856

# PRELIMINARY TESTS OF A

# DYNAMIC MODEL OF URBAN GROWTH

# PRELIMINARY TESTS OF A DYNAMIC MODEL OF URBAN GROWTH

by

VALERIE ANN PRESTON, B.A.

Research Report

Submitted to the Faculty of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Arts

McMaster University

May 1974

MASTER OF ARTS (1974)
(Geography)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Preliminary Tests of a Dynamic Model of Urban Growth

AUTHOR: Valerie Ann Preston

SUPERVISOR: Dr. Leslie J. King and Dr. George J. Papageorgiou

NUMBER OF PAGES: vi, 101

SCOPE AND CONTENTS:

In this paper, empirical tests of a dynamic urban growth model are discussed. It is assumed that population change in any urban region is a function of the population size of the urban centres in the system and of the distances between them. A set of linear equations is simultaneously estimated by a least squares procedure. The parameters of the model; the equilibrium population of each urban region, the rate of natural increase, and the propensity to migrate between urban regions, are claculated from the regression coefficients. By estimating a series of equations at different times, a set of parameter estimates are obtained. The parameter estimates fluctuate erratically. Recommendations for further research include the redefinition of the model, and of the urban system.

ACKNOWLEDGEMENTS

This report is the result of many people's research and their contributions are gratefully acknowledged. Thanks are extended to my supervisors, Dr. Leslie J. King and Dr. George J. Papageorgiou, for their assistance and patience. The suggestions of Dr. Kovarick and Dr. Stephens are gratefully acknowledged. I must also thank my friends in the mathematics department and the geography department for their encouragement and good humour. Particularly, thanks are due to Vivian who wrote the computer algorithm, and to Rick who always listened.

## TABLE OF CONTENTS

LIST OF TABLES

# CHAPTER I

# INTRODUCTION

## I.1 INTRODUCTION

Recently, Batty (1972, 44) identified two major problems of present models of urban growth. He suggested that these models need to be disaggregated to include microeconomic factors, and that the concept of a static equilibrium must be replaced by that of dynamic disequilibrium. He argued that dynamic models increase understanding of the complex processes of change underlying urban structures, as well as providing a better basis for short-term forecasting.

This research is addressed to the second problem identified by Batty. A dynamic model of population growth has been developed by Papageorgiou (1971), in which the rate of growth of the population of an urban region is a function of the deviation of the population from the equilibrium population of the region. The changing disequilibrium of each urban region determines its rate of population growth. It is an aggregated model in which the populations of urban regions are treated as masses of homogeneous individuals, distinguished only by their location in a particular region (Papageorgiou, 1971).

In this report, the results of several empirical tests of this model are discussed. Regression analysis of population and employment data for Southwestern Ontario is used to answer several questions concerning the adequacy of the model and the behaviour of different parameters of the urban system through time. How well does the model describe the patterns of population growth of an urban system? Are the parameters of the model stable? If they are not stable, what trends are evident in their behaviour? Can different patterns of growth be identified at different levels of the urban hierarchy?

A brief survey of the literature of population growth models is presented in the remainder of this chapter, while the problems of defining and of estimating the model are discussed in Chapters Two and Three. The results of the empirical tests are presented in Chapter Four, and Chapter Five includes recommendations for further modifications of the model. This is a preliminary testing of a very complex model. The results should be judged with consideration for the paucity of available information concerning the dynamic behaviour of urban systems.

## I.2 LITERATURE REVIEW

According to Czamanski (1964), there are three models of urban growth. The first is the economic base multiplier model in which the growth of export industries generates increased employment opportunities in both export and service industries. Population growth is caused by the increased flow of migrants, attracted to these economic opportunities. Similarly, in the interregional and regional input-output model, it is assumed that the transmission of economic impulses between different sectors of the economy generates patterns of increased employment among the regions in the urban system. The resulting migration causes population growth. Alternatively, in gravity and potential models, it is assumed that population growth reflects the relative size and accessibility of each centre within the system (Czamanski, 1964).

A link between these two explanations of population growth is available in the innovation diffusion literature. If economic growth results from the diffusion of technological innovations, as Schumpeter has suggested, (Schumpeter, 1967), the process of diffusion of these innovations will give rise to subsequent patterns of economic and population growth. It has been shown that the pattern of diffusion of an innovation is largely explained by the size of the urban centres among which it spreads, and by their distances from the origin of the innovation (Hagerstrand, 1965; Pred, 1971). Thus, the link between the gravity models and the economic explanations of population growth is clear.

In this literature survey, both theoretical and empirical studies based on the economic base multiplier model, and on the regional and interregional input-output model will be discussed.
Then, a short discussion of the pertinent aspects of the theory of the diffusion of innovations will be presented. Finally, the contribution of potential and gravity models to the explanation of population growth will be reviewed.

### I.2.1 Economic Explanations

Several variations of the economic base multiplier model have been developed. They all share certain characteristics. It is assumed that two types of industries contribute to the growth of employment of urban centres. These are 'complementary industries', which realize economies of scale by locating nearby associated industries, for instance, chemical plants which locate close to oil refineries, and 'urban oriented industries' which exist to serve an urban population (Czamanski, 1964, 179). Thompson (1968) has emphasized that a well developed infrastructure will attract additional industries to an urban centre. In these models, it is assumed that the urban centre is moving towards a state of equilibrium, but neither the mechanism by which the system reaches equilibrium, nor the definition of equilibrium are clearly defined (Czamanski, 1964, 181). Results of Czamanski's empirical tests of his simple five equation model indicated that population growth is generated by increased employment. On the other hand, the implicit assumption that the numerical values of the parameters may change through time but their relative importance will remain constant was not supported by his results. This problem has been partically resolved in Paelinck's model of urban growth in which population growth is dependent upon the growth of employment and the level of public expenditure. In this differential equation model, the values of the parameters are functions of time. He was able to identify four stages of urban growth during which the values were relatively stable, but between which they varied considerably (Paelinck, 1970).

Although these theoretical models of the growth of an urban centre provide some information concerning the interrelationship between a few variables, the mathematical difficulties of incorporating more variables into these models restricts their application. Simulation models of the urban growth process have been developed to overcome this problem. In simulation models, the mathematical equations do not have to be solved, instead the behaviour of the system and of it's parameters under different circumstance, defined by the numerical values with which the iterative process is begun, may be observed.

Forrester has developed one simulation model in which both negative and positive feedback loops are incorporated, however, his model, like the previously described theoretical models, is aspatial. It is assumed that the population will increase as some function of employment growth without consideration of the final location of this increased population (Batty, 1972b). Two other models, the Lowry and TOMM models, allocate increased population to different zones of the urban centre, by treating time and space as discrete variables (Goldner, 1971). The problems of defining the appropriate zones and time periods and correctly classifying industries, have not yet been solved. Furthermore, all these simulation models are static and the relative importance of the parameters which are introduced at the beginning of the iterative process remain unchanged throughout the analysis. Certainly, in Forrester's model in which urban growth is simulated for periods of two hundred years, this is an unreasonable assumption. Despite these problems, the simulation models of urban growth do incorporate many variables which seem to influence the rate of population growth. The problems of choosing these variables and of defining the relationships between them will not be solved however, until a better theoretical explanation of urban growth is available.

In the models of the growth of a single centre which have been previously discussed, space is dichotomized into the urban centres and the world outside, between which it is assumed there is no significant interaction. Similarly, in the interregional and regional input-output models of urban growth, an urban system within which the population growth of each centre is dependent largely upon the employment growth resulting from the transmission of economic impulses among the urban centres in the system, is isolated from the world outside. How can this system of interdependent regions be identified?

Although the spatial arrangement of urban centres is explicitly considered in central place theory, the effects of other economic activities, such as manufacturing, are disregarded. Empirical evidence indicates that regional networks of interaction reflect the distribution of manufacturing and other economic activities. Hodge concludes that the growth rate of employment in Canadian metropolitan areas was best explained by the employment structure of each area. If the employment structure is diversified, with a high percentage of the labour force employed in manufacturing, the rate of employment growth will increase more rapidly (Hodge, 1972). King, Casetti, and Jeffrey (1971) hypothesized that the composition of three groups of American cities having similar patterns of unemployment reflected the industrial mix of each centre, the patterns of spatial and structural linkages among them, and the population size of each city. Similarly, after analysing such characteristics as the demographic structure of each city, the state of their housing markets, and the accessibility of each centre, Golant (1972) concluded, that two regional systems can be identified in central Canada. One system (which is incoherent and poorly developed) is focussed on Montreal, the other is dominated by Toronto. King's analysis of the social and economic characteristics of Ontario and Quebec urban centres also indicated that these two regional systems had different patterns of growth (King, 1966). Siegel and Woodyard (1971) found that the rate of population growth of urban centres in Ontario was explained by different factors depending upon their position in the urban hierarchy.

The empirical evidence is conflicting but some of the factors which central place theory postulates should be important explanatory variables, such as population size, accessibility, and employment in service industries do contribute significantly to explanations of the patterns of population and employment growth. These empirical studies indicate that other variables, such as industrial mix, are also significant. Urban systems within which patterns of population change and employment change are similar can be isolated even though the structural links among the centres within these systems are not yet analysed.

Urban systems have also been defined by the analysis of input-output tables in which the linkages among industrial sectors are measured by coefficients which express the contribution of each industry's inputs to another industry's outputs. Isard and Schooler (1955) analysed the petrochemical industry, using input-output tables and larger scale tables describing the linkages among all sectors of the American economy have been published. A recent analysis of the economy of Nova Scotia (Czamanski, 1972) indicated that the Nova Scotia economy is an open system in which the imports of goods and capital were significant. Czamanski commented that the volume of leakages from an economic system decreases as the input-output table is aggregated although the volume of leakages from the Nova Scotia economy was significant despite the low level of aggregation of the input-output table used (Czamanski, 1972). Although the industrial linkages of an economy are specified by this analysis, the geographical organization of this system is not considered. Furthermore, the urban system, as well as the industrial sectors of the intput-output table, must be defined before the analysis is undertaken. These arbitrary choices can seriously influence the conclusions drawn from the input-output table, as Czamanski demonstrated by defining three different input-output tables for Nova Scotia (Czamanski, 1972).

Perroux has suggested that economic growth is caused by rapid growth in an industrial sector which transmits this growth to other sectors by its strong backward and forward linkages. A growth pole is defined as a propulsive industry which increases the rate of growth of a complex of related industries through backward and forward linkages. These industries are clustered in economic space which is defined by the relationships between industrial sectors, but is independent of geographical space (Perroux, 1964, 130). The translation of this notion of a growth pole, as a sectoral cluster in economic space, into the notion of a geographically clustered set of industries from which economic impulses spread through an urban system has not been successful. According to Darwent (1969, 21) and Lasuen (1971, 8) urbanization and industrialization economies are the reasons for geographical clusters of industries, however, these two phenomena are poorly understood and difficult to measure. Darwent (1969, 21) comments that growth pole theory does not explain how the location and spatial distribution of urban centres affect the transmission of growth impulses.

In both the economic base multiplier model and the inter-regional input-output model it is assumed that population growth is the result of increasing employment, but the creation of new jobs is a capital investment extending over long periods of time (Czamanski, 1964, 197). It is also implicitly assumed that the urban system is stable, tending towards a constant state of equilibrium. Thus, the urban system must be stable over the long periods of time which are required for the creation of new employment opportunities. Yet, King found that the dimensions of the Canadian urban system changed significantly in ten years (King, 1967). A theory of urban growth must be developed in which the stability of the urban system is not assumed. A more dynamic model of economic growth, in which the origins and processes of the growth of employment and population are more clearly specified is needed.

### I.2.2 Interaction Explanations

It has been suggested that economic growth results from the initial advantage of a producer who introduces a new product. Thompson emphasizes that urban centres with a well developed infrastructure provide an ideal environment within which technological innovations may be launched. Once the innovation has been introduced production techniques are standardized so that the product may be produced in smaller branch plants. The urban centre where the innovation was introduced has enjoyed the additional income and demand generated by the innovation (Thompson, 1968). Thus, economic growth may be viewed as the process of diffusion of successful innovations. Two patterns of innovation diffusion operating at different spatial scales have been identified.

In his seminal study of the spread of new farm practices, Hagerstrand (1967) demonstrated that the diffusion of an innovation followed a distance decay function. The rate of adoption decreased with distance from the origin of the innovation, farm research stations, while the date of adoption was delayed farther from the research stations. At a regional scale, Hagerstrand analysed the spread of scientific information in Europe. He suggested that regional information networks could be identified. At the national level,information spread first among the capitals of Europe, then among the regional capitals of each country, from which it diffused throughout each country (Hagerstrand, 1965). A similar pattern of diffusion is observed in the flow of telephone calls among urban centres in Quebec and Ontario (Simmons, 1970b). Pred (1971) found that diseases did not simply spread around the american ports where they entered the country but instead they could be traced down the river systems to the largest urban centres from which they diffused into the surrounding countryside. Simmons (1970a) has commented that the volume of interaction between Canadian provinces, which is measured by interprovincial commodity and information flows as well as migration, is largely explained by the size of the provinces and their relative distances apart. The importance of distance is emphasized by the linear configuration of the inhabited parts of Canada. Regional networks dominate these patterns of interaction. Bannister (1974) has recently demonstrated that patterns of change in southern Ontario spread outwards from each centre to its nearest neighbours, rather than down the urban hierarchy.

In the migration literature it is suggested that the volume and composition of migration streams can be explained by three factors; the characteristics of the migrant, of the origin, and of the potential destinations (Isard, 1960, 54). D. S. Thomas (1938) commented that generalizations about migrants were impossible but she concluded that young adults were most likely to migrate. Olsson (1965a) found that several personal characteristics, including income, and age were good predictors of the length of migration. Other socioeconomic characteristics, such as occupation, duration of resident, stage in the life cycle, and level of education have been used by other researchers in England and Canada to explain both the length of migration and the migrant's initial propensity to migrate (Simmons and Baker, 1972; Stone, 1971; Cordey-Hayes and Greave, 1973).

Similarly, the characteristics of the origin and of the destination of migration streams have been investigated. Lansing and Mueller (1967) concluded that migrants were not pushed out of their origins by poor economic conditions, but migrated to areas where they perceived economic opportunities for increased wages and salaries. Nelson (1959) emphasized that migrants chose destinations about which they had information from friends and relatives. Olsson (1965a) found that the length of migration was best explained by the population size of both the migrant's origin and destination and the distance between them. Other measures of the attraction of different urban centres such as economic diversity, and centrality within the urban system, have been suggested by Woodyard (1970), Hill (1970), and Cordey-Hayes (1973). Yet, Cordey-Hayes and Greave (1973, 8) commented that population size is an adequate surrogate by which the intrinsic attractiveness of an urban centre may be measured.

Empirical studies indicate that the major motivation for migration is the desire to enjoy improved economic opportunities which are most likely in larger cities. Thus, the volume and composition of migration streams may be accounted for by the population size of the migrant's origin and subsequent destination, as well as the distance between them. According to Pred (1973), these factors also influence the pattern of diffusion of innovations which result in economic and population growth.

The gravity model is a model of interaction in which population size and distance are explicitly considered. Zipf first proposed this model, remarking that population centres attracted migrants in relation to their values of P/D because of the costs of obtaining information about events at a distance and the costs of migration. He notes that the number of passengers travelling from an urban region i to a destination j is best calculated from the following formula,

$$I_{ij} = CP_i \, P_j/D_{ij} \qquad \text{(Zipf, 1949, 396)}$$

This model of interaction has been widely applied in the geographical literature, where several modifications have been proposed. Olsson (1965b, 55) emphasized that the measurement of population size must reflect only the populations which are actually interacting, therefore, population sizes have been weighted to express the propensity of each population to interact. Exponents have been applied to distance measures to express the decreasing friction of each unit of distance as the distance between regions i and j increases. The most general gravity model is,

$$I_{ij} = \frac{w_i P_i \cdot w_j P_j}{d_{ij}^{b_{ij}}} \qquad \text{(Olsson, 1965b, 56)}$$

Another modification suggested by Stouffer was that the number of people moving a certain distance is proportional to the number of opportunities at that distance and inversely proportional to the intervening opportunities between them (Olsson, 1965b, 64). Olsson (1965b, 65) comments that Stouffer simply replaced the measurement of physical distance with a measure of social distance. Furthermore, no operational definition of intervening opportunities is available.

In their reviews of gravity models both Olsson and Isard note that the definitions of population and distance in the gravity model have not been clearly specified. Intuitively, it seems that measures of social distance, rather than straight line distance, are more accurate reflections of the functional distance between centres, however, the results obtained using social distance are not significantly different from those obtained using straight line distances. Perhaps the extra work involved in computing social distances is not justified (Olsson, 1965b, 43). Similarly, the problem of defining the appropriate population measure has not been resolved.

Olsson (1965b, 27) noted that the gravity model provides an adequate measure of the interaction between centres only for interaction over long distances. There are mathematical problems caused by a small denominator in the gravity model. Also, the economic rationale for the inverse relationship between population and distance seems to apply only when the costs of overcoming distance are significant.

