NORTHEASTERN UNIVERSITY

ANALYTICS OF MULTIREGIONAL POPULATION DISTRIBUTION POLICY

A DISSERTATION
SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
for the degree

DOCTOR OF PHILOSOPHY
Field of Urban Systems Engineering and Policy Planning

by

Frans Jozef C. Willekens

Evanston, Illinois
June 1976
Abstract

Analytics of Multiregional Population Distribution Policy

Frans Willekens

The purpose of this study is to explore some of the analytical problems of population distribution or human settlement policy. It extends the work of A. Rogers on spatial population dynamics to the policy domain.

The study is composed of three parts: I. Introduction: Theory and Facts of Population Distribution Policy; II. Spatial Dynamics of Structural Change in Demographic Analysis; III. Optimal Migration Policies.

Part I is an introduction to population distribution policy analysis. It discusses the basic theoretical issues and some practical features of government intervention. It is claimed that a theory of population distribution policy should be based on three pillars: migration theory, theory of externalities, and theory of government intervention. It is shown how they may combine to a single theory of settlement policy. Settlement policies in various countries of East and West show striking similarities in goals and implementations. These common features are identified and used as a framework of comparison.

Part II is a contribution to the knowledge of spatial population dynamics. No matter what the goals and means are, a population policy ultimately will result in changing the structural parameters of the demographic system, i.e., the age-specific fertility, mortality and migration rates. Applying the technique of matrix differentiation, sensitivity functions are derived which link changes in the important multi-regional demographic statistics, such as life-table statistics and population growth and stable population characteristics to changes in age-specific rates. In addition it is shown how the discrete and continuous model of population growth may be reconciled.

Part III deals with static and dynamic migration policy models. Migration is the most important element of population distribution that the policy-maker may control. In this study attention is focused on linear models in the Tinbergen formulation and in the state-space format. It is shown how they may be derived from demographic and demometric models by adding a new dimension: the goals-means relationship of population distribution policy.
The unifying element in the exposition of policy models is the matrix multiplier. Its rank and structure form the basis for a useful classification scheme, containing all the possible linear policy models.

The fundamental questions of quantitative migration policy may be expressed in terms of existence and design. This enables us to apply recent findings of systems theory and the theory of optimal control to migration policy. The existence problem is whether arbitrarily specified levels of target variables can be reached by the existing set of instruments. The conditions that must be satisfied for a population system to be controllable are formulated in a number of existence theorems. They relate to the rank of the matrix multiplier. The design procedure of optimal policies is dictated by the rank and the structure of the matrix multiplier. It is shown that, apart from mathematical programming techniques, the minimizing properties of generalized inverses are relevant in solutions of policy models with singular matrix multipliers. Although our general treatment encompasses most policy models, attention is focused on models for which solutions may be expressed analytically, such as the initial period control problem and the linear-quadratic control problem.
Acknowledgements

When I left Belgium for Northwestern, it was with a clear objective in mind: studying quantitative methods and computer techniques in urban and regional planning in an interdisciplinary framework.

I think that this objective has been achieved. A pivotal role has been played by Professor Dr. A. Rogers. He enabled me to come to Northwestern by obtaining a research assistantship, and helped to design a curriculum of courses in the departments of civil engineering, industrial engineering and management science, economics, and geography. During the development of this dissertation, I have benefited from his close cooperation. The discussions on spatial population dynamics and policy have always been rich sources of scientific experience, and I am extremely grateful to him for this opportunity.

Next I would like to thank my teachers at Northwestern for introducing me to their areas of expertise. In particular, I am indebted to the dissertation committee members: Professors Dr. J. Blin, Dr. G. Peterson and Dr. W. Pierskalla.

In addition there are a great number of people who contributed indirectly to the achievement of the objective. I limit myself to mentioning only two of them. Professor Dr. G. Boddez of the University of Leuven, Belgium, gave the initial stimulus to continue my education in the United States. Professor Dr. E. Tollens of the National University of Zaire helped me through the application procedures and kept motivating me. Without his cooperation, I would not have succeeded.
This study has mainly been written at the International Institute for Applied Systems Analysis, Laxenburg, Austria. The intellectual atmosphere and the scientific services at IIASA have largely stimulated my work. Financial support has been provided in terms of a research assistantship in the Human Settlements and Services Area. Additional financial aid was given by the Brabant Regional Economic Council, Belgium.

The burden of typing this manuscript and the preceding sequence of drafts was borne by Linda Samide at IIASA. She performed the difficult task of transforming my confusing handwriting into a final copy with great skill and good humour.

Last but not least, I want to thank my wife Maria and my son Jan. Maria always has understood my intellectual ambition and was willing to pay for it in terms of long lonely days and of weekends at home. Jan, on the other hand, kept wondering why my work was more important than his games. On a few occasions, it was hard to come up with a reason.
PART I
INTRODUCTION

CHAPTER

1. POPULATION DISTRIBUTION POLICY: THEORY AND FACTS
       1.1.1. Migration Theory
       1.1.2. The Theory of Externalities
       1.1.3. Theory of Government Intervention
       1.2.1. Population Distribution Policy: Goals
       1.2.2. Population Distribution Policy: Means
PART II

SPATIAL DYNAMICS OF STRUCTURAL
CHANGE IN DEMOGRAPHIC ANALYSIS

CHAPTER

2. IMPACT OF CHANGES IN AGE-SPECIFIC RATES ON LIFE TABLE FUNCTIONS

2.1. The Multiregional Life Table

2.2. Sensitivity Analysis of Life Table Functions

3. IMPACT OF CHANGES IN AGE-SPECIFIC RATES ON THE POPULATION PROJECTION

3.1. The Discrete Model of Multiregional Demographic Growth

3.2. Sensitivity Analysis of the Population Projection

4. IMPACT OF CHANGES IN AGE-SPECIFIC RATES ON STABLE POPULATION CHARACTERISTICS

4.1. The Multiregional Stable Population

4.2. Sensitivity Analysis of the Stable Population

Appendix: Matrix Differentiation Techniques
PART III
OPTIMAL MIGRATION POLICIES

CHAPTER

5. OPTIMAL MIGRATION POLICIES: A CONCEPTUAL FRAMEWORK

5.1. The Tinbergen Paradigm

5.2. Survey of Policy Models

5.2.1. Classification of Policy Models According to the Rank of the Matrix Multiplier

5.2.2. Classification of Policy Models According to the Structure of the Matrix Multiplier

5.2.3. From the Tinbergen Model to the Optimal Control Model

6. REPRESENTATION AND EXISTENCE THEOREMS OF MIGRATION POLICIES

6.1. State-Space Representation of Demometric Models

6.1.1. The State-Space Model

6.1.2. Solution of the State-Space Model

6.1.3. State-Space Representation of the Reduced Form Model

6.2. Existence Theorems of Migration Policies

6.2.1. Existence Theorem in the Tinbergen Model

6.2.2. Existence Theorems in the State-Space Model

7. DESIGN OF OPTIMAL MIGRATION POLICIES

7.1. Design in the Tinbergen Framework

7.1.1. The Matrix Multiplier is Nonsingular and of Rank N
CHAPTER

7.1.2. The Matrix Multiplier is Singular and of Rank N

7.1.3. The Matrix Multiplier is Singular and of Rank K

7.2. Design in the State-Space Framework: Fixed Targets at the Planning Horizon

7.2.1. Initial Period Control

7.2.2. Linear Feedback Control

7.2.3. Horizon Constrained Optimal Control

7.3. Design in the State-Space Framework: Trajectory Optimization

7.3.1. Specification of the Optimal Control Model

7.3.2. The Discrete Minimum Principle

7.3.3. The Linear-Quadratic Control Problem

8. CONCLUSION

8.1. Theoretical Issues of Population Distribution Policy

8.2. Dynamics of Structural Change in Demographic Analysis

8.3. Optimal Migration Policies

8.4. Recommendations for Future Research

Appendix: The Linear-Quadratic Control Model: Solution of the Two-Point Boundary-Value Problem
PART I
INTRODUCTION

Governments all over the world are adopting explicit policies to guide the growth and the distribution of their populations. The realization that land and environment are not free goods, but are scarce resources to be conserved, is accelerating this trend. Direct government control substitutes for the invisible hand of the "laissez faire" system, in an era of growing awareness of the divergence between the way individuals behave and the way the community wants them to behave.

Governments that do not have a settlement policy are being implored to do so. A survey on public attitudes toward population distribution issues, conducted for the Commission on Population Growth and the American Future, revealed that over half of the people see population distribution as a serious and continuing problem and favor Federal action (Mazie and Rawlings, 1972). The outcry is the loudest at the local and regional levels, where the negative effects of the current distribution are felt directly. City, county and even state governments are forced by the voters to adopt no-growth or slow-growth strategies (Scott, Brewer and Miner, 1975). These growth issues at the local level result in distribution issues at a higher level (Alonso, 1973; p. 205).

Although the political debate has been going on vigorously, there has been only limited effort to understand the basic problems of redistributing the population.
Morrison (1972a) and Alonso (1972) have illustrated the difficulties of implementing a population distribution policy. Hidden or implicit policy measures may counteract explicit policies. These hidden factors affecting migration are overlooked in most policy formulations, especially in those instances where government intervention is problem oriented and fragmented.

"Population distribution is the territorial aspect of a highly connected and interdependent social system," writes Alonso (1972, p. 636). To intervene in such a system we need to know how the spatial population system works, and how it is affected by social, economic and cultural factors. The mechanism of spatial demographic growth has been studied by Rogers (1968, 1971, 1975). He shows how important population characteristics emerge from the interaction of fertility, mortality and migration. A change in population characteristics, desired or not, may be traced back to changes in these three basic forces. Before spelling out a policy, the policy-maker should, therefore, have an idea of what the sensitivity of a demographic characteristic is to changes in these basic forces.

A fundamental feature of population distribution policy is that it does not occur in a vacuum. In most instances, it is subordinate to social and economic policies. It has been said that population policy is a policy that tries to eliminate the demographic causes of the problems to be solved (Davis, 1971; p. 7). Most governments are reluctant to define policy goals in terms of demographic
variables, such as the population level or the level of in- and out-migration. Instead, population policy is intended to achieve ultimate non-demographic goals and it makes use of non-demographic, i.e., social, economic and legal, instruments. One must, therefore, know how population distribution and migration relate to socio-economic conditions. This interdependence has been studied in migration theory, and has been represented by what Rogers has called demometric models.¹

Demographic and demometric models are the building blocks of policy models. They describe the behavior of the system undisturbed by policy intervention. To transform them into normative models, we must add a new dimension: the goals-means relationship of population distribution policy. This involves the definition of desired population characteristics and of means to achieve the desired features. That is where the policy-maker comes in.

This study is composed of three parts: Part I is an introduction to the problem of population distribution policy analysis. It discusses some of the basic theoretical issues and some practical features of government intervention. In the literature on policy models, it is assumed that a policy-maker or government defines the goals and the means, and implements the policy. But how valid is this assumption? What is the rationale for government intervention? What can it do better than the invisible

¹Demometrics is the science of identifying, estimating and testing demographic models.
hand? What are the social consequences of the individual's migration decision? These and related questions are treated in a theoretical context. Eventually they will be the basic issues of a theory of population distribution policy. The second section looks at the facts of settlement policy. It focuses on the common characteristics of government intervention in a number of countries.

Part II is a contribution to the knowledge of spatial population dynamics. No matter what the goals and means are, a population distribution policy will ultimately result in changing the structural parameters of the demographic system, i.e., the age-specific fertility, death and migration rates. The policy-maker should be aware of the hidden effects of these changes on the population system. In Part II, we derive a set of sensitivity functions for the important multiregional demographic statistics, such as life table statistics, and population growth and stable population characteristics. The unifying technique used is matrix differentiation.

Part III deals with migration policy models. Migration is the most important element of population distribution that the policy-maker may control. The starting point is a demographic and demometric model, written as a system of simultaneous equations. The goals-means relationship is introduced into these models following the Tinbergen paradigm. We assume that a policy-making body defines the objectives and the range of instruments. The Tinbergen paradigm allows us to treat a wide variety of policy models in a common framework. Most of the
attention in Part III is devoted to two central issues in policy modeling:

a) Does a policy exist that leads to the predefined goals, i.e., does the policy model have a solution? This is known as the existence problem.

b) What is the optimal policy, given the goals-means relationship and given the internal dynamics of the spatial system? This is referred to as the design problem.

To answer those questions, we apply the recent findings of mathematical system theory and of the theory of optimal control to migration policy analysis.
An eventual theory of population distribution policy will be based on three pillars: a migration theory, a theory of externalities, and a theory of government intervention. None of these theories is fully developed yet, and a theory of population distribution policy has not even been mentioned as such. But despite the under-development of the theory, several governments all over the world already have population distribution policies or are moving in that direction.

It is the purpose of this chapter to discuss some theoretical and practical problems of population distribution policy. In this first section, we deal with some fundamental issues of migration theory, the theory of externalities and the theory of government intervention. The treatment is pragmatic and completeness is, therefore, not pursued. We restrict ourselves to issues that are likely to be relevant in the construction of a theory of population distribution policy, and relate them to each other in order to provide a basis for such a theory.

The second section of this chapter deals with the facts, the actual formulation and implementation of settlement policies by governments all over the world. Although the particular form of the policy differs from one country to another, there are striking similarities in goals and means. It is to these general characteristics that we will direct our attention.
1.1. POPULATION DISTRIBUTION POLICY: THEORETICAL ISSUES

Migration theory attempts to explain why people move, and why the migration flows have their current patterns. Basically, it comes down to understanding the behavioral characteristics of the migrant. How are decisions to move and the choice of location made? Migration theory suggests that an individual will move if he is sufficiently dissatisfied with his current location, and that he will move to a place where he expects to be better off.

From welfare theory, one knows that society as a whole gains if the migrant is better off, without making anyone else worse off. However, the real world does not work so conveniently. Someone may be forced to move as a result of a sudden unexpected status of unemployment, or by an urban renewal program that was designed without his participation. On the other hand, the consequences of one's decision to migrate may extend much further than one is aware of. These situations, or imperfections of the decision-making structure, cause an individual's optimum to diverge from what is best for society. They call for regulation and provide the rationale for migration policy. The fundamental role of migration policy is to compensate for imperfections in the decision-making structure of the migrant. Most of the imperfections are included in what economists have labeled externalities. The concept of externality has received much attention in

\[\text{In economics, decisions are made on the market. Imperfections in the decision-making structure are, therefore, equivalent to market imperfections.}\]
recent years. Originally it was developed to handle by-products of the productive process for which no market exists, such as pollution, noise, and the like, in general equilibrium analysis in welfare theory. As early as 1920, Pigou devoted an extensive discussion of these spillover effects.

Although externalities have frequently been considered to be a sufficient basis for government intervention in private decision-making, they have not been integrated into a single theory of policy. The theory of economic policy, for example, has traditionally been directed towards the goals-means relationship (Tinbergen (1963), Theil (1964), Fox, Sen Gupta and Thorbecke (1972), Pindyck (1973), Friedman (1975)). No attention is given to questions concerning why the government should intervene in the first place, how the private and social welfares relate to each other, and what the character and the optimal level of intervention should be. These issues have been dealt with by Baumol and Oates (1975) in their "Theory of Environmental Policy." The authors approach the policy problem from the perspective of general equilibrium theory. The unique objective is Pareto optimality. It is the government's task to assure that the optimality conditions are met by internalizing externalities where they exist. The level of abstraction of their work is intended to provide theoretical insight into the complex problem of policy.

In this study, an attempt is made to integrate both approaches to the theory of policy. The theoretical underpinning of the theory of policy is provided by the theory of externalities. For the implementation and
evaluation of the policy, the Tinbergen paradigm appears to be more useful and is treated in Part III.

1.1.1. Migration Theory

During the past two decades, scientists of several disciplines have been attracted by the phenomena of distribution and redistribution of people over space. The observed human settlement system is the result of two flow phenomena, natural increase and migration. While the study of natural increase has always been a privilege of demographers, joined recently by economists, migration research has been multidisciplinary from the beginning. Planners, demographers, economists, geographers and sociologists have addressed migration-related questions. The different orientations and inclinations of scholars have resulted in a great diversity of migration studies, which has not yet been integrated into a unique interdisciplinary approach. Demographers have typically looked upon migration as a component of population change; economists have examined it as a mechanism enabling an individual to adjust to new situations and enabling the labor market to adjust when disturbed from its equilibrium position; geographers have been primarily interested in the description and explanation of the spatial patterns of mobility; and sociologists have focused on the study of motivation, of the relation between migration and social structure, and of the assimilation of migrants in new communities. A feeling for the state-of-the-art is given by the recent surveys and bibliographies of Welch (1970), Byerlee (1972), Gould (1974),

In contrast to the extensive enquiry on internal migration, which was mainly empirically oriented, little effort has been devoted to the synthesis of this fragmentary knowledge into a general migration theory. The following sections are intended to give a general idea of the approaches followed in the construction of a migration theory. Not surprisingly, most researchers have attempted to build up a theory from the empirical observations. Unlike other social sciences such as economics and sociology, only a very few have adopted the deductive approach.

a. Migration theory: inductive approach

The first attempt at formulating a migration theory dates back almost a century. Following an empirical study on population movements, first in Britain and later in twenty countries, Ravenstein (1885 and 1889) formulated the observed empirical regularities as "Laws of Migration." These came as a reaction to an earlier study of Farr (1876), which claimed that migration was random. The gravity type laws formulated a crude answer to the questions why people migrate, what the migrant's characteristics are, and what the pattern of internal migration is.

The Ravenstein work has been extended along two paths: (i) extension and reformulation of the list of empirical regularities, and (ii) expression of the regularities in gravity type models. Along the first path, Bogue (1959; pp. 499-501) came up with extensive lists of situations
affecting migration. He explicitly stated that the lists are nothing more than a framework for migration analysis, and should not be interpreted as laws or theory. Lee (1966) provides a more explicit attempt at theory formulation. Migration is the result of a decision-making process. Lee classifies the factors that enter directly into the decision-making into four sets:

i) Push factors: factors associated with the area of origin;

ii) Pull factors: factors associated with the area of destination;

iii) Intervening factors: obstacles associated with the movement itself;

iv) Personal factors: characteristics of the potential migrant, that determine the way in which he perceives and evaluates migration as a personal project.

These factors constitute the context, or motivational structure, as Taylor (1969; p. 132) calls it, out of which the decision to migrate finally crystallizes. Lee uses this structure to formulate nineteen hypotheses about the volume of migration, the migration directions and the characteristics of the migrants. Central to many of these hypotheses is the observation that migration is an adjustment to changes in personal and economic conditions.

Ravenstein (1885; p. 198) stated

In forming an estimate of this displacement, we must take into account the number of natives of each county which furnishes the migrants, as also the population of the towns or districts which absorb them.... Migrants enumerated in a certain centre of absorption will consequently
grow less with the distance proportionately to the native population which furnishes them.

This formulation constitutes, in fact, a gravity model: the interaction, e.g., migration between two regions is proportional to the product of their populations and is inversely proportional to the distance between them. The gravity model, developed by Zipf (1946) and later by others, is based on the hypothesis that people interact at a distance as physical particles, obeying Newton's law. This physicist's interpretation of social phenomena has influenced migration research very strongly. Several authors have fitted the gravity model to migration data. In order to increase the performance of the model, exponents and weighting coefficients have been introduced (Isard, 1960; pp. 493-568).

However, the gravity model currently is only an empirical law involving interaction at a distance. It fails to provide an understanding of why there should be such interaction.

There have been several attempts to give a theoretical basis to the gravity model. The distance concept has been related to intervening opportunities by Stouffer (1940 and 1960). In this theory, people move because of opportunities and "the number of persons going a given distance is directly proportional to the number of intervening opportunities" (Stouffer, 1940; p. 846). Isard (1969; pp. 871-874) formulates an abstract "effective distance" concept, which is a resultant of geographic distance and of the economic, social and political distances affecting the interaction of people. A very similar concept is developed by Brown and Horton (1970; p. 76) under the label "functional distance."
The "mass" variable also has been given a theoretical underpinning. Instead of total population, one has used nonagricultural employment and total civilian employment to approximate the availability of economic opportunities.

Large efforts have been devoted in the past several years to determine the factors that enter the migration decision. Following the inductive approach, such efforts have tried to recover, from surveys and reported census data, the causes of observed migration behavior. Attitude surveys go to the heart of the problem by asking migrants directly about their motives\(^3\). Most researchers, however, rely on aggregate data on migration and socio-economic conditions, reported by the Census Bureau or other institutions. Using correlation and regression techniques, they try to infer the determinants of migration. However, the primary purpose of most regression models of migration reported in the literature is a description and an ad hoc explanation of empirical regularities. This search for empirical uniformities is not enough to construct a migration theory. What is required, is a synthesis of these uniformities into coherent interpretations of behavior (Morrison, 1972a; p. 289).

\(^3\)Morrison (1972a) summarizes the findings of the three most important surveys in the United States: the Current Population Survey (1946), the Bureau of Labor Statistics Study (1963) and the Lansing and Mueller Survey (1967).
b. Migration theory: deductive approach

The antipode of the previous approach is the deductive analysis of migration. Instead of focusing on the empirical regularities that result from the actions of the migrants, it may be fruitful to take a distant view at the migrant, and to ignore everything but the essentials. What we would see then is an abstract individual with no other characteristics other than what are common to all migrants. Such creatures, for example, are the "homo economicus" and the "homo sociologicus." Their characteristics constitute the basics of economic and sociological theory, respectively. The "homo migrans" would have some factors in common with both the "homo economicus" and the "homo sociologicus," but he would also have unique characteristics that make him identifiable as a separate entity.

It is beyond the scope of this study to attempt to derive the properties of the "homo migrans." It would constitute the migration theory. However, some initial work along these lines will be reviewed. This pioneering research has been mainly inspired by economic theory. Consequently the "homo migrans" resembles the "homo economicus" and seeks continuously to maximize economic advantages.

Sjaastad (1962) treats the migrant as an investor, and migration as an investment project. It is assumed that the potential migrant will undertake the project, i.e., migrate, only if the benefits exceed the costs, discounted at a proper rate which expresses his time preference. This
approach also assumes that the number of alternative locations to move to is limited, and that the migrant has full information on the costs and benefits involved.

An extension of this cost-benefit model is due to Wolpert (1965). According to Wolpert, an individual assigns a "place utility" to his current place of residence which represents the social, economic, and other costs and benefits derived from that location. Alternative locations are also assigned utilities based on anticipated costs and benefits. The "place utility" concept measures the degree of satisfaction or dissatisfaction at a given location. However, unlike Sjaastad, Wolpert restricts the range of alternative locations to what he calls the migrant's "action space." The first condition for a place to be in the "action space" is that the migrant has sufficient information about it to assign "place utilities." The migrant then behaves according to the principle of utility maximization. He chooses the location which gives him the greatest "place utility," subject to the constraints imposed by the "action space." This approach has been used and extended by Brown and Moore (1971) and by Speare, Goldstein and Frey (1974, Chapter 7).

Common to the cost-benefit or utility maximization approaches to migration theory is the private character of the migrant's behavior. It is assumed that he has control over all the variables entering his place utility function, and that his decision has an impact on only himself. Let \( \{\xi\}^A \) and \( \{\xi\}^B \) represent the features of place utility or the residential bundle of individuals A and B, respectively. Then the place utility of A and B
may be written as functions of these vectors:

\[ u^A = u^A(\{x\}^A) \]  \hspace{1cm} (1.1)

\[ u^B = u^B(\{x\}^B) \]  \hspace{1cm} (1.2)

where \( u^A \) is the place utility of person A, and \( u^A(\cdot) \) represents his place utility function. In this simple two-person society, the social welfare is a function of A and B's utility, i.e.,

\[ w = w(u^A,u^B) = w\left[u^A(\{x\}^A),u^B(\{x\}^B)\right]. \]  \hspace{1cm} (1.3)

However, the migrant does not live in a vacuum. Persons A and B are members of the same society and continuously influence each other, even if they do not realize it. Their private behavior does not take into account these hidden influences. The incompatibility between the private behavior of the migrants and their collective consequences has given rise to numerous issues of migration policy. The theoretical framework to discuss these issues is summarized in the next section.

1.1.2. The Theory of Externalities

In the early literature on externalities, the concept was never clearly defined. Bator (1958) includes most major sources of market failure in the concept. Buchanan and Stubblebine (1962) take a much more narrow view. "Pareto relevant externality," as they call it, is present whenever, in competitive equilibrium, the marginal conditions
of optimal resource allocation are violated. The disadvantage of such a definition, is that it does not tell us what externalities are, only what their consequences are.

Although no uniform definition exists, most formulations include two conditions that must be satisfied for an externality to occur (Baumol and Oates, 1975; pp. 17-18).

**CONDITION 1**: An externality exists, if someone, say A, is affected by a decision made by someone else, say B, without giving particular attention to its impact on A's welfare. In other words, an externality exists whenever someone is affected by a decision in which he takes no part, either directly or indirectly.

Note that if B deliberately does something to affect A's welfare, no externality is involved (Mishan, 1969; pp. 342-343). For example, if an individual moves out because he wants to get away from the people he grew up with, no external effects can be assigned to those left behind.

The first condition determines a type of externality, which Meade (1973; p. 27) labels "externalities due to a shared variable," i.e., the otherwise independent decision-makers A and B share a particular variable in their utility function. It does not mean, for example, that both A and B base their decision to move on the same features of place utility or on the same residential bundle. An externality arises when the residential bundle enjoyed by B affects the place utility of A. Let \( \{x^B\} \) be a vector representing the elements of the residential bundle of B, and let \( U^B(\cdot) \) be the place utility function of B. The place utility of
B is then:

\[ U^B = U^B(\{\hat{x}\}^B) . \]  

(1.4)

Equation (1.4) states that the utility of B is uniquely determined by its own residential bundle. Unlike B, A's place utility depends not only on his own residential bundle \( \{\hat{x}\}^A \), but also on B's residential bundle, i.e.,

\[ U^A = U^A(\{\hat{x}\}^A,\{\hat{x}\}^B) . \]  

(1.5)

In the migration theory, it has been assumed that A has control over all the variables entering his utility function, and that the decision of B only affects himself. However, (1.5) illustrates how a set of variables may be imposed on the utility function of one individual by a unilateral decision of another. When externalities are taken into account, the social welfare function differs substantially from (1.3),

\[ W = W(U^A, U^B) = W\left[U^A(\{\hat{x}\}^A,\{\hat{x}\}^B), U^B(\{\hat{x}\}^B)\right] . \]  

(1.6)

It is now clear that the maximization of (1.3) does not lead to the social optimum if A's utility function is (1.5), i.e., if externalities occur. The derivation of conditions that maximize (1.6), subject to constraints, is a subject of the theory of externalities (see Holterman, 1972; Baumol and Oates, 1975, Chapter 4).
Instead of exploring the theory further, we will continue with the second condition.

**CONDITION 2:** The benefit or the burden A experiences as a result of B's decision is not fully compensated for by either party.

The second condition relates externalities to market failure. However, market failure is not a sufficient condition for externalities to be present. Recent authors, such as Baumol and Oates (1975; p. 18) and Meade (1973; p. 20), make a clear distinction between the first and the second condition. They argue that Condition 1 is necessary and sufficient for an externality to arise. However, the externality may be internalized by the market mechanism or by a common non-market action of both parties involved, so that the external effect no longer exists. Therefore, the role of the second condition is to ensure that the externality persists, once it has been created. It is also clear that the second condition is necessary to prevent the Pareto optimum from being achieved. It is only when externalities are not cleared on the market or by a mutual agreement of the parties involved, that they result in misallocation of resources and government intervention is needed to install the Pareto optimum.

The reasons why the market fails to clear the externalities are related to the costs of market organization. The transaction costs of externalities are so high that the existence of a full market is no longer worthwhile. The transaction costs originate from two sources (Musgrave, 1959; p. 86 and Arrow, 1971; p. 19):
a. Exclusion cost.

It may be infeasible, technically or economically, to exclude people (non-buyers) from the use of a product. A classical example is the apple-grower, who is unable to influence the number of bees of his neighboring bee-keeper, that pollinate his apple trees.

In the population distribution context, the exclusion principle may not work for constitutional reasons. In the United States, people have the constitutional right to move. Local governments or citizen groups may not prevent anyone from coming to live in their localities. Some have tried hard but their attempts have been declared unconstitutional.

For example, a quota system that limited the total number of new residential units permitted to be constructed each year, was adopted by Petaluma, California, but in 1974 a federal court ruled that "the basic constitutional rule is that no city can regulate its population growth numerically so as to preclude residents from any other area from traveling into and establishing residence there" (International City Management Association, 1975; p. 270). Another classic example is Mount Laurel, New Jersey, in which the zoning ordinance was declared unconstitutional in 1975 (Lewis, 1975; p. 4).

Because of the impossibility of keeping migrants out of the destination areas, the externalities caused by the inflow cannot be internalized.

b. Cost of communication and information.

The perfect market paradigm is based on the hypothesis of full information on prices and quantities. When the
necessary information to conclude the market transactions is lacking, then the market cannot perform its function.

The inability of the migration mechanism to relocate labor resources efficiently is primarily explained by the high transaction costs involved. Suppose someone moves to another region in response to job vacancies. His success in finding a suitable job, i.e., his expected utility, will depend on the decisions to migrate made by other people, for the same reason. Since it is impossible to exclude people from responding to job vacancies, the market failure cannot be remedied. On the other hand, the migration decision is based on the available information about alternative localities. The diffusion of information takes time and is costly. This also causes the labor market to deviate from a perfect market.

Symptoms of the persistence of market imperfections are the excessive volume of migration compared with the vacancies to be filled, and the continuation of sizeable regional wage and unemployment rate differentials.

There is a close relation between the cost of information and uncertainty (Arrow, 1971; p. 12). Whenever the information about the outcome of a decision is incomplete, uncertainty is involved. The decision to migrate is, in fact, made under uncertainty. Under these circumstances, there are two extreme possibilities to cope with the problem of risk: full protection against uncertainty of the final outcome, and absence of protection against uncertainty of the final outcome. If a migrant is uncertain as to the final outcome of his move, he can make contracts contingent to the occurrence of possible outcomes.
Real-world counterparts of these theoretical contingent contracts include an employment contract and a rent agreement the migrant signs before his move. Some countries lacking skilled labor, such as South Africa, used to provide the immigrant, who did not have a job at the time of immigration, with a contingent contract in which the government promised to pay travel costs and a daily allowance from the date of arrival to the date a job was found. With these markets for contingent contracts, a competitive equilibrium will arise under the same general hypotheses as in the absence of uncertainty (Radner, 1968).

This brings us to the strategies to deal with externalities and market imperfections in order to reinstall Pareto optimality.

1.1.3. The Theory of Government Intervention

An externality occurs when someone is affected by a decision in which he took no part. A consequence of externalities is the reduction of social welfare. An obvious way to improve matters is to rearrange the institutions of society in such a way that the affected person does become a party in reaching any decision which seriously affects his interests. In other words, one may seek to internalize externalities.

Internalization of externalities requires collective action or joint decision-making to compensate for market failure. Such institutions for joint decision-making may take several forms: family, corporation, legal contracts, local and federal government, multinational ruler. Arrow
(1971; p. 22) also includes the norms of social behavior, including ethical and moral codes.

A fundamental issue in the institutionalization of collective action is how joint decisions should be reached. It has already been indicated that the party imposing externalities, B, and the party suffering or enjoying them, A, may internalize the externalities by negotiation and compensation. In the extreme case, they may even form a merger to produce a single-decision unit, a practice not uncommon in the business world. If unrestricted bargaining is feasible and costless, then the equilibrium which will be reached is Pareto efficient, for it is in the interest of the parties to switch from a Pareto inefficient allocation to a suitably chosen Pareto efficient one.

In many cases, free and unlimited negotiation is impossible. The people imposing the externalities and/or the people affected may be too large a group for the members to engage in direct bargaining. The joint decisions must then be made on another basis. If one adopts the rule that no action will be undertaken except with the unanimous agreement of the members of the institution set up to internalize the externalities, it is unlikely that anything will ever be done. Actions, which make any single member worse off will be difficult to take.

An alternative form of a collective decision-making rule is to abide by the result of a majority vote. In this case, however, there are considerable possibilities for a paradox (Arrow, 1971; p. 21). Several of the major problems
of majority voting may be solved by assigning different weights to the voters. But the allocation of weights may in itself introduce an asocial element in the collective decision process.

A final possibility to internalize externalities is that the decisions are taken by some authoritarian manager, controller, or guardian, on behalf of the parties involved. Such a regime may be set up on a voluntary basis by the parties in a constituent joint decision to appoint a governor and to accept his decisions. In a complex society, the governor may transfer some responsibilities to other persons and agencies, creating a governmental structure, but the principle remains the same. Instead of a unique governor, a body of representatives may be elected to make the social choices. The kind of collective decision rule that will be installed depends largely on the society. But what is essential is that the government is considered here as a device to internalize externalities and hence to ensure that individuals behave so as to maximize social welfare. It should be remembered that the justification of a government is not merely the existence of externalities, but the impossibility of internalizing them by direct negotiation between the parties involved. If unlimited negotiation is feasible and free, no government intervention can be justified on economic grounds (Turvey, 1963; p. 313).

In the policy models developed in Part III of this study, it is assumed that a government or policy-maker exists and that it performs its duty of enhancing social
welfare well\textsuperscript{4}. This duty includes, in the first place, the formulation of social preferences or policy goals. The preference system may be formulated as a vector of relevant variables with their desired values (targets) at a planning horizon. It may also be expressed as a function of the relevant variables, which represent an ordering of alternative social states according to their desirability. This function is the social welfare function or the social choice rule, as Heal (1973; p. 27) prefers to call it.

The main difference between both approaches is that the welfare function takes the trade-offs between the preferences or target variables into consideration. The concept of trade-off resembles the concept of a consumer's marginal rate of substitution. It is the amount by which the value of one variable of the welfare function must be increased to compensate for a unit decrease in the value of another variable. In a democratic society, the heaviest political discussions have to do with the setting of trade-offs or, in more common terms, priorities. In most instances, there is no substantial disagreement on what variables to include in the preference system, but the issue is where to put the priorities.

There has been considerable theoretical work done in economics on the construction of the social welfare function from individual preference functions (Arrow, 1963; Sen, 1971).

\textsuperscript{4}This assumption is invalid in several practical situations, and has given impetus to the development of a theory of government failure, instead of market failure (Posner, 1974; p. 336). The discussion of this topic would lead to political issues beyond the scope of this study. We will, therefore, stick to the view of the government as servant of the public interest.
We will not explore this work here. Instead, we turn to the second duty of the government, namely, the provision of policy instruments.

Government intervention can take either of two forms:


The government can take the externality imposing activities over and run them as government enterprises. Illustrations of this institutional internalization of externalities are police, defense, public administration, and infrastructure construction. Government take-over is justified only if the private sector is unable to internalize the externalities even with government assistance. The take-over of railroads and utilities is a classical illustration. In their effort to distribute economic development more evenly over the national territory, several Western European countries have set up regional policy programs to stimulate the location of growth industries in lagging regions. To compensate for the external disadvantages associated with such a location, the government introduced tax reductions and subsidies. The limited number of firms that reacted positively to the government programs, indicates that the negative externalities involved could not be offset by the government assistance. Recently, some countries have been moving in the direction of government take-over, in the sense that a public corporation invests directly in lagging regions by the construction of state-owned enterprises.

b. Government regulation and taxation.

The government can leave the activities to private initiative, but control them in such a way as to counteract any externalities involved. Control may consist of
administrative regulation by the auctioning of licenses for example, and of taxes and subsidies. Taxes and subsidies are meant to change the price system. In equilibrium, the market price of a good is equal to its marginal cost and is also equal to its marginal utility. A tax on externality-causing activities is, in fact, a marginal cost of engaging in them. In perfect competition, the producer accepts the market price as given. Assuming increasing marginal costs, the imposition of a tax on the producer will result in a decrease of the output until the marginal cost (+ tax) equals the market price and the marginal utility of the consumer. A subsidy to the individual or group suffering the externalities works in the opposite direction.

Two fundamental questions arise with the imposition of a tax or subsidy.

b.1. The first question is who should be taxed, the producer or the consumer of the externalities. Consider, for example, the fast growing communities in retirement havens near big cities. Because of the heavy in-migration, they must accommodate too many people in too short a time. This has given rise to large amounts of external disadvantages, such as rising property taxes and a changing social and political structure of the community and has resulted in defensive actions by citizen group and by local and state governments. The basic issue which always comes up is: are local governments entitled to limit their growth and to impose a burden on the potential in-migrant? We may reformulate this issue as follows: should the in-migrant be forced to compensate the residents of the destination area, e.g., by paying a tax, or should the
residents be forced to compensate the migrant if he does not make his move? The answer to this question depends on who owns the amenities of a region, which is a problem of the distribution of property rights. If the amenities belong to the residents, it is clear that an in-migrant must compensate them for the negative externalities he causes, i.e., he must pay for the loss in amenities. However, if the amenities belong to the migrant or to all the people of the country who may consider moving in, then the additional burden of the in-migration must be borne by the residents of the destination area. Put otherwise, does the right to move dominate over the right to protect the amenities? It is a question that has kept the many local communities, which have attempted to control their growth, very busy.

One man's right to move to the city may infringe upon the right of those already there to preserve the amenities that brought them in the first place. Must Santa Barbara become Los Angeles? Or may it act to stop 'progress' that would alter if not destroy its unique balance of land, sea, and city? Must the state of Vermont stand mute while the inexorable pressure of population turns it from its past towards New York's most pleasant bedroom community?

In this statement, mentioned by Finkler (1975; p. 128), Woods touches the heart of the issue. How property rights finally will be allocated is more a social and ethical problem, rather than an economic one (Meade, 1973; p. 59). On a number of occasions, courts have upheld the basic rights of freedom of movement and

---

5 For an extended discussion of the issue, see Scott, Brewer and Miner (1975, Chapter 10).
have forced local governments to change legislation which was set up to preserve local amenities. We have already mentioned Petaluma and Mount Laurel. In 1973, a Circuit Court invalidated an impact tax imposed by Broward County, Florida, on all new residents to pay for the construction of additional bridges and roads and to lessen traffic congestion. The case is reviewed in great detail by Janis (1975; pp. 290-295) to illustrate who should pay the costs of population growth. Other such cases are reviewed by Fielding (1975) and Franklin (1975). Because of the fundamental character of this issue, the assignment of property rights must play an important role in a theory of population distribution. Discussing one instrument of population distribution policy, zoning, Tarlock (1975; p. 228) states: "The development of an adequate theory of zoning should start from recent development in welfare economics regarding property rights theory."

b.2. Once we know who should be taxed, the question is how much? The social optimum is achieved when the social cost of an additional unit of output is equal to its social benefit. In this situation, not all of the externality-causing activities will be stopped, but the cost to society of further reducing the externalities will exceed the marginal damage done by the externalities. There is thus a social optimal level of externalities. The taxation which reduces the externalities to this optimal level, by increasing the marginal cost to the externality-producing activity, is the optimal tax amount (Baumol and Oates, 1975; p. 136).
This theoretical search for the optimal level of government intervention is difficult, if not impossible, to apply in practice. Baumol and Oates (1975; pp. 136-137) confess:

Because we are unable to measure social welfare, and because we do not know the vector of inputs and outputs that characterize 'the optimum,' we simply do not know whether a given change in the tax rate has moved us toward that optimum or has even been able to improve matters. There seems to be no way in which we can get the information necessary to implement the Pigouvian tax-subsidy approach to the control of externalities.

The local governments that have considered taxing the new residents have based the amount of the tax on some proxy measures of the negative externalities involved. Broward County, for example, imposed a $9.56 tax on each 1,000-square-foot dwelling built on a one-acre plot (Janis, 1975; p. 294). California Tomorrow, a citizen group, proposed to levy a lump sum tax of $1,000 on each new resident from out of state (Forbes, 1975; p. 59).

For the operational policy models, to be discussed in Part III, we have assumed that the government is able to express the vector of variables that characterize "the optimum." In addition, it will be assumed that (a) the kind of government intervention, i.e., the instrument variables, are selected by the policy-maker in the best interests of all the people; (b) the marginal impact of each instrument variable on the target variables, or on the value of the welfare function is known, i.e., we know whether the application of an instrument has moved us closer to the
optimum, and by how much; and (c) the range of available instruments is limited.

