, Toronto City was the major looser of households (see table V.10) for the 1971-76 time period. The estimated equilibrium number of households for the observed number of new house completions was 125057, relatively close to the actual equilibrium number of households, and still below the real number of households.

The equilibrium number of households depends on the initial distribution of households in the CMA as well as the estimates for

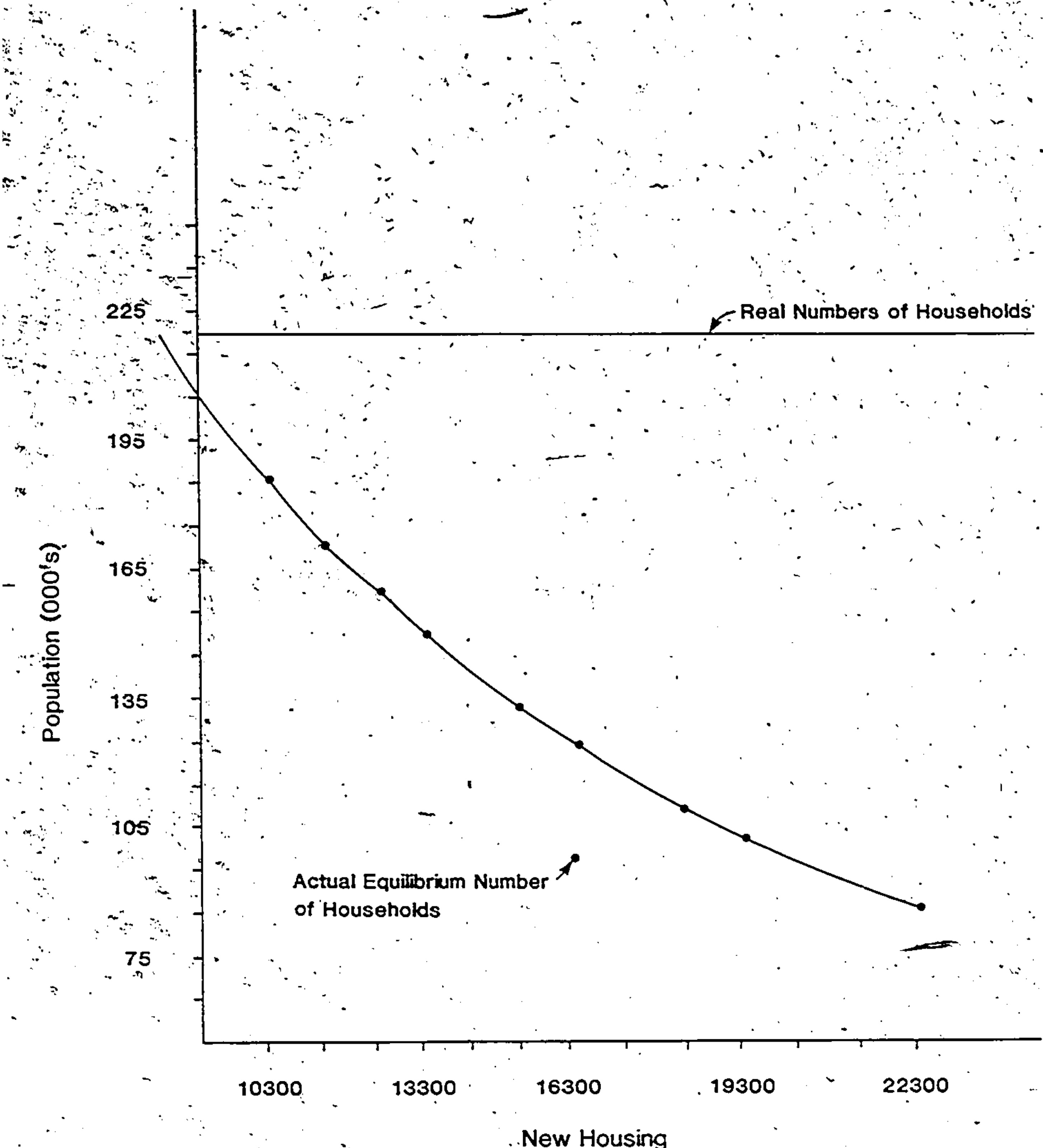

Figure 2. TORONTO CITY-ACTUAL & ESTIMATED EQUILIBRIUM

the probabilities to migrate between zones. In this case the initial distribution of households in the CMA is kept constant while the probabilities (p_{ij}) change as the value of the "New Housing" variable changes.

Referring to the analysis in section 2 since the coefficient of "New Housing" (corresponding to r) was found to be positive, when the value of "New Housing" increases, the probabilities to migrate into Toronto City from the other zones in the system increase whereas the probabilities to migrate from Toronto City to all the other zones remain constant.

The relationship between new house completions and equilibrium number of households, calculated by the program, appears to be a simple monotonically decreasing function, as shown in figure 2. The estimated values of the equilibrium number of households is consistently lower than the real number of households for all the range of the "New Housing" variable studied. These results seem to indicate that, other things being equal, when the number of new house completions in Toronto City gets sufficiently small (approximately 9000) the tendency for Toronto City will be reversed. In other words, Toronto City will have the tendency to gain in number of households.

This result is certainly counter-intuitive and is related not only in the direction of change of the probabilities as "New Housing" increases but also on the magnitude of the change. This is particularly so given the fact that the probabilities to move between zones other than Toronto City are all decreasing as "New Housing" increases. The complex interplay between changes in these probabilities and the effect on the equilibrium distribution is, at best, not well understood at this stage. The topic is left as an open ended research question. It is anticipated that considerable amount of computer time will be required in order to disentangle the various effects.

CONCLUSION

The basic endeavour of the research undertaken in this thesis has been to put forward a general formal framework within which migration can be studied. Ideally, we should have explored the theoretical and empirical possibilities offered by the proposed framework exhaustively. The magnitude of the task, however, and limitations in data availability forced us to leave many questions without an answer. The conclusion, presented here then, summarizes our basic arguments and findings and compiles a research agenda of issues yet to be investigated.

The basic set of ideas through which migration is viewed makes up what we call a conceptual framework. A specific conceptualization of migration gives rise to specific types of migration models that can be either descriptive or formal. In this thesis we are interested in the latter.

The objective of the literature review in chapter II is an attempt to classify the migration literature into conceptual frameworks. This classification, although fuzzy, allows to conveniently relate our conceptualization to the existing frameworks. Three frameworks, the schematic, the economic and the behavioural, were identified. Although traditionally studies in what we termed "the economic framework" have dealt with interegional migration of the macro type, a recent tendency within this framework is towards more behavioural approaches and the usage of microdata. In the same time several advances have been made within the behavioural framework. The conceptualization we adopt could be termed behavioural. It has its roots in probabilistic choice theory and is expressed by the nested logit model.

Another characteristic of the adopted conceptualization is that it is evolutionary in natute. As such it is considered as a generalization of the existing dynamic Marcov type models of migration. This aspect of the framework is modeled by the master equation. The link between the macroevolution of the system, expressed by the master equation, and the microbehaviour of individual decision makers in the system, expressed by the logit model, is provided by the migration rates or equivalently by the probability of an individual migrating between two zones in the system.

The detailed derivation and the underlying assumptions of the master equation and the nested logit model are discussed in detail in chapters III and IV respectively. A deterministic approximation to the master equation is also discussed in chapter III. This is expressed by what was termed the associated deterministic system which is a dynamic set of equations, similar to the usual Marcov type equations, except that the migration rates are not constant.