This hypothesis has been supported by empirical studies in which the gravity model has been fitted to migration data. Olsson fitted Swedish migration data to the logarithmic form of the gravity model,

$$\log I_{ij}/P_iP_j \quad = \quad \log K - b\log D_{ij} \qquad \text{(Olsson, 1965b, 35)}$$

Investigating the values of the distance exponent, b, he concluded that the friction of distance was much larger for places at lower levels in the hierarchy. Similarly, Lowry attempted to explain the patterns of migration flows by weighting population figures both by the hourly manufacturing wage in each region and by its rate of unemployment. He estimated the following equation,

$$\log M_{ij} \quad = \quad a_0 + a_1\log L_i + a_2\log L_j + a_3\log D_{ij} + a_4\log U_i$$
$$+ a_5\log U_j + a_6\log w_i + a_7\log w_j \qquad \text{(Lowry, 1966, 15)}$$

Distance was not a significant predictor of total interaction between pairs of places, however, it significantly contributed to the explanation of the volume of flows in each direction (Lowry, 1966).

Modifying the Rogers matrix model of interregional migration, Wilson (1972) developed a gravity model of migration flows which incorporates birth and survival rates. He suggests that migration and demographic models can be disaggregated and applied to small areal units such as the 'analysis zones of some urban study area'. Yet, the major assumption of the gravity model is that the population is an undifferentiated mass. Warntz (1965, 5) commented,

> "In essence, the assumption is that people exert an influence at a distance which in many instances varies directly with the size of the population and inversely with the distance from it ... A population is positionally most accessible at that place it occupies, and other things being equal, an increase in distance serves to decrease accessibility".

Thus, Isard (1960, 515) suggests that it is not valid to disaggregate population masses into groups having different propensities to interact.

Recently, two explanations of the gravity model have been advanced. Using statistical mechanics, Wilson (1968) proved that the distribution of flows generated by the gravity model was the most probable distribution given the analyst's knowledge of the urban system. Alternatively, Neidercorn and Bechtolt (1969) demonstrated that the distribution of trips which maximizes the utility of trip-making of a homogeneous population was best described by the gravity model. Suzanne Evans showed that in the limit the gravity model generated the same distribution of trips as the Herbert Stevens bid rent model in which the utility of all householders was maximised (Senior and Wilson, 1973). In both explanations, the definitions of the budget and time constraints of the gravity model strongly influenced the solution. These results underline Olsson's (1956b, 50) comment that a theoretical explanation of the gravity model is possible,

> "for those interactions which are reflexive in the sense that they depend on maintenace of earlier personal contact. That kind of interaction can easily be exemplified by telephone calls and migration, where the cost of maintaining the contact is a clear function of the distance separating the two interacting objects".

Yet, the gravity model, like the economic models of population growth, is a static equilibrium model in which the system of centres and the patterns of interaction of these centres are assumed to be stable throughout the period of analysis. Batty (1972a) has recently attempted to define a dynamic model of population growth using simulation techniques. Peaker (1971) has defined the conditions under which balanced and unbalanced growth will occur by analysing the growth of capital and labour in a two region model. He concludes that once a pattern of unbalanced growth is initiated, the system will not return to a state of balanced growth. These models and the economic model of population growth developed by Paelinck (1970) are examples of the very few attempts to dynamically model population growth.

The simplicity of the gravity model is one of its major advantages. Operational definitions of the variables can be established, the data required by the model are available, and the model is mathematically tractable. The failure of the economic explanations of population growth is partially caused by the large number of variables included in these models. The theory of economic growth is not well developed, therefore, the importance of different variables cannot be determined nor can the operational definitions of these variables be stated. The complex mathematical relationships among these variables cannot be defined, nor are they mathematically tractable. Thus, two conceptual frameworks are available within which the process of population growth may be analysed. The operationalization of either of these approaches is not easy, but it is at least possible to obtain results with the gravity model.

## I.3 SUMMARY

This cursory review of the literature of urban growth has indicated three major areas where further research is required. Both the economic and interaction explanations of urban growth require better theoretical explanations of the actual process of growth, how it is initiated and how it is spread so that the variables can be better defined. Another major problem is the lack of knowledge of the patterns of interaction among urban regions. How are growth impulses transmitted from one region to another? The gravity model measures the volume of flows in geographical space, while the input-output table measures the interaction among economic sectors, however, the relationship between these two aspects of urban growth must be investigated. The third major question which remains unanswered is how urban systems behave through time. There has been some investigation of the long term changes in the urban system, under the assumption that the system tends to equilibrium, while short run changes have not been extensively studied (King, 1966; Berry and Horton, 1970, Chatper 3).

The research which is discussed in the remainder of this report addresses itself to this last question. What are the characteristics of the dynamic behaviour of an urban system. It is assumed that the system reacts to the equilibrium state which is a function of time. As this literature review has indicated, a clear explanation of the economic causes of population growth is not available, therefore, a variation of the gravity model based on a small number of naive assumptions concerning the process of population change has been adopted. Three hypotheses will be tested in this analysis.

Does the gravity model accurately describe the patterns of population growth of an urban system? The simple assumptions underlying the gravity model are more likely to successfully describe the patterns of population growth than the more complex economic explanations because the variables are more easily specified. Also, the theoretical explanation of the gravity model which emphasizes the importance of information flows is more applicable to the analysis of migration flows.

Secondly, how do the parameters of the model behave through time? Do the parameters vary systematically or do they fluctuate randomly? From the behaviour of the parameters of the model, the reaction of the urban system to the changing equilibrium state may be inferred.

The third question concerns the variation in growth patterns at different levels in the urban hierarchy. Simmons (1970b) and Woodyard and Siegel (1971) identified different patterns of interaction and employment growth at different levels of the hierarchy. Thus, it is hypothesized that the parameter values of the model will have different patterns of variation at different levels of the urban hierarchy.

The dynamic behaviour of an urban system is very poorly investigated. Any additional information about both the process of urban growth, and the methodology necessary for its analysis which is yielded by this study will be valuable. At the least, this research will indicate problems involved with this approach to the problem of population change.

## CHAPTER II

## DERIVATION OF THE MODEL

### II.1 ASSUMPTIONS

The fundamental assumption of this model is dynamic disequilibrium. It is assumed that the rate of change of the population of an urban region is regulated by deviations from the region's state of equilibrium. In most models of spatial structure, an equilibrium pattern at one point in time is simulated and the patterns generated under different equilibrium conditions are then compared (Batty, 1972b, 152). In this model attention is focussed upon the process whereby the system reacts to different equilibrium states by assuming that the state of equilibrium changes through time. A state of equilibrium prevails when there is no migration either into or from an urban region. Thus, the rate of population change of any urban region during a single time period is some function of the deviation of its population size from the equilibrium population (Papageorgiou, 1971). If the equilibrium population is exceeded, the rate of population change will decrease, alternatively the rate of change will increase if the deviation from equilibrium increases positively.

In the short run, the rate of population change of an urban region is primarily accounted for by migration, although natural increase, the surplus of births over deaths, contributes significantly to population change in the long run (Isard, 1960, 53). In this model, the pattern of population change over short periods of time will be analysed since the state of equilibrium which governs the system's behaviour is constantly changing. Consequently, changing migration patterns significantly contribute to the explanation of changing patterns of population change. The measurement of the volume of migration streams is affected by the definitions of the urban system and of the urban regions within that system. It is implicitly assumed that an urban system can be identified, however, no criteria for its definition are established.

## II.2 DERIVATION

In the most general case, migration between the urban system and the world outside, as well as migration among the urban regions within the system is considered. Thus,

$$1. \qquad \dot{P}_i = n_i(P)+\sum_{j\neq i}(\gamma_{ji}s_{ji}^{-v_{ji}}(\bar{\bar{P}}_i(P)-P_i)P_j-\gamma_{ij}s_{ij}^{-v_{ij}}(\bar{\bar{P}}_j(P)-P_j)P_i) + m_i(P)$$

where $\dot{P}_i$ is the rate of population change per unit time, $n_i(P)$ is the natural increase of region i, $m_i(P)$ is the net migration between region i and the world outside, and the middle term, $\sum_{j\neq i}(\gamma_{ji}s_{ji}^{-v_{ji}}(\bar{\bar{P}}_i(P)-P_i)P_j -\gamma_{ij}s_{ij}^{-v_{ij}}(\bar{\bar{P}}_j(P)-P_j)P_i)$, measures the net migration to region i from all the other urban regions in the system. Net migration is dependent upon the difference between the equilibrium population, $\bar{\bar{P}}_i(P)$, and the actual population at each centre, $P_i$, as well as the propensity to migrate from region i to any other region, $\gamma_{ij}$. The volume of the migration stream from region i to region j also varies inversely with the distance between them, $s_{ij}$, transformed by the elasticity of distance, $v_{ij}$. The rate of change of population of region i is defined by equation 2 if a closed system in which there is no interaction between the system and the world outside is assumed,

$$2. \qquad \dot{P}_i = n_i(P) + \sum_{j \neq i} (\gamma_{ji} s_{ji}^{-v_{ji}} (\bar{\bar{P}}_i(P) - P_i) P_j - \gamma_{ij} s_{ij}^{-v_{ij}} (\bar{\bar{P}}_j(P) - P_j) P_i) \; .$$

Before either of these equations may be empirically tested, the functions, $n_i(P)$, $m_i(P)$, and $\bar{\bar{P}}_i(P)$ must be specified. Two definitions of each of these functions have been developed. In the first instance, the natural increase of region i, $n_i(P)$, is assumed to be a constant proportion of the total population of the region. Similarly, the equilibrium population is a constant proportion of the population, throughout the time period over which it is estimated. Thus,

$$3. \qquad n_i'(P) = \beta_i P_i$$

$$4. \qquad \bar{\bar{P}}_i'(P) = \bar{\bar{P}}_i, \text{ a constant}$$

where $\beta_i$, is a constant. Alternatively, the equilibrium population of region i is the demographic potential of the region which is defined as,

5. $\bar{P}_i''(P) = \delta_i+\varepsilon_i\sum_j P_j s_{ij}^{-\mu}$ $\qquad s_{ii} = 1$

in which $\delta_i$ and $\varepsilon_i$ are constants which must be estimated for each region. In equation 6, the equilibrium population depends upon both the distribution and arrangement of population within the urban system. The natural increase of region i, $n_i(P)$, also varies when it is defined according to equation 6.

6. $n_i''(P) = \beta_i(\bar{P}_i-P_i)P_i$ .

$\bar{P}_i$ is the population of region i corresponding to no migration between the region and the world outside. $\bar{P}_i$, is not necessarily the region's equilibrium population, $\bar{\bar{P}}_i$.

The net migration from the world outside is first defined as a constant proportion of the deviation of the actual population from the equilibrium population.

7. $m_i'(P) = \gamma_i(\bar{\bar{P}}_i(P)-P_i)$

where $\gamma_i$ measures the propensity to migrate to region i from the world outside the urban system. Alternatively, the net migration is defined as,

8. $m_i''(P) = \gamma_{.i}(\bar{\bar{P}}_i(P)-P_i)-\gamma_{i.}(P_i-\bar{P}_i)P_i$

Here, the net migration from the world outside is a function of the propensity to migrate both to region i from the world outside. The deviation between $P_i$, and $\bar{P}_i$ regulates the amount of net migration between the region and the world outside. It is assumed that the world outside acts as a sink, capable of absorbing any number of migrants.

From these alternatives, a family of equations defining the rate of population change of a region can be produced (see Table 1). Unfortunately, after examination of these equations only the following proved mathematically tractable,

$$9. \qquad \dot{P}_i = \beta_i P_i + \sum_{j \neq i} (\gamma_{ji} s_{ji}^{-v_{ji}} (\bar{\bar{P}}_i - P_i) P_j - \gamma_{ij} s_{ij}^{-v_{ij}} (\bar{\bar{P}}_j - P_j) P_i) \quad .$$

This is the most simple model of the population change of an urban region in a closed system, where the equilibrium population is constant over the period of estimation and the natural increase is a constant proportion of the population. A difference equation was derived from equation 11,

$$10. \qquad \Delta P_i = \beta_i P_{it} + \sum_{j \neq i} (\gamma_{ji} s_{ji}^{-v_{ji}} (\bar{\bar{P}}_i - P_{it}) P_{jt} - \gamma_{ij} s_{ij}^{-v_{ij}} (\bar{\bar{P}}_j - P_{jt}) P_{it})$$

where,

$$11. \qquad \Delta P_i = P_{i,t+1} - P_{i,t}$$

## TABLE 1

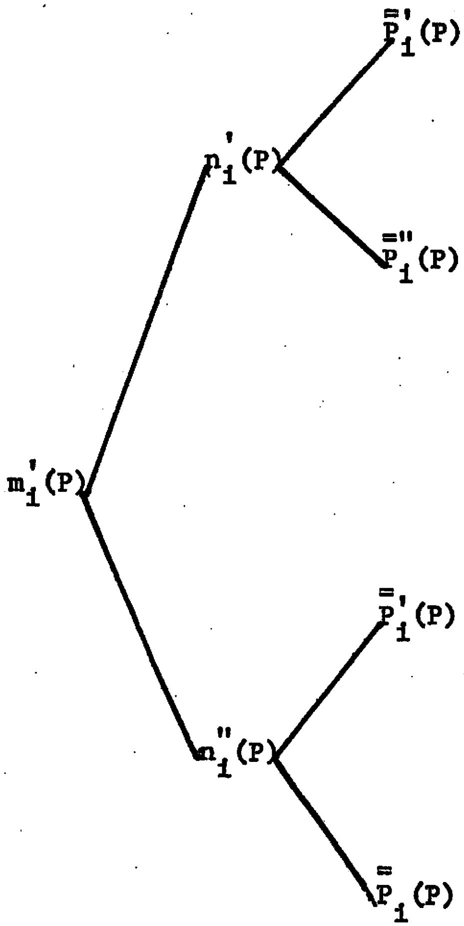
## FAMILY OF EQUATIONS

CLOSED SYSTEM

- $n'_i(P)$
  - $\bar{\bar{P}}'_i(P)$: $$\dot{P}_i = \beta_i P_i + \sum_{j\neq i}(\gamma_{ji} s_{ji}^{-v_{ji}}(\bar{\bar{P}}_i - P_i)P_j - \gamma_{ij} s_{ij}^{-v_{ij}}(\bar{\bar{P}}_j - P_j)P_i)$$
  - $\bar{\bar{P}}''_i(P)$: $$\mathring{P}_i = \beta_i P_i + \sum_{j\neq i}(\gamma_{ji} s_{ji}^{-v_{ji}}((\delta_i + \epsilon_i \sum_j P_j s_{ij}^{-\mu}) - P_i)P_j - \gamma_{ij} s_{ij}^{-v_{ij}}((\delta_j + \epsilon_j \sum_j P_j s_{ij}^{-\mu}) - P_j)P_i)$$
- $n''_i(P)$
  - $\bar{\bar{P}}'_i(P)$: $$\dot{P}_i = \beta_i(\bar{P}_i - P_i)P_i + \sum_{j\neq i} + \gamma_{ji} s_{ji}^{-v_{ji}}(\bar{\bar{P}}_i - P_i)P_j - \gamma_{ij} s_{ij}^{-v_{ij}}(\bar{\bar{P}}_j - P_j)P_i)$$
  - $P''_i(P)$: $$\dot{P}_i = \beta_i(P_i - P_i)P_i + \sum_{j\neq i}(\gamma_{ji} s_{ji}^{-v_{ji}}((\delta_i + \epsilon_i \sum_j P_j s_{ij}^{-\mu}) - P_i)P_j - \gamma_{ij} s_{ij}^{-v_{ij}} ((\delta_j + \epsilon_j \sum_j P_j s_{ij}^{-\mu}) - P_j)P_i)$$

Table 1 Continued.....

Table 1 Continued...




$$\dot{P}_i = \beta_i P_i + \sum_{j\neq i}(\gamma_{ji}s_{ji}^{-v_{ji}}(\bar{\bar{P}}_i - P_i)P_j - \gamma_{ij}s_{ij}^{-v_{ij}}(P_j - P_j)P_i) + \gamma_i(\bar{\bar{P}}_i - P_i)$$

$$\dot{P}_i = \beta_i P_i + \sum_{j\neq i}(\gamma_{ji}s_{ji}^{-v_{ji}}((\delta_i + \varepsilon_i\sum_j P_j s_{ij}^{-\mu}) - P_i)P_j - \gamma_{ij}s_{ij}^{-v_{ij}} ((\delta_j + \varepsilon_j\sum_j P_j s_{ij}^{-\mu}) - P_j)P_i) + \gamma_i((\delta_i + \varepsilon_i\sum_j P_j s_{ij}^{-\mu}) - P_i)$$

$$\dot{P}_i = \beta_i(\bar{P}_i - P_i) + \sum_{j\neq i}(\gamma_{ji}s_{ji}^{-v_{ji}}(\bar{\bar{P}}_i - P_i)P_j - \gamma_{ij}s_{ij}^{-v_{ij}}(\bar{\bar{P}}_j - P_j)P_i) + \gamma_i(\bar{\bar{P}}_i -$$

$$\dot{P}_i = \beta_i(P_i - P_i) + \sum_{j\neq i}(\gamma_{ji}s_{ji}^{-v_{ji}}((\delta_i + \varepsilon_i\sum_j P_j s_{ij}^{-\mu}) - P_i)P_j - \gamma_{ij}s_{ij}^{-v_{ij}} ((\delta_j + \varepsilon_j\sum_j P_j s_{ij}^{-\mu}) - P_j)P_i)$$

OPEN SYSTEM

Table 1 Continued......

Table 1 Continued...