These assumptions are necessary to operationalize the process of social welfare maximization. They allow us to express the policy problem of pursuing the social optimum in concrete terms: the maximization of the social welfare function, subject to constraints imposed by the availability of instruments and by the social system to be regulated, or, in other cases, the pursuing of specified values of a set of target variables. The description of the policy problem in this way conforms completely with the "standard" formulations of Heal (1973; p. 5), Fox, Sengupta and Thorbecke (1972; p. 11) and others.

The role of the government in internalizing externalities has been our focus thus far. We now examine current practice to see how governments in different countries perform their task of guarding the public interest in matters of spatial population distribution.

1.2. POPULATION DISTRIBUTION POLICY: FACTS

Bergman (1974; pp. 11-12) summarizes the essence of population policy as follows:

Population policy can be defined as government action toward objectives which involve the influencing of population characteristics. Thus population policy might be visualized as government actions which influence population characteristics influencing size, composition, distribution, and rates of growth. However, these steps are not taken in a vacuum or exclusively to influence population characteristics merely for the sake of changing them alone. Rather they are taken with reference to other characteristics or conditions of a community or society in which change is sought, such as the expansion of opportunities for employment, education, shelter, health; or the contradiction of these opportunities;
or the security of the community, internal or external; or other priorities envisioned by a society and its government. And these actions can be taken either to the advantage or the disadvantage of a certain population group such as the young, the old, the rich, the poor, ethnic majorities, religious minorities, and the like.

An important characteristic of population policy is that it is intended to achieve ultimate, non-demographic goals. People judge a population trend to be good or bad only in the light of its presumed social and economic consequences. Therefore, whatever the demographic objective or target, population policy is always viewed as instrumental to a non-demographic goal. "A population policy is a policy that tries to eliminate the demographic causes of the problem to be solved," says Davis (1971; p. 7). Thus, population policy is an integral part of a broad economic and social development policy.

Because of this intimate connection between population growth and distribution and the overall economic development and quality of life, population distribution policies can never be evaluated in the narrow framework of direct intervention into the internal flow of people. Alonso (1973; p. 205) goes even further by saying that

... we should avoid what is often done, setting as policy goals arbitrary demographic rates (no growth or fast growth) or particular geographic patterns of distribution (dispersal or concentration). These rates and patterns are not proper goals, although they may be important instruments for advancing the real goals of material efficiency, of equity or fairness, of ecological integrity, and of high quality of life. They are used as goals because they are easy to grasp and they avoid the real questions which are hard.
The real goals and the most effective instruments of a population distribution policy are to be found outside the framework of population. Alonso (1972; p. 638) puts it as a paradox:

It is a curious paradox that most present explicitly territorial policies are thought to be ineffective, while it appears that many other policies and programs, whose intent was not originally territorial, powerfully affect the distribution of population and economic activity. Striking examples of these implicit policies are in the field of defense spending. Through their secondary effects, they constitute a "de facto" policy of population distribution. Morrison (1975; p. 103), therefore, urges that any intervention in the settlement system begin with an assessment of such "hidden" policies.

It is impossible to determine the goals and the means of population distribution policies without studying the general economic and social policy. No country has a separate policy aimed only at the internal redistribution of the population. Human settlement programs are usually part of a regional economic policy or of a land use policy or physical planning program. One has to consult official publications dealing with either of these government activities to find statements related to population distribution policies.

In the following sections, we summarize the goals and the means of settlement policy of several countries of the world. The focus is on the general characteristics common to most countries, and not on the characteristics particular to each country.
1.2.1. **Population Distribution Policy: Goals**

The purpose of the formulation of a goal set is to circumscribe in more practical terms what is meant by a social optimum. Alonso (1972; pp. 637-638) considers four ultimate goals of a population distribution policy: efficiency, equity, environment and life style. A similar classification if given by Hoover (1972; p. 654) and Berry (1975; pp. 70-71).

A population distribution policy is not directed to any one of these goals, but rather pursues several of them simultaneously. Based on the distribution of the weights between the four goals, we may distinguish three broad types of settlement policies:


b. Enhancement of the interregional mobility of the factors of production.

c. Solution to problems of urban growth and decline.

Efficiency considerations are predominant in the first two types; the third is related more to the goals of environment and life style.


Efficiency has always been an important goal in settlement policy. Such an aim requires that the population be located where it can contribute most effectively to national production. The first policies in the United States which influenced the migration and the distribution of population were designed for land development. The Northwest Ordinance of 1787 and the Homestead Act of 1862 were part of a national policy to settle the western
territories. The primary aim was not to distribute people over space but to seek out new growth opportunities. Such "exploitative opportunity-seeking" policies, as Berry (1975; pp. 75-77) calls them, produced large migrations, not only in the United States, but also in South Africa (the Trek) and several other countries. Currently, "exploitative opportunity-seeking" policies are directing people to Siberia in the USSR and to Amazonia in Brazil.

b. Enhancement of the interregional mobility of the factors of production.

Efficient allocation of people over space implies the elimination of areas of labor surplus and areas of labor shortage. However, the efficiency of the spatial distribution of labor cannot be evaluated in a vacuum. It depends on the distribution of capital. Maximum output is only achieved when the marginal productivity of labor and of capital is equal at all locations. Consequently, a situation in which area A has higher returns to labor but lower returns to capital than area B might be improved either by movement of labor from B to A, or by movement of capital from A to B, or by both.

National governments have applied both strategies to improve efficiency in interregional production. The first policy of labor migration started in Britain, following a report of the Industrial Transference Board in 1928 (Willis, 1974; p. 35). The policy of industrial transference sought to attract industry to the Special Areas and to stimulate the movement of labor out of the Depressed Areas. During the period before the Second World War, the emphasis was on bringing the people to the jobs. Since 1945,
regional policy is concentrating on taking work to the workers. The epitome of this was the Local Employment Act of 1960. The rationale behind this switch was the increasing importance of structural unemployment in old industrial centers, especially in mining areas, the resistance among the population to move to other places, and the reluctance of the government to depopulate areas. Other European countries have also adopted a policy of "jobs to people," realizing the high social costs associated with large-scale out-migration.

These regional policies are not only based on efficiency considerations. The policies have also been justified on equity grounds. The tradeoffs between efficiency and equity considerations are difficult to assess in policy-making, since poverty and unemployment are closely linked.

After the Second World War, an additional element has entered labor mobility policy. Labor became an increasingly heterogenous good. Because of the diversification in skills, labor shortage and unemployment has been coexisting in several regions. For this and other reasons, relocation assistance in many countries is no longer restricted to migrants from depressed areas, nor to unemployed people, but also includes aid to key workers. The Key Workers Scheme in Britain, for example, is designed to help employed workers who transfer to key posts in new or expanding establishments in development or intermediate areas (Willis, 1974; p. 41). A similar scheme exists in the Netherlands and in France (Klaassen and Drewe, 1973; Chapter 2). The emphasis remains, however, on workers who are unable to find a
suitable job at home (O.E.C.D., 1967; p. 8). In Britain, the major part of government assistance to migrants comes under the Employment Transfer Scheme. Eligibility is based on the lack of early prospects of suitable and regular work and the inability to learn a skilled job in the home area (Willis, 1974; p. 41).

In the United States, no labor mobility policy exists. Between 1965 and 1969, Labor Mobility Demonstration Projects were conducted under the auspices of the Department of Labor (Hansen, 1973; pp. 39-63). The purpose was to assess the effectiveness of worker relocation assistance for reducing unemployment. To be eligible for assistance, a person had to be involuntarily unemployed and without local prospects for a fixed job or be a member of a farm family with an income of less than $1,200 per year.

c. Solution to problems of urban growth and decline.

There is an increasing awareness among people of the social costs associated with metropolitan growth. Congestion, pollution, high taxes, lack of open space, alienation, and crime all are viewed as symptoms of excessive growth, and they impose high external costs on the residents. Therefore, urban growth and the distribution of population and economic activities over space have become a primary concern. Godschalk (1975; p. 338) remarks that

Traditional comprehensive planning methods are being redefined in response to a concern with supply factors in growth. Population change is seen as a dependent rather than as an independent variable in the growth equation. The independent variables are environmental and institutional resources of the state. Land is thus a resource to be conserved, rather than a commodity for trading.
Concern about the increasing costs for social overhead capital, the environment, the quality of life, and equity is causing governments all over the world to adopt settlement strategies. The objectives are mostly in terms of restraining or even stopping the growth of cities or conurbations. In this section, we review the major objectives of population distribution policy in Europe and in the United States. In European countries the objectives are stated by the national government. The United States is moving toward a national settlement policy, but explicit national goals are not yet defined. On the contrary, local governments and even state governments have very clear growth objectives, which may ultimately force the country as a whole to adopt a national population growth policy.

a. Settlement policy objectives in Europe.

In Europe, national settlement policies are dominated by the growth disparity between the capital or other such conurbation and the rest of the country. For example, the population distribution in Britain is aimed at reducing the growth of London by decentralizing the settlement system and slowing down the urbanization of the Southeast by stopping the "North-South drift." In France the issue is the concentration of people and economic activities in Paris. It has been thirty years since Gravier (1947) published his book "Paris et le Désert Francais," but the debate he initiated is still going on.

Nowhere have equity considerations been so important in the formulation of a settlement policy as in France. At the same time that Paris has been unable to control its
growth, other French cities have remained relatively small and underdeveloped. They lack the capacity to stimulate regional development. Since the Fourth Plan (1961-1964), the main goal of French national settlement policy has been to turn back Paris' cumulative growth process in favor of other existing cities, the "metropoles de'équilibre" (Rousselot, 1975; p. 180). A dominant role in France's settlement policy is played by Datar, an Interministerial Land Use Group, created in 1963 and responsible to the Prime Minister. Since the creation of Datar, the central government has become more directly involved in urban policy (D'Arcy and Jobert, 1975; p. 304).

In the Federal Republic of Germany population distribution as a national goal has been formulated in the First Report of the Federal Government on Physical Planning in 1963 (ter Heide, 1971; p. 2995). It is stated that decentralization of industry and population from the areas where both are heavily concentrated is to be encouraged. The congested regions are located along the Rhine; the underpopulated areas are in the North and the Southeast of the country.

Population distribution policy statements in the Netherlands may be found in the Reports on Physical Planning. The first Report was issued in 1960 and focused on the over-urbanization of the three western provinces containing nearly half of the Dutch population. In the second Report (1966), the over-urbanized area includes the entire South, shifting the attention to the North-South divergence. It became a national goal to stimulate the population growth in the North. In contrast to other countries, the Netherlands have stipulated target populations
for each of five regions for the year 2000, as indicated in Table 1 (Drewe, 1971, p. 145; Drewe, 1974, p. 712).

Table 1


<table>
<thead>
<tr>
<th>REGION</th>
<th>1965</th>
<th>1973</th>
<th>Continued Trend</th>
<th>Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td>1.30</td>
<td>1.46</td>
<td>2.25</td>
<td>3.00</td>
</tr>
<tr>
<td>East</td>
<td>2.20</td>
<td>2.56</td>
<td>4.00</td>
<td>4.75</td>
</tr>
<tr>
<td>West</td>
<td>5.70</td>
<td>6.15</td>
<td>8.50</td>
<td>11.50</td>
</tr>
<tr>
<td>South</td>
<td>2.60</td>
<td>2.91</td>
<td>4.75</td>
<td></td>
</tr>
<tr>
<td>Southwest</td>
<td>0.30</td>
<td>0.32</td>
<td>0.50</td>
<td>0.75</td>
</tr>
<tr>
<td>Netherlands</td>
<td>12.10</td>
<td>13.40</td>
<td>20.00</td>
<td>20.00</td>
</tr>
</tbody>
</table>

The most striking feature of the second Report was the target population of three million in the North. Drewe (1971; p. 150) computed that this would mean an additional net in-migration into the North of 60,000 people every five years starting in 1965. The realized net migration between 1969 and 1974 was only 13,600 people (Drewe, 1975; p. 1073). The government was, therefore, forced in the third Report to call these targets of population distribution "too ambitious," and to cancel its commitment to 3 million people in the North (Netherlands, 1975; p. 4). The objectives now are more realistic in terms of reducing out-migration and reversing the current migration trend.
in Stockholm and two other metropolitan areas in the South. In Finland and Ireland, people have poured into Helsinki and Dublin, respectively, or have left the country entirely. In Poland, the situation is different. The existence of strong provincial capitals has been assuring nearly constant regional population shares (Korcelli, 1975; p. 60). The Warsaw agglomeration accounts only for 5 percent of the national population. The settlement policy, stipulated in the National Plan of Physical Development (1974-1990), provides for a relatively faster growth of middle-sized cities in North and Central Poland.

b. Settlement policy objectives in the United States.

In few countries is the debate on settlement policies as documented and as open as in the United States. Morrison (1972b; p. 22) sees three functional transformations underway in the nation:

1. Evolution toward a metropolitan society:
   concentration of people in a few large metropolitan regions.

2. Rearrangement of population and economic activities within metropolitan areas: suburbanization combined with racial segregation.

3. Obsolescence of many areas, both urban and rural.

These changes have given rise to national problems of efficiency, equity, environment and living conditions. A national settlement policy which would alleviate these problems is in formation.\(^6\)

\(^6\)For a summary of how the debate evolved, see Wingo (1972).
The national political debate on population distribution and urban growth started in 1967 when the Department of Agriculture wrote in *Communities of Tomorrow* that "... our metropolitan areas have more people and problems than they can cope with," (Wingo, 1972; p. 6). More specific was the Advisory Commission on Intergovernmental Relations (1968; pp. 129-130), which concluded after a study of urbanization and regional growth that there was an immediate need for a national policy to guide the location of people and economic activity. This report initiated a wide range of studies and recommendations by national task forces, professional societies and others. An assessment of these studies is beyond the scope of this work. However, summary articles of the various positions have been collected and published by the Urban Land Institute (Scott, Brewer and Miner, 1975).

An important step in the direction of a national settlement policy was the Housing and Urban Development Act of 1970. By Title VII of this Act, every two years, the President, using the capacity of the Domestic Council to gather and analyze information, must submit to Congress a Report on Urban Growth. The purpose behind this legislation was to require the President to assume responsibility.

---


for defining goals, setting policy instruments, and coordinating the implementation of a national urban policy. The topics to be treated by the Report, are well defined (Ashley, 1975; pp. 411-412). However, there is still a long way to go. The 1972 Report states explicitly that it "... makes no claim to present a single, comprehensive national growth policy for the United States," (Ashley, 1975; p. 412). The 1974 Report, summarized by Selvaggi (1975) shows a slow evolution of willingness to move toward a growth policy. The Report provides a wide ranging, generalized discussion of growth issues. Such issues relate to problems of coastal areas, metropolitan growth, the balance between the individual right to find a home and a job and the rights of the community to determine local growth, and ways to increase rural growth. It is important to note that the 1974 Report defines growth in the context of economic growth and an increase in quality of life of all the people. Growth is not defined as territorial or physical change and development. Selvaggi (1975; p. 440) concludes:

This semantic issue is important because, in the future, national 'growth' policy-making is more likely to derive from efforts primarily to increase national economic growth and individual well-being, and only secondarily to induce changes in the location of people or jobs.

In contrast to the national debate, the policy debate on population distribution at the local level is spirited, particularly in destination areas of recent migration flows. "There is a new mood in America," writes the Task Force on Land Use and Urban Growth (1975; p. 85). Citizens all over the country, but mainly in fast growing suburban
Other local governments have also attempted to set a ceiling on the population level or growth rate. They include the San Francisco Association of Bay Area Governments; Orange County, Pacific Beach, and South San Francisco in California, the Denver Regional Council of Governments; Palm Beach County, Florida; and Carson City, Nevada (Carter et al., 1975; p. 350). The city of Los Angeles, now with 2.8 million people, has considered a new zoning plan that calls for no more than 5 million, instead of the 10 million allowed under the old plan (Forbes, 1975; p. 57). Several governments have restricted local growth by zoning, taxation, building permits and the like. Some are more implicit, although clear, in expressing their anti-growth feeling. Oregon's governor, for example, makes his point very clear by saying, "Come and visit us again, but for heaven's sake don't come here to live" (Finkler, 1975; p. 125).

1.2.2. Population Distribution Policy: Means

To achieve the objectives put forward in settlement policies, governments have come up with a wide variety of instruments. Initially fragmentary and uncoordinated, the instruments become combined in a program or a strategy as a country moves closer toward an explicit settlement policy.

In this section, we first review the general strategies to population distribution of various countries. Although the particular form of the strategy varies from country to country, the general characteristic is what Rodwin (1970; p. 4) calls concentrated decentralization. We then turn to the instruments developed to implement the settlement strategies.
a. Strategies of population distribution.

The optimal settlement system is seen as a hierarchy of large existing cities, middle-sized cities and small towns. To complete the hierarchy, three strategies may be followed: the creation of new cities, the diversion of growth to existing growth centers, and a combination of both.

To divert the growth from Paris, the French settlement policy of the Sixth Plan (1971-1975) calls for eight major cities, the so-called "métropoles d'équilibre" and several intermediate centers, the "villes moyennes" of 20,000 to 100,000 inhabitants (Rousselot, 1975; p. 180). Nine new towns are to be built, five of which are to be in the Paris region. The objectives of these towns in terms of population and industrial growth have been fixed in the Plan.

In the Netherlands (1966; p. 86), the settlement strategy adopted by the second Report on Physical Planning is labeled "concentrated deconcentration." In Poland, the optimal settlement system is known as "moderate poly-centric concentration" (Korcelli, 1975; p. 57). In Hungary, the existing cities and towns are, for policy purposes, classified as lower-grade, middle-grade and higher-grade centers (Koloszár, 1975; p. 69). The functions of each position in the hierarchy are well defined. The settlement policy in Australia is characterized by the commitment to free-standing new cities to fill in the hierarchy and to decentralize metropolitan growth (Logan and Wilmoth, 1975; p. 145). Australia closely follows the practice of Britain, where over 1.7 million people
live in thirty-one New Towns (Sharpe, 1975; p. 320).

In the USSR, the strategy adopted by the General Scheme of Population Distribution is the development of Group Settlement Systems (Kudinov, 1975; pp. 28-30). A Group Settlement System is a large regional cluster of cities at the center of which is a metropolis of one to three million people. The center is surrounded by a three-level hierarchy; cities of 500,000 to one million people, cities of 100,000 to 500,000 people and towns of 10,000 to 100,000 people. Each Group Settlement System has a radius of 100-150 km, and is organized around a single integrated industrial complex. The plan is to organize, within the next twenty years, twenty-seven Group Settlement Systems around the largest and most important cities in the country. Most of the Systems will be in the East. To complement existing centers, four hundred new towns will be built.

b. Instruments of population distribution.

New towns and growth centers do not automatically distribute the population in an optimal manner. The settlement strategy is a broad framework within which specific measures to implement the strategy are developed and evaluated. Policy instruments to guide the population distribution may be found in direct, semi-direct, and hidden policies.

b.1. Direct population distribution policies.

Direct policies are measures that affect the population directly. In the first instance, there are the traditional population policy instruments, designed to
affect fertility and mortality. Regional variation in legislation or in the effectiveness of birth control and disease control programs have a direct consequence for the population distribution. A discussion of the instruments of fertility and mortality policy is beyond the scope of this study, however.

The direct instruments of population distribution may be divided into two categories: government regulation of migration and taxation and subsidization of the migrant.

i. Government regulation.

Since the right-to-move clause is included in the constitution in several countries, governments are constrained in regulating internal migration by direct regulation. In such countries migration permits may not be used to control internal migration. However, at the international level, migration permits in the form of visas and work permits are being used to control migration. It is the major instrument used by European countries to keep guest workers out during periods of depression.

Direct government regulation of internal migration occurred in the United States during the Great Depression of the 1930's. Border guards in California turned away the dispossessed tenant farmers from Oklahoma and Arkansas, until this was declared to be unconstitutional (Forbes, 1975; p. 56).

The most widespread illustration of government regulation is land use legislation. It is one of the major instruments

---

that the states and local governments have to control their growth. It is expected that land use regulation will become a major instrument to implement a population distribution policy. The reason is that states are getting "the lion's share of the government authority to implement growth policies and are likely to play the pivotal role in effecting growth policies" (Patton and Patton, 1975; p. 320).

At the federal level, there is a consensus that the responsibility for a national growth policy should be decentralized to the states (Godschalk, 1975; p. 329). At the same time that states are acquiring a larger share of national responsibility, they are reassuming powers previously delegated to local governments. Hawaii, Vermont, Florida and Oregon are examples of states that have taken over certain local growth control powers, in order to assert a state-wide interest (Godschalk, 1975; pp. 332-337). The population growth policies in these states are based on land use control. Federal legislation may well stimulate this trend, which has been called "a quiet revolution in land use control" (Bosseman and Callies, 1971). A national land use bill, requiring the states to develop a land use plan and giving them some other tools to control their growth, is currently before Congress. A similar bill was defeated in 1974 by the House, after passage by the Senate (Kirk, 1975; pp. 419-420).

---

9A complete list of instruments available to the states to control their settlement system is provided in the Land Development Code of the American Law Institute, and is summarized by Dunham (1975). Local government techniques for managing population growth are surveyed by Einsweiler et al. (1975) and Carter et al. (1975).
ii. Taxes and subsidies.

Direct taxation of the migrant has been considered by California Tomorrow, a citizen's group (Forbes, 1975; p. 59). Its plan recommended that a $1,000 fee would be charged to each new resident from out-of-state, and collected as part of the state income tax process, to discourage excessive in-migration. However, taxation of migrants and negative measures in general are difficult to implement since they conflict with the right to move.

Of more importance as a population distribution measure is assistance to the migrant. Direct government aid to migrants is intended to reduce their costs of moving and to compensate them for external disadvantages. This policy has wide application in the migration of labor from depressed areas to areas of high employment. The objective is to reduce unemployment and to improve the functioning of the labor market. Efficiency and equity considerations, therefore, underlie this policy.

The characteristics of government assistance to migrants are well documented for several countries (O.E.C.D., 1967; Wander, 1971; Klaassen and Drewe, 1973, Chapter 2). We restrict ourselves to a consideration of their general features.

Government aid to migrants falls into three categories: financial assistance, information and training. Information and training are given at the pre-departure level. Most West European countries and Canada have an interregional
clearance system, where job vacancies and applications are reported to the local and regional employment offices which place the workers (Wander, 1971; p. 3032). In France, additional information on working and living conditions in the labor demand areas is given through lectures and films. All countries in Western Europe and North America have training facilities to increase the skill of the migrants. In Britain there is, in addition, the "Nucleus Labor Force Scheme," which assists unemployed workers who are temporarily transferred to their new employer's parent factory for training.

Financial assistance consists of travel assistance, starting allowances, separation allowances when the migrant has to leave his family behind, removal grants, and installation allowances (O.E.C.D., 1967; Wander, 1971, p. 3033; Hansen, 1973, pp. 30-62). These instruments to enhance labor mobility were also part of the Labor Mobility Demonstration Projects conducted in the United States by the Department of Labor from 1965 to 1969. During the entire program, 14,221 workers were relocated, many of them high school graduates. A detailed account of the program is given by Hansen (1973, Chapter 2).

In most countries with labor mobility programs, the unemployed only receive financial aid if they do not move to the big cities or to overurbanized areas. (An exception is Sweden, where migration to urban areas is also subsidized.) This additional element of selectivity in migration assistance is in line with the general strategy to limit the growth of big cities in favor of smaller growth centers.
b.2. Indirect population distribution policies.

Although direct assistance to the migrant may be effective, it usually involves only a fraction of the total number of migrants. Most people move without direct assistance. They react to opportunities provided in the destination area. The spatial distribution of economic opportunities may be influenced by government policy measures. Several governments have adopted the strategy of assisting the interregional mobility of capital instead of labor. The attempt to bring jobs to the people in depressed areas is beneficial not only to these areas by reducing out-migration, but also is beneficial to the old industrial centers and cities by alleviating their congestion.

In general, the implementation of such policies consists of defining development areas and designating new industrial centers or existing growth centers in these areas. Incentives are then provided to private industry to locate in these areas. The incentives generally take the form of investment grants, loans, tax exemptions or other privileges with respect to capital depreciation, transport charges, tariffs, and the like. These measures have been supplemented by disincentives to locate in metropolitan areas. In most countries of Western Europe the location of factories and offices is regulated by building permits. Such Industrial Development Certificates, as they are called in Britain, are often refused in overurbanized areas (Cameron, 1972; p. 707). In Paris there is a special tax on industrial and office space (Rousselot, 1975; p. 183). Recently, several governments have initiated an additional measure to promote a more balanced settlement system, namely the dispersion of
government offices away from the capital. How this works in Britain, Sweden and the Federal Republic of Germany is reviewed by Friedly (1974, Chapter 8). A complete inventory of instruments to guide the location of the industrial and the service sector in Britain, France, Italy, West Germany, Sweden, Spain, the Netherlands, the United States and Canada is given in Hansen (1974).

b.3. Hidden population distribution policies.

Hidden policies do not enter the population distribution policy making process. However, because of their powerful but undirected influence, Morrison (1975; p. 103) states that "Efforts to intervene in any nation's system of urban settlement must begin with an assessment of such hidden policies." Alonso (1972; p. 645) claims that direct policies to modify the settlement system have been generally ineffective and sometimes even counter-productive because policy makers have underestimated or misjudged the consequences of hidden or implicit policies.

The best known illustrations of hidden policies are the defense expenditures of the United States. They are concentrated in the richest and fastest growing counties. Therefore, defense spending is a "de facto" spatial growth policy (Morrison, 1972c; p. 318). Other hidden policies are the interstate highway program, federal welfare programs, federal income tax deductions, revenue sharing, environmental conservation programs, and agricultural policies.

An interesting study of hidden policies is that of Schon (1968). He appraises the effects of government policies on rural-urban migration and finds that food
stamps, the minimum wage law, and acreage allotments encourage migration, whereas rural electrification, unemployment insurance and farm credit discourage migration.
References


PART II

SPATIAL DYNAMICS OF STRUCTURAL
CHANGE IN DEMOGRAPHIC ANALYSIS

What is the impact on the multiregional population system of changes in its structural parameters? That is the fundamental question we ask and try to answer in this part. The long run effects of a population policy depend on its ability to change the structural parameters of the demographic system. It is logical, therefore, to try to get to know better the internal dynamics resulting from parameter changes before setting up alternative policy models.

The relevance of such a structural analysis or impact analysis in population dynamics is not restricted to the policy domain, however. It has meaning in any situation where parameter changes may occur or where parametric values may deviate from assigned values. The cause of the change or deviation does not interfere with the final result. It may be an intended policy action or an uncontrollable exogenous factor. For example, parameters may deviate from assigned values because of erroneous measurements or because of imperfections in the data.

The structural parameters which will be considered in this part are age-specific fertility rates, death rates and migration rates. In recent years, there has been a growing interest among demographers in impact analysis (e.g., Keyfitz, 1971, Goodman, 1969, 1971b, Preston, 1974). The general problem is to find how sensitive stationary population characteristics, population projections, and stable population characteristics are to changes in age-specific rates. The
sensitivity of the stable characteristics of population systems undisturbed by migration have received most attention. That most effort has been devoted to the stable population becomes clear if one recalls that the stable population concept was developed as a device which displays the implications for age composition, birth rates, death rates, and growth rates of specified schedules of fertility and mortality, on the assumption that the schedules prevail long enough for other influences to be erased. In actual fact, however, the stable population is never achieved, since the basic schedules change through time. The question of the impact of such changes on the stable population therefore is principally one of theoretical rather than empirical importance.

Two approaches to impact analysis may be distinguished. The first is the simulation approach, or the arithmetic approach as Keyfitz (1971; p. 275) calls it. It is simply the computation of the population projection under the old and the new rates. The difference between the two in the ultimate age distribution and other features gives the impact of changing the rates. Suitable tools for the simulation approach are provided by the model life tables and model stable populations such as those developed by Coale and Demeny (1966) for a single region demographic system and by Rogers (1975; Chapter 6) for a multiregional system. An illustration of this approach has been given by Rogers (1975; pp. 169-172) and Rogers and Willekens (1975; pp. 28-30). Besides its demanding character in terms of computer time, the approach tells us nothing about the complete set of parameters on which the changes in the final results depend. It will be found useful, however, for
verifying the results of the second approach, which is the analytical approach. This procedure derives a general formula for assessing the impact of a particular change in terms of well-known population variables. Such a formula will be designated as a sensitivity function. Partial differentiation will be seen to be the basic ingredient in the analysis of such functions.

In this chapter, impact analysis is performed using the analytical approach. It is assumed that all the functions are differentiable with respect to the variables in which the changes occur. Since multiregional demographic models are formulated in matrix terms, matrix differentiation techniques are applied. And because not much work has been done in the area of matrix calculus, the first section of the Appendix to this part reviews several relevant topics of such a calculus\(^1\).

In order to be able to study the sensitivity of the stable population characteristics, we need an additional piece of information. All stable population features may be expressed as functions of the stable population distribution, the growth ratio of the stable population, and the age-specific fertility, mortality and migration rates. Therefore, the prerequisite to impact analysis of the stable population is a knowledge of the sensitivity of the stable population distribution and the stable growth ratio to changes in the age-specific rates.

\(^1\)All major textbooks on matrix algebra lack a chapter on matrix calculus, although some scattered treatment may occur. The only unified treatment of matrix differentiation that we have found is by Dwyer and MacPhail (1948). A simplified and extended version appeared twenty years later in Dwyer (1967). The formulas given there are general enough to handle differentiation problems in life table functions and in the analysis of population projections over a finite time horizon.
Rogers (1975; p. 128) has shown that the stable growth ratio is the dominant eigenvalue of the growth matrix, and that the stable population distribution is the associated right eigenvector. The problem may, therefore, be reformulated as finding the sensitivity of the dominant eigenvalue and eigenvector to changes in the growth matrix, and the sensitivity of the elements of the growth matrix to changes in the age-specific rates that are used to define it.

The problem of eigenvalue and eigenvector sensitivity has received some attention in the engineering literature (e.g., Cruz, 1970; Part III). An overview of the major relevant results of this literature is given in the second section of the Appendix. It is worth noting at this point that the application of this technique in population dynamics is not restricted to the stable population. This technique is relevant in every situation where the eigenvalues of a particular matrix have some demographic meaning. For instance, Rogers and Willekens (1975; p. 39) state that the dominant eigenvalue of the net reproduction matrix of a multiregional population system represents the net reproduction rate of the whole system. Hence examining the impact on the net reproduction rate of the United States of a change in the net reproduction rate of rural-born women living in urban areas, is a problem of eigenvalue sensitivity analysis.
CHAPTER 2

IMPACT OF CHANGES IN AGE-SPECIFIC
RATES ON LIFE TABLE FUNCTIONS

The concept of a multiregional life table as developed by Rogers (1973 and 1975, Chapter 3) is a device for exhibiting the mortality and migration history of a set of regional cohorts as they age. It is assumed that the age-specific rates describing the mortality and mobility experience of an actual population remain constant, and that the system of regions is undisturbed by external migration.

The first part of this chapter sets out the life table functions. The cohorts we will consider are birth cohorts or radices. Their life history is of special interest because they provide the information required by population projection models. The life table statistics are given by place of birth. In the second part, we combine the life table functions with the matrix differentiation techniques described in the Appendix. This enables us to develop life table sensitivity functions.

2.1. THE MULTIREGIONAL LIFE TABLE

All the life table functions are derived from a set of age-specific death and out-migration rates. Let \( \mathbf{M}(x) \) denote the matrix of observed annual rates for the persons in the age interval from \( x \) to \( x + h \). The length of the interval \( h \) is arbitrary. Without loss of generality, we will consider age intervals of five years. For a \( N \)-region system, \( \mathbf{M}(x) \) is
\[
\tilde{M}(x) = \begin{bmatrix}
\left( M_{1\delta}(x) + \sum_{j \neq 1}^N M_{1j}(x) \right) & -M_{21}(x) & -M_{31}(x) & \cdots \\
-M_{12}(x) & \left( M_{2\delta}(x) + \sum_{j \neq 2}^N M_{2j}(x) \right) & -M_{32}(x) & \cdots \\
-M_{13}(x) & -M_{23}(x) & \left( M_{3\delta}(x) + \sum_{j \neq 3}^N M_{3j}(x) \right) & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{bmatrix}
\]

where \( M_{i\delta}(x) \) is the age-specific annual death rate in region \( i \), and \( M_{ij}(x) \) is the age-specific annual out-migration rate from region \( i \) to region \( j \). It is estimated by the annual number of out-migrants to \( j \) divided by the mid-year population of \( i \).

Let \( \tilde{P}(x) \) be the matrix of age-specific probabilities of dying and out-migrating:

\[
\tilde{P}(x) = \begin{bmatrix}
P_{11}(x) & p_{21}(x) & \cdots & \cdots \\
\vdots & \vdots & \ddots & \ddots \\
\vdots & \vdots & \ddots & \ddots \\
\vdots & \vdots & \ddots & \ddots 
\end{bmatrix}
\]
with $p_{ij}(x)$ being the probability that an individual in region $i$ at exact age $x$ will survive and be in region $j$ at exact age $x + 5$. The diagonal element $p_{ii}(x)$ is the probability that an individual will survive and be in region $i$ at the end of the interval. If $q_i(x)$ is the probability that an individual in region $i$ at age $x$ will die before reaching age $x + 5$, then the following relationship follows

$$p_{ii}(x) = 1 - q_i(x) - \sum_{j \neq i}^N p_{ij}(x) .$$  \hspace{1cm} (2.3)

If multiple transition between two states is allowed during a unit time interval, then $p(x)$ is given by (Schoen, 1975; Rogers and Ledent, 1976):

$$p(x) = \left[ I + \frac{5}{2} M(x) \right]^{-1} \left[ I - \frac{5}{2} M(x) \right] .$$  \hspace{1cm} (2.4)

The probability that an individual starting out in region $j$, i.e., born in $j$, will be in region $i$ at exact age $x$ is denoted by $\hat{j}_{i}(x)$. The matrix containing those probabilities is

$$\hat{j}(x) = \begin{bmatrix}
\hat{j}_{11}(x) & \hat{j}_{12}(x) & \cdots & \cdots \\
\hat{j}_{21}(x) & \hat{j}_{22}(x) & \cdots & \cdots \\
\vdots & \vdots & \ddots & \ddots \\
\end{bmatrix} .$$  \hspace{1cm} (2.5)
By this definition we have that

\[ \hat{\ell}(x) = \hat{\ell}(x - 5) \hat{\ell}(x - 10) \cdots \hat{\ell}(0) = \hat{\ell}(x - 5) \hat{\ell}(x - 5). \]  
(2.6)

Define

\[ \ell(x) = \hat{\ell}(x) \ell(0) \]  
(2.7)

where \( \ell(0) \) is a diagonal matrix of the cohorts of babies born in the \( N \) regions at a given instant in time. Typically, \( j_l^i(0) \) is called the radix of region \( i \) and is set equal to some arbitrary constant such as 100,000. Then \( \ell(x) \) is the matrix of the number of people at exact age \( x \) by place of residence and by place of birth.

Another life table function is the total number of people of age group \( x \), i.e., aged \( x \) to \( x + 5 \), in each region by place of birth:

\[
\ell(x) = \begin{bmatrix}
L_1^1(x) & L_1^2(x) & \cdots \\
L_2^1(x) & L_2^2(x) & \cdots \\
\vdots & \vdots & \ddots \\
\vdots & \vdots & \ddots \\
\vdots & \vdots & \ddots
\end{bmatrix}
\]  
(2.8)

with \( j_L^i(x) \) being the number of people in region \( i \) in age group \( x \) who were born in region \( j \). The element \( j_L^i(x) \) can also be thought of as the total person-years lived in region \( i \) between ages \( x \) and \( x + 5 \), by the people born in
region \( j \). The matrix \( L(x) \) is given by

\[
L(x) = \int_0^5 \mathcal{L}(x + t) \, dt = \left[ \int_0^5 \mathcal{L}(x + 5) \, dt \right] \mathcal{L}(0) .
\]  

(2.9)

Assuming a uniform distribution of out-migrations and deaths over the 5-year age interval, we may obtain numerical values for \( L(x) \) by the linear interpolation

\[
L(x) = \frac{5}{2} \left[ \mathcal{L}(x) + \mathcal{L}(x + 5) \right] ,
\]

or

\[
L(x) = \frac{5}{2} \left[ I + P(x) \right] \mathcal{L}(x) .
\]

(2.10)

Aggregating \( L(x) \) over various age groups, we define the expected total number of person-years remaining to the people at exact age \( x \), as

\[
\mathcal{T}(x) = \sum_{y=x}^{z} L(y)
\]

(2.11)

where \( z \) is the terminal age group. Expressing \( \mathcal{T}(x) \) per individual, we get the matrix of expectations of life of an individual at exact age \( x \):

\[
\mathcal{E}(x) = \mathcal{T}(x) \mathcal{L}^{-1}(x) = \left[ \sum_{y=x}^{z} L(y) \right] \mathcal{L}^{-1}(x) .
\]

(2.12)
A very useful life table function is the survivorship matrix. It is an essential component of the population projection matrix. Rogers (1975; p. 79) has shown that the survivorship matrix

\[
S(x) = \begin{bmatrix}
  s_{11}(x) & s_{21}(x) & \cdots \\
  s_{12}(x) & s_{22}(x) & \cdots \\
  \vdots & \vdots & \ddots \\
\end{bmatrix}
\]

(2.13)

is given by

\[
S(x) = L(x + 5) L^{-1}(x).
\]

(2.14)

The element \( s_{ij}(x) \) denotes the proportion of individuals aged \( x \) to \( x + 4 \) in region \( i \), who survive to be \( x + 5 \) to \( x + 9 \) years old five years later, and are then in region \( j \).

We now have set up the important life table functions, and can proceed to the analysis of their sensitivities to changes in the underlying rates, i.e., in \( M(x) \).

2.2 SENSITIVITY ANALYSIS OF LIFE TABLE FUNCTIONS

The fundamental question posed in this section is: what is the effect on the various life table statistics of a change in the observed age-specific rates? To resolve this question, the life table functions are combined with the matrix differentiation techniques of the appendix.

This section is divided into five parts. Each part starts out with a specific life table function.
The derivative of this function with respect to an element of the matrix of age-specific rates yields the corresponding sensitivity function.

a. Sensitivity of the probabilities of dying and out-migrating

Recall the estimating formula set out in (2.4):

$$\hat{P}(x) = [I + \frac{5}{2} M(x)]^{-1} [I - \frac{5}{2} M(x)] .$$  \hfill (2.4)

In it \(\hat{P}(x)\) only depends on \(M(x)\). Therefore, \(\hat{P}(a)\) is not affected by a change in \(M(x)\) for \(a \neq x\).