We claimed that the proposed models are suitable for any type of migration. We attempted to demonstrate the application of the models on intrametropolitan migration in the Toronto Census Metropolitan Area (CMA). The migration data at our disposal are available for twelve zones in the CMA, for two time periods, 1966-71 and 1971-76, and they refer to migrations of household heads. In retrospect we can say that the zones in our data base are too large and the five year time periods too long to allow a detailed enough exploration of the phenomena. Moreover, the very small number of time periods for which the data are available do not allow any detailed exploration of the dynamic properties of the system. We demonstrate, however, that despite the limitations of our data, the results obtained are meaningful.

In chapter V we provided a descriptive analysis of the data. We also demonstrated how a closed migratory system can be achieved by accounting for household formation and dissolution, since according to our claims of chapter III the dynamics of a closed system are considerably simpler in comparison to those of an open one.

In chapter VI we calibrate the nested logit model to the data. Due to limitations of the data mentioned above, only one part of the model, the destination choice model, provides robust and meaningful statistical results. Out of thirteen independent variables used, four appeared to be consistantly significant in several runs. These were; the logarithm of distance, the percentage of houses built before 1946, the number of new house completions during the study period and the density of the population. Runs with three different age groups demonstrate that the importance of distance is increasing with age whereas that of the other three factors is decreasing.

To explore some of the dynamic properties of the system, in chapter VII, we considered the associated deterministic system to the master equation with fixed migration rates. We argued that in this case, to an observed population distribution and a set of migration rates, through this system of equations, corresponds a unique equilibrium population distribution which indicates the tendency of the migratory system for the moment of observation. Furthermore, we linked this equilibrium population distribution to the social and economic conditions in the system at the time of observation. We show how, other things being equal, the tendency of the system would vary with the number of new house completions in

Toronto City.

The framework proposed in this thesis can be expanded and/or used in several ways. Although for any realistic migratory system the probability distribution over the states of the system is expected to have a low variance, this is no more than a mere speculation and requires verification. Starting from observed states, Monte Carlo simulation techniques can be employed to identify the shape of the probability distribution. If it is unimodal and of low variance, then the mean value equation can be used as a very close, deterministic approximation to the stochastic master equation that models the most likely trajectory of the system.

It has been emphasized within this thesis that ideally the nested logit model should be used with disaggregated survey type of data. In our case we have used the model with aggregate data and we have argued that the inadequacy of the data should not reflect on the model itself. Thus, further research is needed with more disagreegated data to establish the effectiveness of the model in capturing the essense of the phenomena of interest.

Although we have demonstrated that with a simple and plausible assumption, such as IV.2, the effect of the time interval used can be captured in a simple parameter that can be estimated, the same is not true with respect to space. Capturing space effects in any model is not an easy task, since space in a migratory system can vary in terms of the size of the zones used as well as their geometry. The study of the impact of the geometry of space on the results obtained by the nested logit model has recently been undertaken by Boots and Kanaroglou(1987) while the study of the impact of zonal size remains an open research question.

Finally, the effect of social and economic conditions on the equilibrium population distribution needs to be investigated in detail. The results of the impact of the new house completions in Toronto City on the equilibrium population distribution, discussed in section VII.3 are only preliminary and need a more thorough investigation.

APPENDIX 1 TO CHAPTER VI

RELATIONAL VARIABLES FOR THE DESTINATION CHOICE MODEL

This appendix provides information for the relational variables used in the calibration of the destination choice model. These variables represent measures of physical and social distance among the areas in the Toronto CMA migratory system. Only the physical distance variable is available for both the time periods under study. The social distance variables (occupation, education, income, religion and household size) were available only for 1971.

The distance matrix, obtained as explained in section 2 of chapter VI, is shown in table 1. Definitions and discussion for the social status variables follow:

Education - The source for this variable was table 1 in Statistics Canada (1973a). The population of each area was partitioned in six groups:

(1) Less than grade 9

(2) Grade 9 and 10

(3) Grade 11

Area	Toronto	York	East York	North York	Scarborough	Etobicoke	Pickering	Inner York	Albion	Mississauga	Oakville
Toronto	0.00	7.75	6.25	10.75	14.50	14.00	32.50	21.50	25.00	21.25	34.50
York	7.75	0.00	11.25	9.00	18.75	6.75	37.25	18.25	20.50	17.25	32.25
East York	6.25	11.25	0.00	8.75	8.00	18.25	26.50	18.75	31.25	27.25	40.75
North York	10.75	9.00	8.75	0.00	12.75	14.00	30.00	10.75	24.00	25.75	41.25
Scarborough	14.50	18.75	8.00	12.75	0.00	25.25	18.50	18.75	37.50	35.25	48.75
Etobicoke	14.00	6.75	18.25	14.00	25.25	0.00	43.75	20.75	14.50	12.50	28.75
Pickering	32.50	37.25	26.50	30.00	18.50	43.75	0.00	31.25	54.25	53.50	65.50
Inner York	21.50	18.25	18.75	10.75	18.75	20.75	31.25	0.00	25.25	33.00	49.25
Albion	25.00	20.50	31.25	24.75	37.50	11.50	54.25	25.25	0.00	18.75	32.75
Mississauga	21.25	17.25	27.25	25.75	35.25	12.50	53.50	33.00	18.75	0.00	16.25
Oakville	34.50	32.25	40.75	10.25	48.75	28.75	65.50	49.25	32.75	16.25	0.00

Source: See Text

Table 1: The Distance Matrix (Km)

(4) Grade 12 and 13

(5) Some University

(6) University degree

Each of the five groups were further partitioned into two subgroups:

(1) No other training

(2) With other training

Thus we have a total of eleven groups.

The dissimilarity indices computed ranged from 1.4, between Mississauga and Oakville, to 14.8 between York and East York. In fact, York appears to be highly dissimilar to most of the other areas due to its relatively high concentration of people with less than grade 9 schooling.

<u>Household Size</u> - The source of this variable was Statistics Canada (1973b). The households in each area were partitioned in six groups. These were:

(1) 1-person households

(2) 2-person households

(3) 3-person households

(4) 4 and 5 person households

(5) 6 to 9-person households

(6) 10 or more-person households

From the dissimilarity indices computed it appears that North York and Etobicoke were very similar in household size structure with an index of 0.7. The highest dissimilarity (33.6) is recorded between East York, with the smallest households and Pickering, with the highest concentration of large households.

Religion - Source of this variable is table 1 of Statistics Canada (1973a). The need for this variable emanates from the highly visible spatial concentration of ethnic groups in Toronto. It is anticipated that households gravitate to areas with high concentration of households of the same cultural background. During the census data were gathered by asking respondents to give a specific religious body, denomination, sect or community in answer to the question "What is your religion?", even if they did not attend a place of worship. Provision was made for marking "no religion" if the person considered this to be an appropriate answer. All the categories provided were the following:

(1) Anglican

(2) Baptist

(3) Greek Orthodox

(4) Jewish

(5) Lutheran

(6) Pentecostal

(7) Presbyterian

(8) Roman Catholic

(9) Salvation Army

(10) Ukrainian Catholic

(11) United Church

(12) No religion

To what extent religion represents cultural background is probably questionable. Perhaps a better variable in this respect would have been a dissimilarity index based on the ethnic group data provided in the same table. In the 1971 census to collect the ethnic group data, each person was asked to reply to the question: "To what ethnic or cultural group did you or your ancestor (on the male side) belong on coming to this continent?" The end result is 11 ethnic groups. The disadvantage of this classification was thought to be that the Jewish group is not represented and the English, Irish, Scottish and Welsh are lumped together under the title British Isles. We have thus chosen the religion variable. From the dissimilarity indices calculated it appears that Oakville and Scarborough are very similar in religious structure with a dissimilarity index of 4.2. On the other hand Outer York and York appear to be the most dissimilar with an index of 39.8. This is mainly due to two reasons:

(1) A 49% concentration of Roman Catholics in York as opposed to 18% in Outer York.

(2) A 12% concentration of followers of the United Church as opposed to 33% for Outer York.