OPEN SYSTEM

- $m_i''(P)$
  - $n_i'(P)$
    - $\bar{\bar{P}}_i'(P)$

$$\dot{P}_i = \beta_i P_i + \sum_{j\neq i}(\gamma_{ji} s_{ji}^{-v_{ji}}(\bar{\bar{P}}_i - P_i)P_j - \gamma_{ij} s_{ij}^{-v_{ij}}(P_j - P_j)P_i) + \gamma_{i.}(\bar{\bar{P}}_i(P) - P_i) - \gamma_{.i}(P_i - \bar{P}_i)P_i$$

    - $\bar{\bar{P}}_i(P)$

$$\dot{P}_i = \beta_i P_i + \sum_{j\neq i}(\gamma_{ji} s_{ji}^{-v_{ji}}((\delta_i + \epsilon_i \sum_j P_j s_{ij}^{-\mu}) - P_i)P_j - \gamma_{i.} s_{ij}^{-v_{ij}} ((\delta_j + \epsilon_j \sum_j s_{ij}^{-\mu} P_j) - P_j)P_i + \gamma_{i.}((\delta_i + \epsilon_i \sum_j s_{ij}^{-\mu} P_j) - P_i) - \gamma_{i.}(P_i - \bar{P}_i)P_i$$

  - $n_i'(P)$
    - $\bar{\bar{P}}_i(P)$

$$\dot{P}_i = \beta_i(P_i - P_i) + \sum_{j\neq i}(\gamma_{ji} s_{ji}^{-v_{ji}}(\bar{\bar{P}}_i - P_i)P_j - \gamma_{ij} s_{ij}^{-v_{ij}}(\bar{\bar{P}}_j - P_j)P_i) + \gamma_{i.}(\bar{\bar{P}}_i - P_i) - \gamma_{i.}(P_i - \bar{P}_i)P_i$$

    - $\bar{\bar{P}}_i(P)$

$$\dot{P}_i = \beta_i(\bar{P}_i - P_i) + \sum_{j\neq i}(\gamma_{ji} s_{ji}^{-v_{ji}}((\delta_i + \epsilon_i \sum_j P_j s_{ij}^{-\mu}) - P_i)P_j - \gamma_{ij} s_{ij}^{-v_{ij}} ((\delta_j + \epsilon_j \sum_j P_j s_{ij}^{-\mu}) - P_j)P_i) + \gamma_{i.}((\delta_i + \epsilon_i \sum_j s_{ij}^{-\mu} P_j) - P_i) - \gamma_{i.}(P_i - \bar{P}_i)P_i$$

12. $$n_i(P) = \beta_i P_{i,t}$$

13. $$\bar{\bar{P}}_i(P) = \bar{\bar{P}}_i \text{ , a constant}$$

14. $$s_{ij} = s_{ji} \quad .$$

After collecting terms,

15. $$P_i = (\beta_i - \sum_{j \neq i} \gamma_{ij} s_{ij}^{-v_{ij}} \bar{\bar{P}}_j) P_{it} + \sum_{j \neq i} (\gamma_{ji} s_{ji}^{-v_{ji}} \bar{\bar{P}}_i) P_{jt} + \sum_{j \neq i} (\gamma_{ij} s_{ij}^{-v_{ij}} - \gamma_{ji} s_{ji}^{-v_{ji}}) P_{it} P_{jt} \quad .$$

The general form of this equation is,

16. $$P_i = C_{ii} P_{it} + \sum_{j \neq i} C_{ij} P_{jt} + \sum_{j \neq i} C'_{ij} P_{it} P_{jt} \quad ,$$

in which,

17. $$C'_{ij} = \gamma_{ij} s_{ij}^{-v_{ij}} - \gamma_{ji} s_{ji}^{-v_{ji}}$$

and

18. $$C'_{ji} = \gamma_{ji} s_{ji}^{-v_{ji}} - \gamma_{ij} s_{ij}^{-v_{ij}} \quad ,$$

therefore,

$$19. \qquad C'_{ij} = -C'_{ji} \quad .$$

Equation 18 is estimated by least squares analysis to obtain estimates of the matrices, C and C'.[1] The parameter values, $\beta_i$, $\bar{\bar{P}}_i$, and $\gamma_{ij}$, are calculated from these estimates. Before discussing the estimation procedure, a short discussion of the significance and possible interpretations of the model's coefficients will be presented.

## II.3 INTERPRETATION

The contributions of three difference sources of population change are evaluated by the coefficients of equation 15. The first coefficient, $C_{ii}$, measures the population change resulting from the difference between the natural increase of the population of region i and the loss of population to other regions by migration. The number of migrants to region i from all other regions in the system is evaluated by the second set of coefficients, $C_{ij}$. The third term, $C'_{ij}$, is the most difficult term to substantively interpret. In this coefficient, the presence of attributes which influence the propensity to migrate, between any pair of regions is measured. This term cannot be successfully interpreted until the model is empirically tested.

Four parameters will be estimated from this model. The growth of population caused by natural increase, $\beta_i$, the equilibrium population,

[1]All matrices and vectors will be identified by underlining.

$\bar{\bar{P}}_i$, the propensity to migrate between any region in the system, $\gamma_{ij}$, and the friction of distance, $v_{ij}$. According to Batty (1972a) the calibration of parameters such as the friction of distance has been poorly investigated in the geographical iterature. Certainly, numerical values are arbitrarily assigned to these parameters in many urban models. Of these parameters, perhaps equilibrium population is the most poorly defined concept. In this analysis, the equilibrium city size is calculated within the model, rather than subjectively defined. Similarly, the propensity to migrate between urban regions and the friction of distance between them are derived from the model, which may be applied to any urban system. This is a model of population change developed from simple assumptions describing the dynamic behaviour of an urban system. In the next chapter, the estimation of this model by least squares analysis will be discussed.

## CHAPTER III

## ESTIMATION PROCEDURE

### III.1 INTRODUCTION

From equation 18, a system of equations describing the pattern of population change of each urban region can be defined. For a system composed of only three centres, the following system of equations describes the pattern of population change of the urban system at any point in time,

20. $\Delta P_1 = C_{11}P_{1t}+C_{12}P_{2t}+C_{13}P_{3t}+C'_{12}P_{1t}P_{2t}+C'_{13}P_{1t}P_{2t}+\epsilon_1$

21. $\Delta P_2 = C_{21}P_{1t}+C_{22}P_{2t}+C_{23}P_{3t}-C'_{12}P_{1t}P_{2t}+C'_{23}P_{2t}P_{3t}+\epsilon_2$

22. $\Delta P_3 = C_{31}P_{1t}+C_{32}P_{2t}+C_{33}P_{3t}-C'_{13}P_{1t}P_{3t}-C'_{23}P_{2t}P_{3t}+\epsilon_3$

This is a system of dependent equations, since the coefficients, $C'_{12}$, $C'_{13}$, and $C'_{23}$, appear in more than one equation. Originally, these equations were to be estimated by multiple linear regression,[1] however, the dependency of the equations required simultaneous estimation of the coefficients. The simultaneous least squares procedure will be

[1]According to Draper and Smith, (1966, 9), a model is linear or nonlinear in the parameters, therefore this is a second order multiple linear regression model.

described in the next section, then, the derivation of the parameter values from the coefficients will be discussed. Finally, the problems associated with this estimation procedure will be reviewed.

## III.2 LEAST SQUARES ESTIMATION

In a multiple regression model the following equation is estimated,

23. $\underline{Y} = \underline{Xb} + \underline{e}$ ,

in which $\underline{Y}$ is a vector of dependent observations, $\underline{X}$ is a matrix of independent variables, $\underline{b}$ is a matrix of parameter estimates, and $\underline{e}$ is a vector of error terms. The parameter estimates, $\underline{b}$, are chosen such that the sum of the squared deviations of the observed dependent variables, $\underline{Y}$, from the estimated values of the dependent variables, $\hat{\underline{Y}}$, is minimized, according to equation 24

24. $S = \sum_i \sum_j (Y_{ij} - \hat{Y}_{ij})^2$

The solution is obtained by calculating,

25. $\underline{b} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{Y}$

where $\underline{X}'$ is the transpose of $\underline{X}$, and $\underline{X}^{-1}$ is the inverse matrix. The estimated values of Y are calculated according to the following formula,

26. $\underline{\hat{Y}} = \underline{Xb}$

To insure that the constraint, $C'_{ij} = C'_{ji}$, is satisfied, the parameter values of the system of equations must be estimated simultaneously. Initially, it was thought that the matrix of parameter estimates, $\underline{b}$, could be obtained by first defining the normal equations corresponding to each regression equation, forming a matrix of the sums calculated from these normal equations, and solving the resulting system of simultaneous equations.[2] However, a more simple procedure is available. The matrix of independent variables, $\underline{X}$, is first reorganised so that the observations from which each coefficient will be calculated are in one column. In the case of the coefficient, $C'_{12}$, this means the observations for city one and for city two are in one column. Then, the equations are attached one beside each other. Thus, a matrix of independent variables, consisting of twelve columns and 3n rows, for a system of three cities, is defined. This matrix is illustrated in Table 2.

The coefficients of this matrix are estimated according to equation 25. These estimates are then used to calculate the parameter values of the model.

From Equation 15,

$$\Delta P_i = (\beta_i - \sum_{j \neq i} \gamma_{ij} s_{ij}^{-v_{ij}} \bar{\bar{P}}_j) P_{it} - \sum_{j \neq i} (\gamma_{ji} s_{ji}^{-v_{ji}} \bar{\bar{P}}_i) P_{jt}$$
$$+ \sum_{j \neq i} (\gamma_{ij} s_{ij}^{-v_{ij}} - \gamma_{ji} s_{ji}^{-v_{ji}}) P_{it} P_{jt} \quad ,$$

[2]For a more detailed description of this procedure see Appendix I.

TABLE 2

DEFINITION OF MATRIX OF INDEPENDENT VARIABLES

| Y | X |
|---|---|
| | $P_1\ P_2\ P_3\ P_1P_2\ P_1P_3\ P_1\ P_2\ P_3\ P_2P_3\ P_1\ P_2\ P_3$ |
| $Y_{11}$ | $P_{11}P_{21}P_{31}P_{11}P_{21}P_{11}P_{31}$ .................... |
| $Y_{1n}$ | $P_{1n}P_{2n}P_{3n}P_{1n}P_{2n}P_{1n}P_{3n}$ .................... |
| $Y_{21}$ | ........ $P_{11}P_{21}$ ..... $P_{11}P_{21}P_{31}P_{21}P_{31}$ .... |
| $Y_{2n}$ | ........ $P_{1n}P_{2n}$ ..... $P_{1n}P_{2n}P_{3n}P_{2n}P_{3n}$ .... |
| $Y_{31}$ | ............... $P_{11}P_{31}$ ........ $P_{21}P_{31}P_{11}P_{21}P_{31}$ |
| $Y_{3n}$ | ............... $P_{1n}P_{3n}$ ........ $P_{2n}P_{3n}P_{1n}P_{2n}P_{3n}$ |

the set of equations, shown in Table 3, is defined. The parameter values, $\beta_i$, $\bar{\bar{P}}_i$, and $\gamma_{ij}$, can be calculated from this system of equations if it is assumed that the friction of distance, $v_{ij}$, is equal to one for all urban regions included in the system. Estimates of the rate of natural increase, $\beta_i$, are obtained by substitution according to the following formula,

27. $$\beta_i = C_{ii} - \sum_{j \neq i} C_{ji}$$

Similarly, the equilibrium population, $P_i$, can be estimated by solving a system composed of the following type of equations,

28. $$C'_{ij} = C_{ji}\bar{\bar{P}}_i^{-1} - C_{ij}\bar{\bar{P}}_j^{-1} \quad .$$

If the number of centres is larger than three, a nonlinear program must be employed to calculate this parameter, since the system of equations is overdefined. The number of unique elements in the matrix $\underline{C}'$ are,

29. $$N = m(m-1)/2$$

where N is the total number of unique elements, and m is the number of centres in the system. If the number of centres is greater than three, N is greater than m, therefore, the system is overdefined. The objective function of the nonlinear programme minimizes the squared deviations between the estimated values of $\underline{C}'$ and the values calculated according

## TABLE 3

## DEFINITIONS OF COEFFICIENTS

$$C_{11} = \beta_1 - \gamma_{12} s_{12}^{-v_{12}} \bar{\bar{P}}_2 - \gamma_{13} s_{13}^{-v_{13}} \bar{\bar{P}}_3$$

$$C_{22} = \beta_2 - \gamma_{21} s_{21}^{-v_{21}} \bar{\bar{P}}_1 - \gamma_{23} s_{23}^{-v_{23}} \bar{\bar{P}}_3$$

$$C_{33} = \beta_3 - \gamma_{31} s_{31}^{-v_{31}} \bar{\bar{P}}_1 - \gamma_{32} s_{32}^{-v_{32}} \bar{\bar{P}}_2$$

$$C_{12} = \gamma_{21} s_{21}^{-v_{21}} \bar{\bar{P}}_1$$

$$C_{13} = \gamma_{31} s_{31}^{-v_{31}} \bar{\bar{P}}_1$$

$$C_{21} = \gamma_{12} s_{12}^{-v_{12}} \bar{\bar{P}}_2$$

$$C_{23} = \gamma_{32} s_{32}^{-v_{32}} \bar{\bar{P}}_2$$

$$C_{31} = \gamma_{13} s_{13}^{-v_{13}} \bar{\bar{P}}_3$$

$$C_{32} = \gamma_{23} s_{23}^{-v_{23}} \bar{\bar{P}}_3$$

$$C'_{12} = \gamma_{12} s_{12}^{-v_{12}} - \gamma_{21} s_{21}^{-v_{21}}$$

$$C'_{13} = \gamma_{13} s_{13}^{-v_{13}} - \gamma_{31} s_{31}^{-v_{13}}$$

$$C'_{23} = \gamma_{23} s_{23}^{-v_{23}} - \gamma_{32} s_{32}^{-v_{32}}$$

to equation 28. By repeated iterations, estimates of the equilibrium population of each centre are calculated.

After substitution of the estimates of $\bar{\bar{P}}_i^{-1}$, the propensities to migrate may be estimated by the following,

$$30. \qquad \gamma_{ij} = c_{ji} s_{ji} \bar{\bar{P}}_j^{-1}$$

The complete set of equations for a system of three cities is shown in Table 4.

Although the additional assumption that the friction of distance is constant implies only that the transportation system is uniform throughout the urban system, the accuracy of the estimates of the coefficients, $\underline{C}'$ and $\underline{C}$, must be established. In the next section, the accuracy and validity of the least squares procedure will be discussed.

## III.3 EVALUATION OF THE ESTIMATION PROCEDURE

The estimates of the regression coefficients obtained by least squares analysis have several properties. They minimize the error sum of squares, and are linear functions of the dependent variables which provide unbiased estimtaes of $\underline{B}$. The variance of the sampling distribution from which these estimates are drawn is minimized by the least squares procedure. According to Draper and Smith (1966, 59) these properties are independent of the distribution of the error terms. More rigorous assumptions are made in the inferential tests of the significance of these estimates.

TABLE 4

EQUATIONS DEFINING PARAMETERS OF THE MODEL

$$\beta_1 = C_{11}-C_{21}-C_{31}$$

$$\beta_2 = C_{22}-C_{12}-C_{32}$$

$$\beta_3 = C_{33}-C_{13}-C_{23}$$

$$\gamma_{21} = C_{12}s_{21}\bar{\bar{P}}_1^{-1}$$

$$\gamma_{31} = C_{13}s_{31}\bar{\bar{P}}_1^{-1}$$

$$\gamma_{12} = C_{21}s_{12}\bar{\bar{P}}_2^{-1}$$

$$\gamma_{32} = C_{23}s_{32}\bar{\bar{P}}_2^{-1}$$

$$\gamma_{13} = C_{31}s_{13}\bar{\bar{P}}_3^{-1}$$

$$\gamma_{23} = C_{32}s_{23}\bar{\bar{P}}_3^{-1}$$

$$C'_{12} = C_{21}\bar{\bar{P}}_1^{-1}-C_{12}\bar{\bar{P}}_2^{-1}$$

$$C'_{13} = C_{31}\bar{\bar{P}}_1^{-1}-C_{13}\bar{\bar{P}}_3^{-1}$$

$$C'_{23} = C_{32}\bar{\bar{P}}_2^{-1}-C_{23}\bar{\bar{P}}_3^{-1}$$

In the tests of significance of the coefficient of determination, $\underline{R}^2$, which measures the percentage of the variance which is explained by the analysis, and of the hypothesis that the regression coefficients are significantly different from zero, it is assumed that the error terms are normally distributed and are independent of each other. Furthermore, it is assumed that the values of the independent variables are drawn from sampling distributions which are normally distributed. None of these assumptions is satisfied in this model. The sampling distribution of the independent variables cannot be established since only one population value is possible at each point in time.

There is also multicollinearity among the independent variables. According to Johnston (1960, 201)

> "This is the name given to the general problem which arises when some or all of the explanatory variables in a relation are so highly correlated one with another that it becomes very difficult, if not impossible, to disentangle their separate influences and obtain a reasonably precise estimate of their relative effects.

One of the basic assumptions of this model is that the size and spatial arrangement of the urban regions will influence the pattern of population change of each region. Thus, the problem of multicollinearity is implicit in the model, however, this does not significantly affect the numerical estimates of the coefficients. It increases the standard error of estimate, by which the significance of these estimates are tested, and may increase the coefficient of determination. If the estimates are significantly different from zero, despite multicollinearity, Johnston (1960, 202) suggests that the problem may be disregarded.