The derivative of \(\hat{P}(x)\) with respect to an arbitrary element of \(M(x)\) is, by formulas (A.13) and (A.25) of the Appendix,

$$\frac{\delta \hat{P}(x)}{\delta \langle M(x) \rangle} = \frac{\delta [I + \frac{5}{2} M(x)]^{-1}}{\delta \langle M(x) \rangle} \frac{[I - \frac{5}{2} M(x)]}{\delta \langle M(x) \rangle}$$

$$+ [I + \frac{5}{2} M(x)]^{-1} \frac{\delta [I - \frac{5}{2} M(x)]}{\delta \langle M(x) \rangle}$$

$$= - [I + \frac{5}{2} M(x)]^{-1} \frac{\delta [I + \frac{5}{2} M(x)]}{\delta \langle M(x) \rangle} [I + \frac{5}{2} M(x)]^{-1}$$

$$[I - \frac{5}{2} M(x)] + [I + \frac{5}{2} M(x)]^{-1} \frac{\delta [I - \frac{5}{2} M(x)]}{\delta \langle M(x) \rangle}$$

$$= - [I + \frac{5}{2} M(x)]^{-1} \left[ \frac{5}{2} J \hat{P}(x) + \frac{5}{2} J \right]$$

where \(J\) is a matrix of the dimension of \(M(x)\) with all elements zero except for a one on the position of the arbitrary element \(\langle M(x) \rangle\). (This notation is further explained in the Appendix.)
The sensitivity function for $\tilde{\pi}(x)$ therefore is

$$\frac{\delta \tilde{\pi}(x)}{\delta < M(x) >} = - \frac{5}{2} \left[ I + \frac{5}{2} M(x) \right]^{-1} J \left[ \frac{\tilde{\pi}(x) + I}{\tilde{\pi}(x)} \right] . \tag{2.15}$$

After the transformation

$$\left[ \frac{\tilde{\pi}(x) + I}{\tilde{\pi}(x)} \right] = \left[ I + \frac{5}{2} M(x) \right]^{-1} \left[ \left[ I - \frac{5}{2} M(x) \right] + \left[ I + \frac{5}{2} M(x) \right] \right]$$

$$= \left[ I + \frac{5}{2} M(x) \right]^{-1} \left[ I + I \right] = 2 \left[ I + \frac{5}{2} M(x) \right]^{-1}$$

the sensitivity function becomes

$$\frac{\delta \tilde{\pi}(x)}{\delta < M(x) >} = - 5 \left[ I + \frac{5}{2} M(x) \right]^{-1} J \left[ I + \frac{5}{2} M(x) \right]^{-1} . \tag{2.16}$$

b. Sensitivity of the number of people at exact age $a$

A change in $M(x)$ does not affect $\tilde{\lambda}(a)$ for $a \leq x$. Therefore we look only at the case $a > x$. Note that $\tilde{\lambda}(a)$ may be written as

$$\tilde{\lambda}(a) = \tilde{\pi} \left( a - 5 \right) \tilde{\pi} \left( a - 10 \right) \cdots \tilde{\pi}(x) \tilde{\lambda}(x) .$$

Recalling that $M(x)$ only affects $\tilde{\pi}(x)$, we write

$$\frac{\delta \tilde{\lambda}(a)}{\delta < M(x) >} = \tilde{\pi} \left( a - 5 \right) \tilde{\pi} \left( a - 10 \right) \cdots \frac{\delta \tilde{\pi}(x)}{\delta < M(x) >} \tilde{\lambda}(x)$$

$$= - \frac{5}{2} \left[ \tilde{\pi} \left( a - 5 \right) \tilde{\pi} \left( a - 10 \right) \cdots \tilde{\pi}(x + 5) \right] \left[ I + \frac{5}{2} M(x) \right]^{-1} J \left[ \frac{\tilde{\pi}(x) + I}{\tilde{\pi}(x)} \right] \tilde{\lambda}(x) . \tag{2.17}$$
Inserting
\[ I = \left[ I - \frac{5}{2} M(x) \right] \xi(x) \xi^{-1}(x) \left[ I - \frac{5}{2} M(x) \right]^{-1} \]
in (2.17) and substituting for \( \tilde{P}(x) \) yields
\[
\frac{\delta \xi(a)}{\delta \langle M(x) \rangle} = -\frac{5}{2} \left[ P(a - 5) \tilde{P}(a - 10) \cdots P(a + 5) \tilde{P}(x) \xi(x) \xi^{-1}(x) \right]
\left[ I - \frac{5}{2} M(x) \right]^{-1} \xi \left[ P(x) + I \right] \xi(x)
\]
\[
\frac{\delta \xi(a)}{\delta \langle \tilde{M}(x) \rangle} = -\frac{5}{2} \xi(a) \xi^{-1}(x) \left[ I - \frac{5}{2} M(x) \right]^{-1} \xi \left[ P(x) + I \right] \xi(x)
\]
(2.18)
or by (2.10)
\[
\frac{\delta \xi(a)}{\delta \langle \tilde{M}(x) \rangle} = -\xi(a) \xi^{-1}(x) \left[ I - \frac{5}{2} M(x) \right]^{-1} JL(x)
\]
(2.19)

For \( a = x + 5 \), we have
\[
\frac{\delta \xi(x + 5)}{\delta \langle \tilde{M}(x) \rangle} = -P(x) \left[ I - \frac{5}{2} M(x) \right]^{-1} JL(x) = -\left[ I + \frac{5}{2} M(x) \right]^{-1} JL(x)
\]
(2.20)

An interesting formulation of the sensitivity function follows from writing (2.18) as
\[
\frac{\delta \xi(a)}{\delta \langle \tilde{M}(x) \rangle} = \xi(a) \left[ P(x) \xi(x) \right]^{-1} \frac{\delta P(x)}{\delta \langle \tilde{M}(x) \rangle} \xi(x)
\]

\[
= \xi(a) \xi^{-1}(x) \frac{\delta P(x)}{\delta \langle \tilde{M}(x) \rangle} \xi(x)
\]
or

\[
\tilde{\mathbf{L}}^{-1}(a) \frac{\delta \tilde{\mathbf{L}}(a)}{\delta \mathbf{M}(x)} = \tilde{\mathbf{L}}^{-1}(x) \frac{\delta \tilde{\mathbf{P}}(x)}{\delta \mathbf{M}(x)} \tilde{\mathbf{L}}(x) \quad (2.21)
\]

This shows that the relative sensitivity of \( \tilde{\mathbf{L}}(a) \) to changes in \( \mathbf{M}(x) \) is a weighted average of the relative sensitivity of \( \tilde{\mathbf{P}}(x) \), and is independent of \( a \). Consider the first age group and suppose that all regions have the same radices; i.e., \( \tilde{\mathbf{L}}(0) \) is a scalar matrix, i.e., a diagonal matrix with the same diagonal elements. The relative sensitivity of any \( \tilde{\mathbf{L}}(a) \) is then equal to the relative sensitivity of \( \tilde{\mathbf{P}}(0) \).

c. Sensitivity of the number of people in age group 

\( (a, a + 4) \)

What is the impact of a change in \( \mathbf{M}(x) \) on the number of people in age group \( (a, a + 4) \) and on their spatial distribution? It is clear that \( \mathbf{M}(x) \) does not affect \( \tilde{\mathbf{L}}(a) \) for \( a < x \). Therefore, we consider here the case of \( a \geq x \). Recall from (2.10) that

\[
\tilde{\mathbf{L}}(a) = \frac{5}{2} \left[ \tilde{\mathbf{L}}(a + 5) + \tilde{\mathbf{L}}(a) \right]
\]

Differentiating both sides gives

\[
\frac{\delta \tilde{\mathbf{L}}(a)}{\delta \mathbf{M}(x)} = \frac{5}{2} \left[ \frac{\delta \tilde{\mathbf{L}}(a + 5)}{\delta \mathbf{M}(x)} + \frac{\delta \tilde{\mathbf{L}}(a)}{\delta \mathbf{M}(x)} \right]
\]

If \( a = x \), then \( \frac{\delta \tilde{\mathbf{L}}(a)}{\delta \mathbf{M}(x)} = 0 \) and we have
\[
\frac{\delta L(x)}{\delta m(x)} = \frac{5}{2} \frac{\delta l(x + 5)}{\delta m(x)} = \frac{5}{2} \left[ I + \frac{5}{2} M(x) \right]^{-1} JL(x)
\]

(2.22)

which has the following alternative expressions:

\[
\frac{\delta L(x)}{\delta m(x)} = -5 \left[ P(x) + I \right] JL(x)
\]

(2.23)

\[
= -\frac{5}{2} P(x) \left[ I - \frac{5}{2} M(x) \right]^{-1} JL(x)
\]

(2.24)

\[
= -L(x) l^{-1}(x) \left[ I - \frac{5}{2} M(x) \right]^{-1} JL(x)
\]

\[
+ \frac{5}{2} \left[ I - \frac{5}{2} M(x) \right]^{-1} JL(x)
\]

(2.25)

If \( a > x \), we know that \( P(a) \) is independent of \( M(x) \), and therefore

\[
\frac{\delta L(a)}{\delta m(x)} = \frac{5}{2} \left[ P(a) + I \right] \frac{\delta l(a)}{\delta m(x)}
\]

\[
= -\frac{5}{2} \left[ P(a) + I \right] l(a) l^{-1}(x) \left[ I - \frac{5}{2} M(x) \right]^{-1} JL(x)
\]

\[
= -L(a) l^{-1}(x) \left[ I - \frac{5}{2} M(x) \right]^{-1} JL(x)
\]

(2.26)

which may also be written as

\[
\frac{\delta L(a)}{\delta m(x)} = \frac{5}{2} \left[ P(a) + I \right] l(a) l^{-1}(a) \frac{\delta l(a)}{\delta m(x)}
\]
whence, since $\frac{5}{2} \left[ P(a) + I \right] \tilde{\xi}(a) = \tilde{\xi}(a)$,

$$\begin{align*}
\tilde{L}^{-1}(a) \frac{\delta \tilde{L}(a)}{\delta \tilde{M}(x)} &= \tilde{L}^{-1}(a) \frac{\delta \tilde{\xi}(a)}{\delta \tilde{M}(x)}. 
\end{align*}$$

Equation (2.27) indicates that the relative sensitivity of the number of people in age group $(a, \ a + 4)$ is equal to the relative sensitivity of the number of people at exact age $a$ for $a > x$.

d. Sensitivity of the expectation of life at age $a$

We now proceed to deriving the sensitivity function of the most important life table statistic, namely the expectation of life. First consider the sensitivity of $e_x$. Differentiating both sides of (2.12) yields

$$\frac{\delta e_x}{\delta \tilde{M}(x)} = \delta \left[ \sum_{y=x}^{y=x+5} \tilde{L}(y) \right] \tilde{L}^{-1}(x) + \left[ \sum_{y=x}^{y=x+5} \tilde{L}(y) \right] \frac{\delta \tilde{L}^{-1}(x)}{\delta \tilde{M}(x)}. $$

$$\begin{align*}
(2.28)
\end{align*}$$

From (2.22) and (2.26), we see that

$$\begin{align*}
\delta \left[ \sum_{y=x}^{y=x+5} \tilde{L}(y) \right] = & - \left[ \sum_{y=x+5}^{y=x+5} \tilde{L}(y) \right] \tilde{L}^{-1}(x) \left[ I - \frac{5}{2} \tilde{M}(x) \right]^{-1} \tilde{JL}(x) \\
& - \frac{5}{2} \left[ I + \frac{5}{2} \tilde{M}(x) \right]^{-1} \tilde{JL}(x) \\
& = - \left[ \sum_{y=x+5}^{y=x+5} \tilde{L}(y) \tilde{L}^{-1}(x) + \frac{5}{2} P(x) + \frac{5}{2} I - \frac{5}{2} I \right] \\
& \cdot \left[ I - \frac{5}{2} \tilde{M}(x) \right]^{-1} \tilde{JL}(x) 
\end{align*}$$
\[
= - \left[ \sum_{y=x+5}^\infty L(y) \mathcal{L}^{-1}(x) + L(x) \mathcal{L}^{-1}(x) - \frac{5}{2} \mathcal{I} \right] \\
\cdot \left[ \mathcal{I} - \frac{5}{2} \mathcal{M}(x) \right]^{-1} \mathcal{JL}(x) \\
= - \left[ e(x) - \frac{5}{2} \mathcal{I} \right] \left[ \mathcal{I} - \frac{5}{2} \mathcal{M}(x) \right]^{-1} \mathcal{JL}(x) .
\]

Since \( \mathcal{L}(x) \) is independent of \( \mathcal{M}(x) \), we may write (2.28) as follows

\[
\frac{\delta e(x)}{\delta \mathcal{M}(x)} = - \left[ e(x) - \frac{5}{2} \mathcal{I} \right] \left[ \mathcal{I} - \frac{5}{2} \mathcal{M}(x) \right]^{-1} \mathcal{JL}(x) \mathcal{L}^{-1}(x) .
\]

(2.29)

For \( a < x \), we have

\[
\frac{\delta e(a)}{\delta \mathcal{M}(x)} = - \left[ \sum_{y=x+5}^\infty L(y) + L(x) + \sum_{y=a}^x L(y) \right] \mathcal{L}^{-1}(a) .
\]

We know that

\[
\frac{\delta L(y)}{\delta \mathcal{M}(x)} = 0 , \quad \text{for } y < x
\]

and

\[
\frac{\delta \mathcal{L}(a)}{\delta \mathcal{M}(x)} = 0 , \quad \text{for } a < x
\]
Therefore

\[
\frac{\delta e(a)}{\delta <M(x)>} = \frac{\delta \left[ \sum_{y=x}^{y=x+5} L(y) \right] \Downarrow^{-1}(a)}{\delta <M(x)>}
\]

\[
\frac{\delta e(a)}{\delta <M(x)>} = \frac{\delta \left[ \sum_{y=x}^{y=x+5} L(y) \right] \Downarrow^{-1}(x) \Downarrow(x) \Downarrow^{-1}(a)}{\delta <M(x)>}
\]

\[
\frac{\delta e(a)}{\delta <M(x)>} = \frac{\delta e(a)}{\delta <M(x)>} \Downarrow(x) \Downarrow^{-1}(a) \quad (2.30)
\]

\[
\frac{\delta e(a)}{\delta <M(x)>} = - [e(x) - \frac{5}{2} \mathbb{I}] [\mathbb{I} - \frac{5}{2} M(x)]^{-1} \Downarrow L(x) \Downarrow^{-1}(a)
\]

\[
= - e(x) [\mathbb{I} - \frac{5}{2} M(x)]^{-1} \Downarrow L(x) \Downarrow^{-1}(a)
\]

\[
+ \frac{5}{2} [\mathbb{I} - \frac{5}{2} M(x)]^{-1} \Downarrow L(x) \Downarrow^{-1}(a) \quad (2.32)
\]

The second component of the sensitivity function is due to the linear approximation \( \Downarrow(x) = \frac{5}{2} [\Downarrow(x + 5) + \Downarrow(x)] \) of the continuous relationship

\[
\Downarrow(x) = \int_{x}^{x+5} \Downarrow(t) \, dt
\]
Consider the continuous definition of \( \xi(a) \)

\[
\xi(a) = \left[ \int_a^\omega \xi(t) \, dt \right] \xi^{-1}(a)
\]

where \( \omega \) is the terminal age. Differentiating yields

\[
\frac{\delta \xi(a)}{\delta \xi(M(x))} = \left[ \int_a^\omega \frac{\delta \xi(t)}{\delta \xi(M(x))} \, dt \right] \xi^{-1}(a), \quad \text{for } a \leq x
\]

\[
= \left[ \int_x^\omega - \xi(t) \xi^{-1}(x) \left[ I - \frac{5}{2} \xi(M(x))^{-1} \xi\xi(x) \right] \, dt \right] \xi^{-1}(a)
\]

Since \( \xi(t) \) is independent of \( \xi(M(x)) \), if \( t < x \)

\[
\frac{\delta \xi(a)}{\delta \xi(M(x))} = - \left[ \int_x^\omega \xi(t) \, dt \right] \xi^{-1}(x) \left[ I - \frac{5}{2} \xi(M(x))^{-1} \xi\xi(x) \right] \xi^{-1}(a)
\]

\[
= - \xi(x) \left[ I - \frac{5}{2} \xi(M(x))^{-1} \xi\xi(x) \right] \xi^{-1}(a) \quad (2.33)
\]

which is equivalent to the first term of (2.32) with the term \( \xi(x) \) replaced by \( \xi(x) \) in the discrete case. The expression (2.33), written in terms of differentials, is similar to the sensitivity function of the expectation of life, given by Keyfitz (1971, p. 276) for the single-region case

\[
de(a) = - e(x) [dM(x)] \xi(x) \xi^{-1}(a),
\]

where \( e(\cdot) \), \( \xi(\cdot) \) and \( M(\cdot) \) are scalars.
The term \([I - \frac{5}{2} \tilde{M}(x)]^{-1}\) in (2.33) is due to the fact that we consider observed rates where Keyfitz derived the formula using instantaneous rates. If \(\tilde{M}(x)\) contained instantaneous rates, then \(\tilde{M}(x) \equiv 0\) and \([I - \frac{5}{2} \tilde{M}(x)] \equiv I\).

e. Sensitivity of the survivorship proportions

As in the proceeding sections, we treat separately \(\tilde{S}(a)\) for \(a = x\) and for \(a > x\). The survivorship matrix is given by (2.14) as

\[
\tilde{S}(x) = \tilde{L}(x + 5) \tilde{L}^{-1}(x),
\]

(a).

which may be reexpressed as

\[
\tilde{S}(x) = [\tilde{P}(x + 5) + I] \tilde{P}(x) \tilde{L}(x) \tilde{L}^{-1}(x) [\tilde{P}(x) + I]^{-1}
\]

\[
= [\tilde{P}(x + 5) + I] \tilde{P}(x) [\tilde{P}(x) + I]^{-1}.
\]

(2.34)

Differentiating with respect to \(\tilde{M}(x)\) yields

\[
\frac{\delta \tilde{S}(x)}{\delta \tilde{M}(x)} = [\tilde{P}(x + 5) + I] \frac{\delta \tilde{P}(x)}{\delta \tilde{M}(x)} [\tilde{P}(x) + I]^{-1}
\]

\[
+ [\tilde{P}(x + 5) + I] \tilde{P}(x) \frac{\delta [\tilde{P}(x) + I]^{-1}}{\delta \tilde{M}(x)}
\]

\[
= [\tilde{P}(x + 5) + I] \left[ \frac{\delta \tilde{P}(x)}{\delta \tilde{M}(x)} - \tilde{P}(x) [\tilde{P}(x) + I]^{-1} \frac{\delta \tilde{P}(x)}{\delta \tilde{M}(x)} \right] \cdot [\tilde{P}(x) + I]^{-1}
\]
\[
\frac{\delta \mathcal{S}(x)}{\delta \mathcal{M}(x)} = \frac{5}{2} \left[ \mathcal{S}(x) - \mathcal{P}(x + 5) + I \right] \left[ I + \frac{5}{2} \mathcal{M}(x) \right]^{-1} J
\]

\[
= \frac{5}{2} \mathcal{S}(x) \left[ I + \frac{5}{2} \mathcal{M}(x) \right]^{-1} J
\]

\[
- \frac{5}{2} \left[ \mathcal{P}(x + 5) + I \right] \left[ I + \frac{5}{2} \mathcal{M}(x) \right]^{-1} J
\]

\[
= \frac{5}{2} \mathcal{S}(x) \left[ I + \frac{5}{2} \mathcal{M}(x) \right]^{-1} J
\]

Substituting for \( S(x) \) gives

\[
\frac{\delta \mathcal{S}(x)}{\delta \mathcal{M}(x)} = \frac{5}{2} \left[ \mathcal{S}(x) - \mathcal{P}(x + 5) + I \right] \left[ I + \frac{5}{2} \mathcal{M}(x) \right]^{-1} J
\]

\[
= \frac{5}{2} \mathcal{S}(x) \left[ I + \frac{5}{2} \mathcal{M}(x) \right]^{-1} J
\]

\[
- \frac{5}{2} \left[ \mathcal{P}(x + 5) + I \right] \left[ I + \frac{5}{2} \mathcal{M}(x) \right]^{-1} J
\]

\[
= \frac{5}{2} \mathcal{S}(x) \left[ I + \frac{5}{2} \mathcal{M}(x) \right]^{-1} J
\]

\[
- \frac{5}{2} \left[ \mathcal{P}(x + 5) + I \right] \mathcal{P}(x) \left[ I + \frac{5}{2} \mathcal{M}(x) \right]^{-1} J
\]
Since
\[
\frac{5}{2} \left[ \mathbf{P}(x + 5) + \mathbf{I} \right] \mathbf{P}(x) \mathbf{P}(x) = \mathbf{L}(x + 5)
\]

and
\[
\mathbf{L}^{-1}(x) \mathbf{P}^{-1}(x) \left[ \mathbf{I} + \frac{5}{2} \mathbf{M}(x) \right]^{-1} = \mathbf{L}^{-1}(x) \left[ \mathbf{I} - \frac{5}{2} \mathbf{M}(x) \right]^{-1}
\]

where \( \mathbf{L}^{-1}(x) \) may be written as
\[
\mathbf{L}^{-1}(x) = \left[ \frac{5}{2} \right] \mathbf{L}^{-1}(x) \left[ \mathbf{I} + \mathbf{P}(x) \right]^{-1} \left[ \mathbf{I} + \frac{5}{2} \mathbf{M}(x) \right]
\]

we have that
\[
\frac{\delta \mathbf{S}(x)}{\delta \left< \mathbf{M}(x) \right>} = \frac{5}{2} \mathbf{S}(x) \left[ \mathbf{I} + \frac{5}{2} \mathbf{M}(x) \right]^{-1} \mathbf{J}
\]

\[
= \mathbf{L}(x + 5) \mathbf{L}^{-1}(x) \left[ \mathbf{I} + \frac{5}{2} \mathbf{M}(x) \right]^{-1} \mathbf{J} + \frac{5}{2} \mathbf{S}(x) \left[ \mathbf{I} - \frac{5}{2} \mathbf{M}(x) \right]^{-1} \mathbf{J}
\]

\[
= \frac{5}{2} \mathbf{S}(x) \left[ \mathbf{I} + \frac{5}{2} \mathbf{M}(x) \right]^{-1} \mathbf{J} - \frac{5}{2} \mathbf{S}(x) \left[ \mathbf{I} - \frac{5}{2} \mathbf{M}(x) \right]^{-1} \mathbf{J}
\]

\[
- \frac{5}{2} \mathbf{S}(x) \mathbf{P}(x) \left[ \mathbf{I} - \frac{5}{2} \mathbf{M}(x) \right]^{-1} \mathbf{J}
\]

But
\[
\mathbf{P}(x) \left[ \mathbf{I} - \frac{5}{2} \mathbf{M}(x) \right]^{-1} = \left[ \mathbf{I} + \frac{5}{2} \mathbf{M}(x) \right]^{-1}
\]
Therefore

\[ \frac{\delta S(x)}{\delta \langle M(x) \rangle} = - \frac{5}{2} \frac{S(x)}{L} \left[ I - \frac{5}{2} \frac{M(x)}{L} \right]^{-1} J \quad \text{(2.35)} \]

To illustrate the dynamic relationship between the life table statistics, we may express the sensitivity of \( S(x) \) in relation to the sensitivity of other statistics. For example, a combination of (2.35) with (2.26) yields

\[ S^{-1}(x) \frac{\delta S(x)}{\delta \langle M(x) \rangle} = P^{-1}(x) \frac{\delta L(x)}{\delta \langle M(x) \rangle} L^{-1}(x) \]

and a combination of (2.35) with (2.19) gives

\[ S^{-1}(x) \frac{\delta S(x)}{\delta \langle M(x) \rangle} = - \frac{5}{2} P^{-1}(x) \frac{\delta L(x + 5)}{\delta \langle M(x) \rangle} L^{-1}(x) \]

The relative sensitivity of \( S(x) \) may be regarded as a weighted measure of the sensitivities of other statistics.

We now turn to the sensitivity of \( S(a) \) to changes in \( M(x) \) for \( a \neq x \). For \( a > x \) and for \( a < x - 5 \), \( S(a) \) is independent of a change in \( M(x) \). This can easily be seen in equation (2.34) while noting that \( P(a) \) is not affected by \( M(x) \) if \( a \neq x \). The sensitivity of \( S(x - 5) \) to a change in \( M(x) \) is derived next. We begin by writing (2.34) for \( x - 5 \).
\[ S(x - 5) = [P(x) + I] P(x - 5) [P(x - 5) + I]^{-1} \]

\[ \frac{\delta S(x - 5)}{\delta <M(x)>} = \frac{\delta [P(x) + I]}{\delta <M(x)>} P(x - 5) [P(x - 5) + I]^{-1} \]

\[ = - \frac{5}{2} [I + \frac{5}{2} M(x)]^{-1} J [P(x) + I] P(x - 5) [P(x - 5) + I]^{-1} \]

\[ \frac{\delta S(x - 5)}{\delta <M(x)>} = - \frac{5}{2} P(x) [I - \frac{5}{2} M(x)]^{-1} J S(x - 5) \]  \hspace{1cm} (2.36)

\[ = - \frac{5}{2} P(x) [I - \frac{5}{2} M(x)]^{-1} J S(x - 5) \]  \hspace{1cm} (2.37)

The relationship between the sensitivity of \( S(x) \) and of \( S(x - 5) \) is

\[ \frac{\delta S(x)}{\delta <M(x)>} = - \frac{5}{2} S(x) P^{-1}(x) [I + \frac{5}{2} M(x)]^{-1} J S(x - 5) S^{-1}(x - 5) \]

\[ \frac{\delta S(x)}{\delta <M(x)>} = S(x) P^{-1}(x) \frac{\delta S(x - 5)}{\delta <M(x)>} S^{-1}(x - 5) \]  \hspace{1cm} (2.38)

and

\[ \frac{\delta S(x - 5)}{\delta <M(x)>} = P(x) S^{-1}(x) \frac{\delta S(x)}{\delta <M(x)>} S(x - 5) \]  \hspace{1cm} (2.39)
CHAPTER 3

IMPACT OF CHANGES IN AGE-SPECIFIC RATES ON THE POPULATION PROJECTION

Population projection is often carried out under the assumption that an observed population growth regime will remain constant. This implies that the observed age-specific rates will not change over the projection period. (This is a crude assumption and no demographer or planner considers it to be a realistic one. Nevertheless it produces a useful benchmark against which to compare other alternative projections.) In this section, we deal with the question of how sensitive population projections are to changes in age-specific rates. These variations may occur at any point in time. If they occur in the base year, they can be related to observation errors. The sensitivity functions we develop remain exactly the same, no matter what the causes of the variations are.

In the first part, the population growth model is set out as a system of first order linear homogenous difference equations with constant coefficients, as in Rogers (1975, Chapter 5). The second part studies the sensitivity of population growth to changes in observed age-specific rates.

3.1. THE DISCRETE MODEL OF MULTIREGIONAL DEMOGRAPHIC GROWTH

Population growth may be expressed in terms of the changing level of population or in terms of the variation of the number of births over time. In demography, it has been a custom to formulate the discrete model of population growth in terms of total population, while the continuous
version describes the birth trajectory (Keyfitz, 1968; Rogers, 1975). A secondary objective of this and the next chapter is to contribute to the reconciliation of both growth models. We will formulate population growth in the discrete time domain. However, several particularities of the continuous model have a discrete counterpart. In this section, it will be shown how the population growth path relates to the trajectory of births.

a. The population model

A multiregional growth process may be described as a matrix multiplication (Rogers, 1975; p. 123):

$$\{\tilde{K}(t+1)\} = g\{\tilde{K}(t)\}$$

(3.1)

where the vector $\{\tilde{K}(t)\}$ describes the regional age-specific population distribution at time $t$, with

$$\begin{bmatrix}
\tilde{K}(t)(0) \\
\tilde{K}(t)(5) \\
\vdots \\
\tilde{K}(t)(z)
\end{bmatrix}$$

and

$$\begin{bmatrix}
K(t)(x) \\
1 \\
K(t)(x) \\
2 \\
\vdots \\
N \\
K(t)(x)
\end{bmatrix}$$

(3.2)

$z$ being the terminal age interval and $N$ the number of regions.

Each element $K_i(t)(x)$ denotes the number of people in region $i$ at time $t$, $x$ to $x + 4$ years old. Note that $t + 1$ represents the next moment in time, i.e., 5 years later than $t$. We consider age-groups and time intervals of 5 years. The operator $g$ is the generalized Leslie matrix
\[
\mathbf{G} = \begin{bmatrix}
0 & 0 & \tilde{B}(\alpha - 5) & \cdots & \tilde{B}(\beta - 5) & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
S(0) & 0 & \tilde{S}(5) & \cdots & \cdots & \cdots & \cdots & \tilde{S}(\gamma - 5) \cdot 0 \\
\end{bmatrix}
\]

with \( \tilde{S}(x) \), the matrix of survivorship proportions, retaining the definition set out in the previous chapter. The first and last ages of childbearing may be denoted by \( \alpha \) and \( \beta \), respectively, and

\[
\tilde{B}(x) = \begin{bmatrix}
b_{11}(x) & b_{21}(x) & \cdots \\
b_{12}(x) & b_{22}(x) & \cdots \\
\vdots & \vdots & \ddots \\
\end{bmatrix}
\]

where an element \( b_{ij}(x) \) denotes the average number of babies born during the unit time interval in region \( i \) and alive in region \( j \) at the end of that interval, per individual living in region \( i \) at the beginning of the interval and \( x \) to \( x + 4 \) years old. The off-diagonal elements of \( \tilde{B}(x) \) are measures of the mobility of children 0 to 4 years old, who were born to a \( x \) year-old parent. It is clear that their mobility pattern is determined by the mobility pattern of the parents.

It can be shown that \( \tilde{B}(x) \) obeys the relationship (Rogers, 1975; pp, 120-121):
\[ B(x) = \frac{1}{2} \mathcal{L}(0) \mathcal{L}^{-1}(0) \left[ \mathcal{P}(x) + \mathcal{P}(x + 5) \mathcal{S}(x) \right] \]

whence

\[ B(x) = \frac{5}{4} \left[ \mathcal{P}(0) + \mathcal{P}(5) \right] \left[ \mathcal{P}(x) + \mathcal{P}(x + 5) \mathcal{S}(x) \right] \]

since

\[ \mathcal{L}(0) = \frac{5}{2} \left[ \mathcal{L}(5) + \mathcal{L}(0) \right] = \frac{5}{2} \left[ \mathcal{P}(0) + \mathcal{P}(5) \right] \mathcal{L}(0) \]

where \( \mathcal{L}(0) \), \( \mathcal{L}(0) \), \( \mathcal{P}(0) \) and \( \mathcal{S}(x) \) are defined in the previous chapter. Here \( \mathcal{P}(0) \) and \( \mathcal{S}(x) \) are given by the life table, and \( \mathcal{P}(x) \) is a diagonal matrix containing the annual regional birthrates of people aged \( x \) to \( x + 4 \). The number of births in year \( t \) from people aged \( x \) to \( x + 4 \) at \( t \) is \( \mathcal{P}(x) \{ K^{(t)}(x) \} \).

The number of births during a five year period starting at \( t \), from people aged \( x \) to \( x + 4 \) at \( t \), is

\[ \frac{5}{2} \left[ \mathcal{P}(x) \{ K^{(t)}(x) \} + \mathcal{P}(x + 5) \{ K^{(t+1)}(x + 5) \} \right] \]

\[ = \frac{5}{2} \left[ \mathcal{P}(x) + \mathcal{P}(x + 5) \mathcal{S}(x) \right] \{ K^{(t)}(x) \} . \]

Of these births, a proportion \( \mathcal{L}(0) [(5\mathcal{L}(0))^{-1}] \) will be surviving in the various regions at the end of the time interval. Because of the special structure of the generalized Leslie matrix, (3.1) may be written as two equation systems:

\[ \{ K^{(t+1)}(0) \} = \frac{\beta-5}{\alpha-5} \mathcal{B}(x) \{ K^{(t)}(x) \} \]

(3.5)
\[ K^{(t+1)}(x + 5) \] \[ = S(x) K^{(t)}(x) \] , \hspace{1cm} (3.6)

for \( 5 \leq x \leq z - 5 \) .

The vector \( \{ K^{(t)}(x) \} \) may be expressed in the form

\[ \{ K^{(t+x/5)}(x) \} = [S(x - 5) \ S(x - 10) \cdots S(5) \ S(0)] \{ K^{(t)}(0) \} \]

\[ = \Lambda(x) \{ K^{(t)}(0) \} \] , \hspace{1cm} \text{say} \hspace{1cm} (3.7)

where we define

\[ \Lambda(x) = \begin{cases} 
1 & \text{for } x = 0 \\
\frac{1}{2} & \text{for } x = 5, 10 \ldots z \\
0 & \text{otherwise} 
\end{cases} = L(x) L^{-1}(0) 
\]

with \( \prod_{y=x-5}^{y=x-5} S(y) = S(x - 5) S(x - 10) \cdots S(5) S(0) \).

The element \( a_{ij}(x) \) of \( \Lambda(x) \) is the proportion of individuals aged 0 to 4 years in region \( i \), who will survive to be \( x \) to \( x + 4 \) years old exactly \( x \) years later, and will at that time be in region \( j \).

b. The birth model

The growth path of the births may be easily derived from the growth path of the population. Recall (3.5), and substitute (3.4) for \( \hat{\beta}(x) \). Then

\[ K^{(t+1)}(0) = \sum_{a=5}^{b-5} \frac{5}{a} [I + \hat{P}(0)] [\hat{F}(x) + \hat{F}(x + 5) S(x)] \{ K^{(t)}(x) \} \]
\[ \begin{align*}
\frac{\beta - 5}{2} \sum_{\alpha = 5}^{\beta - 5} & \left[ \sum_{\alpha = 5}^{\beta - 5} \frac{5}{2} \left[ F(x) + F(x + 5) S(x) \right] \right] K(t)(x) \\
= & \frac{1}{2} \left[ I + P(0) \right] Q(t+1,t) \\
= & \frac{1}{2} \left[ I + P(0) \right] \{ Q(t+1,t) \} ,
\end{align*} \]

(3.8)

where the regional distribution of births during a five-year period starting at \( t \), is denoted by \( Q(t+1,t) \) and is defined as

\[ Q(t+1,t) = \sum_{\alpha = 5}^{\beta - 5} \frac{5}{2} \left[ F(x) + F(x + 5) S(x) \right] \{ K(t)(x) \} . \]

(3.9)

Note that

\[ \{ K(t+1)(0) \} = L(0) \{ Q(t+1,t) \} \]

(3.10)

and

\[ Q(t+1,t) = \{ Q(t+1)(0) \} \]

(3.11)

Substituting

\[ K(t)(x) = A(x) \left\{ \left( t - \frac{x}{5} \right) \right\} , \quad \text{for } t \geq \frac{x}{5}, \]

from in (3.8), we have

\[ \{ K(t+1)(0) \} = \sum_{\alpha = 5}^{\beta - 5} \frac{5}{2} \left[ I + P(0) \right] \left\{ \left( t - \frac{x}{5} \right) \right\} \]

(3.12)
for \( t \geq \frac{X}{5} \),

and, therefore, the growth path of the births may be related to the number of births that occurred some time ago. Substituting (3.10) into (3.12) gives:

\[
\{Q(t+1,t)\} = \sum_{\alpha=5}^{\beta-5} \frac{5}{4} \left[ F(x) + \frac{F(x+5)}{S(x)} S(x) \right] A(x) \left[ I + P(0) \right] \cdot \{ (t - \frac{X}{5}, t - \frac{X}{5} - 1) \},
\]

for \( t \geq \frac{X}{5} + 1 \),

\[
= \sum_{\alpha=5}^{\beta-5} D(x) \{ (t - \frac{X}{5}, t - \frac{X}{5} - 1) \}, \quad \text{say,} \quad (3.13)
\]

since

\[
\{Q(t+1,t)\} = 2[I + P(0)]^{-1} \{K(t+1)(0)\}
\]

and

\[
\{ (t - \frac{X}{5}) \} \left\{ K \right\} \left( 0 \right) = \frac{1}{2} \left[ I + P(0) \right] \{Q \} \left\{ (t - \frac{X}{5}, t - \frac{X}{5} - 1) \}.
\]

Formula (3.13) expresses the growth path of the births, occurring during the period \((t+1,t)\), five years say. The annual number of births is

\[
\{Q(t)\} = \sum_{\alpha=5}^{\beta-5} F(x) \{K(t)(x)\}
\]

\[
= \sum_{\alpha=5}^{\beta-5} F(x) A(x) \left\{ K \right\} \left( 0 \right).
\]
Assuming stationarity, we may express the number of people in the first age group as a function of the births, as in Equation (2.10)

\[ \{K(t)(0)\} = \frac{5}{2} \left[ I + \frac{P(0)}{P} \right] \{Q(t)\} \quad . \quad (3.15) \]

We have that

\[ \{Q(t)\} = \sum_{a-5}^{b-5} \frac{5}{2} F(x) A(x) \left[ I + \frac{P(0)}{P} \right] \{Q_{a-5}(t-x)\} \quad , \quad (3.16) \]

for \( t \geq \frac{x}{5} \)

which is equal to

\[ \{Q(t)\} = \sum_{a-5}^{b-5} \frac{5}{2} F(x) L(x) \{Q_{a-5}(t-x)\} \quad , \quad (3.17) \]

in which we once again relate the number of births at time \( t \) to the number that occurred some time ago.

The relation between (3.17) and (3.13) is implicit in expression (3.15). Substituting (3.8) into (3.15) gives:

\[ \frac{1}{2} \left[ I + \frac{P(0)}{P} \right] \{Q(t,t-1)\} = \frac{5}{2} \left[ I + \frac{P(0)}{P} \right] \{Q(t)\} \]

or

\[ \{Q(t)\} = \frac{1}{5} \{Q(t,t-1)\} \quad . \quad (3.18) \]

This implies that the annual number of births is a simple average of the births during the previous period. Equation (3.17) is an \((b-5)\)-th order difference equation. To derive
a birth growth model analogue to (3.1), we replace (3.17) by a system of (β-5) first order difference equations:

\[
\begin{pmatrix}
\{\hat{\theta}(t)\} \\
\{\hat{\theta}(t-1)\} \\
\{\hat{\theta}(t-\frac{\alpha-5}{5})\} \\
\vdots \\
\{\hat{\theta}(t-\frac{\beta-10}{5})\}
\end{pmatrix}
= \begin{pmatrix}
0 & 0 & \cdots & F(\alpha-5) & L(\alpha-5) & \cdots & F(\beta-5) & L(\beta-5)
\end{pmatrix}
\begin{pmatrix}
\{\hat{\theta}(t-1)\} \\
\{\hat{\theta}(t-2)\} \\
\{\hat{\theta}(t-\frac{\alpha-5}{5})\} \\
\vdots \\
\{\hat{\theta}(t-\frac{\beta-10}{5})\}
\end{pmatrix}
\]

or, in condensed form,

\[
\{\hat{\theta}(t)\} = H\{\hat{\theta}(t-1)\} .
\]

Equation (3.20) relates the births at time t to the births at t-1. Once the birth trajectory is known, the trajectory of the population distribution may be computed by (3.15) and (3.8).

3.2. SENSITIVITY ANALYSIS OF THE POPULATION PROJECTION

Recall the population growth model defined in (3.1):

\[
\{ \hat{\kappa}(t+1) \} = \hat{\varpi}\{ \hat{\kappa}(t) \} .
\]

The assessment of the sensitivity of \{\hat{\kappa}(t+1)\} to changes in age-specific rates \(M(x)\), may be analyzed by means of a two-step process. The first step considers the sensitivity of the growth matrix to changes in age-specific rates. The second step derives a sensitivity function which describes the impact on the population distribution of a change in the
growth matrix. In our sensitivity analysis of life table statistics, we were not concerned with the time when the change in $\tilde{M}(x)$ occurred. The time consideration was irrelevant, since the life table is a static model. For the sensitivity analysis of the population growth, however, it is important to know not only the age group where a change in $\tilde{M}(x)$ occurs, but also the time when the change occurs. We will denote this time by $t_0$. The time at which the change in the population distribution is measured will be denoted by $t_1$.

Besides the change in $\{K(t_1)\}$ due to a change in the age-specific rates at $t_0$, one may also consider the problem of how a unique change in $\{K(t_0)\}$ affects $\{K(t_1)\}$. These are two separate sensitivity problems. In the first, the parameter changes at $t_0$ and remains at his new level thereafter. The second problem, however, is equivalent to a parameter change at $t_0$ only. These two sensitivity problems will be treated separately.

a. Sensitivity of the growth matrix

The growth matrix $G$ is composed of two types of submatrices, $S(x)$ and $B(x)$. The sensitivity on $S(x)$ of changes in $\tilde{M}(x)$, as given in Section 2.2, appears only in the two age groups, $x$ and $x-5$:

$$\frac{\delta S(x)}{\delta \tilde{M}(x)} = -\frac{5}{2} S(x) [I - \frac{5}{2} \tilde{M}(x)]^{-1} J$$

(b. Sensitivity of the transition matrix

$$\frac{\delta S(x - 5)}{\delta \tilde{M}(x)} = -\frac{5}{2} P(x) [I - \frac{5}{2} \tilde{M}(x)]^{-1} J S(x - 5)$$
\[
\frac{\delta S(a)}{\delta <M(x)>} = 0 \quad \text{for } a > x \quad \text{or} \\
\quad \quad \quad \quad \quad \quad \text{for } a < x - 5 .
\]

The sensitivity function of \( B(x) \) remains to be derived.