Occupation - The source of this variable was table 2 in Statistics Canada (1973b). According to 1971 Census, Occupation refers to the specific kind of work a person did in his/her job. Data relate to the respondent's job in the week prior to enumeration if he/she had a job during that week or the job of longest duration since January 1, 1970 if not employed in that week. The detailed data are grouped within three levels of classification, the Occupation Major Groups, Minor Groups and Unit Groups. The Major Groups are as follows:

Group 11	Managerial, Administrative and Related Occupations
Group 27	Teaching and Related Occupations
Group 31	Occupations in Medicine and Health
Group 21	Occupations in Natural Sciences, Engineering and Mathematics
Group 23	Occupations in Social Sciences and Related Fields
Group 25	Occupations in Religion

Group 33 Artistic, Literary, Recreational and Related Occupations

Group 41 Clerical and Related Occupations

Group 51 Sales Occupations

Group 61 Service Occupations

Group 71 Farming, Horticultural and Animal Husbandry Occupations

Group 73 Fishing, Hunting, Trapping and Related Occupations

Group 75 Forestry and Logging Occupations

Group 77 Mining and Quarrying including Oil and Gas Field Occupations

Group 81/82 Processing Occupations

Group 83 Machining and Related Occupations

Group 85 Product Fabricating, Assembling and Repairing Occupations

Group 87 Construction Trades Occupations

Group 91 Transport Equipment Operating Occupations

Other (for males) Materials Handling and Related Occupations, n.e.c., Other Crafts and Equipment Operating Occupations and Occupations Not Elsewhere Classified

Other (for females) - Same groups as for males, plus Groups 87 and 91

In the data provided groups 21, 23, 25 and 33 were lumped into one group. The same happened with groups 71, 73, 75, 77 and 83, 85. Thus, we have thirteen groups for males and only eleven groups for females, since groups 87 and 91 are included in the group "other", a total of twenty-four groups.

The dissimilarity indices calculated, indicate that Scarborough and Etobicoke are very similar in the occupational structure of their populations (dissimilarity index 4.2). At the other extreme the highest dissimilarity (index 21.2) is between York and Inner York. This is mainly due to the higher concentration of managerial, administrative and sales occupations (groups 11 and 51 make up 19% of the occupations) in Inner York as opposed to York (with 7% concentration in these occupations).

Income - The source of these data is table 3 of Statistics Canada (1973b). Income in this case refers to household income which is defined as the sum of the incomes received by all members of the household 15 years and over from all sources during the calendar year 1970. Households in each area are classified in eight income categories, as follows:

(1) Under $1,000

(2) $1,000 - $2,999

(3) $3,000 - $4,999

(4) $5,000 - $6,999

(5) $7,000 - $9,999

(6) $10,000 - $14,999

(7) $15,000 - $19,999

(8) $20,000 and over

On the basis of the income dissimilarity indices calculated we can loosely classify the areas into three groups. The first group consisting of North York, Etobicoke, Inner York, Mississauga and Oakville exhibit a higher concentration of high income households. The dissimilarity index between all of them is less than 7 with a minimum 2.1 between Inner York and Etobicoke. The second group consists of Scarborough, Pickering and Albion with a higher concentration of medium size incomes. The dissimilarity index between them is less than 5.3. Finally, the rest of the areas (Toronto City, York and East York) exhibit the highest concentration of low incomes. This classification is very much dependent on the size of the areas. At the census track level, for example, one can find areas in Toronto City of high income concentrations.

We now turn to describe the destination specific variables used in the destination choice model.

APPENDIX 2 TO CHAPTER VI

DESTINATION SPECIFIC VARIABLES FOR THE DESTINATION CHOICE MODEL

This appendix provides a description of the way the destination specific variables for the destination choice model are measured. Brief comments related to the values obtained for the variables and the hypothesized relationship to the dependent variable are also provided. The values of the variables for the two time periods are in tables 1 and 2, at the end of the appendix.

Population Share - The population share of a destination was used as a gross surrogate of availability of amenities, public facilities and entertainment opportunities. Thus, the sign of the corresponding coefficient is expected to be positive. The population share as opposed to the absolute population was used in order to facilitate temporal comparisons. Table 2 of chapter V was used to compute the population share for 1966 and 1971. The results for the eleven areas are shown in tables 2 and 3. The most notable change that has occurred between 1966 and 1971 is that the share of Toronto has decreased and that of North York and Mississauga has increased.

Population Density - The population density was perceived as a measure of overcrowding and its implications for an area. Therefore, the corresponding parameter is expected to appear with a negative sign.

To compute the population density, the land area in square miles of each zone was obtained from table 4 of Statistics Canada (1968). We note here that the land area of Albion was not available in this table and it was roughly estimated by comparison of the area of Albion to the area of the surrounding regions in the Toronto Metropolitan area map. The land area of the regions was then divided into the 1966 and 1971 population, provided in table V.2 to obtain the density of the population in persons per square mile. Because the population of the metropolitan area grew between 1966 and 1971, the 1966 densities computed with the above method were consistently lower than the 1971 densities. To allow comparison of the densities and the estimated coefficients between 1966 and 1971, the 1966 densities were multiplied by the ratio of the 1971 metropolitan population to the 1966 metropolitan populations. The end result from these operations is shown in tables 2 and 3. The density has declined substantially in Toronto and York while that of North York and Mississauga has increased.

Housing Value - This is the median value of owner occupied dwellings available from table 2 of Statistics Canada (1973b). This variable is available only for 1971.

The sign of the estimated coefficient for this variable will depend on the type of households that make up the migration stream.

New House Completions - The source of this variable is the monthly records on new house completions maintained by Canada Mortgage and Housing Corporation (CMHC). The values shown in tables 2 and 3 are the aggregations of these monthly figures for the five-year time period in each case (e.g. June 1966 to May 1971). The number of new house completions in Toronto, East York and North York has decreased for 1971-76 by comparison to that for 1966-71. By contrast, the number of new house completions has increased for all the other zones with most of the increase for Inner York, Albion and Mississauga. It is thus obvious that development of new housing increased in the suburbs and decreased in the Inner City. It is expected that this variable will correlate positively with the migration stream.

Housing Age - Table 2 of Statistics Canada (1973b) provides data about the number of occupied houses per zone built before 1946. This information allows us to calculate for each zone the percentage of houses built before 1946, by dividing the total number of houses built before 1946 by the total number of houses available in each zone in 1971. The result of this operation is shown in column 6 of table 3. It is not surprising that the highest percentage of older housing is in Toronto City followed by York and East York. By contrast the lowest percentage of older housing is in North York, Scarborough and Mississauga which represent communities mostly established after 1946.

Unfortunately the same type of data are not provided in the 1966 census. To derive the same variable for 1966 we make the assumption that in 1966 the same number of houses built before 1946 existed in each zone as the number found in 1971. In essense we assume that no houses built up to 1946 were demolished in the time period 1966-71. This would probably not be a bad assumption if the demolition had taken place evenly over the metropolitan area. In this case the variance of the variable would be preserved and only the value of the estimated coefficient would be lower than what it should be. It is, however, unlikely that the demolition was evenly distributed since as we saw from table 3 there is a much higher percentage of older housing in the inner city as opposed to that of the suburbs.

With the above problems in mind, in each zone the total number of houses found in the 1971 census to have been built before

1946 was divided by the total number of houses found in the 1966 census and the resulting ratio was multiplied by 100. The resulting figures were consistently higher than those shown in table 3 because the number of occupied dwellings in every zone in 1966 was lower than that of 1971. To make the two time periods comparable the 1966 number of occupied dwellings in each zone was multiplied by the ratio of the 1971 occupied dwellings to the 1966 occupied dwellings for the whole metropolitan area. The calculation of the percentages was then done with these new figures. The end result is shown in column five of table 2.