Autocorrelation is a more serious problem. If the model is accurate, the residuals should not be spatially or temporally independent. It is assumed that the pattern of population growth at one centre is accounted for by the pattern of growth of other centres. Similarly, the pattern of residuals through time is not random, since the population change in one time period is explained by the size of the population in the previous time period. If the population change in one time period is seriously underestimated, it is likely that subsequent estimates will also be low because the model does not include some factor which is influencing the pattern of population growth. Although, the temporal autocorrelation of the residuals can be measured by a Durbin-Watson test, and the degree of spatial autocorrelation by a statistic developed by Cliff and Ord, (Curry, 1972, 133), their effects can be removed only by redefining the dependent variable. Curry (1972) has also commented that the distance parameters of the gravity model, estimated by regression analysis, reflect both the friction of distance and the historical development of the landscape where the interaction is occurring. Similarly, Barry (1972b) has commented that high coefficients of determination do not adequately measure the goodness of fit of a regression plane, instead the sensitivity of parameters, such as the mean trip length, are better tests of the degree of fit. He suggested that a model in which there are x variables and y parameters should be tested $x^y$ times to determine the sensitivity of the parameter estimates (Batty, 1972b, 164).

Furthermore, since these equations have been estimated simultaneously how are the degrees of freedom of inferential tests determined? The coefficients have been estimated for n time periods but for three cities, and hence, are the degrees of freedom, 3n-1, or n-1? There is no discussion of this problem in the geographical applications of regression analysis.

To some extent, the assumptions of the estimation procedure are incompatible with the assumptions of the theoretical model, but the effects of such problems as multicollinearity and autocorrelation can be measured. Certainly, least squares estimation of these coefficients provides estimates which minimize the deviations of the coefficients from the regression plane.

In the next chapter, the problems of data selection will be discussed before the empirical results are presented.

# CHAPTER IV

# EMPIRICAL ANALYSIS

Empirical tests of the model have been restricted by the limited amount of available computer space. The patterns of population change of small urban systems composed of only three regions have been analysed. The model has been tested on three sets of data; population and employment data for a system composed of Hamilton, London, and Toronto, and employment data for Hamilton, Brantford, and Kitchener. The selection of data and the results of these empirical tests will be discussed in the remainder of this chapter.

## IV.1 DATA

The analyses are constrained by the availability of data. The model should be tested with demographic data, however, this information is not available at frequent intervals in Canada. Consequently, employment data which are reported monthly have been substituted for population data.

The use of employment data to estimate patterns of migration and population change is well documented. Lowry (1966) employed two variables, the size of the civilian labour force and the number of military personnel in American cities, as surrogates for population size. Similarly, in the report, Urban Canada, patterns of population growth were estimated by first predicting the size of the labour force in Canadian cities, and then calculating the population size as a percentage of the labour force (Lithwick, 1970). Although, population data would be preferable within the structure of the model, employment data provide a useful surrogate for this information.

The total employment in six industrial groups was a surrogate for population size.[1] The data were collected from Publication No. 72-002, Employment, Earnings, and Hours, published monthly by Statistics Canada. Information is collected from establishments having fifteen or more employees, however, in 1970, the industrial composite of the major industrial groups accounted for 56.7% of the total estimated Canadian employment (Statistics Canada, 1970). Although, the coverage of this series has increased, only eighteen Ontario cities have been covered continuously between January 1958, and December 1971.[2] A twenty year time period allows the estimation of the model over time periods of varying length so that the rate of change of the optimum population may be observed. The use of employment data also requires that the parameter, $\beta_i$, be redefined as the rate of growth of employment in basic industries which service the urban population, rather than the rate of natural increase.

Some population data were obtained from census information for the years, 1951, 1956, 1961, 1966, and 1971 (Ontario Population Statistics, 1971). Quarterly estimates were interpolated between these five points. Consequently, the population values within the five year

---

[1]These industries are manufacturing, construction, transportation and communication, trade, finance, and services (Statistics Canada, 1970).

[2]These cities are Ottawa-Hull, St. Catharines, Toronto, Hamilton, Brantford, Kitchener, London, Windsor, Thunder Bay, Peterborough, Oshawa, Niagara Falls, Kingston, Guelph, Sudbury, Timmins, Sarnia, Sault Ste. Marie (Statistics Canada, 1970).

intervals increase equally at each time period. Since population change is the dependent variable, and the population sizes are the independent variables, the problems of multicollinearity and autocorrelation are increased with these data. Nevertheless, an empirical test using these data was performed. The choice of the sets of three cities was governed by the availability of the employment data. Sets of three cities were chosen from the group of eighteen cities which were continuously reported between 1958 and 1971.

Two urban hierarchies have been defined in Southern Ontario (Simmons, 1972; Brummell, 1972). Using data describing the volume of telephone calls between urban centres in Ontario, Simmons (1972) identified systems of urban centres by assigning a centre to the hinterland of a larger centre if the largest proportion of its telephone calls were placed to that centre. Thus, Sarnia, St. Thomas, and Stratford are included in the network which focusses upon London, which is, itself, part of the larger regional network centred upon Toronto. Six regional centres focus upon Toronto, London, St. Catharines, Kitchener, Owen Sound, Windsor, and Brantford. The proportion of a centre's calls which are accounted for by its largest flow vary significantly, for instance, 24.9% of Windsor's long distance residential calls end in Toronto, while 54.8% of Barrie's residential calls are destined for Toronto (Simmons, 1972, 206).

Brummell defined a hierarchy on the basis of scores on two principal components, which measured the population size of each centre and its variety of retail establishments. All the urban centres were assumed to be within the urban field of Toronto, therefore, centres such as Windsor which is within the urban field of Detroit, and Ottawa which is strongly linked to Montreal, were excluded. A discriminant analysis of the preliminary groups identified four levels of this hierarchy. Toronto is a fourth order centre, while London and Hamilton are third order centres within its urban field.

Unfortunately, the employment data do not cover many of the urban centres at the lower levels of either of these hierarchies. Consequently, Hamilton, London, and Toronto were chosen as the initial set of cities. A larger proportion of the interaction originating in Hamilton and London should end in Toronto than in other urban centres. Similarly, a second set of cities, Hamilton, Brantford, and Kitchener was chosen. Hamilton is the closest third order centre to these two second order centres, therefore, a significant proportion of their total interaction flows towards Hamilton.

The problems of shifting municipal boundaries were avoided by the use of employment and population data which were collected throughout metropolitan Toronto and throughout metropolitan Hamilton. London, Brantford, and Kitchener are smaller centres, therefore statistics are compiled for the municipalities themselves.[3]

The analysis of the patterns of population and employment change of these two sets of urban centres should identify different patterns of change at different levels of theurban hierarchy.

## IV.2 REGRESSION ANALYSIS

The results of the least squares analysis of the employment data

[3]The distances which were calculated from the Ontario 1971 Official Road Map are listed in Appendix IV.

will be discussed first.

IV.2.1 Employment Data

To observe the behaviour of the model's parameters through time, a series of regression analyses were performed. The data was analysed for time periods of varying lengths to determine how rapidly the parameters changed.

The results of the regression analyses are very disappointing. Although more than half the coefficients of determination are significantly different from zero, the level of explanation of the analyses are very low. A larger proportion of the variance is explained when the length of time over which the equations are estimated is increased. This is not a consistent trend, as the coefficients of determination for the analysis of five year time periods indicate. Only when quarterly data for the fifteen year period are analysed does the level of explanation increase significantly for both urban systems. The high coefficient of determination for the Brantford, Kitchener, Hamilton system in time periods which include 1966, is caused by large fluctuations in employment resulting from strike activities (see Table 5). Despite the low levels of explanation, the regression coefficients are generally significantly different from zero.[4] Yet, the patterns of residuals are very clustered when plotted against the estimated depend t variable, $\hat{Y}$. Durbin Watson tests of the residuals indicate no consistent pattern of autocorrelation. The poor explanatory power of the model may be caused by autocorrelation of the residuals and of the data. The data are

---

[4]A complete list of the estimated coefficients is in Appendix II.

TABLE 5

LEVEL OF EXPLANATION OF REGRESSION ANALYSES

| | Hamilton London Toronto | | | Brantford Hamilton Kitchener | | |
|---|---|---|---|---|---|---|
| Time | $R^2$ | sig. | No. of sig. coefficients | $R^2$ | sig. | No. of sig. coefficients |
| 1958-60 | .22534 | .05 | 4 | .16937 | | 8 |
| 1961-63 | .39370 | .01 | 11 | .17534 | | 4 |
| 1964-66 | .1791 | | 4 | .54766 | .01 | 12 |
| 1967-69 | .14287 | | 12 | .07435 | | 12 |
| 1970-72 | .13436 | | 11 | .23512 | .05 | 11 |
| 1958-61 | .24437 | .01 | 12 | .11130 | | 11 |
| 1962-65 | .22927 | .05 | 12 | .22332 | .05 | 12 |
| 1966-69 | .12797 | | 11 | .41950 | .01 | 12 |
| 1958-62 | .26704 | .01 | 12 | .15916 | | 12 |
| 1963-67 | .14370 | .05 | 12 | .51225 | .01 | 12 |
| 1968-72 | .09970 | | 12 | .1530 | .05 | 11 |
| 1958-72 | .46705 | .01 | 12 | .44837 | .01 | 12 |

TABLE 6

DURBIN WATSON TESTS OF RESIDUALS

| Time | Hamilton London Toronto | Brantford Hamilton Kitchener |
|---|---|---|
| 1958-60 | positive autocorrelation | autocorrelation of independent variables and of residuals |
| 1961-63 | autocorrelation of independent variables and of residuals | positive autocorrelation |
| 1964-66 | no autocorrelation | no autocorrelation |
| 1967-69 | autocorrelation of independent variables and of residuals | no autocorrelation |
| 1970-72 | no autocorrelation | no autocorrelation |
| 1958-61 | no autocorrelation | autocorrelation of independent variables and of residuals |
| 1962-65 | no autocorrelation | positive autocorrelation |
| 1966-69 | no autocorrelation | no autocorrelation |
| 1958-62 | positive autocorrelation | no autocorrelation |
| 1963-67 | no autocorrelation | no autocorrelation |
| 1968-72 | no autocorrelation | autocorrelation of independent variables and of residuals |
| 1958-72 | no autocorrelation | positive autocorrelation |

rounded off to thousands, because the methods of collection and of compilation of the information have varied over the fifteen year time period, therefore, monthly changes are very small compared to the total size of the labour force. The small range of the dependent and independent variables violates the assumption of heteroscedacity. This may contribute to the low level of explanation of the analyses.

The level of explanation of the analyses may also be decreased by the constraint that,

31. $$C'_{ij} = -C'_{ji} .$$

A series of analyses of the Hamilton, London, and Toronto data indicated that the coefficient of determination increased when the constraint was removed. Unfortunately, the regression equations still accounted for less than half of the variance.

The model has failed to adequately describe the pattern of employment growth of either urban system. The results might be improved by the inclusion of a larger number of centres, or by the use of population data which fluctuate less erratically than the employment data. The regression coefficients, estimated by these analyses show no consistent trends. The third term in the regression equation seems to contribute less significantly to the explanation of variance, since its coefficients, $C'_{ij}$, are generally smaller and less significant. Perhaps, the removal of this term might increase the level of explanation of the analyses.

TABLE 7

RESULTS OF UNCONSTRAINED REGRESSION ANALYSES

| | Hamilton London Toronto | |
|---|---|---|
| Time | $R^2$ | sig. |
| 1958-60 | .39661 | .01 |
| 1961-63 | .52639 | .01 |
| 1964-66 | .17723 | |
| 1967-69 | .23581 | .01 |
| 1970-72 | .22525 | |
| 1958-61 | .38934 | .01 |
| 1962-65 | .30976 | .01 |
| 1966-69 | .20508 | .05 |

IV.2.2 Population Data

A similar series of regression analyses of the population data explained a much larger proportion of the variance as the values in Table 8 indicate. The plots of the residuals against the dependent variables indicate that these results are misleading. The dependent observations are interpolated quarterly measures of population change. These values are interpolated from a straight line, therefore, the population change in any quarter is constant everywhere along the line. The plots of the residuals are clustered around these five values of the dependent variable. Yet, the Durbin Watson tests do not identify any patterns of autocorrelation (Table 8).

Despite the model's failure to explain a significant proportion of the variance of the employment data, and the clustered patterns of the residuals of the population data, the majority of the beta coefficients are significant. There are no significant patterns in either set of beta coefficients, except that some coefficients equal zero. This indicates that the pattern of population change of some urban centres is not explained by the population size of other centres, however, there are very few examples of coefficients which equal zero (Appendix II).

IV.3 PARAMETER ESTIMATES

The parameter values were calculated according to the method described in Chapter 3, however, the results are inconclusive.

TABLE 8

RESULTS OF REGRESSION ANALYSES OF POPULATION DATA

| Time | $R^2$ | Hamilton London Toronto sig. | No. of sig. coefficients p=.05 | Durbin Watson Test |
|---|---|---|---|---|
| 1951-53 | .9997 | .01 | 10 | positive autocorrelation |
| 1954-56 | .9996 | .01 | 9 | no autocorrelation |
| 1957-59 | .995 | .01 | 2 | autocorrelation of both residuals and independent variables |
| 1960-62 | .999 | .01 | 8 | autocorrelation of residuals and independent variables |
| 1963-65 | .999 | .01 | 9 | autocorrelation of residuals and independent variables |
| 1966-68 | .9997 | .01 | 7 | no autocorrelation |
| 1951-55 | .9997 | .01 | 10 | positive autocorrelation |
| 1956-60 | .9995 | .01 | 8 | no autocorrelation |
| 1961-65 | .9994 | .01 | 10 | autocorrelation of residuals and independent variables |
| 1966-70 | .994 | .01 | 11 | no autocorrelation |
| 1951-60 | .996 | .01 | 9 | positive autocorrelation |
| 1961-70 | .995 | .01 | 12 | autocorrelation of residuals and independent variables |

### IV.3.1 Equilibrium Population

The equilibrium populations calculated from the regression analyses of employment data do not seem to follow a consistent trend (Table 9). The equilibrium population of Toronto is generally larger than the equilibrium population of all other centres, however, this is not true for the time period, 1967-1969. The actual size of the equilibrium populations are also too large. They should measure the equilibrium size of the labour force, but are of the order of magnitude of the total population. Furthermore, the standard deviations of these distributions are very large, indicating that the values do not cluster around the mean. When the equilibrium population of each centre is regressed against time, specifically the middle month of the interval over which the least squares analysis was performed, the level of explanation is very low (Table 10). The equilibrium population fluctuates erratically, instead of changing systematically through time. This conclusion is reinforced by the nonsensical values calculated from the regression analyses of the population data. These results are shown in Table 11. The negative equilibrium population values indicate that these parameter estimates do not change systematically. Again, the standard deviations are large, and the results of a simple regression analysis if the equilibrium population estimates against time explain a lower proportion of the variance than the previous estimates (Table 12).

The parameter estimates for each urban centre change with the least squares analysis from which they are calculated. The correlation between the population estimated from employment data and that estimated using population data is between ten and twenty percent for Toronto,

## TABLE 9

### EQUILIBRIUM POPULATIONS BASED ON EMPLOYMENT DATA
### (000's)

| Time | Toronto | Hamilton(1)* | Hamilton(2)* | London | Brantford | Kitchener |
|---|---|---|---|---|---|---|
| 1958-60 | 1118.3 | 277.7 | 383.2 | 230.8 | 8.8 | 87.7 |
| 1961-63 | 2795.9 | 7.1 | 1892.1 | 239.6 | 9.3 | 10.2 |
| 1964-66 | 463.9 | 60.9 | 473.9 | 132.4 | 50.9 | 71.6 |
| 1967-69 | 1275.3 | 71.8 | 204.8 | 99.8 | 101.9 | 1803.3 |
| 1970-71 | 2825.3 | 374.1 | 134.6 | 66.8 | 52.2 | 67.6 |
| 1958-61 | 581.9 | 243.0 | 211.6 | 47.3 | 31.0 | 120.7 |
| 1962-65 | 1048.6 | 142.3 | 175.0 | 95.5 | 39.1 | 105.9 |
| 1966-69 | 1009.8 | 289.2 | 190.8 | 84.0 | 36.1 | 132.0 |
| 1958-62 | 721.4 | 171.9 | 41.6 | 59.3 | 10.1 | 9.8 |
| 1963-67 | 1081.9 | 123.9 | 185.0 | 115.5 | 3.1 | 51.8 |
| 1968-72 | 1240.0 | 219.9 | 76.2 | 101.2 | 55.1 | 62.3 |

*Hamilton(1) are the equilibrium population values calculated from the estimation of Toronto, Hamilton, and London. Hamilton(2) are the optimum population values from the estimation of Hamilton, Brantford, and Kitchener.