Recall from (3.4) that

\[
B(x) = \frac{5}{4} [\tilde{P}(0) + I] [\tilde{F}(x) + \tilde{F}(x + 5) \tilde{S}(x)] , \quad (3.4)
\]

where \( \tilde{B}(x) \) depends on the age-specific death and out-migratic rates through \( \tilde{S}(x) \) and \( \tilde{P}(0) \), and on the age-specific fertilit rates \( \tilde{F}(x) \) and \( \tilde{F}(x + 5) \). Consider the partial derivative of \( \tilde{B}(x) \) with respect to \( \tilde{M}(x) \):

\[
\frac{\delta B(x)}{\delta <M(x)>} = \frac{5}{4} \frac{\delta [\tilde{P}(0) + I]}{\delta <M(x)>} \tilde{F}(x) + \frac{5}{4} \frac{\delta [\tilde{P}(0) + I]}{\delta <M(x)>} \tilde{F}(x + 5) \tilde{S}(x) \\
+ \frac{5}{4} [\tilde{P}(0) + I] \frac{\delta S(x)}{\delta <M(x)>} . \quad (3.21)
\]

Since \( \tilde{P}(0) \) is affected by a change in \( \tilde{M}(x) \) only if \( x = 0 \), and because for this case \( \tilde{F}(x) \) and \( \tilde{F}(x + 5) \) are 0, (3.21) reduces to

\[
\frac{\delta B(x)}{\delta <M(x)>} = \frac{5}{4} [\tilde{P}(0) + I] \tilde{F}(x + 5) \frac{\delta S(x)}{\delta <M(x)>} \quad (3.22)
\]

which, by (2.35), is
\[
\frac{\delta B(x)}{\delta <M(x)>} = - \frac{5}{4} \left[ P(0) + I \right] F(x + 5) \frac{5}{2} S(x) \left[ I - \frac{5}{2} \tilde{M}(x) \right]^{-1} \tilde{J} \\
= - \frac{25}{8} \left[ P(0) + I \right] F(x + 5) S(x) \left[ I - \frac{5}{2} \tilde{M}(x) \right]^{-1} \tilde{J} .
\] (3.23)

Since a change of \( \tilde{M}(x) \) affects \( \tilde{S}(x - 5) \), it also affects \( \tilde{B}(x - 5) \)

\[
\frac{\delta B(x - 5)}{\delta <M(x)>} = - \frac{25}{8} \left[ P(0) + I \right] \tilde{F}(x) P(x) \left[ I - \frac{5}{2} \tilde{M}(x) \right]^{-1} \tilde{J} \tilde{S}(x - 5)
\] (3.24)

\[
= - \frac{25}{8} \left[ P(0) + I \right] \tilde{F}(x) \left[ P(x) + I \right] \tilde{J} \tilde{S}(x - 5) .
\] (3.25)

The sensitivity of \( \tilde{B}(x) \) with respect to \( \tilde{F}(x) \) and \( \tilde{F}(x + 5) \) also may be derived easily:

\[
\frac{\delta B(x)}{\delta <\tilde{F}(x)>} = \frac{5}{4} \left[ P(0) + I \right] \tilde{J}
\] (3.26)

\[
= \frac{5}{2} \tilde{L}(0) [5\tilde{L}(0)]^{-1} \tilde{J}
\]

and

\[
\frac{\delta B(x - 5)}{\delta <\tilde{F}(x)>} = \frac{5}{4} \left[ P(0) + I \right] \tilde{J} \tilde{S}(x - 5) .
\] (3.27)
Thus the impact of a unit change in the fertility matrix $\mathbf{F}(x)$ on the element $\mathbf{B}(x)$ is $\frac{5}{2}$ times the proportion of newborn babies that will be alive at the end of the time interval.

Having derived sensitivity functions for the elements of the growth matrix, we now can proceed to the question of how changes in the growth matrix affect the growth of the population. This is sometimes called trajectory sensitivity.

b. Sensitivity of the population trajectory

Recall the population growth equation

$$\{K(t+1)\} = G\{K(t)\}.$$  \hspace{1cm} (3.1)

Since $G$ is assumed to be constant over time, the population distribution at time $t_1$ is given by

$$\{K(t_1)\} = G^{t_1-t_0}\{K(t_0)\}.$$  \hspace{1cm} .

We assume that the change in the growth matrix occurs at $t_0$. Without loss of generality, we may set $t_0$ equal to zero, and $t_1$ equal to $t$. Then

$$\{K(t)\} = G^t\{K(0)\}.$$  \hspace{1cm} .

The sensitivity of $\{K(t)\}$ to a change in $G$ is

$$\frac{\delta\{K(t)\}}{\delta G} = \frac{\delta\{G^t\}}{\delta G} \{K(0)\}.$$  \hspace{1cm} .
The sensitivity of $\zeta_t$ to a change in $\zeta$ is given by (A.24) of the Appendix. Applying this result, yields:

$$\frac{\delta \{\zeta(t)\}}{\delta \zeta} = \sum_{i=0}^{t-1} \zeta^i J_{\zeta}^{t-1-i} \{K(0)\} \quad . \quad (3.28)$$

A related problem might come up in policy making. Under the growth model (3.1), the population distribution which yields a specified distribution at time $t$ is given by

$$\{K(0)\} = [G_t]^{-1} \{K(t)\} \quad .$$

If $\{K(0)\}$ deviates much from the actual population distribution, the policy maker may consider changing some elements of the growth matrix through policy measures. The impact on $\{K(0)\}$ is

$$\frac{\delta \{K(0)\}}{\delta \zeta} = - [G_t]^{-1} \frac{\delta G_t}{\delta \zeta} [G_t]^{-1} \{K(t)\}$$

$$= - [G_t]^{-1} \left[ \sum_{i=0}^{t-1} \zeta^i J_{\zeta}^{t-1-i} \right] [G_t]^{-1} \{K(0)\}$$

$$= - \sum_{i=0}^{t-1} [G(t-i)]^{-1} J[G(i+1)]^{-1} \{K(0)\} \quad . \quad (3.29)$$

If, by some means, an optimal growth matrix is defined which leads a population $\{K(0)\}$ to a desired $\{K(t)\}$, the next problem is to find out under what conditions variations in $\zeta$ do not affect $\{K(t)\}$. Such specific conditions are
derived by Tomović and Vukobratović (1972; p. 138). They will not be discussed here. This and similar problems of trajectory insensitivity or invariance are receiving an increasing attention in system theory and optimal control theory. For a review of some applications in the social sciences, see Erickson and Norton (1973).

The next section addresses the topic of the sensitivity of population growth to changes in the population distribution at a certain point in time. This will be called the analysis of small perturbations around the growth path.

c. Perturbations around the population growth path

The impact on \( \{K(t)\} \) of a change in \( \{K(0)\} \) is very simple in the time-invariant equation system (3.1). Applying the results of vector differentiation of the Appendix gives:

\[
\frac{\delta \{K(t)\}}{\delta \{K(0)\}'} = \frac{\delta \{G^t\} \{K(0)\}}{\delta \{K(0)\}'} = G^t.
\]

(3.30)

where \( \{K(0)\}' \) is the transpose of \( \{K(0)\} \).

Equation (3.30) relates changes in the state vector at time \( t \) to changes in the state vector at time zero. If the growth matrix is time-dependent, then this problem cannot be solved analytically, and one must rely on simulation. An illustration of such a situation is when the model incorporates a feedback loop, i.e., the growth matrix at time \( t \) depends on the state vector at time \( t \). An application of feedback models to urban analysis is given by Forrester (1969). Nelson and Kern (1971) have simulated the impact of small perturbations around the trajectory for a Forrester-type of urban model.
d. Sensitivity of the sequence of births

The sensitivity analysis of the growth matrix of the system trajectory and of perturbations around the trajectory could be repeated with the growth model (3.20). There are no real differences in methodology. The growth matrix now is simpler, and the state vector is the spatial distribution of the births. We will only consider the impact on the births sequence of a change in births at time zero where the birth sequence is described by

\[ \{ \hat{Q}(t) \} = H^t \{ \hat{Q}(0) \} , \]  

(3.31)

with \( H \) given by (3.20).

Suppose that a change occurs in the first sub-vector of \( \{ \hat{Q}(0) \} \), and that the impact is measured on the first sub-vector of \( \{ \hat{Q}(t) \} \), then the sensitivity coefficients are given by the submatrix \([H^t]_{11}\). Since new-born babies only affect the births sequence if they reach the reproductive ages, \([H^t]_{11}\) is 0 for \( t < \frac{5-\alpha}{5} \).

Another approach to sensitivity analysis of the births sequence may be more convenient, especially if, at the same time, one is interested in the sensitivity of the growth path of the whole population. This approach is based on the relationship

\[ \{ Q(t) \} = F \{ K(t) \} = F^t \{ K(0) \} \]  

(3.32)

where \( F \) is the matrix of age-specific fertility rates

\[ F = [0 \quad 0 \quad F(\alpha) \quad \cdots \quad F(\beta-5) \quad 0 \cdots] \]
A change in the growth matrix $G$ affects $\{q^{(t)}(t)\}$ in the following sense

$$\frac{\delta \{q^{(t)}(t)\}}{\delta <G>} = \frac{\delta F}{\delta <G>} \{K(t)\} + \sum \frac{\delta \{K(t)\}}{\delta <G>} .$$

If the change occurs in the mortality or migration, but not in the fertility, then

$$\frac{\delta \{q^{(t)}(t)\}}{\delta <G>} = F \frac{\delta \{K(t)\}}{\delta <G>} = F \sum_{i=0}^{t-1} G^i \sum J G^{t-1-i} \{K^{(0)}\} . \quad (3.33)$$

This chapter dealt with the sensitivity analysis of demographic growth. It has been shown that demographic growth may be expressed equally well in terms of births as in terms of population. This analogy will be extended in the next chapter while discussing the sensitivity of stable population characteristics.
CHAPTER 4

IMPACT OF CHANGES IN AGE-SPECIFIC RATES ON STABLE POPULATION CHARACTERISTICS

The stable population concept provides a major framework for analysis in mathematical demography. It has proved to be a helpful device in understanding how age compositions and regional distributions of populations are determined. The premise upon which the concept is based is the property that a human population tends to "forget" its past. This property is called ergodicity. The regional age compositions and regional shares of a closed multiregional population are completely determined by the recent history of fertility, mortality and migration to which the population has been subject. It is not necessary to know anything about the history of a population more than a century or two ago in order to account for its present demographic characteristics (Lopez, 1961). In fact, the regional shares, the age compositions and the sequence of births can be calculated from no more than a specified sequence of fertility, mortality and migration schedules over a moderate time interval.

Therefore, a particularly useful way to understand how the age and spatial structure of a population are formed and its vital rates determined, is to imagine them as describing a population which has been subjected to constant fertility, mortality and migration schedules for an extended period of time. The population that develops under such circumstances is called a stable multiregional population. Its principal characteristics are: constant regional age compositions and regional shares; constant
regional annual rates of birth, death and migration; and a fixed multiregional annual rate of growth that also is the annual growth rate in each region. Such multiregional stable populations have been studied by Rogers (1973, 1974, 1975).

The first section of this paragraph is an exposition of the major characteristics of stable populations. It is customary in mathematical demography to distinguish between a discrete and a continuous model of population growth, and the stable populations associated with them. The reason is mainly historical. The discrete model, which expresses the population growth as a matrix multiplication using a discrete time-variable and a discrete age-scale, derives largely from the work of Leslie (1945). The Leslie model is, in fact, a system of homogenous first-order difference equations, similar to (3.1). The continuous model uses a continuous time-variable and a continuous age-scale, and in its modern form originates from the work of Lotka (1907) and Sharpe and Lotka (1911). Lotka's work starts out with the population growth equation provided by Malthus (1798), which is, in fact, a homogenous first-order differential equation. Although in the literature the formulations of the continuous and the discrete model of growth seem very different, they are closely related. Goodman (1967) and Keyfitz (1968) have provided insights in the reconciliation of both growth models.

We focus in this section on the discrete model of population growth. However, we shall frequently refer to aspects of the continuous model that can be developed as well for the discrete case.
The second part of this section deals with the sensitivity analysis of the most important stable population statistics: the stable population distribution and the stable growth ratio. Demetrius (1969), Keyfitz (1971), Goodman (1971), Coale (1972) and Preston (1974), among others, have addressed this problem for a single region population without migration. Most take the continuous version of the stable population as a vehicle for sensitivity analysis. Demetrius and Goodman, however, use the discrete version. Their approach is our starting point for the sensitivity analysis. However, there are fundamental differences between the formulation of a single region and a multiregion stable population which necessitate other tools for analysis. One such tool is the eigenvalue and eigenvector analysis derived in the Appendix. An alternative approach, which starts out from the characteristic equation as in Keyfitz (1971), is also provided. This enables us to derive sensitivity functions that are similar to their single-region counterparts.

4.1. THE MULTIREGIONAL STABLE POPULATION

As in the previous chapter, we distinguish between the population model and the birth model. They are two equivalent formulations for population dynamics.

a. The population model

Recall the discrete model of population growth that was set out in (3.1). It may be written as

\[
\{k(t)\} = c^t\{k(0)\} .
\]

(4.1)
Consider the asymptotic properties of (4.1) when $t$ gets large. Such properties have been studied by Keyfitz (1968), Sykes (1969), Feeney (1973), Le Bras (1973) and Pollard (1973; pp. 39-46), among others. Rogers (1975; pp. 124-129) extends the arguments of Le Bras, Feeney, and Sykes to a multiregional system. The key element in the analysis is the Perron-Frobenius theorem. It establishes that any nonnegative, indecomposable, primitive square matrix has a unique, real, positive eigenvalue, $\lambda_1$ say, that is larger in absolute value than any other eigenvalue of that matrix. With this dominant eigenvalue are associated a right and left eigenvector, both with only positive elements. The growth operator is nonnegative and decomposable. However, $G$ may be partitioned, yielding a square submatrix, $\tilde{W}$ say, which is indecomposable and which is similar to $G$, and which therefore has the same eigenvalues. The matrix $\tilde{W}$ is primitive if the fertility of two adjacent age groups are positive in each and every region, i.e., if in (3.3) two consecutive matrices, $\tilde{B}(x)$ are positive (e.g., see Rogers (1975; pp. 124-129)). The dominant eigenvalue and the two associated eigenvectors have demographic meaning. The dominant eigenvalue of $G$ represents the stable growth ratio of the population. The associated right eigenvector gives the stable age- and region-specific population distribution, while the corresponding left eigenvector gives the spatial reproductive values. Therefore, the sensitivity of the growth ratio of the stable population to changes in the growth matrix is a problem of eigenvalue sensitivity. The sensitivity of the stable population distribution may be translated into eigenvector sensitivity.
We have seen, in the previous chapter, that because of the particular structure of $\underline{G}$, the growth equation may be written as:

$$\{K(t+1)(0)\} = \sum_{a=5}^{\beta-5} B(x) \{K(t)(x)\}$$  \hspace{1cm} (3.5)

$$\{K(t+1)(x + 5)\} = S(x) \{K(t)(x)\} \ .$$  \hspace{1cm} (3.6)

At stability, the characteristic value equation holds. Thus

$$\{K(t+1)\} = \underline{G}(\underline{K}(t)) = \lambda \{\underline{K}(t)\} \ .$$  \hspace{1cm} (4.2)

where $\lambda$ is the dominant eigenvalue of $\underline{G}$. Therefore,

$$\{K(t+1)(x + 5)\} = S(x) \{K(t)(x)\} = \lambda \{\underline{K}(t)(x + 5)\} \ ;$$  \hspace{1cm} (4.3)

hence

$$\{K(t)(x + 5)\} = \frac{1}{\lambda} \ S(x) \ {\underline{K}(t)(x)} \ .$$  \hspace{1cm} (4.4)

Combining (4.4) with (3.6), we have

$$\{K(t)(x)\} = \lambda \ \frac{x}{5} \ \underline{A}(x) \ {\underline{K}(t)(0)} \ .$$  \hspace{1cm} (4.5)

where $\underline{A}(x)$ is defined by (3.6).

The single-region analogue to (4.5) may be found in Goodman (1967; p. 543, and 1971; p. 340), Demetrius (1969; p. 133) and Cull and Vogt (1973; p. 647), among others.
Equation (4.3) gives the number of people in each age group and region in terms of the regional distribution of the people in the first age group. Now we derive an expression for the stable growth path of the population in the first age group. By (4.3) and (3.5) we may write:

\[
\{K^{(t+1)}(0)\} = \lambda \{K^{(t)}(0)\}
\]

\[
= \sum_{a=5}^{\beta-5} B(x) \{K^{(t)}(x)\} .
\]

Substituting for (4.5) and deleting the superscript, gives

\[
\lambda \{K(0)\} = \sum_{a=5}^{\beta-5} B(x) \lambda^{x/5} A(x) \{K(0)\} ,
\]

which is the expression given by Rogers (1975; p. 140).

It may be replaced by

\[
\left[ \sum_{a=5}^{\beta-5} \lambda^{x/5} B(x) A(x) - I \right] \{K(0)\} = \{0\} .
\]

Equation (4.7) is the discrete version of equation (4.7) in Rogers (1975; p. 93).

The matrix

\[
\hat{\phi}(x) = B(x) A(x)
\]

is the discrete formulation of the multiregional net maternity function, and

\[
\bar{\varphi}(\lambda) = \sum_{a=5}^{\beta-5} \lambda^{x/5} \hat{\phi}(x)
\]

(4.9)
is the corresponding discrete multiregional characteristic matrix.

The stable growth ratio $\lambda$ is the number that gives $\mathbf{\Psi}(\lambda)$ a characteristic root of unity. The vector $\mathbf{\xi}(0)$ is the associated eigenvector. An equivalent formulation is

$$|\mathbf{\Psi}(\lambda) - I| = 0.$$  \hspace{1cm} (4.10)

Condition (4.10) may also be derived in a different way. The idea is to reduce the growth matrix $\mathbf{G}$ to its generalized companion form. The notion of companion form of a matrix occupies a central place in system theory. See, for example, Wolovich (1974; p. 79) and Barnett (1974; p. 671). Kalman (1969; p. 44) considers several companion forms. Two commonly used forms are

$$\mathbf{M} = \begin{bmatrix} m_1 & m_2 & m_3 & \cdots & m_{Z-1} & m_Z \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \\ \end{bmatrix} \quad \text{and} \quad \mathbf{N} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \\ \vdots & \vdots & \vdots & \cdots & \cdots & 1 \\ m_Z & m_{Z-1} & \cdots & \cdots & m_1 \\ \end{bmatrix}$$

The companion form arises when a dynamic system is written as a linear differential or difference equation of the $Z$-th order. The elements of the first row of $\mathbf{M}$ or last row of $\mathbf{N}$, respectively, are the coefficients of the characteristic equation. Recall that the growth equation (3.1) is a system of $Z$ linear first-order difference equations, where $Z$ is the
number of age groups. Each system of linear first-order difference equations may be transformed into one linear difference equation of the $Z$-th order, and vice versa. This transformation corresponds to a change in the coordinate system. For example, (3.19) is a companion form, arising from the $(\beta-5)$-th order difference equation (3.17). Instead of scalar elements, (3.19) has submatrices as elements. Barnett (1973; p. 6) has called this form a generalized companion matrix. A transformation of a single region population growth matrix into a companion matrix of form $\hat{M}$ is given by Pielou (1969; p. 37). Wu (1972) sets up a transformation to both forms $\hat{M}$ and $\hat{N}$. In fact

$$\hat{M} = \hat{N},$$  

(4.11)

where

$$\hat{E} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix}.$$  

The transformation of the multiregional growth matrix $\hat{G}$ into a generalized companion matrix $\hat{G}$ may be expressed as

$$\hat{G} = \hat{H} \hat{G} \hat{H}^{-1}$$  

(4.12)
where

\[
\tilde{G} = \begin{bmatrix}
\begin{array}{cccc}
\tilde{A}(Z) & 0 & \cdots & 0 \\
0 & \tilde{A}(Z-5) & \cdots & \\
\vdots & \vdots & \ddots & \\
0 & \cdots & \cdots & \tilde{A}(0)
\end{array}
\end{bmatrix}
\]

with \( \tilde{A}(x) \) as defined by (3.6), and where

\[
\tilde{G} = \begin{bmatrix}
\begin{array}{ccccccc}
0 & 0 & [\tilde{B}(10) \tilde{A}(10)][\tilde{B}(15) \tilde{A}(15)] & \cdots & [\tilde{B}(\beta-5) \tilde{A}(\beta-5)] & \cdots & 0 \\
\tilde{I} & 0 & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \tilde{I} & \cdots & \cdots & \cdots & \cdots & \cdots \\
\vdots & \vdots & \ddots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \tilde{I} \\
\end{array}
\end{bmatrix}
\]

(4.13)

Since (4.12) is a similarity transformation, it implies that \( \tilde{G} \) and \( \hat{\tilde{G}} \) have the same eigenvalues. They may be found by solving

\[
|\tilde{G} - \lambda \tilde{I}| = 0
\]

(4.14)

or

\[
|\hat{\tilde{G}} - \lambda \tilde{I}| = 0.
\]

(4.15)

Kenkel (1974; pp. 319-322) shows that (4.15) may be reduced:

\[
|\hat{\tilde{G}} - \lambda \tilde{I}| = |\lambda^{\frac{Z+1}{5}} \tilde{I} - \lambda^{\frac{Z}{5}} \tilde{B}(0)\tilde{A}(0) - \lambda^{\frac{Z-1}{5}} \tilde{B}(5)\tilde{A}(5) | \cdots
\]

\[
- \lambda^{\tilde{B}(Z-5)} \tilde{A}(Z-5) - \tilde{B}(Z)\tilde{A}(Z) |
\]

(4.16)
Dividing by $\frac{Z+1}{5}$, and since $\mathbf{B}(x) = 0$ for $x < \alpha - 5$ and for $x > \beta - 5$, we have that

$$|\hat{G} - \lambda I| = \left| \sum_{a-5}^{\beta-5} \lambda^{-\left(\frac{x}{5}+1\right)} B(x) \mathbf{A}(x) - I \right|$$  \hspace{1cm} (4.17)

which is condition (4.10). Wilkinson (1965; p. 432) labels (4.17) as the generalized eigenvalue problem.

The generalized companion matrix provides a mathematical tool to link (4.10) to (4.14). Since (4.10) is the discrete version of the condition in the continuous model that the stable growth rate must give the characteristic matrix an eigenvalue of unity, the companion matrix has a role in the reconciliation of the discrete and the continuous models of demographic growth.

The eigenvector of $\hat{G}$ and $G$ are related as

$$\mathbf{K} = \mathbf{H}(\mathbf{K}) \hspace{1cm} (4.18)$$

b. The birth model

The birth trajectory may be described by (3.20):

$$\mathbf{Q}(t) = \mathbf{H}(\mathbf{Q}(t-1)) \hspace{1cm} (3.20)$$

Since all the elements of $\mathbf{H}$ are nonnegative, we may apply the Perron-Frobenius theorem and derive expressions for $\lambda$ analogue to (4.10) and (4.14). However, there is a third formulation of the condition that $\lambda$ must satisfy. It draws on the relationship between $\{\mathbf{K}(0)\}$ and $\{\mathbf{Q}\}$, the births in
the stable population:

$$\{k(0)\} = \frac{1}{2} \frac{5}{2} [I + P(0)][Q] \ ,$$

(4.19)

which has its origin in (3.15). Substituting this into (4.6) and introducing $B(x)$ yields

$$\sum_{\alpha=5}^{\beta-5} \lambda \left(\frac{x+1}{2}\right) \frac{5}{4} \left[I + P(0)\right] \left[F(x) + F(x + 5) S(x) \right] \left[A(x) \frac{5}{2} [I + P(0)][Q]\right] = \frac{1}{2} \frac{5}{2} [I + P(0)][Q] \ .$$

(4.20)

Multiplying both sides by $\lambda \frac{5}{2} \frac{2}{5} [I + P(0)]^{-1}$ gives

$$\sum_{\alpha=5}^{\beta-5} \lambda \left(\frac{x+1}{2}\right) \frac{1}{2} \left[F(x) + F(x + 5) S(x) \right] \left[A(x) \frac{5}{2} [I + P(0)][Q]\right] = \{Q\} \ .$$

(4.21)

But

$$\frac{5}{2} [I + P(0)] = L(0)$$

and

$$A(x) L(0) = \bar{L}(x)$$

where $\bar{L}(x)$ is the number of years lived in the age group $x$ to $x + 4$ by unit regional radices. Therefore (4.21) becomes

$$\begin{bmatrix} \sum_{\alpha=5}^{\beta-5} \lambda \left(\frac{x+1}{2}\right) \frac{1}{2} \left[F(x) + F(x + 5) S(x) \right] \bar{L}(x) \end{bmatrix} \{Q\} = \{Q\} \ .$$

(4.22)
The matrix

$$\theta(\lambda) = \sum_{\alpha=5}^{\beta-5} \lambda \left( \frac{x+1}{\alpha+2} \right) \frac{1}{2} \left[ F(x) + F(x + 5) \right] \lambda(x)$$

is very close to the numerical approximation of the continuous characteristic matrix, given by Rogers (1975; p. 100):

$$\psi(r) = \sum_{\alpha=5}^{\beta-5} e^{-r(x+2.5)} F(x) \lambda(x)$$

(4.24)

where $\lambda = e^{5r}$ and $F(x) \equiv \frac{1}{2} \left[ F(x) + F(x + 5) \right] \lambda(x)$. The stable growth rate $\lambda$ is the solution of

$$\left| \sum_{\alpha=5}^{\beta-5} \lambda \left( \frac{x+1}{\alpha+2} \right) \frac{1}{2} \left[ F(x) + F(x + 5) \right] \lambda(x) - 1 \right| = 0.$$

Once the stable distribution of births is known, the stable population distribution can be computed by means of (4.19) and (4.5).

4.2. SENSITIVITY ANALYSIS OF THE STABLE POPULATION

To perform a sensitivity analysis of the stable population, we may apply the eigenvalue and eigenvector sensitivity functions, derived in the Appendix, directly to the growth matrix. Another approach starts out from the generalized eigenvalue problem, expressed in (4.17) and (4.22). This approach is more related to the sensitivity analysis in the single-region case. There is a crucial difference, however. For a single-region growth matrix, the companion form is
composed of scalars. The elements of the first row are the coefficients of the characteristic equation, a scalar polynomial. The characteristic equation of the multiregional growth matrix is a matrix polynomial. Its analysis is much more complicated. Both approaches will be discussed here.

a. Sensitivity analysis with the whole growth matrix

The sensitivity of the eigenvalue to changes in the matrix is given in the Appendix by (A.56):

$$d\lambda_i = \left\{ \xi_i \right\}_1 \left\{ \gamma_i \right\}_1' * d\lambda$$  \hspace{1cm} (A.56)

where $\left\{ \xi_i \right\}_1$ and $\left\{ \gamma_i \right\}_1$ are the right and left normalized eigenvector of $\lambda_i$, respectively, associated with the root $\lambda_i$.

Let $A = \lambda\gamma$, the multiregional growth matrix, and denote the eigenvectors by $\left\{ k \right\}$ and $\left\{ \gamma \right\}$, respectively. When the eigenvectors are not normalized, the formula becomes

$$d\lambda = \frac{1}{\left\{ \gamma \right\}_1 \left\{ k \right\}_1} \left[ \left\{ k \right\} \left\{ \gamma \right\}_1' \right] * dG$$  \hspace{1cm} (4.25)

where

$$\left\{ k \right\}_1 = \begin{bmatrix} \{ k(0) \} \\ \{ k(5) \} \\ \vdots \\ \{ k(Z) \} \end{bmatrix}$$

$$\left\{ \gamma \right\}_1 = \begin{bmatrix} \{ \gamma(0) \} \\ \{ \gamma(5) \} \\ \vdots \\ \{ \gamma(Z) \} \end{bmatrix}$$

The inner product is

$$\left\{ \gamma \right\}_1' \left\{ k \right\}_1 = \sum_{x=0}^{Z} \\{ \gamma(x) \}_1' \{ k(x) \}_1$$
In the single-region case, the inner product

$$V = \{v\} \{K\} = \sum_{x=0}^{Z} v(x) K(x)$$

is the total reproductive value of the stable population. If the eigenvectors are normalized, then $\{v\} \{K\} = 1$, and $v(x) K(x)$ is the reproductive value of age group $x$, as a fraction of the total reproductive value.

If one applies formula (A.59), other useful relationships may be derived

$$d\lambda = [\text{tr} \ R(\lambda)] \ R(\lambda) * \ dG$$  \hspace{1cm} (A.59)$$

where $\tilde{R}(\lambda)$ is the adjoint matrix of $[G - \lambda I]$ and $G$ is the growth matrix. The single-region analogue of (A.59) is derived by Demetrius (1969; p. 134). Morgan (1966; p. 198) has shown that $\text{tr} \ R(\lambda)$ is equal to the first derivative of the characteristic equation of $G$. Based on this result, it can be shown that for the single-region case, the following equality holds:

$$\frac{\delta g(\lambda)}{\delta \lambda} = \text{tr} \ R(\lambda) = \frac{1}{\lambda} A$$  \hspace{1cm} (4.26)$$

where $A$ is the mean age of childbearing of the stable population and $g(\lambda)$ is the characteristic equation of $G$. This result is similar to the one derived by Goodman (1971; p. 346) and Keyfitz (1968; p. 100).

Formula (4.25) and (A.59) are particularly useful to study the interaction of the population distribution and the
distribution of the reproductive values. Goodman (1971) and Demetrius (1969) illustrate this for a single-region system. Consider, for example (4.25), and let $t = \{\bar{y}\}'\{\bar{K}\}$. Written in component terms, (4.25) is

\[
\begin{bmatrix}
\{K(0)\} \\
\{K(5)\} \\
\vdots \\
\{K(Z)\}
\end{bmatrix}
\begin{bmatrix}
\{y(0)\}' \\
\{y(5)\}' \\
\vdots \\
\{y(Z)\}'
\end{bmatrix}
\begin{bmatrix}
0 \\
0 \\
\ddots
\end{bmatrix}
\begin{bmatrix}
\bar{d}_B(0) \\
\bar{d}_B(5) \\
\vdots \\
\bar{d}_B(Z)
\end{bmatrix}
= \frac{1}{t}
\begin{bmatrix}
\{x(x)\} \{y(0)\}'
\end{bmatrix}
\begin{bmatrix}
\bar{d}_B(x)
\end{bmatrix}
\]

(4.27)

The impact on $\lambda$ of a change in $B(x)$ is

\[
d\lambda = \frac{1}{t} \left[ \{x(x)\} \{y(0)\}' \right] d\bar{B}(x) .
\]  

(4.28)

The impact of a change in $S(x)$ is

\[
d\lambda = \frac{1}{t} \left[ \{x(x)\} \{y(x + 5)\}' \right] d\bar{S}(x) .
\]  

(4.29)

From (4.28) and (4.29), we see that a change in $B(x)$ is equivalent to a change in $S(x)$ if

\[
\left[ \{x(x)\} \{y(0)\}' \right] d\bar{B}(x) = \left[ \{x(x)\} \{y(x + 5)\}' \right] d\bar{S}(x)
\]
or

\[
d\bar{B}(x) = \left[ \{x(x)\} \{y(0)\}' \right]^{-1} \left[ \{x(x)\} \{y(x + 5)\}' \right] d\bar{S}(x)
\]

if the inverse exists.
Since

\[ \{\tilde{K}(x)\} = \lambda ^{-\frac{x}{S}} \tilde{A}(x) \{\tilde{K}(0)\} \]

we have

\[ \text{d}\tilde{B}(x) = \left[\left\{\tilde{K}(0)\{y(0)\}\right\}^{'}\right]^{-1} \left[\begin{array}{cc}
-\frac{x}{S} \tilde{A}(x) & -1 \\
\lambda ^{-\frac{x}{S}} \tilde{A}(x) & \lambda ^{-\frac{x}{S}} \tilde{A}(x)
\end{array}\right] \left[\left\{\tilde{K}(0)\{y(x + 5)\}\right\}^{'}\right] \text{d}\tilde{S}(x) \]

\[ \text{d}\tilde{B}(x) = \left[\left\{\tilde{K}(0)\{y(0)\}\right\}^{'}\right]^{-1} \left[\left\{\tilde{K}(0)\{y(x + 5)\}\right\}^{'}\right] \text{d}\tilde{S}(x) \quad (4.30) \]

Equation (4.30) shows that a change in \( \tilde{B}(x) \) may be translated into a change in \( \tilde{S}(x) \), having the same impact on the growth ratio. It formulates, therefore, a trade-off between fertility change and mortality and migration change. The change in \( \tilde{S}(x) \) to have the same effect as \( \text{d}\tilde{B}(x) \) must be smaller the greater are the reproductive values of the people aged \( x + 5 \) to \( x + 9 \), i.e., \( \{y(x + 5)\} \).

It should be noted that the equivalence only holds for the growth ratio, and not for the stable population distribution and other stable characteristics. The stable populations which result from applying \( \text{d}\tilde{S}(x) \) or \( \text{d}\tilde{B}(x) \) given by (4.30) have the same growth ratio, but all other characteristics are different.

b. Sensitivity analysis with the characteristic matrix

The discrete multiregional characteristic matrix is

\[ \varphi(\lambda) = \beta^{-\frac{5}{S}} \sum \lambda^{-\frac{5}{S}+1} \varphi(x) \quad , \quad (4.9) \]
where the stable growth ratio $\lambda$ is the solution of

$$|\bar{\Psi}(\lambda) - I| = 0 \quad (4.10)$$

What effect does a change in an element of the growth matrix have on $\lambda$? As in the previous section, we distinguish between a change in fertility, as expressed by $B(x)$, and a change in mortality and migration, as expressed by $\bar{S}(x)$. This approach is equally valid to trace through the impact of changing fertility, mortality and migration patterns in the continuous model of demographic growth. Instead of using $\bar{\Psi}(\lambda)$, one then uses its continuous counterpart, given by Rogers (1975; p. 93),

$$\bar{\Psi}(r) = \int_{\alpha}^{\beta} e^{-r x} \bar{\phi}(x) \, dx \quad (4.31)$$

where $r$ is the intrinsic growth rate.

The impact on $\lambda$ of a changing element of $\bar{\Psi}(\lambda)$ is such that the determinant $|\bar{\Psi}(\lambda) - I|$ remains zero. We treat the impact on $\lambda$ of a change in $\bar{B}(x)$ and $\bar{S}(x)$ separately.

b.1. Sensitivity of the growth ratio to changes in fertility

Consider first the derivative of the determinant with respect to an element of $\bar{B}(x)$, denoted by $\langle \bar{B}(x) \rangle$. Applying the chain rule of matrix differentiation, given in the Appendix by (A.30), we get

$$\frac{\delta |\bar{\Psi}(\lambda) - I|}{\delta \langle \bar{B}(x) \rangle} = \text{tr} \left[ \frac{\delta |\bar{\Psi}(\lambda) - I|}{\delta \bar{\Psi}(\lambda)} \cdot \frac{\delta \bar{\Psi}(\lambda)}{\delta \langle \bar{B}(x) \rangle} \right] = 0 \quad (4.32)$$
By (A.35)
\[ \frac{\delta |\tilde{\Psi}(\lambda) - I|}{\delta \tilde{\Psi}(\lambda)} = \text{cof} \left[ \tilde{\Psi}(\lambda) - I \right] \text{.} \] (4.33)

The derivative of the transpose of the characteristic matrix with respect to \( \langle B(x) \rangle \) is
\[ \frac{\delta [\tilde{\Psi}(\lambda)]'}{\delta \langle B(x) \rangle} = \frac{\delta \left[ \sum_{\alpha=5}^{\beta} \frac{\lambda - \frac{(X+1)}{5}}{5} B(x) \ A(x) \right]'}{\delta \langle B(x) \rangle} \text{.} \]

Assume that the change in \( \tilde{B}(x) \) is due to a fertility change, then
\[ \delta \left[ \sum_{\alpha=5}^{\beta} \frac{\lambda - \frac{(X+1)}{5}}{5} [A(x)]' [B(x)]' \right] = \frac{\delta [\tilde{B}(x)]'}{\delta \langle B(x) \rangle} \text{.} \]

\[ = \sum_{\alpha=5}^{\beta-5} \frac{\delta [A(x)]' [B(x)]'}{\delta \langle B(x) \rangle} \frac{\delta \lambda}{\delta \langle B(x) \rangle} + \lambda \frac{\delta [\tilde{B}(x)]'}{\delta \langle B(x) \rangle} \]

where
\[ \frac{\delta \lambda}{\delta \langle B(x) \rangle} = \frac{\delta \lambda}{\delta \lambda} \cdot \frac{\delta \lambda}{\delta \langle B(x) \rangle} \]

\[ = - \left( \frac{X+1}{5} + 1 \right) \frac{\delta \lambda}{\delta \langle B(x) \rangle} \text{.} \]
and

\[
\delta [B(x)]' \over \delta \langle B(x) \rangle = J'.
\]

Therefore

\[
\delta [\tilde{\Psi}(\lambda)]' \over \delta \langle B(x) \rangle = - \left[ \frac{1}{\lambda} \sum_{a-5}^{\beta-5} \left( \frac{x}{5} + 1 \right) \lambda^{-(\frac{x}{5}+1)} [B(x) A(x)]' \right] \frac{\delta \lambda}{\delta \langle B(x) \rangle} + \lambda A'(x) J'. \tag{4.34}
\]

Let

\[
\sum_{a-5}^{\beta-5} \left( \frac{x}{5} + 1 \right) \lambda^{-(\frac{x}{5}+1)} [B(x) A(x)]' = [\Psi(0)]^{-1} - 2. \tag{4.35}
\]

Generalizing the idea of Goodman, \([\Psi(0)]^{-1}\) is the matrix of the average age of mothers of children who are in the 0-th age group in the stable population. It is the discrete approximation of the mean age of childbearing. The matrix \(\Psi(0)\) represents the eventual reproductive value of a female in the 0-th age group in the stable population.