Comparing the calculated figures in tables 2 and 3 we see that the 1966 figures are lower (higher) than the 1971 figures for the inner city (periphery). This is due to the fact that the distribution of the total number of dwellings has changed in the CMA from 1966 to 1971. The percentage of dwellings in the suburbs has increased and that of the inner city has decreased.

It is expected that the estimated parameter will assume a negative value.

<u>Housing Availability</u> - This variable represents the percentage of the CMA dwellings available in a zone. The source for this variable is the households per zone columns of table V.2 for 1966 and 1971.

The values for these variables are shown in columns 6 and 7 of tables 2 and 3, respectively. We can see how the distribution of housing has changed over the metropolitan area, between the two time periods. The share of Toronto City and York has decreased and that of North York and Mississauga has increased. We note that a housing availability variable should include all the houses available in each zone. In our case, however, only the occupied houses are included since the number of vacancies per zone are not known.

It is expected that migrants will gravitate to zones of housing availability. Thus, the estimated coefficient for this variable is expected to be positive.

Employment - Since proximity to work is traditionally considered to be one of the important factors of intraurban mobility the number of available jobs per zone was included as a variable. The source for this variable was a special tabulation from Statistics Canada based on information from the 1971 Census. The variable, therefore, is only available for 1971. This tabulation provides the number of individuals 15 years old and over in the employment labour force residing in the province of Ontario by municipality of work for selected counties of residence. The selected counties include the Toronto Metropolitan area and a sizable area around it, so that the number of persons living outside these counties and working in the Toronto Metropolitan area is negligible. The employment provided this way has been integrated to the level of the zones used in this study. The result is shown in table 3. Not surprisingly 50% of the jobs are in Toronto City.

The coefficient of this variable is expected to appear with a positive sign, although its contribution to the overall goodness of fit of the model is not expected to be significant.

Area	Pop. Share	Pop. Density Per/mile	New Housing Completions	Houses Built Before 1946 (%)	Occupied Dwellings (%)
Toronto	0.32	21339	28214	62.84	31.76
York	0.07	18686	4123	41.41	7.06
East York	0.40	13343	7875	36.41	5.30
North York	0.18	6715	43968	5.98	18.00
Scarborough	0.13	4510	23076	7.58	11.80
Etobicoke	0.12	6350	11411	12.04	12.32
Pickering	0.02	391	2355	24.10	1.56
Inner York	0.03	349	5073	20.68	2.81
Albion	0.03	311	6359	20.06	2.49
Mississauga	0.05	1137	22506	12.43	4.58
Oakville	0.02	549	1092	14.49	2.25

Table 1: 1966 Destination Specific Independent Variables

Area	Pop. Share	Pop. Density Per/mile²	Median Housing Value ($)	New House Completions	Houses Made Before 1946 (%)	Occupied Dwelling (%)	Employment
Toronto	.28	19003	32542	16362	66.54	29.99	524530
York	.06	16459	27981	5632	47.08	6.22	33225
East York	.04	12763	28840	4565	38.36	5.11	18830
North York	.20	7382	36919	33334	5.51	19.52	148770
Scarborough	.13	4720	31218	28564	7.33	12.23	88245
Etobicoke	.11	5908	34117	16177	13.50	10.99	94330
Pickering	.02	406	24050	3190	24.04	1.57	12075
Inner York	.03	388	35311	9479	19.35	3.01	29175
Albion	.03	369	31416	14155	17.60	2.84	30285
Mississauga	.07	1583	36127	27623	9.13	6.27	64166
Oakville	.02	587	31487	5068	14.65	2.22	23550

Table 2 : 1971 Destination Specific Independence Variable

APPENDIX 3 TO CHAPTER VI

RESULTS FOR AGE GROUPS - DESTINATION CHOICE, 1971-76

This appendix presents and discusses the results of the statistical analysis done for 1971-76 and for the following three migrant head age groups:

(1) Migrant household heads of age 15-34

(2) Migrant household heads of age 35-54

(3) Migrant household heads of age 55 and over.

The information from these runs is presented in tables 1, 2 and 3. Observing these three tables we can make the following points.

- From run 2 we can see that distance, when making a move, is an important consideration for all the age groups. From the beta weights, however, we can say that distance is a factor of lesser importance for the young age group rather than the households with heads of over 35 years of age.
- Population density (run 2) is of more importance for younger households. The beta weight drops from the -15.4, -14.0 level to the -4.2 level for the elderly group. This is taken as an indication that the elderly would gravitate towards more densely populated

Variable	1	2	3	4	5	6	7	8	9	10
Distance	-10.60 (-8.10)	-10.40 (-7.20)	-4.30 (-4.70)	(-1.40) (-1.30)	-10.20 (-7.20)	-5.70 (-5.60)	-4.60 (-4.80)	-7.30 (-6.40)	-1.30 (-1.00)	-9.00 (-7.10)
Education	-1.90 (-1.60)	—	—	—	—	—	—	—	—	—
Household Size	0.70 (.90)	—	—	—	—	—	—	—	—	—
Income	1.30 (1.70)	—	—	—	—	—	—	—	—	—
Occupation	0.70 (.70)	—	—	—	—	—	—	—	—	—
Religion	-1.40 (-1.10)	—	—	—	—	—	—	—	—	-1.00 (-.60)
Population	—	—	—	—	-0.70 (-.70)	5.70 (5.00)	—	—	—	—
Housing Value	—	—	—	—	—	—	—	—	—	0.30 (.40)
Housing Age	8.00 (11.00)	10.50 (12.40)	—	-0.60 (-.60)	10.40 (12.60)	—	—	—	—	10.20 (12.40)
New Housing	12.30 (8.90)	14.50 (9.70)	7.10 (4.00)	5.00 (4.20)	10.90 (10.10)	—	3.60 (3.70)	4.00 (2.60)	7.50 (4.50)	11.00 (9.30)
Employment	—	—	—	—	—	—	—	5.50 (3.00)	—	—
Pop. Density	-12.30 (-14.60)	-12.60 (-15.40)	-5.10 (-5.60)	—	-11.00 (-15.00)	-7.50 (-9.10)	-3.90 (-5.90)	-7.00 (-10.20)	—	-11.90 (-15.20)
Housing Avail.	—	—	—	—	—	—	0.30 (.50)	—	—	—
ρ^2	0.793	0.774	0.444	0.330	0.771	0.396	0.449	0.623	0.325	0.772
$\bar{\rho}^2$	0.726	0.679	0.529	0.274	0.678	0.465	0.526	0.620	0.267	0.650
D.F.	90	95	96	96	94	96	95	95	97	93

Table 1: 1971-76 Destination Choice Results Migrant Heads of Age 15 - 34

Variable	1	2	3	4	5	6	7	8	9	10
Distance	-12.50 (-9.90)	-13.10 (-9.50)	-9.30 (-8.00)	-5.10 (-4.10)	-12.90 (-9.50)	-9.70 (-8.00)	-9.20 (-7.90)	-11.00 (-9.00)	-3.70 (-2.80)	[illegible] ([illegible])
Education	-2.10 (-1.80)	—	—	—	—	—	—	—	—	—
Household Size	-1.30 (-1.70)	—	—	—	—	—	—	—	—	—
Income	1.60 (2.10)	—	—	—	—	—	—	—	—	—
Occupation	0.60 (.60)	—	—	—	—	—	—	—	—	—
Religion	-1.60 (-1.30)	—	—	—	—	—	—	—	—	-3.00 (-1.90)
Population	—	—	—	—	-0.30 (-.30)	7.30 (5.30)	—	—	—	—
Housing Value	—	—	—	—	—	—	—	—	—	2.00 (1.40)
Housing Age	6.50 (8.60)	6.70 (8.70)	—	-3.10 (-3.70)	6.50 (8.00)	—	—	—	—	7.10 ([illegible])
New Housing	10.70 (8.10)	11.80 (8.50)	9.40 (4.40)	3.80 (3.10)	8.80 (8.60)	—	9.20 (4.60)	6.70 (3.30)	9.00 (5.10)	[illegible] (7.40)
Employment	—	—	—	—	—	—	—	4.20 (4.30)	—	—
Pop. Density	-11.00 (-13.70)	-11.30 (-14.00)	-8.00 (-7.70)	—	-9.70 (-13.00)	-10.30 (-11.40)	-5.00 (-7.50)	-9.20 (-10.40)	—	[illegible] (-13.40)
Housing Avail.	—	—	—	—	—	—	-0.30 (-.40)	—	—	—
ρ^2	0.807	0.788	0.658	0.518	0.766	0.505	0.651	0.726	0.440	[illegible]
R^2	0.800	0.812	0.771	0.500	0.812	0.720	0.773	0.795	0.419	[illegible]
D.F.	90	95	96	96	94	96	95	95	97	[illegible]