TABLE 10

EQUILIBRIUM POPULATION STATISTICS

| | Mean Value (000's) | Standard Deviation | Regression Equation | $R^2$ |
|---|---|---|---|---|
| Toronto | 1287.48 | 796.5 | 849.3+5.3t | .099 |
| Hamilton(1) | 180.17 | 111.8 | 140.2+5t | .045 |
| Hamilton(2) | 360.8 | 522.8 | 850.2-5.8t | .308 |
| London | 115.6 | 64.1 | 100.1+.4t | .102 |
| Brantford | 36.1 | 29.0 | 40.2-.05t | .067 |
| Kitchener | 229.3 | 523.5 | 404.4-2.1t | .039 |

TABLE 11

EQUILIBRIUM POPULATIONS BASED ON POPULATION DATA

| Time | Toronto | Hamilton | London |
|---|---|---|---|
| 1951-53 | 7143.1 | 190.0 | 64.6 |
| 1954-56 | 285.13 | 234.4 | -5569.7 |
| 1957-59 | 797.7 | 118.7 | -7009.1 |
| 1960-62 | 436.4 | 44.4 | -57.7 |
| 1963-65 | 2771.2 | 378.5 | 453.7 |
| 1966-68 | * | * | -129.9 |
| 1951-55 | 3305.9 | 33.6 | 32.4 |
| 1956-60 | 3601.1 | 4611.2 | 4778.6 |
| 1961-65 | -1333.4 | -1029.2 | -481.7 |
| 1966-70 | 10796.5 | 514.3 | -2096.7 |
| 1951-60 | 451.7 | 283.5 | 24722.9 |
| 1961-70 | 4619.5 | 239.6 | -1419.6 |

*No values were obtained because of zero coefficients.

TABLE 12

EQUILIBRIUM POPULATION STATISTICS

| | Mean Value (000's) | Standard Deviation | Regression Equation | $R^2$ |
|---|---|---|---|---|
| Toronto | 2988.6 | 3537.6 | 1490.3+42.3t | .060 |
| Hamilton | 510.8 | 1417.9 | 705.2-5.5t | .007 |
| London | 1107.3 | 8013.4 | 4076.7-78.1t | .050 |

Hamilton, and London. Although the low level of correlation is not surprising since the estimates are based on such different data sets, the negative correlation between the equilibrium population of Toronto calculated from employment data, and the parameter estimates based on population data indicates that the parameters of the model are not stable. They reflect the limitations of the original data. The poor fit of the linear model, and the problems of multicollinearity and autocorrelation may have seriously biased the coefficients from which the parameters are estimated.

The equilibrium state of an urban centre which is represented by the equilibrium population is dynamic. Yet, it fluctuates randomly rather than changing systematically. These random fluctuations may be caused by the limitations of the data and of the model, or they may represent the actual behaviour of the equilibrium state of an urban centre. Further empirical tests will be necessary before this question may be answered. The estimates of the rates of natural increase of the urban centres fluctuate in a similar manner.

IV.3.2 Rate of Natural Increase

Again the parameter estimates do not vary systematically, but fluctuate erratically, as the figures in Table 13 demonstrate. The standard deviations of the estimates of the rate of increase in employment are very large, and there is no linear trend of the parameter estimates through time, as the results in Table 14 indicate. The rates of increase of employment in Toronto and Hamilton are decreasing, while the rate of growth of employment has increased in the smaller urban

TABLE 13

RATE OF EMPLOYMENT GROWTH

| Time | Toronto | Hamilton(1)* | Hamilton(2)* | London | Brantford | Kitchener |
|---|---|---|---|---|---|---|
| 1958-60 | 4.15 | .973 | .177 | -1.48 | 1.004 | -.011 |
| 1961-63 | -.06 | 3.36 | .301 | .88 | -.54 | -.52 |
| 1964-66 | -.12 | .865 | .49 | -.411 | .24 | -1.08 |
| 1967-69 | .07 | -.31 | -.17 | .993 | .03 | .30 |
| 1970-72 | -.11 | -.35 | -.08 | 25.98 | -.023 | .15 |
| 1958-61 | -.26 | .15 | -.10 | 2.76 | .44 | -.19 |
| 1962-65 | -.37 | .53 | .58 | 3.13 | .08 | -1.22 |
| 1966-69 | -.06 | -.34 | 1.96 | 1.53 | 2.03 | -.79 |
| 1958-62 | -.25 | .23 | -.16 | -2.49 | .08 | .23 |
| 1963-67 | -.10 | .33 | .26 | .51 | 1.42 | .99 |
| 1968-72 | -.10 | -.37 | -.19 | 2.23 | .32 | -.31 |

*The values for Hamilton(1) are calculated from the regression analysis of the urban system composed of Hamilton, London, and Toronto. The values for Hamilton(20 are calculated from the analysis of the system composed of Hamilton, Brantford, and Kitchener.

TABLE 14

RATE OF INCREASE OF EMPLOYMENT STATISTICS

| | Mean Value (000's) | Standard Deviation | Regression Equation | $R^2$ |
|---|---|---|---|---|
| Toronto | .2543 | 1.29 | .99-.009t | .244 |
| Hamilton(1) | .4605 | 1.08 | 1.36-.011t | .323 |
| Hamilton(2) | .2784 | .63 | .26+.001t | .002 |
| London | 3.06 | 7.79 | -4.42+.09t | .126 |
| Brantford | .463 | .74 | .41+.001t | .002 |
| Kitchener | -.225 | .65 | -.25+.0003t | .0005 |

centres. The very high rate of increase of London may indicate the effects of decentralisation away from the large centres of Hamilton and Toronto. Again the behaviour of the estimates for Hamilton depends upon the definition of the urban system. In an urban system composed of Toronto, London, and Hamilton, the rate of employment growth of Hamilton decreases through time. On the other hand, the estimate of the rate of employment growth derived from the analysis of the second urban system increases through time, although the relationship is very weak. This parameter, like the equilibrium population, is not stable but changes frequently as a function of the original data, and of the pattern of population change.

Similarly, the rate of natural increase calculated from a least squares analysis of population data changes erratically, as the values in Table 15 indicate. The rate of natural increase is much smaller than the rate of growth of employment, as a comparison of Tables 16 and 14 indicates. The time series of the rates of natural increase cannot be described by a linear trend.

Again a comparison of the parameter estimates for Hamilton, London, and Toronto demonstrates the sensitivity of this parameter. Not only do the absolute values of the parameter change as a function of the data set, but the rate of increase changes from an increasing rate of employment to a decreasing rate of population growth. This contradictory behaviour suggests that employment data may not be an adequate surrogate for population data.

TABLE 15

RATE OF NATURAL INCREASE

| Time | Toronto | London | Hamilton |
|---|---|---|---|
| 1951-53 | -.07 | -.27 | .57 |
| 1954-56 | .33 | -.47 | -1.59 |
| 1957-59 | -.59 | .19 | .10 |
| 1960-62 | -.00 | -.36 | .30 |
| 1963-65 | .00 | -.20 | .17 |
| 1966-68 | .27 | -1.79 | -.52 |
| 1951-55 | -.04 | -.02 | .27 |
| 1956-60 | .21 | -.32 | -.99 |
| 1961-65 | .03 | -.49 | .19 |
| 1966-70 | -.05 | -1.29 | -.37 |
| 1951-60 | .02 | -.00 | -.03 |
| 1961-70 | -.14 | .21 | .24 |

TABLE 16

RATE OF NATURAL INCREASE STATISTICS

| | Mean Value (000's) | Standard Deviation | Regression Equation | $R^2$ |
|---|---|---|---|---|
| Toronto | .064 | .141 | .072-.0001t | .004 |
| Hamilton | -.094 | .659 | -.218+.0002t | .0004 |
| London | -.402 | .589 | .157-.0049t | .308 |

### IV.3.3 Propensity to Migrate

The propensities to migrate have a similar pattern of change. The values change significantly through time, without any underlying pattern.[5] They also vary significantly between data sets, as the average propensities to migrate indicate (Table 17). The standard deviations of the distributions from which these mean values are derived are very large (Table 18). The propensities to migrate are not reflexive, and they depend upon the migrant's origin and destination. Like the other parameter estimates, the propensities to migrate change dramatically depending upon the set of data from which they are derived. Both the absolute values and the signs of the estimates change, as well as the relative ordering of the parameter estimates.

Thus, the parameter estimates are very sensitive and do not conform to any discernible pattern.

## IV.4 SUMMARY

The results of this analysis are very disappointing. The model has not adequately described the pattern of population growth of these urban systems. Nor do the parameter estimates provide any evidence that urban systems move steadily towards a state of equilibrium. From these results it appears that the urban system reacts to large and small deviations from the equilibrium state, rapidly and erratically. Indeed, it seems that the behaviour is oscillatory. A plot of the equilibrium population estimates of Toronto and London indicated that this parameter

---

[5] A complete list of the propensities to migrate is in Appendix III.

TABLE 17

MEAN PROPENSITIES TO MIGRATE

| ORIGIN | DESTINATION | | | | |
|---|---|---|---|---|---|
| | Toronto | Hamilton | London | Brantford | Kitchener |
| Toronto | X | -.4733<br>.0312 | .5069<br>-.0078 | - | - |
| Hamilton | .1905<br>-.0766 | X | -3.3147<br>.1138 | .0246 | .4588 |
| London | .0885<br>.0202 | 10.4137<br>.0011 | X | - | - |
| Brantford | - | .5569 | - | X | -.6919 |
| Kitchener | - | -.1920 | - | 2.2451 | X |

The underlined values are derived from a least squares analysis of population data.

TABLE 18

STANDRAD DEVIATIONS OF MIGRATION PROPENSITIES

| ORIGIN | DESTINATION<br>Toronto | <br>Hamilton | <br>London | <br>Brantford | <br>Kitchener |
|---|---|---|---|---|---|
| | | .054 | .048 | | |
| Toronto | X | 1.139 | .897 | - | - |
| | .2316 | | .262 | .975 | 1.140 |
| Hamilton | .185 | X | 4,624 | | |
| | .062 | .212 | | | |
| | .718 | 10.414 | X | - | - |
| Brantford | - | 1.311 | - | X | 2,800 |
| Kitchener | - | .535 | - | 5.293 | X |

The underlines estimates are calculated from a least squares analysis of population data.

fluctuated in a wavelike manner through time. Perhaps the data would be better fitted by a nonlinear estimation procedure.

The behaviour of parameters at different levels of the urban hierarchy have not been compared because the results are very confused. Any comparisons would be meaningless at this stage in the analysis. A better understanding of the basic process of population change is necessary before these comparisons can be made.

Several factors have contributed to these poor results. Only very small urban systems have been analysed. If the urban system of Ontario is well developed as Simmons (1972) suggests, the links among these urban regions cannot be adequately describing by analysing only parts of the system. There is evidence that Hamilton is not well integrated into the urban system, and that its behaviour is anomalous (Bannister, 1974). In this analysis, Hamilton has been included in every urban system. Perhaps the empirical results would improve if this centre was excluded.

Furthermore, Bannister (1974) has recently suggested that the principal mode of change in southern Ontario follows a nearest neighbour pattern, rather than a pattern of hierarchical diffusion. The urban systems analysed in this report have been chosen to reflect the levels of the urban hierarchy. If these urban systems were redefined to reflect the interaction between nearest neighbours, better results might be obtained. Although Brantford and Kitchener are the largest centres near Hamilton, Simmons (1972) analysis of telephone calls indicates that these three urban centres belong to different regional networks.

An additional problem is the poor data. The employment data are too detailed, and the frequent small fluctuations contribute to the poor fit of the regression equation. Alternatively, the population data are not available at frequent intervals. Bannister (1974) employed census material collected between 1891 and 1971. There are significant problems caused by different collection and compilation methods, as well as by the changing definitions of urban centres over this long time period. Despite these problems, these data might be better explained by the model.

In the final chapter, a general critique of the model will be presented, with recommendations for its improvement. Although this analysis has not been successful, it does provide information about the problems involved in modelling patterns of change within urban systems.

CHAPTER V

# RECOMMENDATIONS AND CONCLUSIONS

The results of this empirical analysis have not supported the original hypotheses concerning the dynamic behaviour of urban systems. This study was intended as a pilot project which would identify major problems associated with the analytical technique and as well provide additional information about the dynamic behaviour of the Ontario urban system. The analysis has demonstrated that the least squares estimation procedure does not adequately describe the patterns of population change of the Ontario urban system. Although the model has not been tested $x^y$ times as Batty (1972) suggested, where x is the number of independent variables in the model, and y is the number of parameters, the sensitivity of the parameters is apparent. Analysing another dynamic system, Peaker (1971) noted that the parameters of the system were very sensitive. Once a pattern of unbalanced growth had begun, the system would never return to balanced growth. The sensitivity of these parameters suggests a similar situation. The parameter estimates change significantly when the data from which the regression coefficients are estimated vary.

The definition of the urban system is crucial, but here the problems of defining urban systems have not been resolved. Once a system was defined, correctly, the amount of leakage from the system could be calculated. The trial and error method of defining urban systems must be improved before this model can be adequately tested.

Similarly, better data are needed describing the pattern of population change of urban centres. It has not been determined at what scale of analysis this model can be applied. Data must not provide too detailed information, similar to the employment data, nor should they cover such a long period of time that short term changes cannot be analysed. The model must be tested using different time series data to determine if the model best describes short run or long run changes in the urban system. It is apparent that the model does not adequately fit the employment data. Perhaps the employment data should be seasonally adjusted to remove some of the fluctuations. This could be accomplished by moving averages, or by the analysis of second or third differences rather than first differences. Alternative population data should also be tested. The problems of multicollinearity and autocorrelation in the time series employed in this analysis may have biased the coefficients and the resulting parameter estimates. This can only be determined by further tests with population data. Transformation of the time series data to remove autocorrelation might also improve the model's goodness of fit.

Similarly, the model must be tested at different levels of aggregation. First, the computer programme must be revised to include more urban regions in the analysis. The use of procedures for sparse matrices would allow the inclusion of more urban centres in the model. Alternatively, the model could be run on a computer having a larger memory. The inclusion of more urban regions in the model should increase the level of explanation of the regression analysis because more of the links by which growth impulses are transmitted will be considered.

The analysis might also be performed on systems which are chosen to test Bannister's (1974) hypothesis that the major mode of change follows a distance decay pattern rather than a pattern of hierarchical diffusion. This could be accomplished by analysing the patterns of change of small regional systems, for instance, London and the small towns which surround it such as Blenheim, Chatham, St. Thomas, and Stratford would comprise such a regional system.

At the same time the analysis should be performed at different levels of aggregation. Perhaps, the model is appropriate for the analysis of the patterns of change of larger population masses, such as provinces. It is necessary to establish the appropriate definitions of the variables; population size, and distance. The results of this analysis in which the estimates calculated using employment data are opposite in direction, and relative size, from the estimates derived from population data suggest that the employment data are not good surrogates for population data, as originally hypothesized. Similarly, the distance measure in this analysis is the highway distance between urban centres. A more realistic measure of distance might be time because the distances between urban regions are very short in this urban system. If none of these suggestions improves the performance of the model, perhaps the model should be redefined.

It has been suggested that another constraint should be introduced into the model. Consider a closed urban system and assume the total population change in the system over each interval of time and therefore, the rates of natural increase are known. Assign the increase in population to each centre in the urban system according to the following formula,

32. $\Delta P_{i,N} = \Delta P_{.} (P_{i,t}/\sum_{j} P_{j,t})$

where

33. $\Delta P_{.} = \Delta P_{.,t+1} - \Delta P_{.,t}$

34. $\Delta P_{.} = \sum_{i} \Delta P_{i}$

such that, $\Delta P_{i,N}$ is the change in population of urban region i caused by natural increase, $P_{.}$ is the total population change of all the centres in the urban system, and $P_{i,t}$ is the population size of centre i at time t. Then, a new dependent variable is defined, such that,

35. $P_{i}' = P_{i} - P_{i,N}$

This new measure of population measures only the population change caused by migration between the urban regions in the system. Perhaps this additional constraint will increase the level of explanation of the model.

Alternatively, there is a family of models which await investigation. Before some of these models can be empirically tested, a method of nonlinear estimation of the coefficients must be developed. Perhaps some of the equations could be linearized. It seems likely that those models in which the optimum population is defined as the demographic potential might more adequately describe the pattern of population change of an urban system. These models might also be simulated to determine the bounds of the parameter estimates. If these limiting values can be established, the solution of the equations defining the parameter estimates is facilitated.

A dynamic, deterministic model of population change has been tested. The results of this analysis underline the need for further investigation of the process whereby urban systems develop and evolve. Several questions have not been examined in this study. In one of the variations of the model, it is assumed that the system can reach equilibrium with the world outside, however, the conditions of this equilibrium are not clearly defined. Does the system reach equilibrium with the world outside when all the urban centres of the system approach the equilibrium state? Can the system be at equilibrium with the world outside if the centres of which the system is composed have not achieved equilibrium. Before these questions can be answered the model must be improved.