Substituting (4.33), and (4.34) in (4.32) gives

\[
\text{tr cof} \left[ \tilde{\Psi}(\lambda) - I \right] \left[ - \frac{1}{\lambda} [\Psi(0)]^{-1} \frac{\delta \lambda}{\delta \langle B(x) \rangle} + \lambda A'(x) J' \right] = 0
\]

The single region counterpart of (4.35) is given by Goodman (1971; p. 346).
which may be written as

$$\frac{1}{\lambda} \cof [\bar{\psi}(\lambda) - I] * \bar{V}^{-1}(0) \frac{\delta \lambda}{\delta <B(x)>} = \lambda^{-(X/S + 1)} \cof [\bar{\psi}(\lambda) - I] * [A'(x) J']$$

Pre-multiplying both sides with $[\cof [\bar{\psi}(\lambda) - I]]^{-1}$ yields

$$\frac{1}{\lambda} I * [\bar{V}^{-1}(0)] \frac{\delta \lambda}{\delta <B(x)>} = \lambda^{-(X/S + 1)} I * A'(x) J'$$

But $I * [\bar{V}^{-1}(0)]$ is nothing else than $\text{tr} [\bar{V}^{-1}(0)]$. Therefore, we have

$$\frac{\delta \lambda}{\delta <B(x)>} = [\text{tr} \bar{V}^{-1}(0)]^{-1} \lambda^{X/S} \text{tr} [A'(x) J'] \quad (4.37)$$

By (A.32) of the Appendix,

$$\frac{\delta \lambda}{\delta B(x)} = [\text{tr} \bar{V}^{-1}(0)]^{-1} \lambda^{X/S} \sum_{kl} \text{tr} [A'(x) J_k J_l]$$

$$= [\text{tr} \bar{V}^{-1}(0)]^{-1} \lambda^{X/S} A'(x) \quad (4.38)$$

In a single-region system, (4.38) reduces to

$$\frac{\delta \lambda}{\delta b(x)} = v(0) \lambda^{X/S} a(x) \quad (4.39)$$

where $b(x)$, $v(0)$ and $a(x)$ are scalars. Formula (4.39) is identical to the sensitivity function given by Goodman (1971;
p. 346), and equivalent to the ones derived by Demetrius (1969; p. 134), Keyfitz (1971; p. 277), Emlen (1970) and others. Note that \( \lambda \frac{X}{5} \tilde{A}(x) \) is the eventual expected number of people in age group \( x \) to \( x + 4 \), per individual in the 0 - 4 age group. In other words, \( \lambda \frac{X}{5} \tilde{A}(x) \) describes the age composition of the stable population.

b.2. Sensitivity of the growth ratio to changes in mortality and migration

The impact on \( \lambda \) of a change in \( \tilde{S}(x) \) may be derived in a way similar to the above arguments. First, note that

\[
\frac{\delta |\tilde{\Psi}(\lambda) - I|}{\delta \tilde{S}(x)'} = \text{tr} \left[ \frac{\delta |\tilde{\Psi}(\lambda) - I|}{\delta \tilde{\Psi}(\lambda)} \cdot \frac{\delta [\tilde{\Psi}(\lambda)]'}{\delta \tilde{S}(x)'} \right] = 0 .
\]

(4.40)

The derivative of \([\tilde{\Psi}(\lambda)]'\) with respect to an element of \( \tilde{S}(x) \) is

\[
\frac{\delta [\tilde{\Psi}(\lambda)]'}{\delta \tilde{S}(x)'} = \frac{\delta \left[ \sum_{\alpha=5}^{\beta-5} \lambda \frac{X+1}{5} B(x) \tilde{A}(x) \right]'}{\delta \tilde{S}(x)'}
\]

(4.41)

\[
= \sum_{\alpha=5}^{\beta-5} \frac{B(x) \tilde{A}(x)'}{\delta \tilde{S}(x)'} \cdot \frac{\delta \lambda}{\delta \tilde{S}(x)} \cdot \frac{-(\frac{X+1}{5})}{\frac{\delta \lambda}{\delta \tilde{S}(x)}} + \sum_{\alpha=5}^{\beta-5} \lambda \frac{-(\frac{X+1}{5})}{\frac{\delta \lambda}{\delta \tilde{S}(x)}} \cdot \frac{\delta \tilde{A}'(x)}{\delta \tilde{S}(x)} \cdot \frac{B'(x)}{\tilde{S}(x)}
\]

\[
+ \sum_{\alpha=5}^{\beta-5} \lambda \frac{-(\frac{X+1}{5})}{\frac{\delta \lambda}{\delta \tilde{S}(x)}} \cdot \frac{\delta \tilde{B}'(x)}{\delta \tilde{S}(x)} \cdot \frac{\tilde{A}'(x)}{\tilde{S}(x)} .
\]

(4.42)
The derivatives are

\[
\frac{-(\frac{x}{2}+1)}{\delta \lambda} \frac{\delta \lambda}{\delta \langle S(x) \rangle} = \frac{-(\frac{x}{2}+1)}{\delta \lambda} \cdot \frac{\delta \lambda}{\delta \langle S(x) \rangle} = -\left(\frac{x}{2} + 1\right) - \frac{x+2}{3} \frac{\delta \lambda}{\delta \langle S(x) \rangle}.
\]

(4.43)

To derive an expression for

\[
\sum_{\alpha=5}^{\beta-5} \frac{-(\frac{x}{2}+1)}{\lambda} \frac{\delta \lambda}{\delta \langle S(x) \rangle} \frac{\delta \lambda}{\langle S(x) \rangle} \frac{\partial A'(x)}{\partial x} B'(x) \]

(4.44)

recall that

\[A'(x) = \langle S'(0) \rangle \langle S'(5) \rangle \ldots \langle S'(x - 5) \rangle.\]

Therefore, a change in \langle S(x) \rangle affects \langle A'(y) \rangle if \( y > x \). For example,

\[
\frac{\delta \langle A'(y) \rangle}{\delta \langle S(x) \rangle} = \langle S'(0) \rangle \langle S'(5) \rangle \ldots \langle S'(x - 5) \rangle \langle S'(x + 5) \rangle \ldots \langle S'(y - 5) \rangle
\]

\[= A'(x) \langle [A'(x + 5)]^{-1} \langle A'(y) \rangle.\]

(4.45)

Applying this result, (4.44) reduces to

\[
\sum_{y=x+5}^{\beta-5} \frac{-(\frac{x}{2}+1)}{\lambda} A'(x) \langle [A'(x + 5)]^{-1} \rangle A'(y) B'(y).
\]

(4.46)
To compute the third element of (4.41), we need

\[
\frac{\delta B'(x)}{\delta<\tilde{S}(x)>} = \frac{5}{4} \frac{\delta S'(x)}{\delta<\tilde{S}(x)>} \tilde{F}'(x+5) \left[ P(0) + \frac{1}{5}\right]
\]

\[
= \frac{5}{4} \tilde{J}' \tilde{F}'(x+5) \left[ P(0) + \frac{1}{5}\right].
\]

Therefore (4.42) becomes

\[
\frac{\delta[\tilde{\Psi}(\lambda)]'}{\delta<\tilde{S}(x)>} = \left[ - \sum_{\alpha=5}^{\beta+5} \frac{\lambda}{(\alpha+1)} \right] \left[ B(x) \tilde{A}(x) \right] ' \frac{\delta \lambda}{\delta<\tilde{S}(x)>}
\]

\[
+ \frac{5}{4} \tilde{A}'(x) \tilde{J}' \tilde{F}'(x+5) \left[ P(0) + \frac{1}{5}\right].
\]

where by (4.35)

\[
\sum_{\alpha=5}^{\beta-5} \frac{\lambda}{(\alpha+1)} \left[ B(x) \tilde{A}(x) \right] ' = [\Psi(0)]^{-1}.
\]

Substituting (4.48) in (4.40) gives

\[
\frac{\delta |\tilde{\Psi}(\lambda) - \mathbb{I}|}{\delta<\tilde{S}(x)>} = \text{tr} \text{ cof } [\tilde{\Psi}(\lambda) - \mathbb{I}] \left[ \frac{1}{\lambda} [\Psi(0)]^{-1} \frac{\delta \lambda}{\delta<\tilde{S}(x)>}
\]

\[
+ \sum_{\alpha=5}^{\beta-5} \frac{\lambda}{(\alpha+1)} \tilde{A}'(x) \tilde{J}' [\tilde{A}'(x+5)]^{-1} \tilde{A}'(y) \tilde{B}'(y)
\]

\[
+ \frac{5}{4} \frac{\lambda}{(\alpha+1)} \tilde{A}'(x) \tilde{J}' \tilde{F}'(x+5) \left[ P(0) + \frac{1}{5}\right].
\]

\[= 0\]
which is equivalent to

\[
\left[ \frac{1}{\lambda} \text{tr} \left[ \mathcal{V}(0) \right]^{-1} \right] \frac{\delta \lambda}{\delta \mathcal{S}(x)} = \frac{\beta - 5}{\lambda} \sum_{y=x+5} \lambda^{-\frac{Y}{5} + 1} \text{tr} \left[ \mathcal{A}'(x) \mathcal{S}'[\mathcal{A}'(x + 5)]^{-1} \right. \\
\left. \mathcal{A}'(y) \mathcal{B}'(y) \right] + \frac{5}{4} \lambda^{-\frac{X}{5} + 1} \text{tr} \left[ \mathcal{A}'(x) \mathcal{J}' \mathcal{F}'(x + 5) [\mathcal{F}(0) + \mathcal{I}] \right]'
\]

or

\[
\frac{\delta \lambda}{\delta \mathcal{S}(x)} = \left[ \text{tr} \mathcal{V}^{-1}(0) \right]^{-1} \left[ \frac{\beta - 5}{\lambda} \sum_{y=x+5} \lambda^{-\frac{Y}{5}} \text{tr} \left[ \mathcal{B}(y) \mathcal{A}(y) [\mathcal{A}(x + 5)]^{-1} \mathcal{J} \mathcal{A}(x) \right] \\
+ \frac{5}{4} \lambda^{-\frac{X}{5}} \text{tr} \left[ [\mathcal{F}(0) + \mathcal{I}] \mathcal{F}(x + 5) \mathcal{J} \mathcal{A}(x) \right] \right]
\]

(4.49)

and

\[
\frac{\delta \lambda}{\delta \mathcal{S}(x)} = \left[ \text{tr} \mathcal{V}^{-1}(0) \right]^{-1} \left[ \frac{\beta - 5}{\lambda} \sum_{y=x+5} \lambda^{-\frac{Y}{5}} \mathcal{B}(y) \mathcal{A}(y) [\mathcal{A}(x + 5)]^{-1} \mathcal{A}(x) \\
+ \frac{5}{4} \lambda^{-\frac{X}{5}} [\mathcal{F}(0) + \mathcal{I}] \mathcal{F}(x + 5) \mathcal{A}(x) \right]
\]

(4.50)

If the effect on \( \mathcal{S}(x) \) of \( \mathcal{S}(x) \) is negligible, as Goodman (1971; p. 343) assumes, and which is so in the continuous case, then

\[
\frac{\delta \lambda}{\delta \mathcal{S}(x)} = \left[ \text{tr} \mathcal{V}^{-1}(0) \right]^{-1} \left[ \frac{\beta - 5}{\lambda} \sum_{y=x+5} \lambda^{-\frac{Y}{5}} \mathcal{B}(y) \mathcal{A}(y) \mathcal{A}^{-1}(x) \mathcal{S}^{-1}(x) \mathcal{A}(x) \right]
\]

(4.51)
The single-region analogue of (4.51) is

\[
\frac{\delta \lambda}{\delta s(x)} = \frac{v(0)}{s(x)} \left[ \beta - 5 \sum_{y=x+5}^{\beta-5} \frac{-y}{5} b(y) a(y) \right]
\]

which is identical to formula (35) of Goodman (1971; p. 346), and equivalent to expressions provided by other authors.

The expression

\[
v(0) \sum_{y=x+4}^{\beta-5} \frac{-y}{5} b(y) a(y) = v(x)
\]

is defined by Goodman as the eventual reproductive value of an individual in the \( x, x + 4 \) age interval. Generalizing this concept to the multiregional case, we define the matrix of eventual reproductive values per individual in the \( x, x + 4 \) age group, by place of birth and by place of residence, to be

\[
\tilde{v}(x) = \sum_{y=x+4}^{\beta-5} \frac{-y}{5} \tilde{b}(y) \tilde{a}(y)
\]

The sensitivity function (4.51) becomes

\[
\frac{\delta \lambda}{\delta s(x)} = [\text{tr} \; \tilde{v}^{-1}(0)]^{-1} \tilde{v}(x+5) \tilde{a}^{-1}(x) \tilde{s}^{-1}(x) \tilde{a}(x)
\]
APPENDIX

MATRIX DIFFERENTIATION TECHNIQUES

The purpose of this appendix is to provide the necessary mathematical tools to perform sensitivity analysis of structural change in multiregional demographic systems. The basic notion is that of matrix differentiation. Neudecker (1969; p. 953) defines matrix differentiation as the procedure of finding partial derivatives of the elements of a matrix function with respect to the elements of the argument matrix. Although not much has been written on matrix differentiation and the technique is not covered in most textbooks on matrix algebra, this appendix does not intend to be complete. It only covers the techniques applied in this study.

The appendix is divided into two parts. The first part deals with the derivatives of matrix functions. It is mainly based on the work of Dwyer and MacPhail (1948) and Dwyer (1967). The second part develops several expressions for the sensitivity of the eigenvalues and the eigenvectors of a matrix with respect to change in its elements. The behavior of the eigenvalues under perturbations of the elements of a matrix has been studied by Lancaster (1969; Chapter 7), among others, under the heading of perturbation theory. In this theory, qualitative measures of eigenvalue sensitivity are developed, in the sense that upper and lower bounds to eigenvalue changes are formulated. Perturbation theory, however, does not provide us with sensitivity functions defining the exact change of eigenvalues and eigenvectors under changing matrix elements.
An eigenvalue sensitivity function was derived by Jacobi in 1846 and has been applied and extended in the systems theory and design literature.

A.7. DIFFERENTIATION OF FUNCTIONS OF MATRICES

Let \( \tilde{X} \) be an \( P \times Q \) matrix with elements \( \tilde{y}_{ij} \), and let \( \tilde{X} \) be an \( M \times N \) matrix with elements \( \tilde{x}_{kl} \). Dynyer makes a distinction between the position of an element in the matrix and its value. The symbol \( \tilde{x}_{kl}^{(k,l)} \) is used to indicate a specific \( k, l \)-element of \( \tilde{X} \). Its scalar value is \( x_{kl} \). Less formally, \( \tilde{x}_{kl}^{(k,l)} \) may be replaced by \( <X> \). Therefore, \( <X> \) is an arbitrary element of the matrix \( \tilde{X} \). As in conventional notation \( \tilde{X}' \) denotes the transpose of \( \tilde{X} \) and \( \tilde{X}^{-1} \) is the inverse of \( \tilde{X} \).

The relevant results of matrix calculus are given below. To introduce some notation, we start out with the differentiation of a matrix with respect to its elements. We follow this with the differentiation of a matrix with respect to a scalar, and the differentiation of a scalar function with respect to a matrix. The most important scalar function is the determinant. The tools provided in the section on the differentiation of matrix products are frequently used in performing sensitivity analysis of multiregional systems. Also of great importance is the derivative of the inverse. The next section gives some chain rules of matrix differentiation. Vector calculus and matrix calculus are closely related, since a vector is a matrix with only one row or one column. The formulas for vector differentiation, however, have a different appearance and are less
complex. Therefore, a separate section will be devoted to vector differentiation.

A.1.1. Differentiation of a matrix with respect to its elements

The derivative of a matrix $X$ with respect to the element $X_{kl}$ is

$$
\frac{\delta X}{\delta X_{kl}} = J_{kl} \quad (A.1)
$$

where $J_{kl}$ denotes an $M \times N$ matrix with zero elements everywhere except for a unit element in the $k$-th row and $l$-th column.

Similarly

$$
\frac{\delta X'}{\delta X_{kl}} = J'_{kl} \quad (A.2)
$$

where $J'_{kl}$ is an $N \times M$ matrix with all elements zero except for a unit element in the $l$-th row and $k$-th column.

Instead of considering the derivative of a matrix with respect to an element, one may also consider the derivative of a matrix-element with respect to the matrix.

$$
\frac{\delta <Y>_{ij}}{\delta Y_{ij}} = K_{ij} \quad (A.3)
$$
where $K_{ij}$ is a $p \times q$ matrix with zeroes everywhere except for a unit element in the $i$-th row and $j$-th column.

Similarly

$$\frac{\delta <Y>_{ij}}{\delta Y} = K'_{ji}$$  \hspace{1cm} (A.4)

For convenience, the subscripts will be dropped. For example, $<X>$ will denote an arbitrary element of $X$ and $J$ a matrix with all elements zero except a unit element on the appropriate place determined by the location of $<X>$.

A.1.2. Differentiation of a matrix with respect to a scalar and of a scalar with respect to a matrix

Let $Y(a)$ be a matrix function of the scalar $a$. The derivative

$$\frac{\delta Y(a)}{\delta a}$$  \hspace{1cm} (A.5)

is a matrix with elements $\frac{\delta y_{ij}}{\delta a}$. Each element of $Y(a)$ is differentiated.

The derivative of a matrix function with respect to a matrix is denoted by

$$\frac{\delta f(X)}{\delta X}$$  \hspace{1cm} (A.6)
and is a matrix with elements

\[ \frac{\delta f(X)}{\delta <X>_{ij}} \]. \quad (A.7)

Two important matrix functions are considered: the determinant and the trace. We begin with the assumption that \( X \) is a square matrix.

a. Determinant

The determinant of the square matrix \( X \) can be evaluated in terms of the cofactors of the elements of the \( i \)-th row (Rogers, 1971; p. 81):

\[ |X| = x_{i1} x_{i1}^c + x_{i2} x_{i2}^c + \cdots + x_{iN} x_{iN}^c. \]

It can easily be seen that

\[ \frac{\delta |X|}{\delta <X>_{ij}} = x_{ij}^c \quad (A.8) \]

where \( x_{ij}^c \) is the cofactor of the element \( |X|_{ij} \). And

\[ \frac{\delta |X|}{\delta X} = \text{cof} \, X = [\text{adj} \, X]' \]

where \( \text{cof} \, X \) is the matrix of cofactors, and \( \text{adj} \, X \) is the adjoint matrix of the matrix \( X \). But if \( X \) is nonsingular,

\[ \text{cof} \, X = |X| [X']^{-1}. \quad (A.9) \]
Equation (A.8) may be written as

$$\frac{\delta |X|}{\delta X} = |X||X'|^{-1}.$$  \hfill (A.10)

This formula is well known in matrix theory and can also be found in Bellman (1970; p. 182).

It should be noted that if $X$ is symmetric

$$\frac{\delta |X|}{\delta <X>_{ij}} = 2X^C_{ij} \quad \text{for } i \neq j$$

$$= X^C_{ij} \quad \text{for } i = j$$ \hfill (A.11)

b. Trace

The trace of the square matrix $X$ is the sum of its diagonal elements, and

$$\frac{\delta \text{tr}(X)}{\delta <X>_{ij}} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}$$ \hfill (A.12)

with

$$\frac{\delta \text{tr}(X)}{\delta X} = \mathbf{I}$$

where $\mathbf{I}$ is the identity matrix.
A.1.3. Differentiation of matrix products

Let \( \tilde{U} \) and \( \tilde{V} \) be two matrix functions of the matrix \( \tilde{X} \). The derivative of their product \( \tilde{Y} = \tilde{U} \tilde{V} \) with respect to \( \tilde{X} \) is

\[
\frac{\delta \tilde{Y}}{\delta \tilde{X}} = \frac{\delta [\tilde{U} \tilde{V}]}{\delta \tilde{X}} = \frac{\delta \tilde{U}}{\delta \tilde{X}} \tilde{V} + \tilde{U} \frac{\delta \tilde{V}}{\delta \tilde{X}}. \tag{A.13}
\]

The derivative of a product of three matrices is

\[
\frac{\delta \tilde{Y}}{\delta \tilde{X}} = \frac{\delta [\tilde{U} \tilde{V} \tilde{W}]}{\delta \tilde{X}} = \frac{\delta \tilde{U}}{\delta \tilde{X}} \tilde{V} \tilde{W} + \tilde{U} \frac{\delta \tilde{V}}{\delta \tilde{X}} \tilde{W} + \tilde{U} \tilde{V} \frac{\delta \tilde{W}}{\delta \tilde{X}}. \tag{A.14}
\]

These general formulas may be applied to various cases. Some cases of interest are listed below. The matrices \( \tilde{A} \) and \( \tilde{B} \) are constant, i.e. independent of \( \tilde{X} \). The matrices \( \tilde{J} \), and \( \tilde{K} \) are as defined in A.1.1.

\[
\begin{align*}
\frac{\delta \tilde{Y}}{\delta \tilde{X}} &= \tilde{A} \tilde{J} + \tilde{J} \tilde{A} + \tilde{K} \tilde{X} + \tilde{X} \tilde{K}.
\end{align*}
\]
\[ \frac{\delta [X^n]}{\delta <X>} \approx JX^{n-1} + \sum_{s=1}^{n-2} X^s JX^{n-1-s} + X^{n-1} J \]

(A.23)

or, if we write \( X^0 = I \), then

\[ \frac{\delta [X^n]}{\delta <X>} = \sum_{s=0}^{n-1} X^s JX^{n-1-s} \]

(A.24)

The derivative of the power of a square matrix can readily be computed using these formulas.

The derivative of an inverse follows. By definition

\[ XX^{-1} = I \]

Therefore

\[ \frac{\delta [XX^{-1}]}{\delta <X>} = \frac{\delta I}{\delta <X>} = 0 \]

but

\[ \frac{\delta [XX^{-1}]}{\delta <X>} = \frac{\delta X}{\delta <X>} X^{-1} + X \frac{\delta [X^{-1}]}{\delta <X>} \]
It follows that

\[
\frac{\delta [X^{-1}]}{\delta <X>} = -X^{-1} J X^{-1} \quad .
\]  
(A.25)

An application of this result is

\[
\frac{\delta XAX^{-1}}{\delta <X>} = JAX^{-1} - XAX^{-1} JX^{-1} \quad .
\]  
(A.26)

So far we have considered the derivative \( \frac{\delta Y}{\delta <X>} \) where \( Y \) is a matrix product and \( <X> \) is an arbitrary element of \( X \). The result is a matrix of partial derivatives. But what is the formula for \( \frac{\delta Y}{\delta X} \), where \( X \) represents the full matrix? This question has been studied by Neudecker (1969). Its solution involves the transformation of a matrix into a vector and the use of Kronecker products. For example, let \( Y = AXB \) and one is interested in the derivative of \( Y \) with respect to \( X \).

If \( Y \) is of order \( P \times Q \), define the \( PQ \) column vector \( \text{vec } Y \) (denoted this way to distinguish it from the vector \( \{ y_i \} \)) where

\[
\text{vec } Y = \begin{bmatrix}
\{ y_{\cdot 1} \} \\
\{ y_{\cdot 2} \} \\
\vdots \\
\{ y_{\cdot Q} \}
\end{bmatrix} \quad .
\]
In a similar way, one can construct vec \( \tilde{X} \). Neudecker shows that

\[
\text{vec} \ (\text{AXB}) = [B' \otimes A] \text{ vec } \tilde{X} \tag{A.27}
\]

where \( \otimes \) denotes the Kronecker product. Equation (A.27) may be differentiated using the formulas for vector differentiation:

\[
\frac{\delta \text{ vec } [\text{AXB}]}{\delta \text{ vec } \tilde{X}} = [B' \otimes A]' .
\]

Since the transpose of a Kronecker product is the Kronecker product of the transposes, we have\(^3\)

\[
\frac{\delta \text{ vec } [\text{AXB}]}{\delta \text{ vec } \tilde{X}} = B \otimes A' . \tag{A.28}
\]

We will not explore the various formulas for \( \frac{\delta Y}{\delta \tilde{X}} \) further since they are not explicitly used in this study.

A.1.4. Chain rules of differentiation

Let \( f(\tilde{Y}) \) be a scalar function of \( \tilde{Y} \) and let \( \tilde{Y} \) be a matrix function of \( \tilde{X} \).

\(^3\)For an exposition of the properties of Kronecker products or direct products, see Lancaster (1969; pp. 256-259).
Then

$$\frac{\delta f(Y)}{\delta <X>} = \frac{\delta f(Y)}{\delta <\sim>_{k\ell}} \cdot \frac{\delta <Y>_{\sim}}{\delta <X>}$$  \hspace{1cm} (A.29)

$$\frac{\delta f(Y)}{\delta <X>} = \text{tr} \left[ \frac{\delta f(Y)}{\delta _{~y}^{~}} \cdot \frac{\delta y'}{\delta _{~x}^{~}} \right].$$  \hspace{1cm} (A.30)

If \( Y \) is a matrix function of a scalar \( a \), i.e. \( Y(a) \), the formula becomes

$$\frac{\delta f(Y)}{\delta a} = \text{tr} \left[ \frac{\delta f(Y)}{\delta _{~y}^{~}} \cdot \frac{\delta y'}{\delta a} \right].$$  \hspace{1cm} (A.31)

Consider also the derivative

$$\frac{\delta f(Y)}{\delta X} = \frac{\delta f(Y)}{\delta _{~y}^{~}} \cdot \frac{\delta <Y>_{\sim}}{\delta X}.$$  \hspace{1cm} (A.32)

Several interesting applications arise. For example, let

\( f(Y) = |X - \lambda I| \), where \( X \) may be the population growth matrix. Then

$$\frac{\delta |X - \lambda I|}{\delta <X>} = \text{tr} \left[ \frac{\delta |X - \lambda I|}{\delta _{~y}^{~}} \cdot \frac{\delta [X - \lambda I]'}{\delta _{~y}^{~}} \right].$$

$$\frac{\delta |X - \lambda I|}{\delta <X>} = |X - \lambda I| \text{ tr} \left( [X - \lambda I]^{-1} J' \right).$$  \hspace{1cm} (A.33)
\[
= \text{tr} \left[ [\text{cof}(\tilde{X} - \lambda \mathbb{I})] \tilde{J}' \right]
\]

and

\[
\frac{\delta |X - \lambda \mathbb{I}|}{\delta X} = \sum_{k \ell} \frac{\delta |X - \lambda \mathbb{I}|}{\delta [X - \lambda \mathbb{I}]_{k \ell}} \cdot \frac{\delta [X - \lambda \mathbb{I}]'_{k \ell}}{\delta X}
\]

(A.34)

\[
= \sum_{k \ell} |X - \lambda \mathbb{I}| \cdot [X - \lambda \mathbb{I}]_{k \ell}^{-1} \quad [X - \lambda \mathbb{I}]_{k \ell} \quad \tilde{J}' \quad \ell
\]

\[
\frac{\delta |X - \lambda \mathbb{I}|}{\delta X} = |X - \lambda \mathbb{I}| \cdot [X - \lambda \mathbb{I}]_{k \ell}^{-1} = \text{cof} [X - \lambda \mathbb{I}] \quad (A.35)
\]

where \( \text{cof} [X - \lambda \mathbb{I}] \) is the cofactor matrix of \( [X - \lambda \mathbb{I}] \).

If \( \tilde{Y}(r) \) is a function of the scalar \( r \), then

\[
\frac{\delta |\tilde{Y}(r)|}{\delta r} = \text{tr} \left[ \frac{\delta |\tilde{Y}(r)|}{\delta [\tilde{Y}(r)]} \cdot \frac{\delta [\tilde{Y}(r)]'}{\delta r} \right]
\]

\[
= \text{tr} \left[ |\tilde{Y}(r)| \cdot [\tilde{Y}(r)]'^{-1} \cdot \frac{\delta [\tilde{Y}(r)]'}{\delta r} \right]
\]

and since \( \text{tr} \tilde{A} \tilde{B} = \text{tr}[\tilde{A}\tilde{B}]' = \text{tr} \tilde{B}'\tilde{A}' \)

\[
\frac{\delta |\tilde{Y}(r)|}{\delta r} = |\tilde{Y}(r)| \cdot \text{tr} \left[ \frac{\delta [\tilde{Y}(r)]}{\delta r} \cdot [\tilde{Y}(r)]^{-1} \right]
\]

(A.36)
Formula (A.36) is not only of interest in a study of the sensitivity of the determinant of a polynomial matrix, but is also useful in order to compute the determinant, as shown by Emre and Hüseyin (1975; p. 136). An application of (A.36) which is relevant is

\[
\frac{\delta |A - \lambda I|}{\delta \lambda} = - \frac{|A - \lambda I| \: \text{tr}(A - \lambda I)^{-1}}{\delta \lambda} . \tag{A.37}
\]

This formula can also be found in Newbery (1974; p. 1016). Finally, consider the application, where \( f(Y) = \text{tr}(AXB) \), whence

\[
\frac{\delta f(Y)}{\delta X} = \frac{\delta \text{tr}(AXB)}{\delta X} = A'B' . \tag{A.38}
\]

A.1.5. Vector differentiation

Vectors may be considered as matrices with only one row or one column, and the rules for matrix differentiation may be applied. But the derivative of a vector or of a vector equation has a simpler form than the matrix analogue. It is, therefore, worthwhile to list the formulas for vector differentiation separately. Two cases are considered: the derivative of a scalar function with respect to a vector and the derivative of a vector function with respect to a vector.
a. Differentiation of a scalar function with respect to a vector

Consider the general scalar function \( f(\{\tilde{x}\}) \), where \( \{\tilde{x}\} \) is the argument vector. Some relevant formulations of \( f(\{\tilde{x}\}) \) and their derivatives are listed below.

\[
\begin{align*}
\frac{\delta f(\{\tilde{x}\})}{\delta \{\tilde{x}\}} \\
\{a\}'\{\tilde{x}\} & \to \{a\} \\
\{x\}'\{\tilde{x}\} & \to 2\{x\} \\
\{x\}'A\{x\} & \to A\{x\} + A'\{x\}
\end{align*}
\]

(A.39) 
(A.40) 
(A.41)

b. Differentiation of a vector function with respect to a vector

Let \( \{f(\{\tilde{x}\})\} \) denote a column vector of scalar functions \( f_i(\{\tilde{x}\}) \), where \( \{\tilde{x}\} \) is the argument vector and \( \{f(\{\tilde{x}\})\} \) represents a system of equations. For example, let \( \{f(\{\tilde{x}\})\} \) be a system of linear equations in \( \{\tilde{x}\} \), then

\[
\frac{\delta A\{x\}}{\delta \{\tilde{x}\}_i} = \{a_i\}
\]

(A.42)

where \( \{a_i\} \) is the \( i \)-th column of \( A \).

The derivatives of \( \{f(\{\tilde{x}\})\} \) with respect to all the elements of the argument vector form a matrix if the argument
vector is a row vector. For example

\[
\frac{\delta A(x)}{\delta \{x\}'} = A
\]  \hspace{1cm} (A.43)

The determinant \[
\left| \frac{\delta \{f(\{x\})\}}{\delta \{x\}'} \right|
\] is known as the Jacobian or functional determinant.

Corresponding to the chain rule of matrix differentiation, one may formulate the chain rule of vector differentiation. Let \{y\}, \{x\} and \{z\} be vectors. It can be shown that

\[
\frac{\delta \{y\}}{\delta \{x\}'} = \frac{\delta \{y\}}{\delta \{z\}'} \cdot \frac{\delta \{z\}}{\delta \{x\}'}.
\]  \hspace{1cm} (A.44)

A.2. DIFFERENTIATION OF EIGENVALUES AND EIGENVECTORS OF MATRICES

The topic of eigenvalue sensitivity has received most attention in the engineering literature. The design engineer is interested in identifying the impact of changes in the parameters of a system on the system's performance. There is a vast literature on sensitivity analysis in design\(^4\). Although most of this literature is not related to the problem in this study, some relevant elements are repeated here. We will separate the eigenvalue sensitivity problem and the eigenvector

\[^4\text{See Cruz (1973) and Tomović and Vukobratović (1972) for example.}\]
sensitivity problem. The former has received considerable attention, while the latter has been very much neglected.

A.2.1. Differentiation of the eigenvalue with respect to the matrix elements

The method which follows is described by Faddeev and Faddeeva (1963; p. 229) and can also be found in Van Ness et al. (1973; p. 100) and in Tomović and Vukobratović (1972; pp. 196-197). The assumption underlying the method is that all the eigenvalues of the matrix are distinct. Let \( A \) be such a matrix. Consider the equation

\[
A \{ \xi \}_i = \lambda_i \{ \xi \}_i \quad (A.45)
\]

where \( \lambda_i \) is the \( i \)-th eigenvalue of \( A \) and \( \{ \xi \}_i \) is the right eigenvector associated with \( \lambda_i \).

Taking the partial derivatives of both sides with respect to an element of \( A \), \( \langle A \rangle \) say, gives

\[
\frac{\delta A}{\delta \langle A \rangle} \{ \xi \}_i + \frac{\delta \{ \xi \}_i}{\delta \langle A \rangle} = \frac{\delta \lambda_i}{\delta \langle A \rangle} \{ \xi \}_i + \lambda_i \frac{\delta \{ \xi \}_i}{\delta \langle A \rangle} \quad (A.46)
\]

If the real matrix \( A \) is transposed, the eigenvalues will not change. However, a new set of eigenvectors will be formed: the left eigenvectors, denoted by \( \{ y \}_j \). The scalar product of each of the terms of (A.46) with \( \{ y \}_j \) is:
\[(\mathcal{J}_{\xi_i}, \{\nu\}_j) + \left(\mathcal{A} - \frac{\delta \xi_i}{\delta \langle \xi \rangle}, \{\nu\}_j\right) =\]

\[\lambda_i \left(\frac{\delta \xi_i}{\delta \langle \xi \rangle}, \{\nu\}_j\right) + \frac{\delta \lambda_i}{\delta \langle \xi \rangle} \left(\xi_i, \{\nu\}_j\right)\]

(A.47)

where \( \mathcal{J} \) has the same meaning as in section A.1. If \( i \) is taken equal to \( j \), and use is made of the relationship

\[\mathcal{A}'\{\nu\}_j = \lambda_j \{\nu\}_j\]

(A.48)

then (A.47) becomes

\[\left(\mathcal{J}_{\xi_i}, \{\nu\}_i\right) + \left(\mathcal{A} - \frac{\delta \xi_i}{\delta \langle \xi \rangle}, \{\nu\}_i\right) =\]

\[\left(\frac{\delta \xi_i}{\delta \langle \xi \rangle}, \mathcal{A}'\{\nu\}_i\right) + \frac{\delta \lambda_i}{\delta \langle \xi \rangle} \left(\xi_i, \{\nu\}_i\right).\]

(A.49)

Since

\[\left(\mathcal{A} - \frac{\delta \xi_i}{\delta \langle \xi \rangle}, \{\nu\}_i\right) = \left(\frac{\delta \xi_i}{\delta \langle \xi \rangle}, \mathcal{A}'\{\nu\}_i\right),\]
we may write

\[
\frac{\delta \lambda_i}{\delta <A>} = \left( \{J\{\xi\}_i, \{\psi\}_i\} \right) \quad \text{(A.50)}
\]

Expression (A.50) represents the sensitivity of the eigenvalues of \( A \) with respect to an element of \( \tilde{A} \).

If the eigenvectors are normalized such that their inner product is unity, i.e.

\[
\left( \{\xi\}_i , \{\psi\}_i \right) = 1
\]

then

\[
\frac{\delta \lambda_i}{\delta <A>} = \left( J\{\xi\}_i, \{\psi\}_i \right) = \left( \{\psi\}_i , J\{\xi\}_i \right) \quad \text{(A.51)}
\]

It can be shown that (A.51) is equivalent to

\[
\frac{\delta \lambda_i}{\delta <A>} = \text{tr} \left[ \{\xi\}_i \{\psi\}_i ' \right] J \quad \text{(A.52)}
\]

or

\[
\frac{\delta \lambda_i}{\delta <A>} = \{\xi\}_i \{\psi\}_i ' * J \quad \text{(A.53)}
\]

where * denotes the inner product of two matrices\(^5\).

---

\(^5\)The inner product \( A * B \) is defined as \( \sum_i \sum_k a_{ik} b_{ki} \).

The result is equal to \( \text{tr}[AB] \).
The structure of (A.52) is very similar to (A.33) of the previous section. The derivative of $\lambda_i$ with respect to the whole matrix $\mathbf{A}$ is

$$
\frac{\delta \lambda_i}{\delta \mathbf{A}} = \{\bar{\xi}_i\}_i \{\bar{\nu}_i\}_i \ .
$$

(A.54)

The matrix $\{\bar{\xi}_i\}_i \{\bar{\nu}_i\}_i$ is the adjoint matrix of $[\mathbf{A} - \lambda I]$, normalized such that the trace is equal to one\(^6\). The sensitivity of the eigenvalue is sometimes expressed in terms of differentials

$$
d\lambda_i = \{\bar{\nu}_i\}_i \ [d\mathbf{A}] \{\bar{\xi}_i\}
$$

(A.55)

or

$$
d\lambda_i = \left[\{\bar{\xi}_i\}_i \{\bar{\nu}_i\}_i\right]^* \ d\mathbf{A} \ .
$$

(A.56)

The computation of the sensitivity of $\lambda_i$ requires that the left and right eigenvectors be known.

If the eigenvectors are not normalized, the sensitivity function is

$$
\frac{\delta \lambda_i}{\delta \mathbf{A}} = \frac{1}{\{\bar{\nu}_i\}_i \{\bar{\xi}_i\}_i} \left[\{\bar{\xi}_i\}_i \{\bar{\nu}_i\}_i\right]
$$

(A.57)

\(^6\) $\text{tr}[[\bar{\xi}_i]_i [\bar{\nu}_i]_i]$ is equal to $\{\bar{\nu}_i\}_i \{\bar{\xi}_i\}_i$ which is equal to one for normalized $v$ eigenvectors.
where \([\{\xi\}_i \{\nu\}_i]')\) is the adjoint matrix of \([\tilde{A} - \lambda \mathbf{1}].\) Denoting the adjoint matrix by \(\tilde{R}(\lambda_i), \) (A. 51) may be written as

\[
\frac{\delta \lambda_i}{\delta \tilde{A}} = \left[ \text{tr} \tilde{R}(\lambda_i) \right]^{-1} \tilde{R}(\lambda_i) \tag{A.58}
\]

and (A.56) becomes

\[
\delta \lambda_i = \left[ \text{tr} \tilde{R}(\lambda_i) \right]^{-1} \tilde{R}(\lambda_i) * d\tilde{A} \tag{A.59}
\]

Equation (A.59) is exactly the sensitivity formula given by Morgan (1973; p. 76). The matrix \(\tilde{R}(\lambda_i)\) can be efficiently computed by means of the Leverrier algorithm, described by Faddeev and Faddeeva (1963; p. 260) and Morgan (1973; p. 76). This is particularly interesting since the rows of \(\tilde{R}(\lambda_i)\) are left eigenvectors and the columns are right eigenvectors. For a formal proof that (A.59) is identical to (A.56), see Mac Farlane (1970; pp. 413-419).

Formulas (A.54) and (A.58) have the benefit that they are easily computed. For analytical purposes, however, it would be beneficial to have an expression linking the change in the eigenvalue directly to a change in \(\tilde{A},\) and to the original value of \(\tilde{A}\) and of the eigenvalues. Such an expression is derived by Rosenbrock (1965; p. 278):
\[ d\lambda_i = \frac{\text{tr} \left[ \prod_{r \neq i} (A - \lambda_i I) \, dA \right]}{\prod_{r \neq i} (\lambda_i - \lambda_r)} \]  \hspace{1cm} (A.60)

### A.2.2. Differentiation of the eigenvector with respect to the matrix elements

Recall equation (A.47):

\[
\left( \mathcal{J}^\star \{ \xi \}_i, \{ \psi \}_j \right) + \left( A \frac{\delta \{ \xi \}_i}{\delta <A>} , \{ \psi \}_j \right) = \\
\lambda_i \left( \frac{\delta \{ \xi \}_i}{\delta <A>} , \{ \psi \}_j \right) + \frac{\delta \lambda_i}{\delta <A>} \left( \frac{\delta \{ \xi \}_i}{\delta <A>} , \{ \psi \}_j \right)
\]

(A.47)

For \( i \neq j \), we have

\[
\left( \{ \xi \}_i , \{ \psi \}_j \right) = 0
\]

We have also that

\[
\left( A \frac{\delta \{ \xi \}_i}{\delta <A>} , \{ \psi \}_j \right) = \left( A' \frac{\delta \{ \xi \}_i}{\delta <A>} , \{ \psi \}_j \right) = \lambda_j \left( \frac{\delta \{ \xi \}_i}{\delta <A>} , \{ \psi \}_j \right)
\]

Equation (A.47) may be rewritten as

\[
\left( \mathcal{J}^\star \{ \xi \}_i , \{ \psi \}_j \right) = (\lambda_i - \lambda_j) \left[ \frac{\delta \{ \xi \}_i}{\delta <A>} , \{ \psi \}_j \right]
\]

\[
\left( \frac{\delta \{ \xi \}_i}{\delta <A>} , \{ \psi \}_j \right) = \frac{\left( \mathcal{J}^\star \{ \xi \}_i , \{ \psi \}_j \right)}{(\lambda_i - \lambda_j)}
\]

(A.61)
Let \[ \frac{\delta \{\xi\}_i}{\delta \{\xi\}_j} = \sum_{j=1}^{N} c_{ij} \{\xi\}_j \] \hspace{1cm} (A.62)

then

\[ \left( \frac{\delta \{\xi\}_i}{\delta \{\xi\}_j}, \{y\}_j \right) = c_{ij} \left( \{\xi\}_j, \{y\}_j \right) \]

and consequently, for normalized eigenvectors

\[ c_{ij} = \frac{\langle y \{\xi\}_i, \{y\}_j \rangle}{\lambda_i - \lambda_j} \quad \text{for } i \neq j \] \hspace{1cm} (A.63)

The element \( c_{ii} \) remains undefined in view of the non-uniqueness of the eigenvector. We may assume that \( c_{ii} = 0 \) without loss of generality.