Table 2: 1971 – 76 Destination Choice Results Migrant Heads of Age [illegible]

Variable	1	2	3	4	5	6	7	8	9	10
Distance	-12.00 (-9.40)	-13.00 (-9.40)	-13.90 (-8.90)	-14.40 (-7.60)	-13.00 (-9.70)	-16.20 (-10.10)	-14.50 (-9.30)	-14.00 (-10.00)	-13.70 (-7.00)	-12.90 (-9.00)
Education	-0.50 (-.60)	—	—	—	—	—	—	—	—	—
Household Size	-0.70 (-.70)	—	—	—	—	—	—	—	—	—
Income	1.90 (2.10)	—	—	—	—	—	—	—	—	—
Occupation	-1.30 (-1.20)	—	—	—	—	—	—	—	—	—
Religion	-2.30 (-1.70)	—	—	—	—	—	—	—	—	-3.10 (-1.70)
Population	—	—	—	—	2.80 (2.50)	13.10 (5.20)	—	—	—	—
Housing Value	—	—	—	—	—	—	—	—	—	1.30 (.90)
Housing Age	1.60 (1.90)	1.40 (1.80)	—	-2.00 (-1.90)	0.90 (1.10)	—	—	—	—	1.60 (1.80)
Low Housing	6.70 (4.50)	7.30 (5.10)	[illegible]	6.90 (3.60)	3.00 (2.90)	—	4.30 (2.80)	8.70 (3.40)	13.70 (4.70)	8.10 (4.40)
Employment	—	—	—	—	—	—	—	3.70 (2.40)	—	—
Pop. Density	-3.60 (-4.30)	-3.50 (-4.20)	-4.30 (-2.70)	—	-4.40 (-5.50)	-11.30 (-6.90)	-4.90 (-4.40)	-6.00 (-4.70)	—	-3.40 (-4.00)
Housing Avail.	—	—	—	—	—	—	-2.60 (2.20)	—	—	—
ρ^2	0.840	0.821	0.819	0.804	0.835	0.818	0.832	0.841	0.757	0.837
R^2	0.776	0.764	0.759	0.714	0.767	0.748	0.763	0.771	0.677	0.776
D.F.	90	95	96	96	94	96	95	95	97	95

Table 3 : 1971 – 76 Destination Choice Results Migrant Heads of Age over 55

areas and give less consideration to the negative aspects of overcrowding in favor of more public facilities and services.

- Areas with a higher level of new house completions such as North York, Mississauga and Scarborough attract younger households. Although the "New Housing" variable is significant for all the age groups (run 2) the beta weight drops from 9.7 for the younger group to 5.1 for the elderly.
- The significance of the "Housing Age" variable is dropping with age and it becomes insignificant for the elderly group. Recall that the "Housing Age" variable has been interpreted as representing the positive aspects of high population densities. Apparently, in this respect, the traditional variable of "Population Share" represents these aspects better for the elderly group as we can see from run 6.
- As expected, employment appears to be of more importance to the younger age groups. Evidence of this statement is provided in run 8. "Employment" is a significant variable with a beta weight dropping from 6.0 for the young age group to 4.3 for the middle age group and to 2.4 for the elderly.
- As pointed out earlier on "Religion" does not appear to be a critical explanatory variable. By looking at run 10 however we see that this variable is insignificant for the younger age group

but significant for the other two groups (t-ratio -3.0 and -3.1). This result tends to suggest that in the 1971-76 time period households with heads of age over 35 would tend to move towards destinations with a higher percentage of people of the same religious background. Perhaps this is due to the fact that the younger age group contains a lower percentage of first generation immigrants.

- Finally from run 10 we see that the "Housing Value" variable becomes significant mainly for households with middle age household heads. The positive sign of the t-ratio is indicative of the fact that most of the moves for the middle age group were towards North York, Mississauga and Inner York, where the more expensive new homes were built. The high percentage of renters among the households with young age heads renders this variable insignificant in table 1.

APPENDIX 4 TO CHAPTER VI

RESULTS FOR AGE GROUPS - DESTINATION CHOICE, 1966-71

This appendix presents and discusses the results of the statistical estimation done for 1966-71 and for the following three age groups of household heads.

(1) Migrant household heads of age 15-34

(2) Migrant household heads of age 35-54

(3) Migrant household heads of age 55 and over.

Within the limited context of the smaller number of explanatory variables we can make the same observations as those of the time period 1971-76.

The importance of the "New Housing" variable is diminishing with age. The variable is significant for all the age groups and the beta weight is falling in run 1 from 8.8 for the 15-34 age group to 6.4 and 5.9 for the other two groups.

The "Population Share" variable becomes a better predictor of the dependent variable with increasing age as we can see from run 4 in tables 1, 2 and 3. The *t*-ratio improves substantially from

Variable	1	2	3	4
Distance	-5.6 (-5.4)	-3.0 (-3.6)	-5.7 (-5.7)	-4.0 (-4.5)
Population	--	--	1.3 (2.5)	4.9 (5.7)
Housing Age	7.0 (11.6)	--	6.7 (11.2)	--
New Housing	9.4 (8.8)	5.4 (3.9)	5.6 (7.5)	--
Pop. Density	-10.0 (-19.6)	-6.1 (-9.1)	-9.0 (-21.1)	-7.2 (-12.9)
# of Variables Used	4	3	5	3
ρ^2	0.602	0.322	0.600	0.339
R^2	0.505	0.440	0.510	0.421
D.F.	96	97	95	97

Table 1: 1966 - 71 Migrant Heads of Age 15 - 34

Variable	1	2	3	4
Distance	-7.9 (-6.9)	-5.6 (-5.9)	-8.4 (-7.4)	-6.8 (-6.5)
Population	--	--	3.1 (5.3)	5.5 (6.2)
Housing Age	5.1 (7.7)	--	5.2 (7.3)	--
New Housing	7.7 (6.4)	5.1 (3.2)	3.3 (3.8)	--
Pop. Density	-9.7 (-16.5)	-7.7 (-10.2)	-9.5 (-19.8)	-9.0 (-13.7)
# of Variables Used	4	3	5	3
ρ^2	0.635	0.447	0.651	0.512
R^2	0.545	0.575	0.578	0.591
D.F.	96	97	95	97

Table 2: 1966 - 71 Migrant Heads of Age 35 - 54

Variable	1	2	3	4
Distance	-7.4 (-6.8)	-5.8 (-5.5)	-9.0 (-8.3)	-8.4 (-6.8)
Population	--	--	5.8 (8.4)	9.4 (7.6)
Housing Age	4.0 (5.7)	--	3.4 (4.1)	--
New Housing	8.0 (5.9)	7.1 (3.7)	0.8 (.9)	--
Pop. Density	-6.1 (-10.2)	-4.8 (-4.6)	-8.0 (-14.7)	-9.4 (-9.5)
# of Variables Used	4	3	5	3
r^2	0.633	0.449	0.610	0.579
$\bar{R}^2$	0.548	0.543	0.625	0.642
D.F.	96	97	95	97

Table 3: 1966 - 71 Migrant Heads of Age over 55

table 1 to table 2. The beta-weight, however, remains approximately constant (6.7, 6.2, 7.6) which indicates that the strength of the relationship does not change with age.