# APPENDIX I

## APPENDIX I

## ALTERNATIVE METHOD OF ESTIMATION

The normal equations are defined by calculating the derivative of the sum of square deviations with respect to each parameter value. Then the derivative is set equal to zero to insure that the minimum solution is obtained. For example, if the regression equation is,

$$Y_{1k} = C_{11}X_{1k}+C_{12}X_{2k}+C_{13}X_{3k}+C'_{12}X_{1k}X_{2k}+C'_{13}X_{1k}X_{3k}$$

the first normal equation would be,

$$\frac{\partial \sum\limits_{k=1} (Y_{1k}-Y_{1k})^2}{\partial C_{11}} = \sum_{k=1}^{n} X_{1k}(Y_{1k}-\hat{Y}_{1k})^2$$

$$\sum_{k=1}^{n} (X_{1k}(Y_{1k}-\hat{Y}_{1k}) = 0 .$$

The expanded form of this equation is,

$$\sum_{k=1}^{n} X_{1k}Y_{1k}-C_{11}\sum_{k=1} X_{1k}^2-C_{12}\sum_{k=1}^{n} X_{1k}X_{2k}-C_{13}\sum_{k=1}^{n} X_{1k}X_{3k}$$

$$-C'_{12}\sum_{k=1}^{n} X_{1k}^2X_{2k}-C'_{13}\sum_{k=1}^{n} X_{1k}^2X_{3k} = 0$$

which is

$$C_{11}\sum_{k=1}^{n} X_{1k}^{\;2}+C_{12}\sum_{k=1}^{n} X_{1k}X_{2k}+C_{13}\sum_{k=1}^{n} X_{1k}X_{3k}+C_{12}\sum_{k=1}^{n} X_{1k}X_{2k}$$

$$+C_{13}^{'}\sum_{k=1}^{n} X_{1k}^{\;2}X_{3k} \quad = \quad \sum_{k=1}^{n} X_{1k}Y_{1k} \quad .$$

A matrix in which these sums are the elements can be formed from these equations, in the following manner.

TABLE 19

MATRIX OF NORMAL EQUATIONS

| | $C_{11}$ | $C_{12}$ | $C_{13}$ | $C_{21}$ | $C_{22}$ | $C_{23}$ | $C_{31}$ | $C_{32}$ | $C_{33}$ | $C_{12}^{'}$ | $C_{13}^{'}$ | $C_{23}^{'}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{11}$ | * | * | * | | | | | | | * | * | |
| $C_{12}$ | * | * | * | | | | | | | * | * | |
| $C_{13}$ | * | * | * | | | | | | | * | * | |
| $C_{21}$ | | | | * | * | * | | | | * | | * |
| $C_{22}$ | | | | * | * | * | | | | * | | * |
| $C_{23}$ | | | | * | * | * | | | | * | | * |
| $C_{31}$ | | | | | | | * | * | * | | * | * |
| $C_{32}$ | | | | | | | * | * | * | | * | * |
| $C_{33}$ | | | | | | | * | * | * | | * | * |
| $C_{12}^{'}$ | * | * | * | * | * | * | | | | * | * | * |
| $C_{13}^{'}$ | * | * | * | | | | * | * | * | * | * | * |
| $C_{23}^{'}$ | | | | * | * | * | * | * | * | * | * | * |

This matrix equals a column vector. The solution of this matrix is the set of coefficient estimates $\underline{C}$ and $\underline{C}'$.

# APPENDIX II

## APPENDIX II

## COEFFICIENT ESTIMATES

### Employment Data for Hamilton London Toronto

| | 1958-60 | 1961-63 | 1964-66 | 1967-69 | 1970-72 | 1958-61 |
|---|---|---|---|---|---|---|
| $C_{11}$ | -.4321 | .51180 | -.43149 | -.56742 | -.38264 | -.31460 |
| $C_{12}$ | 1.50210 | 5.83244 | -.57237 | 34.04212 | 19.71079 | 3.94443 |
| $C_{13}$ | -.09667 | -.64121 | .12588 | -2.52002 | -1.51022 | -.31959 |
| $C'_{12}$ | -.01365 | -.08146 | .01682 | -.31692 | -.17810 | -.05225 |
| $C'_{13}$ | .00181 | .00755 | -.00139 | .02433 | .01423 | .00498 |
| $C_{21}$ | -.37270 | -2.28802 | .98249 | -15.67954 | -8.37136 | -1.66733 |
| $C_{22}$ | -2.56194 | -2.67855 | -.28211 | .16935 | -2.13229 | -2.67344 |
| $C_{23}$ | .15356 | .59034 | -.16093 | 2.53726 | 1.37651 | .41803 |
| $C'_{23}$ | .00112 | -.01344 | .00318 | -.05150 | -.02605 | -.00648 |
| $C_{31}$ | .79372 | 5.13930 | .31429 | 15.93639 | 8.40221 | 2.13240 |
| $C_{32}$ | 4.44780 | -2.27260 | .44297 | -33.27797 | -15.33570 | 1.48413 |
| $C_{33}$ | -.33579 | -.64259 | -.08458 | .03530 | .02516 | -.35564 |

| | 1962-65 | 1966-69 | 1958-62 | 1963-67 | 1968-72 |
|---|---|---|---|---|---|
| $C_{11}$ | -.67206 | -.64719 | -.29430 | -.21392 | -.49655 |
| $C_{12}$ | -6.90998 | 12.16088 | 7.60223 | -1.21107 | 17.16488 |
| $C_{13}$ | .69769 | -.87661 | -.62814 | .14657 | -1.29917 |
| $C'_{12}$ | .08137 | -.09991 | -.09924 | .01834 | -.15575 |
| $C'_{13}$ | -.00675 | .00815 | .00890 | -.00159 | .01212 |
| $C_{21}$ | 31.3178 | -4.85856 | -3.30401 | .98991 | -7.54724 |
| $C_{22}$ | -.14449 | -.00981 | -2.36287 | -.33125 | .22731 |
| $C_{23}$ | -.60967 | .76815 | .72417 | -.14344 | 1.14906 |
| $C'_{23}$ | .01610 | -.01577 | -.01610 | .00314 | -.02407 |
| $C_{31}$ | -1.93547 | 5.16872 | 3.82555 | -.44091 | 7.67154 |
| $C_{32}$ | 10.18411 | -10.62476 | -2.74385 | 2.05373 | -15.76250 |
| $C_{33}$ | -.45310 | .04894 | -.34686 | -.10251 | .04551 |

## COEFFICIENT ESTIMATES

Population Data for London Hamilton Toronto

| | 1951-53 | 1954-56 | 1957-59 | 1960-62 | 1963-65 | 1966-68 |
|---|---|---|---|---|---|---|
| $c_{11}$ | -.43178 | -1.00939 | .23843 | .20998 | -.95925 | -1.79090 |
| $c_{12}$ | .15768 | -.96058 | -.70091 | -.30817 | .54903 | -.20132 |
| $c_{13}$ | .04054 | .21094 | .11432 | .06781 | -.00353 | .16530 |
| $c'_{12}$ | .00211 | .00209 | -.00010 | -.00042 | -.00121 | .00155 |
| $c'_{13}$ | -.00073 | .00000 | -.00005 | -.00014 | .00028 | .00002 |
| $c_{21}$ | .86510 | .53039 | .00000 | .00000 | .00000 | .00000 |
| $c_{22}$ | -.30751 | -.79461 | -.25526 | -.35028 | -.41695 | -.38013 |
| $c_{23}$ | -.02257 | .13946 | .06222 | .10777 | .00029 | -.00006 |
| $c'_{23}$ | .00022 | .00000 | -.00007 | -.00022 | -.00029 | -.00006 |
| $c_{31}$ | -.70286 | -.01080 | -.05290 | -.57371 | .75438 | .00000 |
| $c_{32}$ | .71963 | .16962 | .36220 | .96228 | .03888 | .05763 |
| $c_{33}$ | -.09959 | -.02027 | -.07378 | -.17892 | -.10924 | -.01758 |

| | 1951-55 | 1956-60 | 1961-65 | 1966-70 | 1951-60 | 1961-70 |
|---|---|---|---|---|---|---|
| $c_{11}$ | -.29316 | -.14447 | -1.04115 | .00000 | .36068 | -.09546 |
| $c_{12}$ | .26707 | -.61176 | .34514 | -.39635 | -.55632 | .16668 |
| $c_{13}$ | -.00345 | .12271 | .03372 | .12985 | .09633 | -.01357 |
| $c'_{12}$ | -.00037 | -.00008 | .00018 | -.00246 | -.00134 | -.00106 |
| $c'_{13}$ | -.00015 | .00000 | .00007 | .00005 | .00000 | .00016 |
| $c_{21}$ | .26513 | -.14746 | .55214 | -1.16864 | -.36713 | -.47278 |
| $c_{22}$ | -.35972 | -.45290 | -.39795 | -.15047 | .41907 | -.09144 |
| $c_{23}$ | .04664 | .09085 | .05181 | .15260 | -.07008 | .08368 |
| $c'_{23}$ | -.00003 | .00000 | -.00013 | -.00028 | .00000 | -.00017 |
| $c_{31}$ | .01156 | -.03205 | .00000 | -.12512 | .00176 | .78328 |
| $c_{32}$ | .36307 | .07905 | .24046 | .18309 | .11165 | .16380 |
| $c_{33}$ | -.07938 | -.00052 | -.05647 | -.08430 | -.00951 | -.12320 |

## COEFFICIENT ESTIMATES

Employment Data for Brantford Hamilton Kitchener

| | 1958-60 | 1961-63 | 1964-66 | 1967-69 | 1970-72 | 1958-61 |
|---|---|---|---|---|---|---|
| $C_{11}$ | .16452 | -.41107 | -2.28542 | -.91848 | -.92050 | .35363 |
| $C_{12}$ | -.01141 | -.05649 | -1.77005 | -.90013 | .19368 | .06411 |
| $C_{13}$ | .03720 | .24058 | 3.99637 | 1.59128 | -.23450 | -.06906 |
| $C'_{12}$ | -.00130 | .01073 | .11175 | .05101 | -.00348 | -.00686 |
| $C'_{13}$ | -.00262 | -.02122 | -.20380 | -.07457 | .01541 | .00019 |
| $C_{21}$ | .82522 | .06120 | .08835 | -.10601 | -.12596 | .00987 |
| $C_{22}$ | -.24189 | .06120 | .08835 | -.10601 | -.12596 | -.18301 |
| $C_{23}$ | .09130 | -.17396 | -3.95750 | -.85006 | .37588 | .31982 |
| $C'_{23}$ | .00109 | .00187 | .03668 | .01102 | -.00255 | -.00166 |
| $C_{31}$ | .01432 | -.72819 | -9.01424 | -4.01639 | .89496 | .07304 |
| $C_{32}$ | .07648 | .29613 | 2.17001 | .83126 | -.14847 | .01858 |
| $C_{33}$ | -.11762 | -.50036 | -1.11823 | -.44261 | .00386 | -.19562 |

| | 1962-65 | 1966-69 | 1958-62 | 1963-67 | 1968-72 |
|---|---|---|---|---|---|
| $C_{11}$ | -.71098 | -1.04515 | -.79014 | -1.41813 | -.03898 |
| $C_{12}$ | 1.04730 | .95011 | .09081 | -1.16424 | -.00454 |
| $C_{13}$ | -2.00604 | 1.61081 | -.14799 | 2.62211 | -.02056 |
| $C'_{12}$ | -.06387 | .05802 | .00179 | .07269 | .00145 |
| $C'_{13}$ | .13734 | -.08122 | .00927 | -.13339 | -.00035 |
| $C_{21}$ | -5.89501 | 6.32365 | .00196 | 7.50471 | .00087 |
| $C_{22}$ | .05413 | -.02872 | -.14297 | .0565] | -.16242 |
| $C_{23}$ | 1.91650 | -1.71074 | .60610 | -2.62911 | .55947 |
| $C'_{23}$ | -.02189 | .01605 | -.00287 | .02446 | -.00238 |
| $C_{31}$ | 6.68385 | -3.24016 | .87160 | -4.66587 | .35691 |
| $C_{32}$ | -.51755 | 1.04155 | -.11024 | 1.36590 | -.0332 |
| $C_{33}$ | -1.13031 | -.69453 | -.23094 | .99284 | -.30831 |

APPENDIX III

## APPENDIX III

## PROPENSITY TO MIGRATE

Employment Data for *Hamilton London Toronto

| | 1958-60 | 1961-63 | 1964-66 | 1967-69 | 1970-72 | 1958-61 |
|---|---|---|---|---|---|---|
| $\gamma_{12}$ | -.00247 | -.7450 | .5787 | -12.2490 | -9.7853 | -2.8095 |
| $\gamma_{13}$ | .08091 | .1956 | .0285 | .5247 | .2319 | .0154 |
| $\gamma_{21}$ | .2272 | 63.9186 | -.4197 | 36.969 | 4.1092 | 1.2660 |
| $\gamma_{23}$ | .31022 | -.0008 | .1089 | -.0032 | -.6188 | .0291 |
| $\gamma_{31}$ | .0272 | -3.7838 | -.0497 | -1.0349 | -.1695 | .0552 |
| $\gamma_{32}$ | .0004 | .2809 | -.1385 | .1934 | 2.3509 | 1.0295 |

| | 1962-65 | 1966-69 | 1958-62 | 1963-67 | 1968-72 |
|---|---|---|---|---|---|
| $\gamma_{12}$ | 2.5588 | -4.5125 | -4.3442 | .6684 | -5.8196 |
| $\gamma_{13}$ | -.0776 | .3992 | .2227 | -.7621 | 6.3001 |
| $\gamma_{21}$ | -3.7881 | 3.2805 | 3.4500 | -.7621 | 6.3001 |
| $\gamma_{23}$ | 1.1077 | -1.1994 | -.4336 | .2164 | -1.4490 |
| $\gamma_{31}$ | .2059 | -.1273 | -.1535 | .0497 | -.2481 |
| $\gamma_{32}$ | -.0776 | 1.0427 | 1.3916 | -.1416 | 1.2950 |

*The numbers identify each urban region. Number 1 is Hamilton, 2 is London, and number 3 is Toronto.

## PROPENSITY TO MIGRATE

Population Data for *London Hamilton Toronto

| | 1951-53 | 1954-56 | 1957-59 | 1960-62 | 1963-65 | 1966-68 |
|---|---|---|---|---|---|---|
| $\gamma_{12}$ | .0035 | .1765 | .0000 | .0000 | .0000 | .0000 |
| $\gamma_{13}$ | -.0112 | .1330 | -.0076 | -.1499 | .0310 | .0000 |
| $\gamma_{21}$ | -.0194 | .0135 | .0062 | .4164 | .0944 | - ++ |
| $\gamma_{23}$ | -.0005 | -.7694 | .0191 | -.0926 | .0006 | - ++ |
| $\gamma_{31}$ | .0715 | .0043 | -.0015 | -.1339 | -.0009 | - ++ |
| $\gamma_{32}$ | -.0049 | .0250 | .0220 | .1132 | .0128 | - ++ |

| | 1951-55 | 1956-60 | 1961-65 | 1966-70 | 1951-60 | 1961-70 |
|---|---|---|---|---|---|---|
| $\gamma_{12}$ | .6116 | -.0025 | -.0418 | -.1772 | -.1010 | -.2271 |
| $\gamma_{13}$ | .0004 | -.0010 | .0000 | -.0013 | .0004 | .0193 |
| $\gamma_{21}$ | .8022 | -.0099 | -.0559 | .0147 | -.0018 | -.0092 |
| $\gamma_{23}$ | .0046 | .0008 | -.0076 | .0007 | .0002 | .0015 |
| $\gamma_{31}$ | -.0151 | .0029 | -.0080 | -.0071 | .0004 | .0011 |
| $\gamma_{32}$ | .1594 | .0008 | -.0021 | .0125 | -.0104 | .0147 |

*The numbers identify each urban region. Number 1 is London, number 2 is Hamilton, and number 3 is Toronto.

++There are no values for these parameters because of zero coefficients. For further details see Table 11.

PROPENSITY TO MIGRATE

Employment Data for *Brantford Hamilton Kitchener

| | 1958-60 | 1961-63 | 1964-66 | 1967-69 | 1970-72 | 1958-61 |
|---|---|---|---|---|---|---|
| $\gamma_{12}$ | .0560 | .0351 | 1.4184 | .6771 | .0010 | .0001 |
| $\gamma_{13}$ | .0021 | -.0093 | .4673 | -.7912 | .3738 | .1870 |
| $\gamma_{21}$ | -.0339 | -.0072 | -1.4871 | -.6492 | .0915 | .1785 |
| $\gamma_{23}$ | .2268 | .0192 | .6471 | .2267 | -.0859 | .0659 |
| $\gamma_{31}$ | .1106 | .0355 | 3.3576 | 1.1476 | -.1108 | -.1922 |
| $\gamma_{32}$ | .0062 | -.0614 | -.6734 | -.1604 | .1776 | .0061 |

| | 1962-65 | 1966-69 | 1958-62 | 1963-67 | 1968-72 |
|---|---|---|---|---|---|
| $\gamma_{12}$ | -1.1380 | 3.9483 | .0001 | 1.1148 | .0127 |
| $\gamma_{13}$ | 2.5794 | -8.6209 | .3167 | -1.1451 | .1793 |
| $\gamma_{21}$ | .5222 | 2.4400 | .0464 | -.7751 | -.0376 |
| $\gamma_{23}$ | -.2756 | 3.8370 | -.0555 | .4641 | -.0232 |
| $\gamma_{31}$ | -1.0002 | 4.1363 | -.0757 | 17.4578 | -.1702 |
| $\gamma_{32}$ | .5123 | -1.4790 | -.0084 | -.5408 | .1089 |

*The numbers identify each urban region. Number 1 is Brantford, number 2 is Hamilton, and number 3 is Kitchener.