The computation of the sensitivity of the eigenvector by (A.62) has a disadvantage, since it requires the knowledge of all the eigenvalues and eigenvectors. Another approach that relates the change in a specific eigenvector to the change in \( \tilde{A} \) and to the change in the associated eigenvalue, is given below. Consider the homogeneous equation

\[ [\tilde{A} - \lambda_i I] \{\xi\}_i = \{0\} \] \hspace{1cm} (A.64)

Assume that all the eigenvalues of \( \tilde{A} \) are distinct, and let the first element of \( \{\xi\}_i \), i.e. \( \xi_{1i} \), be equal to 1. We may
now delete the first equation of (A.64). The resulting set forms a linearly independent system of non-homogenous equations of order N-1.

\[
\begin{bmatrix}
a_{21} \\
a_{31} \\
\vdots \\
a_{N1}
\end{bmatrix} + 
\begin{bmatrix}
a_{22} - \lambda_i \\
a_{32} \\
\vdots \\
a_{N2}
\end{bmatrix}
\begin{bmatrix}
a_{23} \\
a_{33} - \lambda_i \\
\vdots \\
a_{N3}
\end{bmatrix}
\begin{bmatrix}
a_{N1} \\
\vdots \\
\vdots \\
a_{NN} - \lambda_i
\end{bmatrix}
\begin{bmatrix}
\xi_{2i} \\
\xi_{3i} \\
\vdots \\
\xi_{Ni}
\end{bmatrix} = 
\begin{bmatrix}
0 \\
0 \\
\vdots \\
0
\end{bmatrix}
\]

or in matrix notation

\[
\{\tilde{a}_1\} + [\tilde{A} - \lambda_i \tilde{I}]\{\tilde{\xi}\}_i = \{\tilde{0}\}
\]  \hspace{1cm} (A.65)

where the bar denotes the order N-1. Because of the non-singularity of \([\tilde{A} - \lambda_i \tilde{I}]\), we have

\[
\{\tilde{\xi}\}_i = - [\tilde{A} - \lambda_i \tilde{I}]^{-1} \{\tilde{a}_1\}
\]  \hspace{1cm} (A.66)

Applying formula (A.13) of section A.1. to (A.66) gives

\[
\frac{\delta\{\tilde{\xi}\}_i}{\delta A} = - \frac{\delta [A - \lambda \tilde{I}]^{-1}}{\delta A} \{\tilde{a}_1\} - [A - \lambda \tilde{I}]^{-1} \frac{\delta \{\tilde{a}_1\}}{\delta A}
\]

\[
= [\tilde{A} - \lambda_i \tilde{I}]^{-1} \frac{\delta [\tilde{A} - \lambda_i \tilde{I}]}{\delta A} [\tilde{A} - \lambda_i \tilde{I}]^{-1} \{\tilde{a}_1\}
\]

\[
- [\tilde{A} - \lambda_i \tilde{I}] \frac{\delta \{\tilde{a}_1\}}{\delta A}
\]
Substituting for \( \{ \tilde{v} \} \) and differentiating \( [ \tilde{A} - \lambda_i \tilde{I} ] \) yields

\[
\frac{\delta \{ \tilde{v} \}}{\delta < \tilde{A} >} = - [ \tilde{A} - \lambda_i \tilde{I} ] \left[ \left[ \frac{\delta A}{\delta < A >} - \frac{\delta \lambda_i}{\delta < A >} \tilde{I} \right] \{ \tilde{v} \} + \frac{\delta \{ \tilde{a}_i \}}{\delta < A >} \right] \tag{A.67}
\]

where \( \frac{\delta \lambda_i}{\delta < A >} \) is computed using (A.51) or an equivalent formula.

Some special cases now may be considered.

a. If the change in \( \tilde{A} \) occurs in the first row, this change has no direct impact on the eigenvector, since \( \tilde{A} \) and \( \{ \tilde{v} \} \) do not include elements of the first row of \( \tilde{A} \). There is an indirect effect on \( \{ \tilde{v} \} \), however, through the change in the eigenvalue.

b. If the change in \( \tilde{A} \) occurs in the first column, i.e. in \( \{ \tilde{a}_1 \} \), then

\[
\frac{\delta \tilde{A}}{\delta < \tilde{A} >} = 0 .
\]

c. If the change in \( \tilde{A} \) occurs not in the first column nor in the first row, then

\[
\frac{\delta \{ \tilde{a}_i \}}{\delta < \tilde{A} >} = 0 .
\]
Besides (A.62) and (A.67), a third method to compute the eigenvector sensitivity may be derived. It is based on the fact that the columns of the adjoint matrix are right eigenvectors and that the rows are left eigenvectors. This technique will not be discussed here.
References


PART III

OPTIMAL MIGRATION POLICIES

In recent years, there has been an increasing interest in the dynamics of spatial demographic growth. Models for multiregional population growth have been developed to describe the growth process and to analyze its impact on future population characteristics (Rogers, 1975). The various economic, social, climatological and cultural forces influencing spatial population growth have been brought together in explanatory demometric models (Greenwood, 1975a). The mathematical demographic models and the demometric models have a common feature. They are designed to describe and to explain the dynamics of the spatial population growth.

Once the dynamics of a phenomena are understood, human nature comes up with the ultimate question: can we control it and how. The models associated with this third concern are population policy models. The subject of migration policy models has been treated by Rogers (1966; 1968, Chapter 6; 1971, pp. 98-108), and more recently, MacKinnon (1975a, 1975b) devotes considerable attention to the design of optimal-seeking migration policy models.

This part of the study is devoted to a methodological analysis of migration policy models. We assume that a demometric or a demographic model, consisting of a system of linear simultaneous equations, has been successfully specified and estimated. Therefore, we do not devote any attention, for example, to identification and estimation procedures. The main thread of the analysis is provided
by the Tinbergen paradigm, to which we will refer frequently. Chapter 5 is a conceptual survey of the various possible policy models. Every model is related back to the original Tinbergen framework. The matrix of impact multipliers, well known in economic analysis, is seen to be of crucial importance to the classification scheme. After the introductory chapter has set the scene, we devote our attention to the two central issues in the theory of policy: the concepts of existence and of design. The existence problem deals with the question whether the system is controllable, i.e., whether a set of arbitrary targets can be achieved at all, given the internal dynamics of the system and given the set of available instruments. The answer to the controllability problem provides input information for the design problem. For the design of an optimal policy, the policy maker may apply a wide range of mathematical programming techniques, given that he has a clear idea of his preferences. To facilitate the discussion of the controllability of dynamic systems in Chapter 6 and of the design of optimal policies in Chapter 7, we introduce in Chapter 6 the state-space representation of demometric models.
CHAPTER 5
OPTIMAL MIGRATION POLICIES:
A CONCEPTUAL FRAMEWORK

There are several analytical differences between a policy model and a conventional demographic or demometric model. The most basic classification of variables in any model consists of two categories: endogenous variables, which are determined within the model, and exogenous variables, which are predetermined. Suppose the population system is linear and may be modeled as

\[ \tilde{\Lambda}\{\tilde{y}\} = \tilde{E}\{\tilde{z}\} \]  

(5.1)

where \{\tilde{y}\} is a \(M \times 1\) vector of endogenous variables,
\{\tilde{z}\} is a \(L \times 1\) vector of exogenous variables,
\(\tilde{\Lambda}\) is a \(M \times M\) matrix of coefficients,
\(\tilde{E}\) is a \(M \times L\) matrix of coefficients.

Equation (5.1) is the reduced form of a population model. The endogenous and the exogenous variables are separated. Assuming that \(\tilde{\Lambda}\) is nonsingular, we obtain

\[ \{\tilde{y}\} = \tilde{\Lambda}^{-1}\tilde{E}\{\tilde{z}\} = \tilde{C}\{\tilde{z}\} \]  

(5.2)

where \(\tilde{C}\) is the matrix of multipliers, i.e. the reduced form matrix. The elements of \(\tilde{C}\) represent the impact on \{\tilde{y}\} of a unit change in \{\tilde{z}\}.

The policy models treated here, will be discussed with reference to (5.2). Tinbergen (1963) proposed a classification of the variables of (5.2) better suited for the
policy problem. His ideas are general enough to encompass the whole range of policy models. Starting from the Tinbergen paradigm, we try to present a unified treatment of various classes of models, which are relevant for population policy.

5.1. THE TINBERGEN PARADIGM

Tinbergen (1963) distinguished two categories of variables in both the endogenous and the exogenous variables. The endogenous variables consist of target variables, which are of direct interest for policy purposes, and other variables which are not. The latter are labeled by Tinbergen as irrelevant variables. However, they may be of indirect interest for policy planning, since their values may in turn influence various target variables. The exogenous variables are divided according to their controllability. Instrument variables are subject to direct control by the policy authorities. Data variables are beyond their control. The latter include exogenously predetermined and uncontrollable variables, as well as lagged endogenous variables. They define the environment in which the levels of instrument variables have to be set. Applying this approach, equation (5.2) may be partitioned to give

\[
\begin{bmatrix}
\{y_1\} \\
\{y_2\}
\end{bmatrix} =
\begin{bmatrix}
R & S \\
P & Q
\end{bmatrix}
\begin{bmatrix}
\{z_1\} \\
\{z_2\}
\end{bmatrix}
\]

where \(\{y_1\}\) is the \(N \times 1\) vector of target variables,
\(\{y_2\}\) is the \((M - N) \times 1\) vector of other endogenous variables,
\(\{z_1\}\) is the \(K \times 1\) vector of instrument variables,
\(\{z_2\}\) is the \((L - K) \times 1\) vector of uncontrollable
exogenous variables and lagged endogenous variables,
\( \tilde{z}, \tilde{z}, \tilde{z}, \tilde{Q} \) are conformable partitions of the model's reduced form matrix.

The value of the target vector is

\[
\{y_1\} = \tilde{R}\{z_1\} + \tilde{S}\{z_2\}.
\] (5.3)

The policy problem, as formulated by Tinbergen, is to choose an appropriate value of the instrument vector \( \{z_1\} \) so as to render the value of the target vector \( \{y_1\} \) equal to some previously established desired value \( \{\tilde{y}_1\} \). The choice of the level of the instrument variables depends on the levels of the uncontrollable variables, represented by \( \{z_2\} \), and on how much they affect the targets.

It is important to keep in mind that the policy model (5.3) is derived from the explanatory model (5.2) by adding a new dimension to (5.2). This new dimension is the goals-means relationship of population policy. The explanatory model may be a pure demographic model, relating population growth and distribution to demographic factors such as fertility, mortality and migration. It may also be a demometric model, which statistically relates spatial population growth to socio-economic variables. Any model may be converted into a policy model if and only if all the target variables of the policy model are part of the set of endogenous variables of the explanatory model and if at least one of the exogenous variables is controllable. Most migration models found in the literature are single-equation models with gross or net
migration as the dependent variable. They serve only a restricted category of policy models, namely those with targets that consist of migration levels and instruments which are socio-economic in nature. Various regional economic models include migration as an exogenous variable. Therefore, they are not suited to become migration policy models if population distribution is the goal. Simultaneous equation models, such as the ones developed by Greenwood (1973, 1975b) and Olvey (1972), are relevant to model population policy problems of all types, because they include demographic and socio-economic variables in both the set of endogenous and the set of exogenous variables. Thus they may be applied in situations where the goals-means relationship consists of demographic, as well as of socio-economic measures. Finally, the multiregional population growth models of Rogers (1975) may be converted to policy models to study purely demographic policy problems, i.e., both targets and instruments are demographic in nature.

Before going into greater detail in our exposition, we would like to stress that the analytical solution of Tinbergen's formulation of the policy problem is restricted to linear policy models. If the model is nonlinear, one can only solve it numerically. The latter approach is denoted by Naylor (1970; p. 263) as the simulation approach, and has been applied extensively by Fromm and Taubman (1968). In this part, we only deal with linear models and do not discuss the simulation approach.
5.2. SURVEY OF POLICY MODELS

Conceptually, any policy model may be related to (5.3). For convenience, we drop the subscript of the target vector.

\[ \{y\} = D(z_1) + S(z_2) \]

(5.3)

Throughout our discussion of policy models, it will be assumed that both the targets and the instruments are linearly independent. The matrix \( D \) then plays a crucial role in policy analysis. The existence of an optimal policy, i.e., a solution to (5.3), depends on the rank of \( D \). The design of an optimal policy, i.e., the assignment of values to the instrument variables, depends on the structure of \( D \), and on the values of its entries. The matrix \( D \) is known in the economic literature as the matrix of impact multipliers. The name refers to the fact that an element \( r_{ij} \) gives the change in the value of the target variable \( i \) when the instrument variable \( j \) is varied by one unit. The ratio \( -r_{ij}/r_{ik} \) is the amount by which the \( j \)-th instrument may be cut down without changing the level of the \( i \)-th target, if the value of the \( k \)-th instrument is increased with one unit. It is, therefore, the marginal rate of substitution between the two instruments (Fromm and Taubman, 1968; p. 109).

It is the purpose of this section to classify relevant policy models without going into technical detail. Detailed treatment will be given later. The survey revolves around the matrix multiplier \( D \) and its characteristics. A first
classification scheme is based on the rank of $R$, or alternatively on the relation between the number of targets and the number of instruments. A second classification scheme relates to the structure of $R$. The structure of $R$ also provides us with a link between the reduced form models and the models of optimal control.

5.2.1. Classification of Policy Models According to the Rank of the Matrix Multiplier

We may distinguish between three categories of policy models: $R$ is nonsingular and of rank $N$; $R$ is singular and of rank $K$; $R$ is singular and of rank $N$. The parameters $N$ and $K$ are, respectively, the number of instruments and the number of targets. An illustration is given by a typical policy model, namely the Theil (1964) model.

a. The matrix multiplier is nonsingular and of rank $N$.

If $R$ is nonsingular, i.e., there are as many instruments as there are targets, then there exists a unique combination of instruments leading to the set of desired targets. Once the targets are specified, the unique instrument vector is given by

$$\{z_1\} = R^{-1}\{y\} - S\{z_2\} \quad \text{(5.4)}$$

The solution to (5.3) is unique, and there is no need for the policy maker to provide any other information than the set of target values.
b. The matrix multiplier is singular and of rank $K < N$.

If the number of instruments is less than the number of targets, however, the system (5.3) is inconsistent and there is no way that all the target values can be reached. This poses an additional decision problem for the policy maker. Does he give up some targets in order to reach others, or does he want to achieve all the targets as closely as possible with the limited resources? In the latter case, the policy maker may also wish to weight the targets differently. If the first alternative is chosen, some targets are deleted, and the instrument vector is given by (5.4). The second alternative often leads to the formulation of a quadratic programming model. If $\{\tilde{y}\}$ is the vector of desired target values, and $\{\hat{y}\}$ is the vector of realized values, then the problem is to minimize the squared deviation between $\{\tilde{y}\}$ and $\{\hat{y}\}$ subject to (5.3), which describes the behavior of the population system. That is,

$$\min \left[ \{\tilde{y}\} - \{\hat{y}\} \right]' A \left[ \{\tilde{y}\} - \{\hat{y}\} \right]$$

subject to

$$\{\hat{y}\} = R\{z_1\} + S\{z_2\}$$

(5.6)

The weight matrix $A$ represents the policy maker's differential preferences towards the targets. The target variables with the highest weights will be forced very close to their desired values. Those with the lowest weights will not.
c. The matrix multiplier is singular and of rank N.

If the number of instrument variables exceeds the number of targets, then there is an infinite number of solutions to (5.3) and, therefore, an infinite number of instrument vectors. To get a unique solution, the policy maker may force the number of instruments to be equal to the number of targets, by deleting some instruments. On the other hand, he may put some constraints on the instruments. There is a wide variety of possible constraints, but we consider only two categories.

   c.1. Some Instruments are Linearly Dependent.

   By making some instruments linearly dependent, the freedom of policy action is reduced in a way such that only one strategy is available to achieve the targets. An illustration of this constraint is the intervention model of Rogers (1971; pp. 99-101). Targets are specified only for the planning horizon, but instruments are available in each time period. In order to get a unique policy, the constraint is introduced that the values of the instruments in all the time periods are linearly related to each other.

   c.2. Introduction of Acceptable Values of the Instruments.

   In most cases, the policy maker has a good idea of what levels of the instrument variables are acceptable politically. Minimizing the squared deviations between the realized and the most acceptable values assures a unique instrument vector.
d. Illustration: the Theil quadratic programming model.

We have described how policy models are related to the rank of the matrix of impact multipliers or, equivalently, to the number of targets and instruments. Only some alternative policy models have been indicated. A wider variety is possible. For example, the targets and the instruments may be constrained at the same time, and these constraints need not to be linear. The objective function (5.5) may not be quadratic, and (5.6) can be supplemented with both equality and inequality constraints. The reader is referred to the mathematical programming literature for such illustrations. The quadratic objective function with linear constraints, however, is common in economic policy analysis. It is based on two assumptions. The first is that the policy maker's preferences are quadratic in targets and controls. The second assumption is that each of the targets depends linearly on all the instruments, the coefficients of these linear relations being fixed and known. The basic structure of this linear quadratic model is due to Theil (1964; pp. 34-35), and may be expressed as

$$\begin{align*}
\min W &= \{a\}' \{z_1\} + \{b\}' \{\hat{y}\} + \frac{1}{2} \left[ \{z_1\}' A \{z_1\} ight. \\
&\left. \quad + \{\hat{y}\}' Q \{\hat{y}\} + \{z_1\}' C \{\hat{y}\} + \{\hat{y}\}' C' \{z_1\} \right]
\end{align*}$$

(5.7)

subject to

$$\{\hat{y}\} = R \{z_1\} + S \{z_2\}$$

(5.3)
where \( \{ \hat{y} \} \) is the vector of realized values of the target variables,
\( \{ z_1 \} \) is the vector of instrument variables,
\( \{ z_2 \} \) is the vector of exogenous variables,
\( A, Q, C \) are weight matrices,
\( R, S \) are matrices of multipliers.

Applications of the Theil model in economic policy literature may be found in Fox, Sengupta and Thorbecke (1972; p. 215), and in Friedman (1975; pp. 158-160). To simplify matters we may suppose that \( \{ a \} = \{ b \} = \{ 0 \} \) and \( C = 0 \). The problem then reduces to

\[
\min \frac{1}{2} \left[ (\hat{y})' Q(\hat{y}) + (z_1)' A z_1 \right] \tag{5.8}
\]
subject to

\[
\{ \hat{y} \} = R \{ z_1 \} + S \{ z_2 \} , \quad \text{where} \tag{5.3}
\]

\( Q \) and \( A \) are weights attached to the target vector and to the instrument vector respectively.

To illustrate the application of the Theil model in migration policy analysis, consider the following problem. The costs of public services are held to be too high because some regions are over-urbanized and are subject to diseconomies of scale, while other areas have insufficient people to reach the threshold needed for an efficient public service system. The high costs in the public sector can, therefore, be related to the inefficient population distribution. To reduce the costs, a migration policy is needed. However,
there is a cost associated with the redistribution of people over space. Assume that the cost function of public services is a quadratic function of the population distribution \( \{ \hat{y} \} \), i.e.

\[
C_p = \{b\}' \{ \hat{y} \} + \{ \hat{y} \}' E \{ \hat{y} \} .
\]  
(5.9)

Assume also that the cost associated with population distribution is quadratic in the vector of the number of people relocated by the policy program, \( \{ z_1 \} \), i.e.

\[
C_m = \{ z_1 \}' P \{ z_1 \} .
\]  
(5.10)

An element \( z_{1i} \) of \( \{ z_1 \} \) is positive if the program attracts people to region \( i \). It is negative if the program has an out-migration effect. On comparing the cost functions with the preference function (5.7), we see that

\[
\{ a \} = \{ 0 \} , \quad c = 0 , \quad o = 2 \hat{s}
\]

and

\[
\hat{s} = 2 \hat{f} .
\]

Since \( \{ z_1 \} \) represents the additional migration, \( \hat{r} = \hat{s} \) in the constraint. The vector of uncontrollable variables is the population distribution in the previous time period, and \( \hat{s} \) is the multiregional population growth matrix.
5.2.2. **Classification of Policy Models According to the Structure of the Matrix Multiplier**

We now turn to the question of how policy models may be related to the structure of the matrix $\mathbf{R}$. The structure determines the nature of the dependence of $\{z_{1}\}$ upon $\{y\}$. Several assumptions may be adopted to simplify the form of $\mathbf{R}$. They have been studied by Tinbergen (1963, Chapter 4), by Fox, Sengupta and Thorbecke (1972; pp. 24-25) and by Friedman (1975; pp. 149-153) among others. We consider four different structures of $\mathbf{R}$: diagonal, triangular, block-diagonal and block-triangular. Our illustration considers the block-triangular multiperiod policy model.

a. The matrix multiplier is diagonal.

If $\mathbf{R}$ is diagonal, then each target variable can be associated with one and only one instrument variable and vice versa. Since $\mathbf{R}^{-1}$ is also diagonal, equation (5.4) implies a series of expressions

$$\bar{z}_{1i} = \frac{1}{r_{ii}} \left[ y_{i} - \sum_{k} s_{ik} z_{2k} \right], \quad i = 1, \ldots, N,$$

each of which may be solved independently. The practical implication of this is that the policy maker can, in such an instance, pursue each target with a single specific instrument, and no coordination between the various policies is required.
d. The matrix multiplier is block-triangular.

Here, as in the case of a triangular $R$, the set of instruments corresponding to any given block can be solved for without any knowledge of the instruments belonging to blocks which are lower in the hierarchy. The overall policy could be decomposed into a hierarchical system of policies.

e. Illustration: the multiperiod policy problem.

An important application of the block-triangular form of $R$ is found in dynamic policy analysis. The models presented thus far have been static, but they are general enough to handle dynamic policy problems as well. If the entries of the target vector and of the instrument vector belong to different time periods, we clearly have a dynamic or multiperiod policy model. Suppose, for example, that a target vector is given for a sequence of time periods from 1 to $T$, say. Then $\{y\}$ is itself composed of vectors, one for each time period. Suppose, moreover, that there exists an instrument vector for each time period. The reduced form model (5.3) now may be expressed as

$$\{y\} = R\{z_1\} + S\{z_2\}$$

(5.11)

where

$$\{y\} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(T)} \end{bmatrix} \quad \{z_1\} = \begin{bmatrix} z_1^{(1)} \\ z_1^{(2)} \\ \vdots \\ z_1^{(T)} \end{bmatrix} \quad \{z_2\} = \begin{bmatrix} z_2^{(1)} \\ z_2^{(2)} \\ \vdots \\ z_2^{(T)} \end{bmatrix}$$
Vector $\{z_1\}$ is of order $KT$, and $\{z_2\}$ and $\{y\}$ are of order $NT$. The submatrix $R_{ij}$ is $N \times K$ and its elements are dynamic policy multipliers which express the impact on the target vector $\{y(t)\}$ in time period $t = i$ of changes in the instrument vector $\{z_1(t)\}$ in time period $t = j$. $R$ is $NT \times KT$; $S$ is $NT \times NT$ and the submatrices $S_{ij}$ are of order $N \times N$. $S$ shows the dynamic effects of predetermined variables on the target variables.

Most policy models assume that policy actions do not influence events which precede them in time and, therefore, generally ignore expectational effects or advance announcement effects. This assumption of unidirectional causality yields a block-triangular $R$ matrix:

$$R = \begin{bmatrix}
R_{0} & 0 & 0 & \cdots & 0 \\
R_{1} & R_{0} & 0 & & \\
R_{2} & R_{1} & \ddots & & \\
\vdots & \vdots & & \ddots & \\
R_{T-1} & R_{T-2} & \cdots & R_{0} & \\
\end{bmatrix} \quad (5.12)$$
where the elements of $R_t$ are dynamic policy multipliers. A triangular $R$ matrix leads to a sequential decision making procedure analogous to that of the static model. The key distinction is that here the sequence is across time, rather than across individual instrument and target variables.

By way of illustration, consider the application of the Theil model in population policy. Assume that there is a time sequence of target population distributions, and a time sequence of vectors of induced migration. Suppose that no tough policy actions are expected by the potential migrants, therefore the population distribution at time $t$ does not depend on the migration policies beyond $t$. Equation (5.11) may, therefore, be written with $R$ being lower block-triangular.

We may reduce the form of this policy model even further. Suppose that the migration policy at time $t$ only affects the population distribution at $t + 1$ directly. The impact on the population distributions at a later time is indirect in the sense that the population distribution at $t + 1$ affects the distribution beyond $t + 1$. This implies the recurrence equation

$$
\{y(t + 1)\} = R_0 \{z_1(t)\} + S_{t+1,t} \{y(t)\}.
$$

(5.13)

The submatrix $S_{t+1,t}$ is the growth matrix of the population between $t$ and $t + 1$. If we assume the growth matrix to be time-independent, i.e. $G = S_{t+1,t}$ for all $t$, we may write

$$
\{y(t + 1)\} = R_0 \{z_1(t)\} + G \{y(t)\}.
$$

(5.14)
Therefore, (5.11) may be reduced to a set of recurrence equations

\[
\{y(1)\} = R_0\{z_1(0)\} + G\{y(0)\}
\]

\[
\{y(2)\} = R_0\{z_1(1)\} + G\{y(1)\}
\]

\[
= R_0\{z_1(1)\} + GR_0\{z_1(0)\} + G^2\{y(0)\}
\]

\[
\vdots
\]

\[
\{y(t)\} = R_0\{z_1(t - 1)\} + G\{y(t - 1)\}
\]

\[
= G^t\{y(0)\} + \sum_{i=0}^{t-1} G^{t-i}R_0\{z_1(i)\} .
\]

In matrix form, we have that

\[
\begin{bmatrix}
\{y(0)\} \\
\{y(1)\} \\
\{y(2)\} \\
\vdots \\
\{y(T)\}
\end{bmatrix} =
\begin{bmatrix}
I \\
G \\
G^2 \\
\vdots \\
G^T
\end{bmatrix}
\begin{bmatrix}
R_0 \\
GR_0 \\
G^2R_0 \\
\vdots \\
G^{T-1}R_0 \\
\vdots \\
R_0
\end{bmatrix}
\begin{bmatrix}
\{y(0)\} \\
\{z_1(0)\} \\
\{z_1(1)\} \\
\vdots \\
\{z_1(T - 1)\}
\end{bmatrix}
\]

(5.15)

Equation (5.14) has received much attention in system theory. It is called the discrete state equation and forms the central component of the discrete version of the state-space model. Stimulated by recent work in system theory and optimal control, an increasing number of authors have adopted the state-space approach to describe dynamic models in the social sciences\(^1\). We have shown how the

---

\(^1\)See, for example, Pindyck (1973), Kenkel (1974) and Chow (1975).
state-space model may be derived conceptually from the reduced form model. How the transformation is done mathematically will be shown later.

By introducing the assumption of unidirectional causality of the population system, we may write the Theil model (5.8), (5.3) as

$$\min J = \frac{1}{2} \left[ (\hat{y})' Q(\hat{y}) + (z_1)' \tilde{A}(z_1) \right]$$  \hspace{1cm} (5.8)$$

subject to

$$\{\hat{y}(t + 1)\} = R_0 \{z_1(t)\} + G\{y(t)\} \hspace{1cm} (5.16)$$

Recall that $Q$ is a $NT \times NT$ matrix, where $T$ is the planning horizon, $N$ is the number of target variables at each period, and $\tilde{A}$ is a $KT \times KT$ matrix, where $K$ is the number of instrument variables.

5.2.3. From the Tinbergen Model to the Optimal Control Model

In this chapter, we started out with the Tinbergen paradigm. The original model, based on this paradigm, was simple in nature, in the sense that the number of instruments was equal to the number of targets and that the optimal policy was the unique solution to a system of linear equations. When the number of instruments and targets differs, the policy maker is confronted with an additional decision problem. He needs to specify his preferences in order to get a unique policy which is optimal. This led us to the Theil model and to the broad application of mathematical programming. When policy problems
shows that the multiperiod Theil problem may be reduced to a linear-quadratic control problem by assuming inter-temporal separability of the objective and unidirectional causality of the population system. If these conditions are not met, one must apply the dynamic generalization of the Theil model (Theil, 1964, Chapter 4).

In control theory, it is normal to denote the target vector \( \{y(t)\} \) by \( \{x(t)\} \), and the control vector \( \{z_1(t)\} \) by \( \{u(t)\} \). In most practical applications, it is also assumed that \( Q(t) = \bar{Q} \) is equal for all time periods up to \( T - 1 \). This assumption is only valid if the preference system and tastes do not change over time. It also implies that the contribution of a certain set of target and control values is independent of when they appear on the time path, since no discounting measure has been introduced. The matrix \( \bar{Q}(T) \) is commonly denoted by \( \bar{F} \). The weight matrices \( \bar{A}(t) \) associated with the instruments or controls are also assumed to be time independent, and are denoted by \( \bar{R} \). The multiplier matrix \( \bar{R}_0 \) is denoted by \( \bar{B} \). To facilitate reference to the optimal control literature, we will adopt this notation in the remainder of this study. The linear-quadratic problem, therefore, is reformulated as

\[
\min \frac{1}{2} \{x(T)\}' \bar{F}\{x(T)\} \\
+ \frac{1}{2} \sum_{t=0}^{T-1} \left[ \{x(t)\}' \bar{Q}\{x(t)\} + \{u(t)\}' \bar{R}\{u(t)\} \right]
\]  

(5.19)

subject to

\[ \{x(t + 1)\} = \bar{G}\{x(t)\} + \bar{B}\{u(t)\} \]
To solve (5.19), one can apply the quadratic programming algorithm to the original Theil problem with the matrices $A$, $Q$, $R$ and $S$ of the appropriate structure. However, if $T$ and $N$ are of some practical magnitude, the scale of the problem becomes immense. In recent years, algorithms have been sought which could solve the general linear-quadratic problem and dynamic problems directly. The optimization of such dynamic systems may be approached from three alternative perspectives:

1. **Variational calculus**, dealing with the problem of finding the function describing the optimal trajectory of the system. The solution of such a problem involves the determination of maxima and minima of functionals (Gelfand and Fomin, 1963).

2. **Dynamic programming**, based on Bellman's principle of optimality (Bellman, 1957).

3. **Optimal control theory**, based on the "maximum principle" derived by Pontryagin and his associates (1962).

A discussion of the three approaches is beyond the scope of this study. The interested reader is referred to the literature. A clear exposition of the relationship between the calculus of functionals and the calculus of functions is given by Connors and Teichroew (1967). How dynamic programming and optimal control theory relate to each other, is discussed by Noton (1972). A fine textbook on applied optimal control is Bryson and Ho (1969).

Optimal control has the broadest field of application. Problems which may be solved by calculus of variations or by dynamic programming, can also be solved by optimal
control. Therefore, we adopt the optimal control approach to the optimization of dynamic population systems. This will enable us in Chapter 7 to derive the optimal solution to the linear-quadratic control problem.
CHAPTER 6
REPRESENTATION AND EXISTENCE THEOREMS
OF MIGRATION POLICIES

In this chapter we deal with constraints (5.6) and (5.16); in other words, with the demometric model representation of the dynamics of a population system, and with the policy model describing the goals-means relationship in migration policy. Nothing will be said about goal-setting or about the selection of optimal values for the instrument variables. These will be considered in the next chapter.

Let us begin with a demometric model in the form of a system of simultaneous linear equations. It is assumed that the model has been specified and that the coefficients have been estimated. The model relates demographic with socio-economic variables, in a manner such as is found in Greenwood (1973, 1975b). We assume that the model is dynamic, i.e., that it contains lagged endogenous and exogenous variables. It is also assumed that the goals-means relationship of migration policy is known, i.e., the target variables and the instrument variables have been separated from the other endogenous and exogenous variables.

We first transform the reduced form of the model into the discrete state-space form. A general solution of the discrete state-space equation is then derived. Next, we consider the question whether arbitrary specified levels of target variables can be reached by the existing set of instruments. The existence theorems which are
derived are related to the rank of the matrix of impact multipliers.

6.1. STATE-SPACE REPRESENTATION OF DEMOMETRIC MODELS

Stimulated by recent work in optimal control and system theory, an increasing number of authors have adopted the state-space approach to describe dynamic models in the social sciences. This section describes the characteristics of the state-space model and the procedure for its solution. Since most demometric models are given in the reduced form, we also consider its transformation to the state-space form.

6.1.1. The State-Space Model

The state-space representation of a linear system is defined by the following set of first order linear difference equations\(^2\):

\[
\{\tilde{x}(t + 1)\} = G(t) \{\tilde{x}(t)\} + B(t) \{\tilde{u}(t)\} \quad (6.1)
\]

\[
\{\tilde{y}(t)\} = C(t) \{\tilde{x}(t)\} + E(t) \{\tilde{u}(t)\} \quad (6.2)
\]

where \(\{\tilde{x}(t)\}\) is an \(N\)-dimensional vector-valued function of time, called the state of the system,

\(\{\tilde{u}(t)\}\) is an \(K\)-dimensional vector-valued function of time, called the input or control to the system,

\(^2\)We shall consider only the discrete state-space model. The continuous version is a set of differential equations. For details see, for example, Director and Rohrer (1972) and Wolovich (1974).
\( \{ y(t) \} \) is an \( \mathbf{P} \)-dimensional vector-valued function of
time, called the output of the system,
\( \tilde{A}(t), \tilde{B}(t), \tilde{C}(t) \) and \( \tilde{E}(t) \) are real-time dependent
matrices of dimension \( \mathbf{N} \times \mathbf{N}, \mathbf{N} \times \mathbf{K}, \mathbf{P} \times \mathbf{N} \) and
\( \mathbf{P} \times \mathbf{K} \), respectively.

If \( \tilde{G}(t), \tilde{B}(t), \tilde{C}(t) \) and \( \tilde{E}(t) \) are constant over time, the
system is time-invariant. In this section, we will only
consider the case where these matrices are constant. Thus,

\[
\{ x(t + 1) \} = \tilde{G}\{ x(t) \} + \tilde{B}\{ u(t) \} \quad \text{(6.3a)}
\]

\[
\{ y(t) \} = \tilde{C}\{ x(t) \} + \tilde{E}\{ u(t) \} \quad . \quad \text{(6.3b)}
\]

The homogenous part of (6.3a):

\[
\{ x(t + 1) \} = \tilde{G}\{ x(t) \}
\]

gives the growth of the system without intervention. The
matrix \( \tilde{G} \) is the growth matrix. The discrete model of
population growth, studied by Rogers (1975; p. 123), is of
this form.

The interpretation of (6.3) as a migration policy model
is straightforward. Suppose \( \{ \tilde{x}(t) \} \) is the interregional
and/or age-specific population distribution. The matrix \( \tilde{G} \)
is the population growth matrix, and \( \{ \tilde{u}(t) \} \) is a vector of
instrument variables, which may range from pure demographic
variables to socio-economic variables. It defines a
policy at time \( t \). The impact of each policy variable on
the population distribution in the next period, is given
by the elements of \( \tilde{B} \). If \( \{ \tilde{u}(t) \} \) has no lagged instrument
variables, and if \( \{u(t)\} \) has no impact on \( \{x(t)\} \), then \( B \) is the matrix of impact multipliers. If the policy is a direct population influencing policy, then \( \{u(t)\} \) is expressed in numbers of people, exactly as \( \{x(t)\} \), and therefore \( B \) is the identity matrix.

In demographic policy problems with socio-economic goals, the target vector is not expressed in terms of population distribution, but in terms of socio-economic variables. The matrix \( C \) transforms the population distribution \( \{x(t)\} \) into the vector \( \{y(t)\} \) of socio-economic target variables, whereas \( E \) gives the direct impact of the policy variables on the new target variables. In fact, \( C \) can be any transformation matrix. For example, suppose \( \{x(t)\} \) is the regional distribution of the population by age. If the policy maker is interested only in the spatial distribution of the total population, then \( C \) will be a consolidation matrix.

### 6.1.2. Solution of the State-Space Model

In order to derive the solution to (6.3a), we write (6.3a) for various \( t \):

\[
\{x(1)\} = C\{x(0)\} + B\{u(0)\}
\]

\[
\{x(2)\} = C\{x(1)\} + B\{u(1)\} = C^2\{x(0)\} + CB\{u(0)\} + B\{u(1)\}
\]

\[
\vdots
\]

\[
\{x(T)\} = C^T\{x(0)\} + \sum_{i=0}^{T-1} C^{T-1-i} B\{u(i)\}.
\]

Therefore the general solution to (6.3) is
\{x(t)\} = G^t(x(0)) + \sum_{i=0}^{t-1} G^{t-1-i} B\{u(i)\} \quad (6.4)

\{y(t)\} = CG^t(x(0)) + \sum_{i=0}^{t-1} CG^{t-1-i} B\{u(i)\} + E\{u(t)\}.

The solution to the homogenous part of (6.3a) is

\{x(t)\} = G^t(x(0)) \quad (6.5)

where $G^t = \Phi(t,0)$ is known as the discrete state-transition matrix. The solution in terms of the state-transition matrix is:

\{x(t)\} = \Phi(t)\{x(0)\} + \sum_{i=0}^{t-1} \Phi(t-1-i) B\{u(i)\} \quad (6.6)

and

\{y(t)\} = CG^t(x(0)) + C \sum_{i=0}^{t-1} \Phi(t-1-i) \ B\{u(i)\} + \ E\{u(t)\} \quad (6.7)

where $\Phi(t) = G^t$.

Consider the system where $E = 0$. Then

\{y(t)\} = CG^t(x(0)) + \sum_{i=0}^{t-1} CG^{t-1-i} B\{u(i)\}.

Let $H(t) = CG^{t-1} B$, then

\{y(t)\} = CG^t(x(0)) + \sum_{j=1}^{t-1} H(t-j)\{u(i)\}.$
and, if \( j = t - i \),

\[
\{ y(t) \} = CG^t\{ x(0) \} + \sum_{j=1}^{t} H(j)\{ u(t - j) \} .
\] (6.8)

If \( \{ x(0) \} \) is the initial population distribution, if \( \{ u(t) \} \) is vector of control or policy actions at time \( t \), and if \( \{ y(t) \} \) describes the population distribution at time \( t \) (in this case, \( C = I \)), then \( H(j) \) is the matrix of dynamic impact multipliers. The element \( h_{rs}(j) \) represents the change of the population in group or region \( r \) at time \( t \) due to a unit change in the \( s \)-th instrument at time \( t - j \). \( H(j) \) can also be thought of as the contribution of the policy action at time \( (t - j) \) to the population distribution at time \( t \). Each matrix \( H(j) \) corresponds to the various submatrices of (5.15), which are not in the first column.

6.1.3. **State-Space Representation of the Reduced Form Model**

The reduced form of a demometric model is

\[
\{ \dot{y} \} = \{ E \{ \dot{z} \} \} ,
\] (5.1)

where \( \{ y \} \) is the vector of endogenous variables, and

\( \{ z \} \) is the vector of predetermined variables consisting of exogenous and lagged endogenous variables.

The general reduced form is
\[ \tilde{A}\{\tilde{y}(t)\} = E_1\{\tilde{y}(t - 1)\} + E_2\{\tilde{y}(t - 2)\} + \ldots + E_r\{\tilde{y}(t - r)\} \]
\[ + D_0\{\tilde{z}_1(t)\} + D_1\{\tilde{z}_1(t - 1)\} + \ldots + D_s\{\tilde{z}_1(t - s)\} \]

(6.9)

where \((t - i)\) indicates a time lag of \(i\) periods. In order to put (6.9) into state-space form, we must define new variables and corresponding equations to replace the reduced form variables that have second order or higher order lags. The procedure is then one of the replacement of an \(r\)-th order difference equation by \(r\) first-order difference equations.