Finally, "Population Density" becomes a weaker variable with increasing age both in terms of the *t*-ratio and the "beta weight": In run 1 the beta weight drops from -19.6 to -16.5 to -10.2.

APPENDIX TO CHAPTER VII

THE SIMULATION PROGRAM

The structure of the computer program we developed is best described with the use of a general algorithm as follows:

```
Start program SSANAL

        Read Pop. Distribution
        Read Migration Matrix
        Read Independent Variables
        Read Estimated Coefficients
        Calculate:
                Actual Migration Frequencies
                Equilibrium Distribution (1)

        LOOP FOR I= -ITER,ITER,1

                Calculate:
                        Predictors (I)
                        Estimated Migration Frequencies
                        Equilibrium Distribution (I)

        ENDLOOP

        Print:

                Equilibrium Distribution (1)
                Equilibrium Distribution (I)

End Program SSANAL
```

The program SSANAL has three sections; input, main body, and output. The input section reads four types of information for subsequent processing. These are:

- The distribution of households in the zones of the study area at the beginning to the time period under consideration
- The origin-by-destination migration matrix
- The most significant and most important of the predictor variables revealed in the statistical analysis
- The coefficients of the predictor variables estimated in the statistical analysis.

The main body processes the input information and accumulates the required output. From the migration matrix and the population distribution the actual migration frequencies are calculated as estimates of the probabilities to migrate from one zone to another. Thus, an element n_{ij} of the migration matrix, representing the number of migrants from j to i, is divided by the population of area j, to obtain an estimate of p_{ij}. The estimated p_{ij} values are then used as coefficients in the linear system of equations (7). The EIGRF routine is invoked from the IMSL package to solve the linear system. The solution represents the equilibrium population distribution that corresponds to the observed migration frequencies and is represented in the algorithm above as "Equilibrium Distribution (1)".

A loop is subsequently entered. In each iteration one or more of the predictor variables are modified and estimates of the probabilities p_{ij} are obtained with the help of the estimated coefficients and equations IV.22 to IV.26. The same procedure as for "Equilibrium Distribution (1)" is then repeated and an "Equilibrium Distribution (I)" is calculated. In the loop, the values of I are modified from -ITER to +ITER in increments of 1. The desired value of ITER is entered in the program as a parameter.

Finally, the accumulated output of "Equilibrium Distribution (1)" and "Equilibrium Distribution (I)" is printed out. Also, for every value of I the corresponding values of the predictor variables are printed.

BIBLIOGRAPHY

Alonso, W. Location and Land Use. Cambridge, Mass.: Harvard University Press, 1964.

Anas, A. and D.S. Dendrinos. "The New Urban Economics: A Brief Survey", In G.~~J.~~ Papageorgiou (ed.) Mathematical Land Use Theory, Heath: Lexington Books, 1976, 23-51.

Anderson, T.R. "Potential Models and Spatial Distribution of Population", Papers and Proceedings of the Regional Science Association, 2 (1956).

Atkinson, R and A. Phipps. Housing Allowance Demand Experiment. Locational Choice. Part II. Neighborhood Change, Contract H-2040R, U.S. Department of Housing and Urban Development, Abt Associates Inc., Cambridge, Mass., August, 1977.

Barrett, F.A. Residential Search Behaviour~~: A~~ Study of Intraurban Relocation in Toronto. Toronto: York University Geographical Monographs no. 1, 1973.

Batty, M. "Operational Urban Models Incorporating Dynamics in a Static Framework", In A. Karlqvist, L. Lundqvist, F. Snickars, J.W. Weibull (eds.), Spatial Interaction Theory and Planning Models, Amsterdam: North Holland.

Ben-Akiva, M. and S.R. Lerman. Travel Behaviour: Theories, Models and Prediction Methods. Cambridge, Mass.: MIT Press, 1985.

Boots, B.N. and P. Kanaroglou. "Incorporating the Effects of Spatial Structure in Discrete Choice Models of Migration", Paper presented at the annual meetings of the Canadian Regional Science Association, McMaster University, Hamilton, Ontario, May 1987.

Brown, L.A. and E.G. Moore. "The Intraurban Migration Process: A Perspective", Geografiska Annaler-B, 52 (1970), 1-13.

Brummell, A.C. "A Theory of Intraurban Residential Mobility Behaviour", Unpublished Ph.D. Thesis, Dept. of Geography, McMaster University, Hamilton, Ontario, 1977.

Butler, E.W., F.S. Chapin, G. Hemmens, E. Kaiser and M. Stegman. "Moving Behaviour and Residential Choice: A National Survey", National Cooperative Highway Research Program, Report No. 81, 1969.

Cadwallader, M.T. "Neighborhood Evaluation in Residential Mobility", Environment and Planning A, 11 (1979), 393-404.

Carrothers, G.A.P. "An Historical Review of the Gravity and Potential Concepts of Human Interaction", Journal of the American Institute of Planners, 22 (1956), 94-103. Reprinted in P. Ambrose (Ed.), Analytical Human Geography, London: Longman, 1970, 226-242.

Chiang, A.C. Fundamental Methods of Mathematical Economics, Second Edition, New York: McGraw-Hill, 1974.

Clark, W.A.V., and J.O. Huff. "Some Empirical Tests of the Duration-of-Stay Effects in Intraurban Migration", Environment and Planning, 9 (1977), 1357-1374.

Cliff, A.D., R.L. Martin and J.K. Ord. "Evaluating the Friction of Distance Parameters in Gravity Models", Regional Studies, 8 (1974), 281-286.

__________. "Map Patterns and Friction of Distance Parameters: Reply to Comments by R.J. Johnston, and by L. Curry, D.A. Griffith and E.S. Sheppard", Regional Studies, 10, (1975), 285-288.

__________. "A Reply to the Final Comment", Regional Studies, 10 (1976), 341-342.

Courchene, T.J. "Interprovincial Migration and Economic Adjustment", Canadian Journal of Economics, 3 (1970), 551-576.

Curry, L. "A Spatial Analysis of Gravity Flows". Regional Studies, 6 (1972), 131-147.

Curry, L., D.A. Griffith and E.S. Sheppard. "Those Gravity Parameters Again", Regional Studies, 9 (1975), 289-296.

Domencich, R. and D. McFadden. Urban Travel Demand: A Behavioural Analysis, Amsterdam: North Holland, 1975.

DaVanzo, J. "Micro-economic Approaches to Studying Migration Decisions", in G.F. DeJong and R.W. Gardner, Migration Decision Making, New York: Pergamon Press, 1981.

dePalma, A. and C. Lefevre. "Individual Decision-Making in Dynamic Collective Systems", Journal of Mathematical Sociology, 90 (1983), 103-124.

Dodd, S.C. "The Interactance Hypothesis: A Gravity Model Fitting Physical Masses and Human Groups", American Sociological Review, (April, 1950).

Duncan, G.L. and S.J. Newman. "Expected and Actual Residential Mobility", Journal of the American Institute of Planners, 42 (1976).

Feeney, G. "Two Models of Multiregional Population Dynamics", Environment and Planning, 5 (1973), 31-43.

Fields, G.S. "Labor Force Migration, Unemployment and Job Turnover", Review of Economics and Statistics, 58 (1976), 407-415.

Friedman, J. Housing Location and the Supply of Local Public Services. Ph.D. Dissertation, Dept. of Economics, University of California, Berkeley, 1975.