APPENDIX IV

# APPENDIX IV

## INTERURBAN DISTANCES (MILES)

| | Hamilton | London | Toronto | Brantford | Kitchener |
|---|---|---|---|---|---|
| Hamilton | X | | | | |
| London | 78 | X | | | |
| Toronto | 42 | 114 | X | | |
| Brantford | 26 | 57 | 65 | X | |
| Kitchener | 36 | 65 | 69 | 26 | X |

All distances were obtained from the Ontario 1971 Official Road Map.

```
C    REGRESSION ANALYSIS DEVELOPED BY V.ZACHOWSKI
      PROGRAM TST (INPUT,OUTPUT,PUNCH,TAPE5=INPUT,TAPE6=OUTPUT)
C
C   Y IS AN N BY 1 VECTOR, X IS AN N BY M MATRIX, B IS AN M BY 1 VECTOR
C
C   N IS THE NUMBER OF ROWS IN Y AND X
C   M IS THE NUMBER OF COLUMNS IN X AND ROWS IN BETA
C   J IS AN INDICATOR.  IF J =1 THEN THE PROGRAMME WILL PRINT XTX,G,AND
C    BETA-HAT.  IF J=2 THE PROGRAMME WILL OMIT PRINTING THEM.
C   JJ=1 SIGNIFIES THAT THE MEAN IS IN THE MODEL.  JJ=2 SIGNIFIES THAT
C   THE MEAN IS NOT IN THE MODEL
C   IF JJJJ=1 THEN AN HYPOTHESIS IS TO BE TESTED IF IS IS ESTIMABLE.
C   IF JJJJ=2 THEN NO HYPOTHESIS IS TO BE TESTED.
C   NJ IN THE SUBROUTINES SIGNIFIES WHICH COLUMN OF X1 IS BEING USED.
C
C   PROBLEM-TO SOLVE Y=XB FOR B
C
      DIMENSION BETA(12),XT(2200),G(144),IROW(2200),ICOL(12),IWORK(12),
     2JWORK(12),XG(144),XGXTX(144),XAGA(144),XBETA(200),XTX(144),B(2200)
     3,Y(200),X(2200),YX(2500),UN(200),X1(2200),SEB(12),BETAT(12),TSS(1)
     4,SSM(1),E(200),Z(200),VB(144),BV(12),RB(144),YXS(13),YXMEAN(13),
     6A(2500),VACO(170),V(13),SD(13),SE(13),R(170),XTY(12)
      READ(5,801) N,M,J,JJ
  801 FORMAT(4I5)
C  READ IN Y AND X ROWWISE
      DO 2 I=1,N
      K=I
      KK=(M-1)*N+I
    2 READ(5,802) Y(K),(X(IX),IX=K,KK,N)
  802 FORMAT(1X,F9.1,6F10.1,/,6F10.1)
C
C  TRANSPOSE OF X,XT, IS FOUND.  XT IS AN M BY N MATRIX.
C
      CALL MTRA(X,XT,N,M,0)
C
C  THE PRODUCTS XTX (XT AND X) AND XTY (XT AND Y) ARE FOUND.
C
      CALL MPRD(XT,X,XTX,M,N,0,0,M)
      CALL MPRD(XT,Y,XTY,M,N,0,0,1)
C
C  THE G-INVERSE OF XTX IS FOUND.
      CALL GINV(XTX,B,G,M,M,IRANK,IROW,ICOL,IWORK,JWORK)
C  A CHECK FOR ERROR IS MADE.
      CALL MPRD(XTX,G,XG,M,M,0,0,M)
      CALL MPRD(XG,XTX,XGXTX,M,M,0,0,M)
      CALL MSUB(XGXTX,XTX,XAGA,M,M,0,0)
C
C  BETA-HAT, THE SOLUTION TO Y=XB IS FOUND.
C  BETA-HAT IS THE PRODUCT OF G AND XTY.
C
      CALL MPRD(G,XTY,BETA,M,M,0,0,1)
      CALL MPRD(X,BETA,XBETA,N,M,0,0,1)
      PRINT 808
  808 FORMAT(1H1,*      Y VECTOR      *,10X,* Y ESTIMATES */)
      DO 10 I=1,N
      K=I
   10 WRITE(6,905) Y(K),XBETA(K)
      PRINT 809
  809 FORMAT (1H-,*      X MATRIX      */)
```

```
      DO 20 I=1,N
      K=I
      KK=(M-1)*N+I
20    WRITE (6,905) (X(IX),IX=K,KK,N)
904   FORMAT(5F20.5)
```

THE ABILITY TO PRINT OUT XTX,G, AND BETA IS PROVIDED BY THE FOLLOWING SECTION OF THE PROGRAMME.

```
      GO TO(11,12)J
 11   PRINT 803
803   FORMAT(1H1,*        XTX         */)
      DO 3 I=1,M
      K=I
      KK=(M-1)*M+I
  3   WRITE(6,905) (XTX(IX),IX=K,KK,M)
905   FORMAT(5F20.5)
      PRINT 804
804   FORMAT(1H-,*   GENERALIZED INVERSE OF XTX   */)
      DO 4 I=1,M
      K=I
      KK=(M-1)*M+I
  4   WRITE(6,904) (G(IG),IG=K,KK,M)
      PRINT 805
805   FORMAT(1H-,*  BETA-HAT,SOLUTION OF Y=XB   */)
      DO 9  I=1,M
      K=I
      PUNCH 910,BETA(K),K
910   FORMAT(F20.5,55X,I5)
  9   WRITE(6,904) BETA(K)
      GO TO 12
 12   CONTINUE
      GO TO (55,65)JJ
 65   XN=N
      NM=N*M
      MM=M
      M=M+1
      M2=M*M
      DO 75 I=1,NM
 75   X1(I)=X(I)
      GO TO 85
 55   CALL CCUT(X,2,UN,X1,N,M,0)

      XN=N
      MM=M-1
      M2=M*M
```

Y AND X1 ARE ADJOINED TO FORM YX AN N BY M MATRIX

```
 85   CALL CTIE(Y,X1,YX,N,1,0,0,MM)
      PRINT 806
806   FORMAT(1H1,*AUGMENTED MATRIX COMPOSED OF Y AND X1*/)
      PRINT 814
814   FORMAT(1H-,* Y  VECTOR         X1  MATRIX*/)
      DO 13 I=1,N
      K=I
      KK=MM*N+I
 13   WRITE(6,815) (YX(IX),IX=K,KK,N)
815   FORMAT(F20.5,5X,5F20.5)
```

THE FOLLOWING PART OF THE PROGRAMME COMPUTES STANDARD STATISTICS FOR COLUMNS OF AN AUGMENTED MATRIX COMPOSED OF Y AND X1,PART OF THE MATRIX X IN COMPUTING THE CORRELATION MATRIX CARE MUST BE TAKEN THAT THE THE DENOMINATOR IN THE CORRELATION EXPRESSION NOT BE ZERO. THE DENOMINATOR WILL ONLY BE ZERO WHEN ONE OF THE VARIANCES INVOLVED IS ZERO WHICH IMPLIES A COLUMN OF CONSTANT TERMS IN THE Y OR X MATRIX. IF Y WERE TO BE COMPOSED OF CONSTANTS, SAY Z, THEN THE MODEL WOULD REDUCE TO Y=ZV, ZV A COLUMN VECTOR COMPOSED OF Z. THE ONLY CONSTANT TERM OR COLUMN IN THE X MATRIX OR DESIGN SPACE IS THE FIRST COLUMN OF ONES INDICATING THE MEAN IS IN THE MODEL. THE PROGRAMME BELOW OMITS THIS COLUMN WHEN CALCULATING THE CORRELATIONS.
IN AN X MATRIX OF ACTUAL OBSERVATIONS IF THE I-TH COLUMN WERE CONSTANT, SAY EACH TERM IN THE COLUMN WAS EQUAL TO ZZ, THEN THE TERM B(I)X(I) IN THE MODEL COULD BE AMALGAMATED WITH THE MEAN TERM IN THE MODEL AND THUS NO COLUMN OF THE INPUT MATRIX WOULD HAVE A ZERO VARIANCE.

UN IS A VECTOR COMPOSED OF ONES. THE VARIANCE,STANDARD DEVIATION AND STANDARD ERROR OF THIS COLUMN OF X ARE ALL ZERO.

```
    CALL STATS(YX,YXMEAN,YXS,A,VACO,V,SD,SE,R,N,XN,M,MM,M2)
    MML=MM-1
    GO TO (555,655)JJ
655 M=MM
555 CALL ANOVA(Y,X1,XTY,BETA,BETAT,TSS,SSM,SSE,SSPE,SSLF,PP,N,M)
    XK=M
    GO TO (35,45)JJ
 35 XMSS=(YXS(1)**2)/XN
    TSSC=TSS(1)-XMSS
    SSMC=SSM(1)-XMSS
    XXMSS=XMSS
    IF(XK.EQ.1.0)GO TO 111
    XMSMC=SSMC/(XK-1.0)
    GO TO 45
111 XMSMC=SSMC
 45 XMSM=SSM(1)/XK
    XMSE=SSE/(XN-XK)
    IF(PP-XK.LT.1.0)GO TO 112
    XMSLF=SSLF/(PP-XK)
112 IF(XN-PP.LT.1.0)GO TO 113
    XMSPE=SSPE/(XN-PP)
113 XMSPE=0.0
    IF(SSPE.EQ.0.0.OR.XN-PP.LT.1.0.OR.PP-XK.LT.1.0)GO TO 115
    FLF=XMSLF/XMSPE
    GO TO 125
115 PRINT 826
826 FORMAT(1H1,*ADEQUACY OF MODEL CANNOT BE CHECKED*/)
125 IF(JJ.GT.1)GO TO 116
    FM1=XXMSS/XMSE
    FMC=XMSMC/XMSE
116 FM2=XMSM/XMSE
    CALL SMPY(G,XMSE,VB,M,M,0)
    CALL DCPY(VB,BV,M,0)
    PRINT 645
645 FORMAT(1H1,*STANDARD ERRORS OF REGRESSION COEFFICENTS*/)
    DO 77 I=1,M
    SEB(I)=SQRT(BV(I))/SQRT(XN)
 77 WRITE(6,904)SEB(I)
    M3=M*M
    DO 67 I=1,M
    K=0
    DO 67 KK=I,M3,M
```

```
      PRINT 69
   69 FORMAT(1H-,*THE CORRELATION MATRIX OF THE REGRESSION COEFFICIENTS*
     1/)
      LL=M-1
      DO 68 I=1,M
      KK=M*LL+I
   68 WRITE(6,891) (RB(IB),IB=I,KK,M)
  891 FORMAT(5F20.5)
      MM1=1
      MM2=M-1
      MM3=N-M
      MM4=PP-XK
      MM5=XN-PP
      MM6=N-1
      WRITE(6,700)
  700 FORMAT(1H1,//15X,* ANALYSIS OF VARIANCE TABLE*//////)
      WRITE(6,710)
  710 FORMAT(4X,* SOURCE      *,4X,* D.F.*,16X,* S.S.*,21X,* M.S.*,20X,*F-
     1STATISTICS*////)
      WRITE(6,720) M,SSM(1),XMSM,FM2
  720 FORMAT(1X,* MODEL*,9X,I7,12X,F14.5,12X,F14.5,15X,F14.5,//)
      IF(JJ.GT.1)GO TO 118
      WRITE(6,730) MM1,XMSS,XXMSS,FM1
  730 FORMAT(4X,* MEAN*,8X,I7,13X,F14.5,13X,F14.5,15X,F13.5,//)
      WRITE(6,740) MM2,SSMC,XMSMC,FMC
  740 FORMAT(4X,* MODEL(CFM)  *,I7,13X,F14.5,13X,F14.5,15X,F13.5,//)
  118 WRITE(6,750) MM3,SSE,XMSE
  750 FORMAT(1X,* RESIDUAL*,6X,I7,12X,F14.5,12X,F14.5,//)
C
C     JJJ=1 SIGNIFIES THAT THE F RATIO XMSLF/XMSPE IS NOT AVAILABLE.
C     JJJ=2 SIGNIFIES THAT THE F RATIO IS AVAILABLE.
C
      JJJ=1
      IF(PP-XK.LT.1.0)GO TO 212
      IF(XN-PP.LT.1.0)GO TO 214
      IF(XMSPE.EQ.0.0)GO TO 216
      WRITE(6,701) MM4,SSLF,XMSLF,FLF
  701 FORMAT(4X,* LACK OF FIT *,I7,13X,F14.5,13X,F14.5,15X,F13.5,//)
      WRITE(6,711)  MM5,SSPE,XMSPE
  711 FORMAT(4X,* PURE ERROR  *,I7,13X,F14.5,13X,F14.5,//)
      GO TO 215
  212 WRITE(6,721) MM4,SSLF
  721 FORMAT(4X,* LACK OF FIT *,I7,13X,F14.5,//)
      IF(XN-PP.GE.1.0)GO TO 213
      WRITE(6,731) MM5,SSPE
  731 FORMAT(4X,* PURE ERROR  *,I7,13X,F14.5,//)
      GO TO 217
  213 WRITE(6,711) MM5,SSPE,XMSPE
      GO TO 217
  214 WRITE(6,741) MM4,SSLF,XMSLF
  741 FORMAT(4X,* LACK OF FIT *,I7,13X,F14.5,13X,F14.5,//)
      WRITE(6,731) MM5,SSPE
      GO TO 217
  216 WRITE(6,741)  MM4,SSLF,XMSLF
      WRITE(6,711) MM5,SSPE,XMSPE
      GO TO 217
  215 KC=PP-XK
      KV=XN-PP
      CALL MDFD(FLF,KC,KV,P,IER)
      P=1.0-P
      JJJ=JJJ+1
```

```
217 PRINT 551
551 FORMAT(1H-,*--------------------------------------------------------------
   1----------------------------------------------------------------------*)
    WRITE(6,760) N,TSS(1)
760 FORMAT(1X,* TOTAL*,9X,I7,12X,F14.5,//)
    IF(JJ.GT.1)GO TO 119
    WRITE(6,770) MM1,XMSS
770 FORMAT(4X,* MEAN*,8X,I7,13X,F14.5,//)
    WRITE(6,780) MM6,TSSC
780 FORMAT(4X,* TOTAL(CFM)  *,I7,13X,F14.5,//)
119 GO TO (218,219)JJJ
219 PRINT 751
751 FORMAT(1H-,*TEST FOR ADEQUACY OF MODEL (IF APPLICABLE)*//)
    WRITE(6,790) FLF,P
790 FORMAT(1H-,*PROBABILITY OF OBTAINING AN F-VALUE EXCEEDING *,F10.5,
   1*      IS  *,F10.5,//)
    PRINT 84
 84 FORMAT(1H-,*OTHER F-PROBABILITIES OF POSSIBLE INTEREST*)
    CALL MDFD(FM2,M,MM3,P2,IER)
    P2=1.0-P2
    WRITE(6,790) FM2,P2
    IF(JJ.NE.1) GO TO 218
    IF(XK.EQ.1.0) GO TO 218
    CALL MDFD(FMC,MM2,MM3,P3,IER)
    P3=1.0-P3
    WRITE(6,790) FMC,P3
218 CONTINUE
    RS=SSM(1)/TSS(1)
    QS=SSM(1)/(TSS(1)-SSPE)
    WRITE(6,500) RS,QS
500 FORMAT(1H-,*COEFF. OF DETERMINATION *,F10.5,*  MODIFIED COEFF. OF
   1DETERMINATION*,F10.5)
    GO TO (22,23)JJ
 22 RSM=SSMC/TSSC
    QSM=SSMC/(TSSC-SSPE)
    WRITE(6,510) RSM,QSM
510 FORMAT(1H-,*COEFF. OF DET. (CFM)    *,F10.5,*  MODIFIED COEFF. OF
   1DET. (CFM)   *,F10.5)
 23 CONTINUE
    IF(XN-PP.LT.1.0.OR.XMSPE.EQ.0.0) GO TO 123
    CALL SRED(Y,XBETA,XMSPE,E,Z,N,SUM,ZZK)
    GO TO 124
123 CALL SRED(Y,XBETA,XMSE,E,Z,N,SUM,ZZK)
124 CONTINUE
    CALL SPLOT(XBETA,E,N,1)
    WRITE(6,220)
220 FORMAT(1H-,*GRAPH OF RESIDUALS VERSUS Y ESTIMATES*/)
    STOP
    END
    SUBROUTINE ANOVA(Y,X1,XTY,BETA,BETAT,TSS,SSM,SSE,SSPE,SSLF,PP,N,M)
    DIMENSION Y(1),X1(1),XTY(1),BETA(1),BETAT(1),TSS(1),SSM(1)
    NN=N-1
    CALL MATA(Y,TSS,N,1,0)
    CALL MTRA(BETA,BETAT,M,1,0)
    CALL MPRD(BETAT,XTY,SSM,1,M,0,0,1)
    SSE=TSS(1)-SSM(1)
    PP=N
    SSPE=0.0
```

```
      SSLF=SSE-SSPE
      RETURN
      END
      SUBROUTINE GINV (A,B,G,M,N,IRANK,IROW,ICOL,IWORK,JWORK)
      DIMENSION A(1),B(1),G(1),IROW(1),ICOL(1),IWORK(1),JWORK(1)
C
C     CHECK DIMENSIONS
      IF(M.LT.1.OR.N.LT.1) GO TO 90
      EPS=1.E-14
C     COPY MATRIX A TO B
      NM=N*M
      DO 10 I=1,NM
   10 B(I)=A(I)
C     FIND RANK , BASIC ROWS, BASIC COLS OF B
      CALL MFGR(B,M,N,EPS,IRANK,IROW,ICOL)
C     SAVE NONSINGULAR MINOR IN B
      DO 20 I=1,IRANK
      DO 20 J=1,IRANK
      IA=IROW(I)+(ICOL(J)-1)*M
      IB=I+(J-1)*IRANK
   20 B(IB)=A(IA)
C     FIND INVERSE OF B
      CALL MINV(B,IRANK,D,IWORK,JWORK)
C     CLEAR G MATRIX
      DO 30 I=1,NM
   30 G(I)=0.
C     INSERT APPROPRIATE ELEMENTS OF B INTO G FOR GENERALISED INVERSE
      DO 40 I=1,IRANK
      DO 40 J=1,IRANK
      IB=J+(I-1)*IRANK
      IG=ICOL(J)+(IROW(I)-1)*N
   40 G(IG)=B(IB)
      GO TO 100
   90 IRANK=-1
  100 RETURN
      END
      SUBROUTINE SPLOT(X1,Y,N,NJ)
      DIMENSION X1(1),Y(1)
      K=(NJ-1)*N+1
      KK=NJ*N
      DO 1 I=K,KK
      CALL PLOTPT(X1(I),Y(I),4)
    1 CONTINUE
      CALL OUTPLT
      RETURN
      END
      SUBROUTINE SRED(Y,XBETA,XME,E,Z,N,SUM,ZZK)
      DIMENSION Y(1),XBETA(1),E(1),Z(1)
      CALL MSUB(Y,XBETA,E,N,1,0,0)
      SX=SQRT(XME)
      CALL SDIV(E,SX,Z,N,1,0)
      PRINT 807
  807 FORMAT(1H1,* THE VECTOR OF RESIDUALS *,20X,* THE VECTOR OF STANDAR
     1DIZED RESIDUALS*/)
      DO 72 I=1,N
      PUNCH 580,E(I),Z(I),I
  580 FORMAT(F15.5,10X,F15.5,35X,I5)
   72 WRITE(6,520) E(I),Z(I)
```

```
520 FORMAT(F15.5,45X,F15.5)
    SUM=0.0
    DO 74 I=1,N
    SUM=SUM+F(I)
 74 CONTINUE
    WRITE(6,560) SUM
560 FORMAT(1H-,* THE SUM OF THE RESIDUALS IS *,F20.5)
    ZKOUNT=0.0
    DO 73 I=1,N
    IF(ABS(Z(I)).GT.2.0) GO TO 530
    GO TO 550
530 WRITE(6,540) I
540 FORMAT(1H0,* THE ABS. VALUE OF Z(*,I3,*) IS GREATER THAN 2*)
    ZKOUNT=ZKOUNT+1.0
550 CONTINUE
 73 CONTINUE
    ZZK=ZKOUNT*100.0/FLOAT(N)
    WRITE(6,570) ZZK
570 FORMAT(1H-,* THE PERCENTAGE OF STANDARDIZED RESIDUALS WITH ABSOLUT
   1E VALUE GREATER THAN 2 IS  *,F20.5)
    RETURN
    END
    SUBROUTINE STATS(YX,YXMEAN,YXS,A,VACO,V,SD,SE,R,N,XN,M,MM,M2)
    DIMENSION YX(1),YXMEAN(1),YXS(1),A(1),VACO(1),V(1),SD(1),SE(1),
   1R(1)
    CALL CSUM(YX,YXS,N,M,0)
    CALL SDIV(YXS,XN,YXMEAN,1,M,0)
    PRINT 816
816 FORMAT(1H-,*MEAN OF EACH COLUMN OF AUGMENTED MATRIX*/)
    WRITE(6,825) (YXMEAN(IS),IS=1,M)
825 FORMAT(1H0,F20.5,5X,5F20.5)
    DO 200 K=1,M
    KK=(K-1)*N+1
    KN=K*N
    DO 200 I=KK,KN
    A(I)=(YX(I)-YXMEAN(K))/SQRT(XN-1.0)
200 CONTINUE
    CALL TPRD(A,A,VACO,N,M,0,0,M)
    PRINT 810
810 FORMAT(1H-,*VARIANCE-COVARIANCE MATRIX*/)
    DO 70 I=1,M
    K=I
    KK=M*MM+I
 70 WRITE(6,811) (VACO(IV),IV=K,KK,M)
811 FORMAT(1H0,5F25.10)
    CALL DCPY(VACO,V,M,0)
    PRINT 813
813 FORMAT(1H-,*VARIANCE OF EACH COLUMN OF AUGMENTED MATRIX*/)
    WRITE(6,825) (V(IV),IV=1,M)
    DO 40 I=1,M
    SD(I)=SQRT(V(I))
    SE(I)=SD(I)/SQRT(XN)
 40 CONTINUE
    PRINT 818
818 FORMAT(1H-,*STANDARD DEVIATION OF COLUMNS OF AUGMENTED MATRIX*)
    WRITE(6,825) (SD(IS),IS=1,M)
    PRINT 819
819 FORMAT(1H-,*STANDARD ERROR OF EACH COLUMN OF AUGMENTED MATRIX*)
    WRITE(6,825) (SE(IS),IS=1,M)
    DO 60 I=1,M
    K=0
```

```
60  R(KK)=VACO(KK)/SQRT(V(I)*V(K))
    PRINT 812
812 FORMAT(1H1,*CORRELATION MATRIX*/)
    DO 80 I=1,M
    K=I
    KK=M*MM+I
 80 WRITE(6,811) (R(IR),IR=K,KK,M)
    RETURN
    END
      6400 END OF RECORD
```