First, let

\[ \{\tilde{z}(t)\} = \begin{bmatrix}
\{\tilde{z}_1(t)\} \\
\{\tilde{z}_1(t - 1)\} \\
\vdots \\
\{\tilde{z}_1(t - s)\}
\end{bmatrix} \]

and

\[ D = [D_0, D_1, \ldots, D_s] \]  

Equation (6.9) then may be simplified to yield

\[ \tilde{A}\{\tilde{y}(t)\} = E_1\{\tilde{y}(t - 1)\} + E_2\{\tilde{y}(t - 2)\} + \ldots + E_r\{\tilde{y}(t - r)\} \]
\[ + D\{\tilde{z}(t)\} \]  

(6.10)
Following Kenkel (1974; pp. 295-299), we define a set of new vectors:

\[
\{y_1(t)\} = \{y(t)\}
\]

\[
\{y_2(t)\} = \{y(t - 1)\}
\]

\[
\vdots
\]

\[
\{y_r(t)\} = \{y(t - r + 1)\}
\]

\[
\{y_r(t - 1)\} = \{y(t - r)\}
\]

Therefore (6.10) becomes

\[
\{y_1(t)\} = A^{-1}_1 E_1 \{y_1(t - 1)\} + A^{-1}_2 E_2 \{y_2(t - 1)\} + \ldots
\]

\[
+ A^{-1}_r E_r \{y_r(t - 1)\}
\]

which may now be rewritten as a recurrence equation of the form

\[
\{x(t)\} = G\{x(t - 1)\} + B\{u(t)\}
\]

(6.11)

where

\[
\{x(t)\} = \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_r(t) \end{bmatrix}
\]

\[
\{u(t)\} = \{z(t)\}
\]
Equation (6.11) is the state-space representation. The submatrices in the first row denote the impact on \( \{y(t)\} \) of the vectors of lagged endogenous variables. The submatrix \( \tilde{A}^{-1}\tilde{D} \) denotes the direct effect on \( \{y(t)\} \) of the exogenous variables. \( \tilde{A}^{-1}\tilde{D}_0 \) is the matrix of impact multipliers. The matrix \( \tilde{A}^{-1}\tilde{D}_1 \) gives the direct effect on \( \{y(t)\} \) of the vector of exogenous variables, lagged by \( i \) periods. These are not total delay multipliers, since \( \{z(t - i)\} \) also affects \( \{y(t)\} \) through its impact on \( \{y(t - k)\} \), \( k = 1, \ldots, i \).

There is another transformation of the reduced form to the state-space form. This transformation has no direct motivation for demometric models, but it facilitates the study of the state-space model. Equation (6.10) may be written as
\[ E_r \{y(t - r)\} = -E_{r-1} \{y(t - r + 1)\} - E_{r-2} \{y(t - r + 2)\} \ldots \]
\[ - E_1 \{y(t - 1)\} - A \{y(t)\} + D \{z(t)\} \tag{6.13} \]

Suppose \( E_r \) is nonsingular, and define the new vectors

\[ \{y_1(t - 1)\} = \{y(t)\} \]
\[ \{y_i(t - 1)\} = \{y(t - 1)\} \]
\[ \vdots \]
\[ \{y_i(t - 1)\} = \{y(t - (i - 1))\} \]
\[ \vdots \]
\[ \{y_r(t - 1)\} = \{y(t - (r - 1))\} \]

The extended version of (6.13) then is

\[ \{x(t)\} = \hat{G} \{x(t - 1)\} + \hat{B} \{u(t)\} \]

where \( \{x(t)\} \) and \( \{u(t)\} \) are as defined in (6.12),

\[
\hat{G} = \begin{bmatrix}
0 & \tilde{I} & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & \tilde{I} & 0 & 0 \\
-\frac{1}{E_r} A & -\frac{1}{E_r} \tilde{E} & -\frac{1}{E_r \tilde{E} - 1} & \vdots \\
0 & \cdots & \tilde{E} & \tilde{E} & \tilde{E} & \tilde{E} & \tilde{E}
\end{bmatrix} \tag{6.14}
\]
The matrices $\mathbf{G}$ and $\hat{\mathbf{G}}$ are generalized companion matrices. In Chapter 4 of this study, we have already introduced the companion matrix in the demographic analysis. It has been indicated that this matrix can play an important role in the reconciliation of discrete and continuous models of demographic growth. Here we have shown that the companion matrix provides the natural link between the reduced form model and the state-space model. A similar link may be formulated between the structural form and the state-space model. A detailed description of the technique is given by Pindyck (1973; pp. 89-94).

6.2. **EXISTENCE THEOREMS OF MIGRATION POLICIES**

It is argued that there are two central issues in the theory of policy. These are the concepts of existence and of design. Existence of policy refers to the controllability of the system or the ability to design any policy at all; design refers to the techniques for designing optimal policies once existence is assured. Although both issues have been recognized for a long time in system theory, policy analysis in the social sciences, led by the theory of economic policy, has focused almost entirely on the design problem. Only Tinbergen (1963) has given considerable attention to
both issues. His policy model is formulated in the reduced form. An alternative representation is the state-space format.

This section is divided into two parts. The first deals with the existence of optimal policies in the Tinbergen framework. The other derives existence theorems for the state-space model. Until very recently, the existence of optimal policies in the state-space framework has not been investigated in the theory of economic policy. Based on findings of system theory, Aoki (1973, 1974, 1975) and Preston (1974) have supplemented Tinbergen's existence theorem with theorems related to state-space economic models.

6.2.1. Existence Theorem in the Tinbergen Model

Recall the Tinbergen model

\[
\{y\} = R\{z_1\} + S\{z_2\} \quad .
\] (5.3)

In the original formulation, (5.3) represented a static policy problem, i.e. the targets and the instruments belonged to the same time period. The model, however, may include lagged variables in the vector of uncontrollable variables \(\{z_2\}\). Contrary to Preston's (1974; p. 65) claim, the Tinbergen model also fits dynamic situations, where the targets and instruments belong to different time periods. This is shown in (5.11). The cornerstone of Tinbergen's theory of policy is the condition for which there exists for any \(\{\tilde{y}\}\) a corresponding unique policy vector \(\{\tilde{z}_1\}\) such that (5.3) is satisfied. In other words, under what conditions has (5.3) a unique solution for \(\{z_1\}\)? The necessary
independent. This implies that the equations of (5.3) are independent. The system (5.3) is consistent, i.e. has a solution if and only if the number of unknowns K is greater than or equal to the number of equations N. But this implies that the rank of \( \mathbf{R} \) is \( N \). If \( K \) is less than \( N \), the rank of \( \mathbf{R} \) is \( K < N \), and the system is inconsistent. The general solution to a consistent system is (Rogers, 1971; p. 258):

\[
\{ \tilde{z}_1 \} = \mathbf{R}^{(1)} \left[ \{ \tilde{y} \} - \mathbf{S} \{ z_2 \} \right] + \left[ \mathbf{I} - \mathbf{R}^{(1)} \end{array} \begin{array}{c} \mathbf{R} \\
\tilde{z} \end{array} \right] \{ \tilde{c} \} \quad (6.15)
\]

where \( \mathbf{R}^{(1)} \) is a generalized inverse of \( \mathbf{R} \), satisfying

\[
\mathbf{R} \cdot \mathbf{R}^{(1)} \cdot \mathbf{R} = \mathbf{R} \,
\]

and \( \{ \tilde{c} \} \) is an arbitrary vector.

If \( K > N \), there exists an infinite number of instrument vectors associated with \( \{ \tilde{y} \} \). However, in most cases, there is only one instrument vector which is most suited to the policy maker's preferences. The design of such a policy vector will be discussed in the next chapter. If on the other hand \( K = N \), then \( \mathbf{R} \) is nonsingular and (5.3) has a unique solution:

\[
\{ \tilde{z}_1 \} = \mathbf{R}^{-1} \left[ \{ \tilde{y} \} - \mathbf{S} \{ z_2 \} \right] \quad . \quad (6.16)
\]
6.2.2. **Existence Theorems in the State-Space Model**

In the previous chapter, the state-space model was derived by introducing the assumptions of unidirectional causality and time independence into the Tinbergen model. Recall

\[ \{x(t + 1)\} = G\{x(t)\} + B\{u(t)\} \]  \hspace{1cm} (6.3a)

\[ \{y(t)\} = C\{x(t)\} + E\{u(t)\} \] \hspace{1cm} (6.3b)

Two existence problems may be distinguished. The first deals with the question of whether there is a sequence of control vectors \(\{u(t)\}, t = 0,\ldots,T-1\), such that a desired target vector can be achieved at a predefined planning horizon \(T\). The second deals with the question whether there exists a sequence of control vectors \(\{u(t)\}, t = 0,\ldots,T-1\), such that any sequence of target vectors \(\{y(t)\}, t = 1,\ldots,T\) can be realized. The first existence problem is known in system theory as state and output controllability; the latter is sometimes referred to as output function controllability. The state and output controllability has received most attention in the literature. Both existence problems will be dealt with below. Two applications will be discussed: the separation of the controllable and the noncontrollable parts of the system, and the achievement of the targets with a minimum number of instruments.
a. State and Output Controllability

The system

\[ \{\dot{x}(t + 1)\} = C\{x(t)\} + B\{u(t)\} \] \hspace{1cm} (6.3a)

is said to be controllable (state controllable) if and only if there exists a control \( \{u(t)\} \) which transfers any initial state \( \{x(t_0)\} \) at any time \( t_0 \) to any arbitrary final state \( \{x(t_f)\} \) at any time \( t_f > t_0 \geq 0 \), (Wolovich, 1974; p. 65). Otherwise, the system is uncontrollable or only "controllable in part," i.e. it may be possible to transfer certain states to any desired final states or to transfer all the states to a position close to the desired states.

The controllability concept assumes that there is no constraint on \( \{u(t)\} \). The only requirement is that there exists a trajectory between the initial and the terminal state.

In order to determine conditions for controllability of (6.3a), consider its solution

\[ \{x(t)\} = C^t\{x(0)\} + \sum_{i=0}^{t-1} C^{t-1-i} B\{u(i)\} \] \hspace{1cm} (6.4)

which may be rewritten as

\[ \begin{bmatrix} \{x(t)\} \\ \{x(t)\} - C^t\{x(0)\} \end{bmatrix} = \begin{bmatrix} B \cdot G \cdot B \cdot G^2 \cdot B \cdot \cdots \cdot G^{t-1} B \end{bmatrix} \begin{bmatrix} \{u(t-1)\} \\ \{u(t-2)\} \\ \vdots \\ \{u(0)\} \end{bmatrix} \] \hspace{1cm} (6.17)
where \( \{x(t)\} \) and \( \{x(0)\} \) are given and \( \sim \) is time-invariant. The dimension of the target vector \( \{x(t)\} \) is \( N \) \( (t = 1, \ldots, T) \) and of the control vector \( \{u(t)\} \) is \( K \) \( (t = 0, \ldots, T-1) \). The matrix

\[
D = [B; GB; G^2B; \cdots; G^{t-1}B]
\]

is therefore of dimension \( N \times Kt \). Equation (6.3a) is controllable, or there exists a solution to (6.17) if the rank of \( D \) is \( N \). If \( t > N \), i.e. if the number of control intervals is greater than the number of targets, then we don't need to consider the whole matrix \( D \) to evaluate the controllability of (6.3a).

According to the Cayley-Hamilton theorem, each matrix satisfies its own characteristic equation. If \( G \) has the characteristic equation

\[
\lambda^N + c_1\lambda^{N-1} + c_2\lambda^{N-2} + \ldots + c_N = 0
\]

then

\[
G^N + c_1G^{N-1} + c_2G^{N-2} + \ldots + c_NI = 0 \quad . \tag{6.18}
\]

Therefore \( G^N \) and any \( G^{N+i} \) \( (i \geq 0) \) is linearly dependent on \( [I; G; G^2; \cdots; G^{N-1}] \). It follows that no extra independent column vectors would be added to \( D \) if there are more than \( N \) control intervals.

This result is formulated as follows:
THEOREM 2: State Controllability Theorem

The dynamic system

\[
\{\tilde{x}(t+1)\} = \tilde{C}\{\tilde{x}(t)\} + \tilde{B}\{\tilde{u}(t)\}
\] (6.3a)

is completely controllable for all \(\{\tilde{x}(t)\} = \{\tilde{z}(t)\}\) if and only if the \(N \times KN\) matrix

\[
\tilde{D} = [\tilde{B}; \tilde{G}\tilde{B}; \cdots; \tilde{G}^{N-1}\tilde{B}]
\] (6.19)

is of rank \(N\). This theorem has been considered by Preston (1974, p. 68) as the dynamic generalization of Tinbergen's theory of policy. Several observations may be made at this point.

a) It is a corollary to the theorem that if a target vector \(\{\tilde{z}\}\) cannot be reached in \(N\) control intervals, it will never be reached. This is important for policy purposes, since it answers the question of how fast a target population distribution, for example, can be achieved.\(^3\)

b) Tinbergen-controllability implies state controllability. If the rank of \(\tilde{B}\) is \(N\), as required for Tinbergen controllability, then the rank of \(\tilde{D}\) is \(N\) also. If the rank of \(\tilde{B}\) is \(N\), then the targets can be reached in only one control interval \((t_1 = t_0 + 1)\).

c) An argument similar to the one leading to Theorem 2, may be used to derive the conditions for output

\(^3\)Under the assumption that no constraints are imposed on the trajectory of control and state variables.
controllability. The system

\[ \tilde{x}(t + 1) = G\tilde{x}(t) + B\tilde{u}(t) \]  \hspace{1cm} (6.3a)

\[ \tilde{y}(t) = C\tilde{x}(t) + E\tilde{u}(t) \]  \hspace{1cm} (6.3b)

is output controllable if and only if there exists a control vector \( \{u(t)\} \) which transfers any initial output \( \{\tilde{y}(t_0)\} \) at time \( t_0 \) to any arbitrary final output \( \{\tilde{y}(t_1)\} \) at any \( t_1 > t_0 \). Wolovich (1974; p. 71) states the condition for output controllability to be

\[
\text{rank } \begin{bmatrix} \cdots & CG & \cdots & CG^{N-1} & B & E \end{bmatrix} = P
\]  \hspace{1cm} (6.20)

if \( P \times 1 \) is the dimension of the output vector. Output controllability is sometimes referred to as reproducibility (Brockett and Mesarovic, 1965; p. 549).

d) The "dual" notion of controllability is observability. The system (6.3) is said to be observable if and only if the entire state \( \{\tilde{x}(t)\} \) can be determined over any finite interval \( [t_0, t_1] \) from complete knowledge of \( \{\tilde{u}(t)\} \) and \( \{\tilde{y}(t)\} \) over the interval \( [t_0, t_1] \) with \( t_1 > t_0 \geq 0 \) (Wolovich, 1974; p. 73). The condition for observability is that the \( MN \times N \) matrix

\[
Q = \begin{bmatrix} C & \cdots & CG & \cdots & CG^{N-1} \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
CG & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
CG^{N-1} & \cdots & \cdots & \cdots & \cdots \\
\end{bmatrix}
\]  \hspace{1cm} (6.21)
be of rank N. Equation (6.3b), written out for the time periods \( t = 0, \ldots, N-1 \), while noting that

\[
\{ x(t) \} = G^t \{ x(0) \} + B \{ u(t - 1) \} ,
\]

gives

\[
\begin{bmatrix}
\{ x(0) \} \\
\{ x(1) \} \\
\vdots \\
\{ x(N - 1) \}
\end{bmatrix} =
\begin{bmatrix}
C \\
CG \\
\vdots \\
CG^{N-1}
\end{bmatrix}
\begin{bmatrix}
\{ 0 \} \\
\{ 0 \} \\
\vdots \\
\{ 0 \}
\end{bmatrix} +
\begin{bmatrix}
\{ 0 \} \\
\{ 0 \} \\
\vdots \\
\{ 0 \}
\end{bmatrix}
\begin{bmatrix}
B \{ u(0) \} \\
B \{ u(0) \} \\
\vdots \\
B \{ u(0) \}
\end{bmatrix} +
\begin{bmatrix}
E \{ u(0) \} \\
E \{ u(1) \} \\
\vdots \\
E \{ u(N - 1) \}
\end{bmatrix}
\]

(6.22)

where \( \{ x(t) \} \) and \( \{ y(t) \} \) are known for \( t = 0, \ldots, N-1 \), and \( \{ x(0) \} \) is unknown. System (6.22) consists of \( MN \) equations in \( N \) unknowns. \( \{ x(0) \} \) can be calculated if \( G \) has rank \( N \). If \( \{ x(0) \} \) is known, the whole sequence of state vectors is known by (6.3a).

The notion of observability might be useful in the study of populations with incomplete data. For example, let \( \{ x(0) \} \) be the spatial distribution of a population by age group, at time \( t = 0 \). Let \( \{ y(t) \} \) be the observed spatial distribution of the total population at time \( t \) and let \( \{ u(t) \} = \{ 0 \} \) for all \( t \). The matrix \( C \) is then a consolidation matrix. Assuming that the condition for observability is met, and that \( G \) is known and remains constant in time, \( \{ x(0) \} \) can be computed from \( \{ y(t) \} \) \( [t = 0, \ldots, N-1] \). If \( G \) is unknown, it may be approximated by some underlying model mortality, fertility and migration schedules.

The problem of controllability and observability has been studied by Vajda (1975) in manpower planning, although
the author does not refer to the concepts and theorems just described and instead focuses on totally different techniques. He uses the simplex algorithm to determine the population distribution from which a given distribution can be obtained, and to find out if a target distribution can be reached from the present distribution in one, two or more steps.

b. **Output Function Controllability**

The controllability concept discussed in the previous section dealt with the existence of a control vector, such that a desired target vector can be achieved at a predefined planning horizon. In practice, policy makers would be interested in not only achieving desired target values, but also keeping them on some desired time trajectory once achieved, or achieving the targets along a desired path. It is not uncommon in politics that short term objectives conflict with long term goals. In designing a policy to achieve the short term objectives, the policy maker includes elements which make the long term goals unattainable. The careful policy-maker, therefore, will design a policy that enables him not only to achieve, for example, a desired population distribution at a certain point in time, but also to control the growth path of the multiregional population system once the target distribution is achieved. A system whose trajectory is controllable is called **output function controllable** or, equivalently, **functionally reproducible** (Brockett and Mesarović, 1965; p. 556).
Recall the dynamic system described by (6.3)

\[ \{\tilde{x}(t + 1)\} = G\{\tilde{x}(t)\} + B\{u(t)\} \]  
(6.3a)

\[ \{y(t)\} = C\{\tilde{x}(t)\} \]  
(6.3b)

where \(\{\tilde{x}(t)\}\) is the \(N \times 1\) state vector,
\(\{u(t)\}\) is the \(K \times 1\) control or input vector, and
\(\{y(t)\}\) is the \(P \times 1\) output vector.

If the target is related to the state of the system, (6.3b) may be deleted, or \(\hat{C}\) may be set identical to the identity matrix. In order to derive the condition for output function controllability, we take \(z\)-transforms in equation (6.3) (Director and Rohrer, 1972; p. 317):

\[ z\{\tilde{x}(z)\} - z\{\tilde{x}(0)\} = G\{\tilde{x}(z)\} + B\{u(z)\} \]

\[ \{y(z)\} = \hat{C}\{\tilde{x}(z)\} \] .

Thus

\[ [zI - G]\{\tilde{x}(z)\} = z\{\tilde{x}(0)\} + B\{u(z)\} \]

\[ \{\tilde{x}(z)\} = [zI - G]^{-1} z\{\tilde{x}(0)\} + [zI - G]^{-1} B\{u(z)\} \]

\[ \{y(z)\} = \hat{C}[zI - G]^{-1} z\{\tilde{x}(0)\} + \hat{C}[zI - G]^{-1} B\{u(z)\} \] .

(6.23)
The $P \times K$ matrix

$$\underline{C}[\underline{zI} - \underline{G}]^{-1} \underline{B} = \underline{H}(\underline{z})$$

(6.24)

is called the discrete transfer matrix (Director and Rohrer, 1972; p. 317)\(^4\). The transfer matrix describes the relationship between the output $\{\underline{y}(t)\}$ and the input $\{\underline{u}(t)\}$ of the system. It is independent of any particular choice of $\{\underline{x}(0)\}$. Equation (6.23) may be written as

$$\underline{H}(\underline{z}) \{\underline{u}(\underline{z})\} = \{\underline{y}(\underline{z})\} - \underline{C}[\underline{zI} - \underline{G}]^{-1} \underline{z}\{\underline{x}(0)\}$$

(6.25)

This allows us to formulate precisely the question of output function controllability and to answer it.

The question of output function controllability is: given any desired $P$-dimensional output vector $\{\underline{y}(t)\}$, defined for all $t \geq t_0$, and the initial state $\{\underline{x}(0)\}$, can the sequence $\{\underline{y}(t)\}$, $t \geq t_0$ be obtained by choosing the appropriate sequence $\{\underline{u}(t)\}$, $t \geq t_0$? The answer to this question is formulated in the following theorem.

**THEOREM 3**: Output Function Controllability Theorem

The system

$$\{\underline{x}(t + 1)\} = \underline{C}\{\underline{x}(t)\} + \underline{B}\{\underline{u}(t)\}$$

\(^4\)The discrete transfer matrix is the analogue of the transfer matrix of continuous models, derived using Laplace transforms:

$$\underline{T}(s) = \underline{C}[\underline{sI} - \underline{G}]^{-1} \underline{B}$$

(6.26)

See Director and Rohrer (1972; p. 303) and Wolovich (1974; p. 101).
is output function controllable if and only if the rank of
the transfer matrix

\[ H(z) = \tilde{C}[zI - \tilde{G}]^{-1} \tilde{B} \]  \hspace{1cm} (6.24)

is equal to \( P \). The control \( \{ \tilde{u}(t) \}, t \geq t_0 \), is unique if \( P \)
is equal to \( K \). The existence theorem, formulated by Wolovich
(1974; p. 164) states that the transfer matrix must have an
inverse, i.e. must be nonsingular. The control sequence he
derives, is, therefore, unique. However, the uniqueness of
\( \{ \tilde{u}(t) \} \) is not a necessary condition for output function
controllability. If \( P < K \), an infinite number of control
sequences leads to the desired output sequence.

The condition for output function controllability may
also be expressed in terms of the matrices \( \tilde{G}, \tilde{B} \) and \( \tilde{C} \) of
the original system (6.3) (Brockett and Mesarović, 1965;
p. 556).

THEOREM 3':

The system

\[ \{ \tilde{x}(t + 1) \} = \tilde{G}\{ \tilde{x}(t) \} + \tilde{B}\{ \tilde{u}(t) \} \] \hspace{1cm} (6.3a)

\[ \{ \tilde{y}(t) \} = \tilde{C}\{ \tilde{x}(t) \} \] \hspace{1cm} (6.3b)

is output function controllable if and only if the
\( PN \times (2N - 1) \) \( K \) matrix
\[
F = \begin{bmatrix}
CB & CGB & CG^2B & \cdots & CG^{N-1}B & CG^NB & \cdots & CG^{2N-1}B \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
0 & CB & CGB & \cdots & CG^{N-2}B & CG^{N-1}B & \cdots & CG^{2N-2}B \\
0 & 0 & CB & \cdots & CG^{N-3}B & CG^{N-2}B & \cdots & CG^{2N-3}B \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
0 & 0 & 0 & CB & CGB & \cdots & CG^{N-1}B & \vdots \\
\end{bmatrix}
\] (6.27)

is of rank PN. (See R. Brockett and M. Mesarović (1965; pp. 556-559) for the formal proof.) Two observations, which are corollaries to theorem 3', may be made at this point.

a) Output function controllability implies output controllability. In a corollary to Theorem 2, the system (6.3) was said to be output controllable if

\[
\text{rank } CQ = [\begin{bmatrix} CB & CGB & \cdots & CG^{N-1}B \end{bmatrix}] = P
\]

The matrix \( CQ \) is the last row of \( F \). Now, if the rank of \( F \) is PN, then the rank of \( CQ \) must be \( P \) and the system is output controllable.

b) A sufficient condition for \( F \) to be of rank PN is that

\[
PN \leq 2(N - 1)K + K
\]
or

\[
P \leq \frac{2N - 1}{N}K
\] (6.28)

for any \( N \). This means that the number of target variables must be less than or equal to the number of instrument variables (Aoki, 1975; p. 295). This leads Aoki to conclude
that the condition for output function controllability is a more proper dynamic generalization of Tinbergen's theory of policy than is the condition for output controllability proposed by Preston (1974; p. 68), because the former contains the original Tinbergen condition that the number of targets cannot exceed the number of instruments.

**c. Separation of Controllable and Non-Controllable Parts of a System**

If a system is not completely state controllable, i.e. the rank of \( D \) is less than \( N \), it is important for policy purposes to determine the controllable part of the system. Two relevant methods are given below. The first is based on the diagonalization of the matrix \( \tilde{G} \). The other method starts directly from the controllability condition.

Assume that the growth matrix \( \tilde{G} \) is primitive, a common assumption in the mathematical demography literature. Then \( \tilde{G} \) has \( N \) distinct eigenvalues, and \( N \) linearly independent eigenvectors. Now, any square matrix of order \( N \) that has \( N \) linearly independent eigenvectors may be diagonalized. Let \( \tilde{P} \) be the modal matrix, formed by stacking the \( N \) eigenvectors side by side. Because the eigenvectors are linearly independent, \( \tilde{P} \) is nonsingular. Equation (6.3a) can be written in its canonical form

\[
\{\tilde{x}(t+1)\} = \tilde{A}\{\tilde{x}(t)\} + \tilde{B}\{u(t)\}
\]  
(6.29)

---

5The condition of distinct eigenvalues is sufficient but not necessary. (Rogers, 1971; p. 412.)
where

\[ \{ \hat{x}(t) \} = \tilde{P}^{-1} \{ \tilde{x}(t) \} \]

\[ \hat{\Lambda} = \tilde{P}^{-1} \tilde{\lambda} \tilde{P} \]

\[ \hat{B} = \tilde{P}^{-1} \tilde{B} , \quad \text{rank} (\hat{B}) = K \leq N . \]

\( \hat{\Lambda} \) is the diagonal matrix of eigenvalues of \( \tilde{G} \).

We now use the result that system controllability is unaffected by any equivalence transformation of the state (Wolovich, 1974; p. 76). The system (6.3a) is controllable if and only if (6.29) is controllable. With \( \hat{\Lambda} \) diagonal, an element \( \hat{x}_i(t + 1) \) is only affected by \( \hat{x}_i(t) \) and is uncoupled from \( \hat{x}_j(t) , j \neq i \). Therefore, a control of \( \hat{x}_i(t + 1) \) requires that

\[ \{ \hat{\beta}_i \}' \{ \hat{u}(t) \} \neq 0 \]

where \( \{ \hat{\beta}_i \}' \) is the \( i \)-th row of \( \hat{B} \). The vector \( \{ \hat{\beta}_i \}' \) must have at least one nonzero element. Preston (1974; p. 69) labels the condition that there exists at least one nonzero element in each row of the transformed instrument coefficient matrix \( \hat{B} = \tilde{P}^{-1} \tilde{B} \) as the coupling criterion. The coupling criterion is an alternative condition for the controllability of the dynamic system (6.3a).

In order to separate the controllable part of a system, it is not necessary to compute all the eigenvalues and eigenvectors. An alternative transformation is given by MacFarlane (1970; pp. 466-469). It starts out from the matrix:
Define $\tilde{S}$ as the $N \times N_k$ matrix obtained by selecting from left to right as many linearly independent columns of $\tilde{D}$ as possible. The column vectors of $\tilde{S}$ span the controllable subspace of the target space, and any vector in this subspace can be expressed as a linear combination of these basic vectors. If the system is controllable, $\tilde{S}$ is of full rank, i.e. $N = N_k$.

If the system is only controllable in part, $N_k < N$. Define any $N \times (N - N_k)$ matrix $\tilde{X}$ such that

$$\tilde{T} = [\tilde{S} \; \tilde{X}]$$

is nonsingular. Then

$$\tilde{T}^{-1} = \begin{bmatrix} Y \\ W \end{bmatrix}$$

$$\tilde{T}^{-1} \tilde{T} = \begin{bmatrix} Y \\ W \end{bmatrix} [\tilde{S} \; \tilde{X}] = \begin{bmatrix} YS & YX \\ WS & WX \end{bmatrix} = \begin{bmatrix} I_{(N_k)} & 0 \\ 0 & I_{(N - N_k)} \end{bmatrix}.$$

Hence, $\tilde{Y}$ and $\tilde{W}$ satisfy the conditions

$$\begin{array}{ll}
\tilde{Y}_S = \tilde{I} & \tilde{Y}_X = 0 \\
\tilde{W}_S = 0 & \tilde{W}_X = \tilde{I}.
\end{array}$$

And the dynamic system

$$\{x(t + 1)\} = \tilde{G}\{x(t)\} + \tilde{B}\{u(t)\}$$

(6.3a)
is transformed to

\[ T^{-1}\{x(t + 1)\} = T^{-1}\tilde{G}T^{-1}\{x(t)\} + T^{-1}B\{u(t)\} \]

\[ \{\hat{x}(t + 1)\} = T^{-1}\tilde{G}\{\hat{x}(t)\} + T^{-1}B\{u(t)\} \]

\[ \{\hat{x}(t + 1)\} = \begin{bmatrix} YGS & YGX \\ ~~~ & ~~~ \end{bmatrix} \{\hat{x}(t)\} + \begin{bmatrix} YB \\ ~~~ \end{bmatrix} \{u(t)\} \quad (6.30) \]

It already has been stated that the controllability of 
(6.3a) is not affected by an equivalence transformation. 
It is also true that the controllable subspace is invariant 
under the operator \( \tilde{G} \). Therefore, for any vector \( \{s_i\} \) in the 
controllable subspace, the vector \( \tilde{G}\{s_i\} \) must lie in the same 
subspace. However, since \( W_\sim = 0 \), the rows of \( \sim W \) are orthog-
onal to the columns of \( \sim S \), and, therefore, to any vector 
lying in the subspace spanned by the columns of \( \sim S \). This 
implies

\[ WGS = 0 \quad \sim \sim \]

The column vectors of \( B \) also belong to the controllable 
subspace spanned by \( \sim S \), so that

\[ W_\sim = 0 \quad \sim \sim \]

We may write

\[ \begin{bmatrix} \{\hat{x}_1(t + 1)\} \\ \{\hat{x}_2(t + 1)\} \end{bmatrix} = \begin{bmatrix} YGS & YGX \\ ~~~ & ~~~ \end{bmatrix} \begin{bmatrix} \{\hat{x}_1(t)\} \\ \{\hat{x}_2(t)\} \end{bmatrix} + \begin{bmatrix} YB \\ ~~~ \end{bmatrix} \{u(t)\} \quad (6.31) \]
It follows that the controllable part of the system is given by

$$\{\hat{x}_1(t + 1)\} = YG\{\hat{x}_1(t)\} + YB\{u(t)\} \quad (6.32)$$

where \(\{\hat{x}_1(t)\}\) has dimension \(N_k \times 1\). The vector \(YG\{\hat{x}_2(t)\}\) can be treated as a known disturbance. The \((N - N_k)\) dimensional subsystem defined by the remaining rows of \((6.31)\), namely

$$\{\hat{x}_2(t + 1)\} = WG\{\hat{x}_2(t)\} \quad (6.33)$$

is completely independent of \(\{u(t)\}\), and therefore is uncontrollable.

d. Achieving the Targets with a Minimal Number of Instruments

Applying the above transformations to uncouple the controllable part of the system from the uncontrollable part, an important question in policy-making may be answered: what is the minimal number of dynamic instruments, needed to steer the system towards a set of targets. Consider \((6.3a)\). Assume that only one instrument \(u_i(t)\) is used in policy implementation. Whether this instrument can transfer the system from \(\{\hat{x}(0)\}\) to \(\{\hat{x}(T)\}\) depends on the controllability of the subsystem

$$\{\hat{x}(t + 1)\} = G\{\hat{x}(t)\} + \{b_i\} u_i(t) \quad (6.34)$$

where \(\{b_i\}\) is the \(i\)-th column of the matrix \(B\). The system \((6.34)\) will be controllable with the \(i\)-th instrument if
the matrix

\[ Q = \{b_1\} \ G \{b_1\} \ G^2 \{b_1\} \ \ldots \ G^{N-1} \{b_1\} \]

is of rank N.

In terms of the coupling criterion, the existence condition is that

\[ \{\hat{b}_i\} = \hat{G}^{-1} \{b_i\} \]

contains N nonzero elements, since the zero elements indicate the noncontrollable part of the system. A zero element occurs in \{\hat{b}_i\} whenever a row of \(\hat{G}^{-1}\) is orthogonal to the vector \{\hat{b}_i\}. The rows of \(\hat{G}^{-1}\) are the normalized left eigenvectors of \(\hat{G}\).\(^6\) Zero elements in \{\hat{b}_i\} are precluded if and only if \{\hat{b}_i\} is linearly dependent on all the N eigenvectors of \(\hat{G}\).\(^7\) Preston (1970; p. 70) refers to this condition as the eigenvector condition. If there exists one instrument that does not violate the eigenvector condition, the system can be controlled by just one instrument. If no instrument satisfies the eigenvector condition, a combination of instruments may still satisfy the coupling criterion, if their nonzero elements mutually offset the zero elements that disqualify them individually. Therefore, the minimal set

\(^6\)The first row of \(\hat{G}^{-1}\) has special meaning in demography. It shows the reproductive value of the population. (Keyfitz, 1968; p. 53.)

\(^7\)This may be compared with the possibility of writing the observed population distribution as a linear combination of the right eigenvectors of \(G\). (Keyfitz, 1968; p. 56.)
of instruments necessary and sufficient for dynamic controllability is equal to the number of columns, $\hat{k}$, of the smallest $\hat{B}_c$ matrix possessing $N$ nonzero rows.

The result that under certain circumstances defined by $\hat{B}$, all the targets can be reached by using only one instrument, is rather intriguing and is totally contrary to the thinking engendered by the Tinbergen framework. It means, for example, that a desired population distribution over $N$ regions can be realized by having a population policy in only one region. The achievement of the target distribution, however, needs time. From looking at the $q$-matrix, it is clear that if there is only one instrument, the objective can only be reached after $N$ periods of time\(^8\). Therefore, there exists a trade-off between the minimal length of the planning horizon and the minimal number of instruments.

If the targets must be reached immediately ($T = 1$), the minimal number of instruments is $N$, since

$$
\hat{Q} = [\hat{G}^0B] = B
$$

(6.35)

must be of rank $N$. Equation (6.35) is the static controllability condition, discussed earlier.

---

\(^8\)It should be remembered that the controllability condition is based on the assumption that no constraints are imposed on the instrument. Constraints would reduce the degrees of freedom associated with dynamic controllability.
CHAPTER 7

DESIGN OF OPTIMAL MIGRATION POLICIES

Any design of optimal policies should begin with a statement of objectives. Thus far we have focused our attention on the description of system dynamics by means of a demometric model. We have answered the question under what conditions it is possible to specify certain objectives or targets and to achieve them by the instruments at hand. Under very specific conditions, there is a unique instrument vector assuring the achievement of the targets. The optimal levels of the instrument variables then follow directly. Under other conditions, however, there is an infinite number of combinations of the instruments that lead to the desired targets. In this case, the policy maker is confronted with an additional decision problem: which alternative set of instruments to choose. This requires the set-up of a cost function or welfare loss function which aggregates the relative costs incurred in the implementation of each instrument. Or the feasible set of instruments may be limited by imposing constraints on them. A further possibility is that the objectives are overstated, i.e. that no combination of instruments can be found that realizes all the targets. The system is uncontrollable and again the policy maker has an additional decision to make: where should he modify his preference system? Is he willing to give up some targets completely in order to achieve the others, or is he satisfied with approximating all the targets without reaching them exactly? This amounts to specifying a welfare function of the target variables of interest. The
coefficients of the welfare function are the trade-offs between the target variables. The specification of the cost and the welfare function is the most difficult and the most socially sensitive task in the policy design process. We have dealt with it in Part I. Here we assume that these functions are given by the policy maker.

This chapter is divided into three sections. The first discusses the design of optimal policies in the Tinbergen framework. It will be shown that in some instances implicit objective functions may be used to derive the optimal policy. The unifying feature of this section is the notion of the generalized inverse. The importance of the minimizing properties of generalized inverses for policy analysis will be illustrated. The other two sections are related to the state-space model and consider time series of controls. The policy problem in which all targets relate to the planning horizon is discussed in the second section. The last section treats the policy design in the case that a target trajectory is given. It applies the theory of optimal control to migration policy problems.

7.1. DESIGN IN THE TINBERGEN FRAMEWORK

From the previous chapter, we know that an optimal policy exists if the rank of the impact multiplier matrix $R$ is equal to the number of targets. The targets may belong to one time period or to different periods. Following Tinbergen, we consider three cases according to the relationship between the number of targets ($N$) and the number of instruments ($K$) or, equivalently, to the rank of the multiplier matrix and its singularity property.
7.1.1. **The Matrix Multiplier is Nonsingular and of Rank N**

Recall equation (5.3):

\[ \{ \bar{y} \} = R\{z_1\} + S\{z_2\} \]  \hspace{1cm} (5.3)

If \( R \) is nonsingular, then the optimal policy is unique and given by (5.4)

\[ \{ \bar{z}_1 \} = R^{-1}[\{ \bar{y} \} - S\{z_2\}] \]  \hspace{1cm} (5.4)

It is clear from (5.4) that the policy depends not only on the target vector, but also on the uncontrollable variables. If \( \{z_2\} \) has some lagged endogenous variables, then the effects of past policies will be felt in the current policy.

The nature of the dependence of \( \{ \bar{z}_1 \} \) upon \( \{ \bar{y} \} \) is associated with different types of structures of the matrix \( R \). They were discussed in Chapter 5. Since there is only one possible set of instruments leading to the target vector \( \{ \bar{y} \} \), no cost or welfare function is needed to distinguish between alternatives.

7.1.2. **The Matrix Multiplier is Singular and of Rank N**

If \( N < K \), there exists an infinite number of instrument vectors which lead to the achievement of a preassigned value of the target vector. The solution set to (5.3) may be represented by

\[ R\{\bar{z}_1\} = \{\bar{y}\} - S\{z_2\} \]

\[ \{\bar{z}_1\} = R^{(1)}[\{\bar{y}\} - S\{z_2\}] + \left[I - R^{(1)}R\right]\{\zeta\} \]  \hspace{1cm} (7.1)
where $\mathbf{R}^{(1)}$ is a generalized inverse of $\mathbf{R}$, satisfying 
\[
\mathbf{R} \mathbf{R}^{(1)} = \mathbf{R}
\]
and $\{\mathbf{z}\}$ is an arbitrary vector.

In order to get a unique instrument vector, one must impose additional conditions on $\{\mathbf{z}_1\}$. Two illustrations are given of how this may be done. Both minimize a function of $\{\mathbf{z}_1\}$ over a constrained set. The first illustration is the formulation of a general mathematical programming problem. The second makes use of the minimizing properties of some types of generalized inverses.

**Illustration a:** Suppose a cost or welfare loss function $f(\{\mathbf{z}_1\})$ has been defined. One wants to minimize this function subject to the dynamic behavior of the system and to some other constraints imposed upon the instrument vector and represented by the vector-valued inequality $g(\{\mathbf{z}_1\}) \geq 0$. The problem then may be formulated as a mathematical programming problem,

\[
\min f(\{\mathbf{z}_1\})
\]

subject to

\[
\{\mathbf{y}\} = \mathbf{z}\{\mathbf{z}_1\} + \mathbf{S}\{\mathbf{z}_2\} \tag{7.2}
\]

\[
g(\{\mathbf{z}_1\}) \geq 0 .
\]

If $g(\{\mathbf{z}_2\})$ and $f(\{\mathbf{z}_1\})$ are both linear, the problem is a linear programming problem and can be solved by the simplex technique.
Illustration b: This illustration is a special case of the problem (7.2). We delete the constraint \( g(\{z_1\}) \geq 0 \), and we let \( f(\{z_1\}) \) be the Euclidean norm defined on \( \{z_1\} \), i.e.,

\[
f(\{z_1\}) = \left[ \{z_1\}' \{z_1\} \right]^{\frac{1}{2}}.
\]

(7.3)

Ben-Israel and Greville (1974; p. 114) prove that the unique solution to this problem is given by

\[
\{z_1\} = R^{(1,4)}_r \left[ \{v\} - S\{z_2\} \right]
\]

(7.4)

where \( R^{(1,4)}_r \) is a generalized inverse satisfying

\[
R^{(1,4)}_r R = R
\]

and

\[
\left[ R^{(1,4)}_r \right]' = \left[ R^{(1,4)}_r R \right].
\]

Because \( R^{(1,4)}_r \) defines a minimum norm solution to (5.3), it is often called the "minimum-norm inverse."