Ginsberg, R.B. "Semi-Markov Processes of Mobility", Journal of Mathematical Sociology, 1 (1971), 233-262.

__________. "Incorporating Causal Structure and Exogenous Information with Probabilistic Models with Special Reference to Choice, Gravity, Migration and Markov Chains", Journal of Mathematical Sociology, 2 (1972), 83-103.

Goldstein, S. "Repeated Migration as a Factor in High Mobility Rates", American Sociological Review, 19 (1954).

Goodman, J.L. Jr. "Housing Consumption Disequilibrium and Local Residential Mobility", Environment and Planning, 8 (1976), 855-874.

Grand, E.K. and J. Vanderkamp. Economic Causes and Effects of Migration: Canada, 1965-71. Ottawa: Economic Council of Canada, 1976.

__________. "The Effects of Migration on Income: A Micro Study with Canadian Data 1965-71", Canadian Journal of Economics, 13 (1980), 381-406.

Graves, P.E. "A Life-Cycle Empirical Analysis of Migration and Climate by Race", Journal of Urban Economics, 6 (1979), 135-147.

__________. "Migration With a Composite Amenity: The Role of Rents", Journal of Regional Science, 23 (1983), 541-546.

Graves, P.E. and P.D. Linneman. "Household Migration: Theoretical and Empirical Results", Journal of Urban Economics, 6 (1979), 383-404.

Greenwood, M.J. "Research on Internal Migration in the United States: A Survey", Journal of Economic Literature, 13 (1975), 397-433.

Haag, G. and W. Weidlich. "A Stochastic Theory of Interregional Migration", Geographical Analysis, 16 (1984), 331-357.

Haken, H. Synergetics, an Introduction. Berlin: Springer-Verlag, 1977.

Hanushek, E.A. and J.M. Quigley. "The Dynamics of the Housing Market: A Stock Adjustment Model of Housing Consumption", Journal of Urban Economics, 6 (1979), 90-111.

Hearon, J.Z. "Theorems of Linear Systems", Annals of the New York Academy of Sciences, 108 (1963), 36-66.

Hicks, J. The Theory of Wages. London: MacMillan, 1932.

Hodges, H.M. Social Stratification, Cambridge, Mass.: Schenkman, 1964.

Horowitz, J. "The Accuracy of the Multinomial Logit Model as an Approximation to the Multinomial Probit Model of Travel Demand", Transportation Research, 145 (1980), 331-341.

Hoyt, H. The Structure and Growth of Residential Neighborhoods in American Cities. Washington, D.C.: USGPO, 1939.

Huff, J.O. and W.A.V. Clark. "Cumulative Stress and Cumulative Inertia: A Behavioural Model of the Decision to Move", Environment and Planning, 10 (1978), 1101-1119.

Jennrich, R.I. and H. Moore. "Maximum Likelihood Estimation by Means of Nonlinear Least Squares", Proceedings of Statistical Computing Section, Washington, D.C.: American Statistical Association, 1975, 57-65.

Johnston, R.J. "Map Pattern and Friction of Distance Parameters: A Comment", Regional Studies, 9 (1975), 281-282.

Joseph, G. "Inter-regional Population Distribution and Growth in Britain - a Projection Exercise", Scottish Journal of Political Economy, 21 (1974), 159-170.

__________. "A Markov Analysis of Age/Sex Differences in Inter-Regional Migration in Great Britain", Regional Studies, 9 (1975), 69-78.

Kanaroglou, P. and Y.Y. Papageorgiou. "Disequilibrium in the Canadian Regional System: Preliminary Evidence, 1961-1983" in D.A. Griffith and R.P. Haining (eds.) Transformations Through Space and Time, Dordrecht, Martinus Nijhoff Publishers, 1986, 212-233.

Kanaroglou, P., K.L. Liaw and Y.Y. Papageorgiou. "An Analysis of Migratory Systems: I. Theory", Environment and Planning A, 18 (1986a), 913-928.

________. "An Analysis of Migratory Systems: II. Operational Framework", Environment and Planning A, 18 (1986b), 1039-1060.

Krumm, R.J. "Regional Labor Markets and the Household Migration Decision", Journal of Regional Science, 23 (1983), 361-376.

Kurtz, T.G. "Strong Approximation Theorems for Density Dependent Markov Chains", Stochastic Processes and Their Applications, 6 (1978), 223-240.

Ledent, J. "Constructing Multiregional Life Tables Using Place-of-Birth-Specific Migration Data. In A. Rogers (ed.), Advances in Multiregional Mathematical Demography, Research Report RR-81-6, Laxemburg, Austria: International Institute for Applied Systems Analysis, 35-49.

Ledent, J. and K.-L. Liaw. "Interprovincial Migration Outflows in Canada, 1961-1983: Characterization and Explanation", Paper presented at the IUSSP Seminar on Internal Migration and Regional Development, (April 1985), Montreal, Canada.

Lee, E.S. "A Theory of Migration", Demography, 3 (1966), 47-57.

Leslie, P.H. "On the Use of Matrices in Certain Population Mathematics", Biometrica, 33 (1945), 183-212.

Liaw, K.-L. "Multistate Dynamics: The Convergence of an Age-by-Region Population System", Environment and Planning A, 12 (1980), 589-613.

________. Multiregional Population Projection: An Analytic Approach, Working Paper WP-81-81. Laxemburg, Austria: International Institute for Applied Systems Analysis, 1981.

Liaw, K.-L. and C.P.A. Bartels. "Estimation and Interpretation of a Nonlinear Migration Model", Geographical Analysis, 14 (1982), 227-245.

Liaw, K.-L. and P. Kanaroglou. "Metropolitan Elderly Outmigration in Canada, 1971-1976: Characterization and Explanation", Research on Aging, A Quarterly of Social Gerontology and Adult Development, 8 (1986), 201-231.

Liaw, K.-L., P. Kanaroglou and P. Moffett. "Metropolitan Outmigration Patterns of Canadian Labour Force Entrants, 1971-76", The Canadian Geographer, 30 (1986), 229-242.

Liaw, K.-L., M.J. Webber, P. Kanaroglou and Y.Y. Papageorgiou. "Interpreting Interregional Migration Data: Net Migration and the Steady-State Distribution", QSEP Research Report No.151, Dept. of Economics, McMaster University, Hamilton, Ontario, Canada, 1985.

Lotka, A.J. "Relation Between Birth Rates and Death Rates", Science, 26 (1907), 21-22.

Lowry, I.S. Migration and Metropolitan Growth: Two Analytical Models. San Francisco: Chandler Publishing, 1966.

Martin, J.K. and D.T. Lichter. "Geographical Mobility and Satisfaction with Life and Work", Social Science Quarterly, 64 (1983), 524-535.

McFadden, D. "Conditional Logit Analysis of Qualitative Choice Behaviour", In P. Zarembka (Ed.) Frontiers in Econometrics, New York: Academic Press, 1974, 105-142.

__________. "The Mathematical Theory of Demand Models", in P. Stopher and A. Meyburg (eds.), Behavioural Travel Demand Models, Lexington, MA: Lexington Books, 1976.

__________. "Modelling the Choice of Residential Location", In A. Karlqvist, L. Lundqvist, F. Snickars and J.W. Weibull (Eds.), Spatial Interaction Theory and Planning Models, Amsterdam: North Holland, 1978.

__________. "Quantitative Methods for Analysing Travel Behaviour: Some Recent Developments", In D. Hensher and P.R. Stopher (eds.), Behavioural Travel Modelling, London: Croom Helm, 1979.

McGinnis, R. "A Stochastic Model of Social Mobility", American Sociological Review, 33 (1968), 712-722.

Michelson, W. Environmental Choice, Human Behaviour and Residential Satisfaction. New York: Oxford University Press, 1977.

Miron, J.R. "Changing Patterns of Household Formation in the Toronto CMA; 1951 to 1976", Research Paper No. 106, Centre for Urban and Community Studies, University of Toronto, 1979.