CDTOT 0430

## BIBLIOGRAPHY


Bannister, G., 1974, "Modes of Change in the Ontario Economy", University of Toronto, mimeo.

Batty, M., 1972a, "Dynamic Simulation of an Urban System", in Wilson, A. G., ed., Patterns and Processes in Urban and Regional Systems, London: Pion Ltd.

Batty, M., 1972b, "Recent Developments in Land-Use Modelling: A Review of British Research", Urban Studies, pp. 151-177.

Baumol, W. J., 1959, Economic Dynamics. An Introduction, New York: The MacMillan Company.

Berry, B. L. and F. E. Horton, 1970, Geographic Perspectives on Urban Systems. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Blalock, H. M., 1969, Theory Construction. Englewood Cliffs: Prentice Hall, Inc.

Bourne, L. S. and R. MacKinnon, eds., 1972, Urban Systems Development in Central Canada, Toronto: Univ. of Toronto Press.

Casetti, E., 1970, "Equilibrium Values and Population Densities in an Urban Setting", Discussion Paper No. 11, Dept. of Geography, Ohio State University, Columbus, Ohio; Ohio State University Press.

Casetti, E., L. J. King and D. Jeffrey, 1971, "Structural Imbalance in the U.S. Urban Economic System, 1960-1965", Geographical Analysis, Vol. 3, No. 3, (July), pp. 239-255.

Christ, C. F., 1966, Econometric Models and Methods. New York: John Wiley and Sons, Inc.

Cordey-Hayes, M. and A. G. Wilson, 1970, "Spatial Interaction", Centre for Environmental Studies, Working Paper No. 57, CESWP57, London.

Cordey-Hayes, M. and D. Greave, 1973, "Migration Movements and the Differential Growth of City Regions in England and Wales", Centre for Environmental Studies, Research Paper No. 1, CESRP1.

Curry, L., 1972, "A Spatial Analysis of Gravity Flows", Regional Studies, Vol. 6, No. 2, (June), pp. 131-147.

Czamanski, S., 1964, "A Model of Urban Growth", Papers and Proceedings of the Regional Science Association, Vol. 13, pp. 177-200.

Czamanski, S., 1972, Regional Science Techniques in Practice; The Case of Nova Scotia, Lexington, Mass: Lexington Books, D.C. Heath and Company.

Darwent, D. F., 1969, "Growth Poles and Growth Centres in Regional Planning - A Review", Environment and Planning, Vol. 1, pp. 5-32.

Denton, F. T., 1971, "A Model for the Variational Transition Probabilities for Simulating Short-Run Labour Force Behaviour in Canada", Working Paper No. 71-16, Dept. of Economics, McMaster University.

Garvin, W. W., 1960, Introduction to Linear Programming. New York: McGraw-Hill Book Co., Inc.

Golant, S. M., 1972, "Regression Models of Urban Growth in Ontario and Quebec", in Bourne, L. S. and R. MacKinnon, eds., Urban Systems Development in Central Canada, Toronto: University of Toronto Press, pp. 117-131.

Goldner, W., 1971, "The Lowry Model Heritage", Journal of the American Institute of Planners, Vol. 37, No. 2, (March), pp. 100-110.

Hagerstrand, T., 1965, "Aspects of the Spatial Structure of Social Communication and the Diffusion of Innovations", Papers and Proceedings of the Regional Science Association, Vol. 16, pp. 27-42.

Hagerstrand, T., 1967, Innovation Diffusion as a Spatial Process. Chicago, Illinois: University of Chicago Press.

Hansen, N. M., ed., 1972, Growth Centres in Regional Economic Development. New York: The Free Press.

Hatt, P. K. and E. J. Reiss, eds., 1957, Cities and Society. New York: The Free Press.

Hill, F., 1970, "Migration in the Toronto-Centered (MTARTS) Region", Centre for Urban and Community Studies, Research Paper No. 18, Univ. of Toronto, Toronto.

Hodge, G., 1972, "Regional and Structural Components of Urban Growth", in Bourne, L. S. and R. MacKinnon, eds., Urban Systems Development in Central Canada, Toronto: Univ. of Toronto Press, pp. 107-117.

Hodgson, M. J., 1972, "Variability in the Growth of Small Urban Areas", in Bourne, L. S. and R. MacKinnon, eds., Urban Systems Development in Central Canada, Toronto: Univ. of Toronto Press, pp. 132-176.

Hudson, J., 1969, "Diffusion in a Central Place Hierarchy", Geographical Analysis, Vol. 1, No. 1, (January), pp. 45-58.

Isard, W., 1960, Methods of Regional Analysis: An Introduction to Regional Science. Cambridge, Mass.: The M.I.T. Press.

Isard, W. and E. W. Schooler, 1955, "Location Factors in the Petro-chemical Industry", Office of Technical Services, U.S. Department of Commerce, Washington, D.C.

Johnston, J., 1960, Econometric Methods. New York: McGraw-Hill Book Co.

Karaska, G. J., 1969, "Manufacturing Linkages in the Philadelphia Economy: Some Evidence of External Agglomeration Forces", Geographical Analysis, Vol. 1, No. 4, pp. 354-369.

Kelly, L. G., 1967, Handbook of Numerical Methods and Applications, U.S.A. Addison-Wesley Publishing Co.

King, L. J., 1966, "Cross-Sectional Analysis of Canadian Urban Dimensions, 1951 and 1961", Canadian Geographer, Vol. 10, pp. 205-224.

King, L. J., 1967, "Discriminatory Analysis of Urban Growth Patterns in Ontario and Quebec, 1951 and 1961", Annals of the American Association of Geographers, Vol. 57, pp. 566-575.

King, L. J., E. Casetti and D. Jeffrey, 1969, "Economic Impulses in a Regional System of Cities: A Study of Spatial Interaction", Regional Studies, Vol. 3, No. 2, pp. 213-218.

King, L. J., 1969, Statistical Analysis in Geography. Englewood Cliffs, New Jersey: Prentice-Hall Ltd.

Lansing, J. B. and and E. Mueller, 1967, The Geographic Mobility of Labour, Ann Arbor, Michigan: Institute for Social Research, University of Michigan.

Lasuen, J. R., 1972, "On Growth Poles", in Hansen, N. M. ed., Growth Centres in Regional Economic Development, New York: Free Press, pp. 20-29.

Lithwick, N. H., 1970, Urban Canada: Problems and Prospects. Ottawa: CMHC.

Lowry, I. S., 1966, Migration and Metropolitan Growth: Two Analytical Models. Institute of Gout and Public Affairs, University of California, San Fransisco, California: Chandler Publishing Co.

Nelson, P., 1959, "Migration, Real Income, and Information", Journal of Regional Science, Vol. 2, pp. 43-74.

Niedercorn, J. H. and B. V. Bechtolt, Jr., 1969, "An Economic Derivation of the 'Gravity Law' of Spatial Interaction", Journal of Regional Science, Vol. 9, No. 2, pp. 273-282.

Olsson, G., 1965a, "Distance and Human Interaction, A Migration Study", Geografiska Annaler, Vol. 47B, pp. 3-43.

Olsson, G., 1965b, "Distance and Human Interaction", Regional Science Research Institute, Bibliography Series No. 2, Philadelphia, Penn.

Ontario Official 1972 Road Map, 1972, Dept. of Transportation and Communications, Government of Ontario, Toronto.

Ontario Population Statistics, 1970, 1971, Dept. of Municipal Affairs, Community Planning Branch, Toronto.

Paelinck, J., 1970, "Dynamic Urban Growth Models", Papers of the Regional Science Association, Vol. 24, pp. 25-37.

Papageorgiou, G. J., 1971, "Spatial Population Growth and Migration", McMaster University, Hamilton, Ontario, mimeo.

Peaker, A., 1971, "Regional Growth and Economic Potential - A Dynamic Analysis", Regional Studies, pp. 49-54.

Perloff, S. and L. Wingo, eds., 1968, Issues in Urban Economics, Washington, D.C.: John Hopkins Press.

Perroux, F., 1955, "Note sur La Notion de 'Pole de Croissance", Economie Appliquée, Vol. 1, pp. 307-320.

Perroux, F., 1964, L'Economie duXX$^{ème}$Siede, Paris, France: Presses Universitaires de France.

Pred, A. R., 1971, "Large City Interdependence and Preelectric Diffusion", Geographical Analysis, Vol. 3, No. 2, (April), pp. 165-181.

Pred, A. R. and G. Tornqvist, eds., 1973, "Systems of Cities and Information Flows", Lund Series in Geography, Series B, Human Geography No. 38, Royal University of Lund, Lund: C.W.K. Gleerup Publishers.

Pred, A. R., 1973, "The Growth and Development of Systems of Cities in Advanced Economies", in Pred, A. R. and G. Tornqvist, eds., Systems of Cities and Information Flows, Royal University of Lund, Lund: C.W.K. Glerrup Publishers, pp. 9-82.

Schumpeter, J. A., 1967, trans. by R. Opie, The Theory of Economic Development. Oxford: Oxford University Press.

Senior, M. L. and A. G. Wilson, 1973, "Explorations and Syntheses of Linear Programming and Spatial Interaction Model of Residential Location", Working Paper 43, Department of Geography, University of Leeds.

Siegel, J., and M. Woodyard, 1971, "Urban Growth and its Relation to the Urban Hierarchy in Central Canada", Centre for Urban and Community Studies, Research Paper No. 52, (December).

Simmons, J. W., 1970, "Interprovincial Interaction Patterns in Canada", Research Paper No. 24, Dept. of Geography, University of Toronto.

Simmons, J. W., 1970, "Patterns of Interaction Within Ontario and Quebec", Centre for Urban and Community Studies, Research Paper No. 41, University of Toronto.

Simmons, J. W., 1971, "Net Migration Within Metropolitan Toronto", Centre for Urban and Community Studies, Research Paper No. 44, University of Toronto.

Simmons, J. W. and A. M. Baker, 1972, "Household Movement Patterns", Centre for Urban and Community Studies, Research Paper No. 54, University of Toronto.

Simmons, J. W., 1972, "Interaction Among the Cities of Ontario and Quebec", in Bourne, L. S. and R. MacKinnon, eds., Urban Systems Development in Central Canada, Selected Papers, Toronto: University of Toronto Press, pp. 198-219.

Statistics Canada, 1957, Employment, Earnings and Hours, formerly, Employment and Average Weekly Wages and Salaries, formerly, Employment and Payrolls, Catalogue Number 72-002, Ottawa: Queen's Printer.

Stone, L. O., 1967, Urban Development in Canada. Statistics Canada, Ottawa: Queen's Printer.

Stone, L. O., 1969, Migration in Canada: Regional. Statistics Canada, Aspects, Ottawa: Queen's Printer.

Stone, L. O., 1971, "On the Analysis of the Structure of Metropolitan Area Metropolitan Streams: A Theoretical Framework with Empirical Glimpses from Canadian and American Census Data", Educational Planning Occasional Paper No. 1/71, Toronto: Ontario Institute for Studies in Education.

Thomas, D. S., 1938, Research Memorandum on Migration Differentials. New York: Social Science Research Council.

Thompson, W. R., 1968, "Internal and External Factors in the Development of Urban Economics", in Perloff, H. S. and L. Wingo, eds., Issues in Urban Economics, Resources for the Future Inc., Washington, D.C.: John Hopkins Press, pp. 43-80.

Warntz, W., 1965, "Macrogeography and Income Fronts", Regional Science Research Institute, Monograph Series Number Three, Philadelphia, Penn.: Regional Science Research Institute.

Wilson, A. G., 1968, "Notes on Some Concepts in Social Physics", Centre for Environmental Studies, Working Paper 4, CES WP4, London.

Wilson, A. G., ed., 1972a, Patterns and Processes in Urban and Regional Systems. London: Pion Ltd.

Wilson, A. G., 1972b, "Multiregional Models of Population and some Implications for a Dynamic Residential Location Model", in Wilson, A. G., ed., 1972a, Patterns and Processes in Urban and Regional Systems, London: Pion Ltd., pp. 217-242.

Woodyard, M., 1972, "Components of the In-Migration Stream in Central Canada".Centre for Urban and Community Studies Research Paper No. 57, University of Toronto.

Zipf, G. K., 1949, Human Behaviour and the Principle of Least Effort. Cambridge, Mass.: Addision-Wesley Publishers Inc.