__________. "Economic Equilibrium in Urban Land Use" in L.S. Bourne (Ed.), Internal Structure of the City: Readings on Urban Form, Growth and Policy, New York: Oxford University Press, 1982, 124-136.

__________. "A Replacement Model of Metropolitan Expansion", Geographical Analysis, 15 (1983), 107-127.

Moore, E.G. "Residential Mobility in the City", Association of American Geographers, Commission on College Geography Resource Paper No. 13, 1972.

Mueller, C.F. The Economics of Labor Migration: A Behavioural Analysis. New York: Academic Press, 1982.

Muth, R.F. Cities and Housing. Chicago: University of Chicago Press, 1969.

Muth, R.F. "Migration: Chicken or Egg?" Southern Economic Journal, 37 (1971), 295-306.

Nicolis, G. and I. Prigogine. Self-Organization in Non-Equilibrium Systems. New York: Wiley, 1977.

Onaka, J. and W.A.V. Clark. "A Disaggregate Model of Residential Mobility and Housing Choice", Geographical Analysis, 15 (1983), 287-304.

Park, R.E., E.W. Burgess and R.D. McKenzie. The City. Chicago: University of Chicago Press, 1925.

Pickles, A. "Models of Movement: A Review of Alternative Methods", Environment and Planning A, 12 (1980), 1383-1404.

Pickvance, C.G. "Life-Cycle, Housing Tenure and Residential Mobility: A Path Analytic Approach", Urban Studies, 2 (1974), 171-188.

Porell, E.W. "Spatial Impacts of Local Housing Programs: An Empirical Application of a Housing Turnover Model", Journal of the American Planning Association, 47 (1981), 59-69.

Porell, F.W. Models of Intraurban Residential Relocation, Boston: Kluwer-Nijhoff Publishing, 1982.

Quigley, J.M., and D.H. Weinberg. "Intraurban Residential Mobility: A Review and Synthesis", International Regional Science Review, 2 (1977), 41-66.

Ravenstein, E.G. "The Laws of Migration", Journal of the Royal Statistical Society, 48 Part 2 (June 1885), 167-227.

__________. "The Laws of Migration", Journal of the Royal Statistical Society, 52 (June 1889), 241-301.

Reilly, W.J. Methods for the Study of Retail Relationships. University of Texas, Bureau of Business Research, Research Monograph No. 4, Nov. 1929.

__________. The Law of Retail Gravitation. New York: G.P. Putman and Sons, 1931.

Rogers, A. "Matrix Analysis of Interregional Population Growth and Distribution", Papers of the Regional Science Association, 18 (1967), 177-196.

__________. Matrix Analysis of Inter-Regional Population Growth and Distribution, Berkeley: California University Press, 1968.

__________. Matrix Methods in Urban and Regional Analysis, San Francisco: Holden-Day, 1971.

__________. Introduction to Multiregional Mathematical Demography, New York: Wiley, 1975.

Rogers, A. and F. Willekens. "Spatial Population Dynamics", Papers of the Regional Science Association, 36 (1976), 3-34.

__________. "The Spatial Reproductive Value and the Spatial Momentum of Zero Population Growth", Environment and Planning A, 10 (1978), 503-518.

Rogerson, P.A. "Prediction: A Modified Markov Chain Approach", Journal of Regional Science, 19 (1979), 469-478.

Roistacher, E. "Residential Mobility" in J.N. Morgons (Ed.) Five Thousand American Families Patterns of Economic Progress, Vol. 2, Institute for Social Research, University of Michigan, 1974.

Rossi, P. Why Families Move. Glencoe, Il.: Free Press, 1955.

Shaw, P.R. Intermetropolitan Migration in Canada: Changing Determinants Over Three Decades. Toronto: Nc Press, 1985.

Sheppard, E.S., D.A. Griffith and L. Curry. "A Final Comment on Mis-Specification and Autocorrelation in Those Gravity Parameters", Regional Studies, 10 (1976), 337-339.

Sjaastad, L.A. "The Costs and Returns of Human Migration", Journal of Political Economy, 70 (1962), 80-93.

Simmons, J.W. Patterns of Residential Movement in Metropolitan Toronto. Dept. of Geography Research Publication, Toronto: University of Toronto Press, 1974.

Speare, A. "Residential Satisfaction as an Intervening Variable in Residential Mobility", Demography, 11 (1974), 173-188.

Speare, A., S. Goldstein and W. Frey. Residential Mobility, Migration and Metropolitan Change. Cambridge, Mass.: Ballinger, 1975.

Statistics Canada. 1966 Census of Canada, Census Track Series, Catalogue 95-620, 1968.

__________. Dictionary of the 1971 Census Terms, 1972.

__________. 1971 Census of Canada, Census Track Series A, Catalogue 95-721, 1973a.

__________. 1971 Census of Canada, Census Track Series B, Catalogue 95-751, 1973b.

__________. 1976 Census of Canada, Population: Geographic Distributions, Catalogue 92-809, 1978.

Stewart, J.G. "An Inverse Distance Variation for Certain Social Influences", Science, (Jan. 1941).

__________. "Demographic Gravitation: Evidence and Application", Sociometry, (May 1948).

Stouffer, S.A. "Intervening Opportunities: A Theory of Relating Mobility and Distance", American Sociological Review, 5 (1940), 845-867.

__________. "Intervening Opportunities and Competing Migrants", Journal of Regional Science, 2 (1960), 1-26.

Tarver, J. and W. Gurley. "A Stochastic Analysis of Geographic Mobility and Population Projections of the Census Divisions in the United States", Demography, 2 (1965), 134-139.

Tiebout, C.M. "A Pure Theory of Local Expenditures", Journal of Political Economy, 64 (1956), 416-424.

Vanderkamp, J. "Migration Flows, Their Determinants and the Effects of Return Migration", Journal of Political Economy, 79 (1971), 1012-1032.

Vining Jr., D.R. "The Spatial Distribution of Human Populations and its Characteristic Evolution Over Time: Some Recent Evidence from Japan", Papers of the Regional Science Association, 35 (1975), 157-178.

Weidlich, W. and G. Haag. Concepts and Methods of Quantitative Sociology. Berlin: Springer-Verlag, 1983.

Weinberg, D.H. Intraurban Household Mobility: A Microeconomic Study of San Francisco Bay Area Households. Ph.D. Dissertation, Yale University, New Haven, 1975.

Weinberg, D.H., J. Friedman and S.K. Mayo. "Intraurban Residential Mobility: the Role of Transactions Costs, Market Imperfection and Household Disequilibrium", Journal of Urban Economics, 9 (1981), 332-348.

Wheaton, W.C. "A Comparative Static Analysis of Urban Spatial Structure", Journal of Economic Theory, 9 (1974), 223-237.

______. "Monocentric Models of Urban Land Use: Contributions and Criticisms", in P. Mieskowski and M. Straszheim (Eds.), Current Issues in Urban Economics, John Hopkins, 1979, 107-129.

Wilson, A.G. "A Statistical Theory of Spatial Interaction Models", Transportation Research, 1 (1967), 253-269.

______. Entropy in Urban and Regional Modelling. London: Pion, 1970.

Wolpert, J. "Behavioural Aspects in the Decision to Migrate", Papers of the Regional Science Association, 15 (1965), 159-169.

Yezer, A.M.J. and L. Thurston. "Migration Patterns and Income Change: Implications of the Human Capital Approach to Migration", Southern Economic Journal, 42 (1976), 693-702.

Zelinsky, W. "The Hypothesis of Mobility Transition", Geographical Review, 61 (1971), 219-149.

Zipf, G.K. "The P_1P_2/D Hypothesis; On the Intercity Movement of Persons", American Sociological Review, (Dec. 1946).