There may be other norms defined on the instrument vector. Suppose the policy maker lists some most acceptable values of the instrument variables \( \{z_1\} \), and wants to minimize the squared deviation between the optimal values and these preassigned values. The policy model is then
\[
\min f(\{z_1\}) = \|\{z_1\} - \{\bar{z}_1\}\| = \left\| \left[ \{z_1\} - \{\bar{z}_1\} \right] \right\|^2 \\
\text{s.t. } \{\bar{z}_1\} = R\{z_1\} + S\{z_2\}.
\]

The optimal solution is given by

\[
\{z_1\} = R^{(1,4)} \left[ \{\bar{z}_1\} - S\{z_2\} \right] + \left[ I - R^{(1,4)} R \right] \{\bar{z}_1\}.
\]

The matrix $R^{(1,4)}$ has a special meaning for policy analysis. An element $r_{ij}^{(1,4)}$ indicates the change in the $i$-th instrument variable required for a unit change in the $j$-th target variable, assuming that $\{z_2\}$, and, in the second case, also $\{\bar{z}_1\}$ remain unchanged. It is, therefore, a multiplier in the economic sense, measuring the relative effectiveness of the $i$-th instrument.

7.1.3. The Matrix Multiplier is Singular and of Rank $K$

If $N > K$, the system (5.3) is inconsistent and no solution exists, i.e. the residual vector $\{r\}$ is nonzero, where

\[
\{r\} = \left[ \{\bar{y}\} - S\{z_2\} \right] - R\{z_1\} = \{\bar{y}\} - \{\bar{z}_1\}
\]

where $\{\bar{y}\}$ is the realized value of the target vector.

In this case, it is common to search for an approximate solution of (5.3), which makes $\{r\}$ closest to zero in some sense. Again two illustrations will be given. As before, the first is a mathematical programming model, namely, a quadratic programming model, and the second applies the minimizing properties of some generalized inverses.
Illustration a: Theil (1964; p. 159) was the first to assume that a policy-maker, confronted with an overstatement of his goals set, i.e. $N > K$, formulates his preferences as a quadratic function of the target and control variables. The Theil model has been given in Chapter 5 without proposing a solution to it. Recall the model

\[
\min W(\{z_1\}) = \{a\}'\{z_1\} + \{b\}'\{y\} + \frac{1}{2}\{z_1\}'\{\alpha\}{\tilde{z}}_1\{z_1\} \\
+ \{y\}'\{\beta\}{\tilde{y}} + \{z_1\}'\{\gamma\}{\tilde{z}}_2 + \{y\}'\{\delta\}'\{z_1\}
\]

s.t. $\{\hat{y}\} = \gamma z_1 + \delta z_2 \tag{5.3}$

where $\alpha$, $\beta$, $\gamma$ are symmetric positive definite weight matrices. This optimization problem may be solved by means of the Lagrangian technique. An alternative method of deriving the optimum consists of using the constraints to eliminate the target vector in the objective function and then minimizing this function unconditionally with respect to the instruments (Theil, 1964; pp. 40-41). This solution procedure is also followed by Friedman (1975; pp. 159).

Substituting the constraint in the objective function gives

\[
W(\{z_1\}) = \kappa_0 + \{k\}'\{z_1\} + \frac{1}{2}\{z_1\}'\kappa\{z_1\} \tag{7.7}
\]

where

\[
\kappa_0 = \{b\}'\delta z_2 + \frac{1}{2}\left[\delta z_2\right]'\beta\left[z_2\right]
\]
\[
\{k\} = \{a\} + R'\{b\} + [C + R'B][S\{z_2\}]
\]

\[
\k = A + R'BR + CR + R'C'.
\]

The first order condition for minimizing \(W(\{z_1\})\) with respect to the instrument vector \(\{z_1\}\) is

\[
\frac{dW(\{z_1\})}{d\{z_1\}} = \{0\} = \{k\} + \k\{z_1\}.
\]

The optimal solution follows immediately

\[
\{z_1\} = -\k^{-1}\{k\}
\]

where \(\k\) and \(\{k\}\) are as defined in (7.7). The second order condition for the minimization of \(W(\{z_1\})\) with respect to \(\{z_1\}\) is that \(\k\) is positive definite. The corresponding value of the target vector is

\[
\{\hat{y}\} = RK^{-1}\{k\} + S\{z_2\}.
\]

It should be noted that a nontrivial solution to (7.9) exists only if \(\{k\}\) is nonzero.

Illustration b: Suppose the policy maker only wants to minimize \(\{r\}\). The model may be considered as a variant of the Theil model.

\[
\min \left[ ((\{\hat{y}\} - \{\hat{y}\})'([\{\hat{y}\} - \{\hat{y}\}]^2)^{1/2}\right]
\]

s.t. \(\{\hat{y}\} = \k\{z_1\} + S\{z_2\}\).
The objective function defines the Euclidean norm of \( \{r_1\} \). Ben-Israel and Greville (1974; p. 104) show that the optimal solution to this problem is given by:

\[
\{z_1\} = \mathbb{R}^{(1,3)}_{\sim} \left[ \{\overline{y}\} - S\{z_2\} \right]
\]  

(7.10)

where \( \mathbb{R}^{(1,3)}_{\sim} \) is the generalized inverse of \( \sim \) satisfying

\[
[\mathbb{R}^{(1,3)}_{\sim} \sim]^T = \mathbb{R}^{(1,3)}_{\sim}
\]

Because of the property that \( \mathbb{R}^{(1,3)}_{\sim} \) minimizes the Euclidean norm of the residual vector, i.e., the sum of squares of the residuals, it is called the "least-squares inverse." An element \( r_{ij}^{(1,3)} \) indicates how much the i-th instrument has to change for a unit change in the j-th target variable, in order to maintain the smallest sum of squared deviations between the realized and the preassigned values of the target variables. The general least-squares solution is

\[
\{z_1\} = \mathbb{R}^{(1,3)}_{\sim} \left[ \{\overline{y}\} - S\{z_2\} \right] + \left[ I - \mathbb{R}^{(1,3)}_{\sim} \sim \right] \{c\}
\]  

(7.11)

where \( \{c\} \) is an arbitrary \( K \times 1 \) vector.

Ben-Israel and Greville note that the least-squares solution is unique only when \( \sim \) is of full column rank. This condition is always satisfied in policy models discussed here, since we have assumed initially that the instruments are linearly independent.
This illustration shows that the least-squares generalized inverse is the solution to a special variant of the Theil model. A similar observation has recently been made by Russell and Smith (1975; p. 143).

7.2 DESIGN IN THE STATE-SPACE FRAMEWORK: FIXED TARGETS AT THE PLANNING HORIZON

Consider the discrete system

\[
\{x(t + 1)\} = \{G\{x(t)\} + B\{u(t)\}\}
\]

(6.3a)

\[
\{y(t)\} = \{C\{x(t)\}\}
\]

(6.3b)

where \(\{x(t)\}\) is the population distribution at time \(t\).

It can be the age distribution, the regional distribution, or both.

\(\{y(t)\}\) is any policy relevant measure dependent on the population distribution.

\(\{u(t)\}\) is the intervention vector, control or instrument vector at time \(t\).

\(G\) is the \(N \times N\) growth matrix without intervention.

\(B\) is the \(N \times K\) dynamic impact multiplier matrix.

\(C\) is the \(P \times N\) conversion matrix.

In the following, we make the simplifying assumption that \(C\) is the identity matrix. The solution to (6.3a) for \(t_0 = 0\) is
\[ \{x(t)\} = G^t \{x(0)\} + \sum_{i=0}^{t-1} G^{t-1-i} B\{u(i)\} \quad . \]  

(6.4)

The policy design problem starts out from (6.4) and seeks to answer the question: what is the sequence of control vectors \{u(i)\}, such that, given the initial condition \{x(0)\} and the assumption of time-invariance of the coefficient matrices, a target vector at the horizon \{x(T)\} will be reached in an optimal manner. The intermediate states are of no importance in this formulation.

Equation (6.17) may be written as

\[
\begin{bmatrix}
\{u(T - 1)\} \\
\vdots \\
\{u(1)\} \\
\{u(0)\}
\end{bmatrix}
= \begin{bmatrix}
G^T \{x(0)\} \\
\vdots \\
G \{x(0)\}
\end{bmatrix}
= D \{\hat{u}\} , \quad \text{say} . \]

(7.12)

(7.13)

The system is state controllable if the \(N \times KT\) matrix \(D\) is of rank \(N\), where \(N\) is the dimension of the target vector \(\{x(T)\}\). The controllability condition implies that \(N \leq KT\). We distinguish two cases: \(N = KT\) and \(N < KT\).

**CASE 1: \(N = KT\)**

In the dynamic policy model, it is the combined magnitude of the number of instruments and the planning horizon that determines the state controllability. In the previous chapter, we saw that a trade-off exists between the minimal length of the planning horizon and the minimal number of instruments. Any target vector may be reached
by only one instrument, provided that the planning horizon
is not less than N. Also, any target vector can be achieved
in only one time period, if the policy maker may handle at
least N instruments\(^9\). If \(N = KT\), and if the instruments
of the different time periods are independent, then \(D\) is
nonsingular, and the unique control sequence is

\[
\{\tilde{u}\} = D^{-1} [(\tilde{x}(T)) - C^T \{x(0)\}], \tag{7.14}
\]

where \(\{\tilde{x}(T)\}\) is the target vector at the planning horizon.

**CASE 2: \(N < KT\)**

If \(D\) has rank \(N\) and is rectangular, then an infinite
number of combinations of the controls leads to the pre-
defined target population. The solution of (7.13) is

\[
\{\tilde{u}\} = D^{(1)} [(\tilde{x}(T)) - C^T \{x(0)\}] + [\tilde{D}^{(1)} \tilde{D} - I] \{c\} \tag{7.15}
\]

where \(\{c\}\) is arbitrary,
and \(\tilde{D}^{(1)}\) is a generalized inverse of \(\tilde{D}\).

In order to find a unique policy, the policy maker may
minimize a cost function of the instruments, he may put
constraints on the instruments, or he may do both. The
introduction of a cost function will be discussed at the
end of this section. First, we deal with the imposition

\(^9\)This is exactly the controllability condition
derived by Tinbergen for a static policy model.
of constraints on the instruments to ensure uniqueness of the instrument vector. The idea is to reduce the degrees of freedom of the policy measures by making the instrument vector at a certain time period depend on the controls exercised at previous time periods. Two reduction methods are distinguished. The first formulates the control vector at time $t$ as a linear combination of the control vector at $t-1$. This implies that the control at $t$ may be directly related to the control vector at the initial time period. Therefore, we call this the initial period control. This method has been developed by Rogers (1966; 1968, Chapter 6; 1971; pp. 98-108) for migration policy purposes. The second reduction method, known as feedback control, makes the control vector at time $t$ a linear function of the state vector at the same time $t$.

7.2.1. Initial Period Control

Suppose that the control vector at time period $t$ is

\[ \{u(t)\} = \tilde{W}\{u(t - 1)\} \]  (7.16)

where $\tilde{W}$ is nonsingular, fixed, and known.

Recall the state-space model of (6.3a):

\[ \{\tilde{x}(t + 1)\} = \tilde{G}\{\tilde{x}(t)\} + \tilde{B}\{u(t)\} \]  (6.3a)

Its solution is given by (6.4). Let $t = T$ be the planning horizon, then
\[
\{x(T)\} = G^T\{x(0)\} + \sum_{i=0}^{T-1} G^{T-1-i} B\{u(i)\}.
\] (7.17)

But

\[
\{u(i)\} = W\{u(i - 1)\} = W^i\{u(0)\}.
\]

Therefore (7.17) becomes

\[
\{x(T)\} = G^T\{x(0)\} + \left[ \sum_{i=0}^{T-1} G^{T-1-i} B W^i \right] \{u(0)\}
\]

and

\[
\{x(T)\} - G^T\{x(0)\} = \left[ \sum_{i=0}^{T-1} G^{T-1-i} B W^i \right] \{u(0)\}.
\] (7.18)

Let

\[
A(T) = \sum_{i=0}^{T-1} G^{T-1-i} B W^i,
\]

then

equation (7.18) may be written as

\[
\{x(T)\} - G^T\{x(0)\} = A(T)\{u(0)\}
\] (7.19)

which is in fact the formulation of the Tinbergen model, with \{x(T)\} the target vector, \{x(0)\} the vector of uncontrollable variables and \{u(0)\} the control vector. The multiperiod problem (7.17) with the target vector given for the planning horizon, and with the control vector at each
The time period being a linear combination of the control vector at the initial time period, is in fact a single-period problem. Only the control at the initial time period must be specified. The existence and the uniqueness of \( \{u(0)\} \) depends only on the rank of \( B \) and is independent of the choice of \( T \). If \( B \) is nonsingular, then the unique and optimal value of \( \{u(0)\} \) is given by

\[
\{u(0)\} = A^{-1}(T) \left[ \{x(T)\} - G^T \{x(0)\} \right].
\]  

(7.20)

A special case of the initial period control is discussed by Rogers (1971; pp. 99-100). Suppose that \( W \) is a scalar matrix, i.e.

\[
W = wI
\]

with \( w \) being a scalar. It means that the controls change in time at a constant rate. Equation (7.18) may be rewritten as

\[
\{x(T)\} - G^T \{x(0)\} = \left[ \sum_{i=0}^{T-1} w^{T-1-i} G^i B \right] \{u(0)\}.
\]

Premultiplying both sides with \( (W - G) \) gives

\[
(W - G) \left[ \{x(T)\} - G^T \{x(0)\} \right] = (W - G) \left[ \sum_{i=0}^{T-1} w^{T-1-i} G^i B \right] \{u(0)\}
\]

\[
= \left[ \sum_{i=0}^{T-1} w^{T-1-i} G^i \right] B \{u(0)\}
\]

\[
= \left[ w^T + w^{T-1} G + w^{T-1} G^2 + \ldots - w^{T-1} G - w^{T-1} G^2 \ldots \right.
\]

\[
- w^0 G^T \right] B \{u(0)\}
\]
\[ \begin{bmatrix} w^T I - G^T \end{bmatrix} \beta(u(0)) \]

Therefore

\[ \begin{bmatrix} \beta(wI - G) \{x(T)\} - G^T \{x(0)\} \end{bmatrix} = \begin{bmatrix} w^T I - G^T \end{bmatrix} \beta(u(0)) \]  

\[ \{x(T)\} = G^T \{x(0)\} + (wI - G)^{-1} \begin{bmatrix} w^T I - G^T \end{bmatrix} \beta(u(0)) \]  

(7.21)

and, given that \( \beta \) is nonsingular,

\[ \{u(0)\} = \beta^{-1} (w^T I - G^T)^{-1} (wI - G) \begin{bmatrix} \{x(T)\} - G^T \{x(0)\} \end{bmatrix} \]  

(7.22)

which is in fact also a single-period problem:

\[ \{\bar{u}\} = \hat{A}(T) \{\bar{x}(T)\} - \{a(T)\} \]  

(7.23)

where

\[ \hat{A}(T) = \beta^{-1} (w^T I - G^T)^{-1} (wI - G) \]

\[ \{a(T)\} = \beta^{-1} (w^T I - G^T)^{-1} (wI - G) G^T \{x(0)\} \]

The special case, \( w = 1 \), is the intervention model of Rogers (1971; pp.99-100) with constant policy.

We now consider two illustrations of the initial period control model. We will assume that \( w \) is equal to the identity matrix. The constant instrument vector is given by

\[ \beta(\bar{u}) = (I - G^T)^{-1} (I - G) \begin{bmatrix} \{x(T)\} - G^T \{x(0)\} \end{bmatrix} \]  

(7.24)
and the target vector is

\[ \{\tilde{x}(T)\} = G^T\{\tilde{x}(0)\} + (\tilde{I} - G)^{-1} (\tilde{I} - G^T) \tilde{B}\{\tilde{u}\} \] (7.25)

The first illustration is the stationary population model and the second is the pure migration model.

**Illustration a: Stationary Population Model.**

In the literature on zero population growth, it is emphasized that an immediate reduction of fertility to replacement level will result in a further increase of the population for at least 50 years in most countries. A policy that would keep the population constant at the current level, would result in an unrealistic fluctuation of fertility and mortality rates over the next decades (Coale, 1972; p. 595).

In the multiregional case, keeping total population as well as the population distribution at the current level implies that \( \{\tilde{x}(T)\} = \{\tilde{x}(0)\} \), hence we have that

\[ \{\tilde{x}(0)\} = G^T\{\tilde{x}(0)\} + (\tilde{I} - G)^{-1} (\tilde{I} - G^T) \tilde{B}\{\tilde{u}\} \]

\[ (\tilde{I} - G^T) \{\tilde{x}(0)\} = (\tilde{I} - G)^{-1} (\tilde{I} - G^T) \tilde{B}\{\tilde{u}\} \]

\[ \tilde{B}\{\tilde{u}\} = (\tilde{I} - G^T)^{-1} (\tilde{I} - G)(\tilde{I} - G^T) \{\tilde{x}(0)\} \]

\[ = \{\tilde{x}(0)\} - (\tilde{I} - G^T)^{-1} G(\tilde{I} - G^T) \{\tilde{x}(0)\} \]

\[ = \{\tilde{x}(0)\} - (\tilde{I} - G^T)^{-1} [G - G^{T+1}] \{\tilde{x}(0)\} \]

\[ = \{\tilde{x}(0)\} - (\tilde{I} - G^T)^{-1} (\tilde{I} - G^T) G\{\tilde{x}(0)\} \]
\[ \tilde{B}(\tilde{u}) = (\tilde{I} - \tilde{G}) \{\tilde{x}(0)\} . \]

If \( \tilde{B} \) is nonsingular then

\[ \{\tilde{u}\} = \tilde{B}^{-1}(\tilde{I} - \tilde{G}) \{\tilde{x}(0)\} . \quad (7.26) \]

If \( \tilde{B} \) is singular, we have

\[ \{\tilde{u}\} = \tilde{B}^{(1)}(\tilde{I} - \tilde{G}) \{\tilde{x}(0)\} + [\tilde{B}^{(1)}\tilde{B} - \tilde{I}] \{\tilde{c}\} . \quad (7.27) \]

The result then may be given as follows: If equation (7.24) has a solution for an arbitrary \( T \), then there exists a constant policy vector \( \{\tilde{u}\} \) which keeps the total population as well as its distribution constant at the current level.

The vector \( \{\tilde{u}\} \) does not depend on the planning horizon, but only on the current population level and distribution.

**Illustration b:** Pure Migration Model.

The procedure to compute the intervention vector is described by Rogers (1971; p. 106) as follows. The migration rates are taken out of the growth matrix and the migration flows are introduced via the control vector \( \{\tilde{u}\} \). The new matrix is \( \tilde{S} \). However, an in-migrant with respect to one region is an out-migrant with respect to another region, and therefore net internal migration must be equal to zero. The instruments are not independent. After computing \( \{\tilde{u}\} \) by (7.24) with the revised growth matrix \( \tilde{S} \), and a target vector \( \{\tilde{x}(T)\} \) by (7.25), some elements of \( \{\tilde{u}\} \) are adjusted such that

\[ \{1\}' \{\tilde{u}\} = 0 \]
where \( \{1\} \) is a vector of ones. A change of an element \( \tilde{u}_i \) implies that the target population of region \( i \) will not be reached. \( x_i(T) \) becomes uncontrollable. After the adjustment procedure, the revised target population is computed using (7.25).

For an illustration of another approach that draws on the controllability concept, consider (7.25) once again:

\[
\{x(T)\} = G^T\{x(0)\} + (I - G)^{-1} (I - G^T) B\{\tilde{u}\}
\] (7.25)

where \( G \) is the unreduced growth matrix. Any linear constraints on \( \{\tilde{u}\} \) may be introduced in (7.25) via \( B \). The idea is similar to the introduction of linear restrictions in the general linear regression model (Johnston, 1972; p. 157). In the general case where \( \{\tilde{u}\} \) is unrestricted, \( B \) is the identity matrix.

Suppose the policy problem is to find \( \{\tilde{u}\} \) such that \( \{x(T)\} = \{\tilde{x}(T)\} \) is the target vector, and such that the level of the fourth control variable is equal to the sum of the first and the third variable, i.e.

\[
u_4 = u_1 + u_3.
\] (7.28)

Equation (7.25) may then be written as:

\[
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
u_1 \\
u_2 \\
u_3 \\
u_4 \\
u_5 \\
\vdots \\
\end{bmatrix}
= 
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
u_1 \\
u_2 \\
u_3 \\
u_4 \\
u_5 \\
\vdots \\
\end{bmatrix}
\] (7.29)
Because of the linear restriction (7.28), $\mathbf{B}$ is no longer of full rank. Since the instrument vector must remain constant in time, (7.25) may be written as

$$\{\mathbf{y}(T)\} = \mathbf{FB}\{\mathbf{\bar{u}}\} \quad (7.30)$$

where

$$\{\mathbf{y}(T)\} = \{\mathbf{\bar{x}}(T)\} - \mathbf{G}^T\{\mathbf{x}(0)\}$$

$$\mathbf{F} = (\mathbf{I} - \mathbf{G})^{-1}(\mathbf{I} - \mathbf{G})^T.$$

Equation (7.29) is equivalent to a static policy problem. By Theorem 1, it has a solution if

$$\text{rank } (\mathbf{FB}) = N.$$ 

Since $\mathbf{I}$, $\mathbf{G}$ and $\mathbf{G}^T$ are nonsingular, $\mathbf{F}$ is nonsingular. Therefore (Lancaster, 1969; p. 45):

$$\text{rank } (\mathbf{FB}) = \text{rank } (\mathbf{B}) = N - 1,$$

and the system is not controllable\textsuperscript{10}. Because the fourth column of $\mathbf{FB}$ is $\{\mathbf{0}\}$, $\mathbf{u}_4$ may be deleted, and (7.30) becomes

\textsuperscript{10}The fourth column of $\mathbf{FB}$ is $\{\mathbf{0}\}$.
\[
\{\tilde{y}(T)\} = \mathbf{F} \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
\tilde{u}_1 \\
\tilde{u}_2 \\
\tilde{u}_3 \\
\vdots \\
\tilde{u}_N \\
\end{bmatrix}.
\] (7.31)

The non-controllable part of the system may be determined by the methods described in Chapter 6. If \{\tilde{u}\} is a vector of in-migrations, it is immediately clear that \(x_4(T)\), or equivalently \(y_4(T)\), cannot be controlled. One may delete \(\tilde{y}_4(T)\) and the fourth row of \(\mathbf{FB}\), giving a new vector \(\{\tilde{y}_1(T)\}\) and a new matrix \(\left(\mathbf{FB}\right)_1\) respectively. The instrument vector \(\{\tilde{u}\}\) is then found as

\[
\{\tilde{u}\} = \left(\left(\mathbf{FB}\right)_1\right)^{-1} \{\tilde{y}_1(T)\}.
\] (7.32)

Entering \(\{\tilde{u}\}\) in (7.31) gives the value of \(\tilde{y}_4(T)\), which will not coincide with the target value.

In the pure migration model the net internal migration must add up to zero. The restriction on \(\{\tilde{u}\}\) is

\[
\{1\}' \{\tilde{u}\} = 0.
\] (7.33)

In a two region case, the people leaving one region must enter the other. The incorporation of this constraint in (7.25) or (7.30) yields

\[
\{\tilde{y}(T)\} = \mathbf{F} \begin{bmatrix}
1 & 0 \\
-1 & 0 \\
\end{bmatrix}
\begin{bmatrix}
\tilde{u}_1 \\
\tilde{u}_2 \\
\end{bmatrix}
\]
hence \( \tilde{u}_2 \) may be deleted. The system is not controllable. If the number of regions is greater than two, and the policy maker is interested in setting a target for only one region, then various combinations of \( \tilde{u}_1 \)'s satisfy the constraint (7.33). At the planning horizon, the population distribution over the other regions depends on the combination chosen initially, i.e. the entries of \( \tilde{B} \).

It has been assumed throughout this section that the policy maker is willing and able to give up an element of his target vector for each linear constraint on the instrument variables. By doing so, he makes it possible to achieve the other target variables exactly. In some situations, it may be more realistic to assume that he wants to approximate the target vector as closely as possible with the restricted instrument vector. The vector \( \{\tilde{u}\} \) which minimizes the deviation between the realized \( \{\tilde{y}(T)\} \) and the target \( \{\tilde{y}(T)\} \) is

\[
\{\tilde{u}\} = \tilde{B}^{(1,3)} \tilde{F}^{-1} \{\tilde{y}(T)\}
\]

(7.34)

where \( \tilde{B}^{(1,3)} \) is the least-squares generalized inverse of \( \tilde{B} \), defined in (7.10).

7.2.2. Linear Feedback Control

Suppose that the intervention vector at time \( t \) is a linear function of the population distribution:

\[
\{u(t)\} = \tilde{Z}\{\tilde{y}(t)\}
\]

(7.35)
Equation (7.35) is known as the linear state variable feedback control law (Wolovich, 1974; p. 195). If the policy authorities choose the value of the policy instruments according to equation (7.35), then their actions cease to represent an external influence, but instead become part of the population system. The feedback control law defines a closed-loop solution to the optimal control problem.

Equation (7.35) is the simplest case of linear state feedback control. It is unrealistic in the sense that it takes all freedom of action out of the hands of the policy-makers. A linear state feedback control of the form

\[ \{u(t)\} = \bar{Z}\{\bar{x}(t)\} + \bar{H}\{\bar{y}(t)\} \quad . \] (7.39)

is certainly more realistic.

Here \(\{\bar{y}(t)\}\) is an external input or a vector of real exogenous variables. The state space representation of the compensated system is obtained by substituting for \(\{\bar{y}(t)\}\) in (6.3)

\[ \{\bar{x}(t + 1)\} = (\bar{G} + \bar{B}\bar{E})\{\bar{x}(t)\} + \bar{B}\bar{H}\{\bar{y}(t)\} \quad . \] (7.40a)

\[ \{\bar{y}(t)\} = (\bar{C} + \bar{E}\bar{Z})\{\bar{x}(t)\} + \bar{E}\bar{H}\{\bar{y}(t)\} \quad . \] (7.40b)

Instead of a state feedback, one can also imagine a linear output feedback

\[ \{u(t)\} = \bar{F}\{\bar{y}(t)\} + \bar{H}\{\bar{y}(t)\} \quad . \] (7.41)
Substituting (7.41) in (6.3), we obtain

$$\{x(t+1)\} = \tilde{G}\{x(t)\} + \tilde{B}_F\{y(t)\} + \tilde{B}_H\{y(t)\}$$

$$\{y(t)\} = \tilde{C}\{x(t)\} + \tilde{E}_F\{y(t)\} + \tilde{E}_H\{y(t)\}$$

$$(\tilde{I} - \tilde{E}_F) \{\tilde{y}(t)\} = \tilde{C}\{x(t)\} + \tilde{E}_H\{y(t)\} .$$

If $$(\tilde{I} - \tilde{E}_F)$$ is nonsingular, then

$$\{\tilde{y}(t)\} = (\tilde{I} - \tilde{E}_F)^{-1} \{\tilde{C}\{x(t)\} + \tilde{E}_H\{y(t)\}\} \quad (7.42)$$

$$\{x(t+1)\} = \left[\tilde{G} + \tilde{B}_F(\tilde{I} - \tilde{E}_F)^{-1} \tilde{C}\right] \{\tilde{x}(t)\}$$

$$+ \left[\tilde{B}_H + \tilde{B}_F(\tilde{I} - \tilde{E}_F)^{-1} \tilde{E}_H\right] \{\tilde{y}(t)\} .$$

If $$\{\tilde{y}(t)\} = \{0\}$$, i.e., the dynamics of the system is governed by a pure output feedback, then we have the closed loop system

$$\{x(t+1)\} = \left[\tilde{G} + \tilde{B}_F(\tilde{I} - \tilde{E}_F)^{-1} \tilde{C}\right] \{x(t)\} . \quad (7.44)$$

There are two noteworthy special cases of (7.44):

i) $$\tilde{E} = 0$$, i.e., the output $$\{\tilde{y}(t)\}$$ depends only on the state vector $$\{\tilde{x}(t)\}$$. Equation (7.44) becomes:

$$\{\tilde{x}(t+1)\} = [\tilde{G} + \tilde{B}_F\tilde{C}] \{\tilde{x}(t)\} .$$
ii) \( E = 0 \) and \( \zeta = I \), i.e., the output vector is equal to the state vector. Equation (7.44) then reduces to the expression for a linear state feedback control law

\[
\{x(t + 1)\} = [\tilde{G} + BF] \{x(t)\}.
\] 

(7.45)

To illustrate the usefulness of the feedback control model for migration policy, we take a policy problem described by Hansen (1974; p. 17). In the last twenty years, central governments of several Western countries have been trying to decrease regional differences in living conditions. A popular strategy to achieve this objective, which is based on equity considerations, was to allocate development funds to lagged regions. The funds allocated by the central government to the regions are a function of its "backwardness." A major indicator of it is the level of out-migration. In order to model this policy, assume that the development funds each region gets at time \( t \) is a linear combination of its level of out-migration and the level of out-migration of all the other regions.

Let \( \{x(t)\} \) be the regional population distribution at time \( t \),

\( \{u(t)\} \) be the regional distribution of the development funds,

\( \{y(t)\} \) be the level of out-migration of the regions.
The dynamics of the multiregional system is described by the state-space model

\[
\begin{align*}
\{x(t + 1)\} &= G\{x(t)\} + B\{u(t)\} \\
\{y(t)\} &= C\{x(t)\}
\end{align*}
\] (6.3a)

(6.3b)

where \(G\) is the population growth matrix,
\(B\) is the matrix of impact multipliers. The element \(b_{ij}\) is the effect of a dollar allocated to region \(j\) at time \(t\) on the population of region \(i\) at time \(t + 1\),
\(C\) is a diagonal matrix of out-migration rates.

The policy may be written as a linear output feedback control law

\[
\{u(t)\} = F\{y(t)\}
\]

where the \(i\)-th row of \(F\) gives the coefficients of the linear combination between \(u_i(t)\) and the regional levels of out-migration. The dynamics of the controlled population system is then given by:

\[
\{x(t + 1)\} = [G + BFC] \{x(t)\}
\] (7.45)

where \(G\) is a nonnegative matrix,
\(B\) has supposedly nonnegative diagonal elements and nonpositive off-diagonal elements,
F describes the trade-offs set by the policy maker. It is realistic to assume that the diagonal elements are positive and most off-diagonal elements are nonpositive. A positive off-diagonal element \( f_{ij} \) would mean that the funds region \( i \) gets increase with the out-migration of region \( j \). This is not unrealistic if the out-migrants of \( j \), who go to \( i \), cause a congestion problem in region \( i \) necessitating additional investments (population responsive policy).

7.2.3. Horizon Constrained Optimal Control

If the number of target variables at the planning horizon is less than the product of the number of instruments and the length of the planning horizon, then there is an infinite number of combinations of controls leading to the desired target variables. Suppose, as before, that the target is the regional population distribution at the horizon \( T \). All the feasible control vectors are given by (7.15), which is the general solution to (7.12).

To arrive at a unique instrument vector, the policy maker may apply the design techniques described under the Tinbergen framework to this multiperiod situation. The first technique is based on the minimizing properties of the generalized inverse. If (7.15) is the general solution to (7.12), then there is a unique solution which minimizes the Euclidean norm of the instrument vector \( \{u\} \). This solution is given by

\[
\{\tilde{u}\} = D^{(1,4)} \left[ \{\tilde{x}(T)\} - G^T\{x(0)\} \right]
\]  
(7.46)
where $D^{(1,4)}$ is the "minimum norm inverse" of $\bar{D}$.

The other approach is to formulate a mathematical programming model, similar to (7.2). However, we have seen in Chapter 5 that by assuming inter-temporal separability of the objectives, and by neglecting the inequality constraint of (7.2), we may write it as an optimal control problem. Assuming, in addition, a quadratic objective functional, the problem becomes identical to (5.19), except for the addition of the horizon constraint,

$$\{\bar{x}(T)\} = \{\bar{x}(T)\} .$$  \hspace{1cm} (7.47)

In the literature, this problem is known as the linear-quadratic control problem with zero terminal error or with a right-hand-side constraint. The solution will be discussed in the next section.

7.3 DESIGN IN THE STATE-SPACE FRAMEWORK: TRAJECTORY OPTIMIZATION

In the models discussed in the previous section, the migration policy objectives were formulated only for the planning horizon. It was assumed that the policy-maker did not care about how the target variables converged to their desired values. In order to identify a unique combination of instruments, we have imposed severe restrictions on the path of the control vector. Now, we broaden the perspective by allowing the policy-maker to define a dynamic preference system, i.e., the targets are defined for each time period instead of only one. The instruments may vary more freely
in the sense that no fixed pattern is imposed. The range
of admissible instruments and their variation, however,
may be constrained for economic, political or stability
reasons. The latter means that the inclusion of the
instruments in the policy maker's preference function is
an appropriate way to avoid an excessive fluctuation of
the values of the instruments over time (Holbrook, 1972;
p. 57).

It has been argued in Chapter 5 that, if the policy-
maker seeks to define a time path of the control vector
out of all the feasible trajectories, such that his dynamic
preference system, expressed in the form of a functional,
is optimized, the policy problem becomes very similar to
the optimal control problem. The models presented earlier
may also be encompassed in this framework. In what follows,
we specify a dynamic policy model using the optimal
control technique. This enables us to list the set of
necessary conditions for optimizing the preference func-
tional. These conditions are known as the Pontryagin
minimum (or maximum) principle.

The optimal control problem specified here covers a
wide variety of dynamic policy problems. Its solution
however is at least quite difficult and its interpretation
is not always easy. A frequently used policy model in the
economic literature is the linear-quadratic model\(^{11}\). It
is characterized by a quadratic objective function and a

\(^{11}\) See, for example, Sengupta (1970), Turnovsky (1971),
1975), and Chow (1970, 1972, and 1975, Chapter 9).
linear constraint. This problem formulation is attractive because it allows one to express a direct relation between the control vector and the target vector at each time period, thereby leading to a simple analytic solution of the optimal control problem. It is also interesting because it is a direct extension of the Theil model to dynamic situations.

7.3.1. Specification of the Optimal Control Model

Policy problems of dynamic systems may be solved by the theory of optimal control. The basic ingredients of a discrete optimal control model are:

1) A set of difference equations that represent the system to be controlled. The system is described by a demometric model in state-space notation

\[
\{\tilde{x}(t + 1)\} = \tilde{F}(\{\tilde{x}(t)\}, \{u(t)\}, t) , \quad t = 0, \ldots, T-1
\]

(7.48)

In control theory, \{\tilde{x}(t)\} is called the state vector and describes the state of the system. The vector \{u(t)\} is the control vector, and \{\tilde{F}(\cdot)\} is a vector-valued function of dimension \(T \times 1\). The equation is known as the state equation or transition equation. Throughout this study, we have dealt with a linear time invariant system, i.e.

\[
\{x(t + 1)\} = \tilde{G}\{\tilde{x}(t)\} + \tilde{B}\{\tilde{u}(t)\} .
\]

(6.3a)

2) A set of constraints on the state and control variables,
\[ g(\{\underline{x}(t)\}, \{\underline{u}(t)\}, t) \geq \{0\} \ . \] (7.49)

where \(g(\cdot)\) is a vector-valued function of dimension \(M\). This function defines the admissible set of state and control variables.

3) A set of boundary conditions. The initial state is given

\[ \{\underline{x}(0)\} = \{\underline{x}_0\} \ . \] (7.50)

We may also require that at the terminal time, or planning horizon, the state vector satisfies the vector-valued function \(\underline{m}(\{\underline{x}(T)\})\) = \(\{0\}\). (7.51)

4) A preference functional, welfare functional, cost functional or a performance index which is to be minimized. The functional may be written

\[ J = K(\{\underline{x}(T)\}) + \sum_{i=0}^{T-1} L(\{\underline{x}(t)\}, \{\underline{u}(t)\}, t) \] (7.52)

The functional reduces all the utilities and disutilities of the controlled dynamic system to a single scalar. All the functions of the cost functional and of the constraints are assumed to be known and to be continuously differentiable with respect to \(\{\underline{x}(t)\}\) and \(\{\underline{u}(t)\}\). Note that the control \(\{\underline{u}(t)\}\) affects the objective functional both directly and indirectly through the value imparted to the states \(\{\underline{x}(t + \lambda)\}\), \(\lambda > 0\).
The optimal control problem is formulated now as the determination of the control sequence \( \{u^*(t)\} \) for \( t = 0, \ldots, T-1 \), and the corresponding trajectory of the state vector \( \{x^*(t)\} \) for \( t = 0, \ldots, T \), such that the constraints (7.48) and (7.49), and the boundary conditions (7.50) and (7.51) are satisfied, and such that the cost functional (7.52) is minimized. The sequence \( \{u^*(t)\} \) is then called the optimal control, and \( \{x^*(t)\} \) the optimal trajectory. In other words, the optimal control problem is to steer a dynamic system, so as to optimize a performance index, subject to constraints. This formulation is very general and explains why the theory pertaining to its solution has found such a wide range of applications, and why it is also has relevance for population policy problems\(^{12}\).

7.3.2. The Discrete Minimum Principle

We now turn to the necessary conditions for optimality. Originally, these conditions were derived by Pontryagin and his associates (1962) for continuous-time systems, described by differential equations. For a thorough statement of the Pontryagin minimum principle, the reader is referred to Athans and Falb (1966). To remain consistent with the other parts of this study, we will state the discrete version of the minimum principle. Several derivations of the discrete time minimum principle have

\(^{12}\)For a survey of applications of optimal control in economic policy planning and of possible extensions, see Athans and Kendrick (1974) and the two special issues of the Annals of Economic and Social Measurement (1972, 1974).
appeared in the literature\(^{13}\). We will state the principle without proof, since it may be found in the literature.

The discrete minimum principle: Suppose the sequence \(\{u^*(t)\}, t = 0, \ldots, T-1\) constitutes an optimal control and \(\{x^*(t)\}, t = 0, \ldots, T\) is an optimal trajectory of the system described by (7.48), and constrained by (7.49), (7.50) and (7.51). In order for \(\{u^*(t)\}, t = 0, \ldots, T-1\) to minimize the cost functional (7.52), it is necessary that there exist a sequence of \(N \times 1\) vectors \(\{\lambda^*(t)\}, t = 1, \ldots, T\), and a sequence of \(M \times 1\) vectors \(\{\nu^*(t)\}, t = 1, \ldots, T\), such that the following conditions hold:

1) The scalar function

\[
H\left(\{x^*(t)\}, \{u(t)\}, \{\lambda^*(t+1)\}, \{\nu^*(t+1)\}\right)
= L\left(\{x^*(t)\}, \{u(t)\}, t\right) + \{\lambda^*(t+1)\}'\{\xi(\{x^*(t)\}, \{u(t)\}, t)\}
- \{\nu^*(t+1)\}'\{g(\{x^*(t)\}, \{u(t)\}, t)\}
\]

is minimized as a function of \(\{u(t)\}\) at \(\{\nu(t)\} = \{u^*(t)\}\) for all \(t = 0, \ldots, T-1\). This implies that

\[
\left.\frac{\delta H}{\delta \{\nu(t)\}\right|_{t} = \{0\}.
\]

The vector \(\{\lambda(t)\}\) is the co-state vector, and \(\{\nu(t)\}\) is the co-constraint vector. With each difference equation (7.48) is associated a co-state vector, and with each

\(^{13}\)See, for example, Wallin (1964), Holtzman (1966) and Pindyck (1973a, 1973b).
constraint (7.49) a co-constraint vector\(^\text{14}\). The function \(H(\cdot)\) is called the Hamiltonian.

2) The dynamics of \(\{x^*(t)\}, \{\lambda^*(t)\}\) and \(\{y^*(t)\}\) are governed by the equations:

\[
\{x^*(t + 1)\} = \frac{\delta H}{\delta \{x^*(t + 1)\}} = \{f(\{x^*(t)\}, \{y^*(t)\}, t)\} \tag{7.55}
\]

\[
\{x^*(0)\} = \{x_0\} \tag{7.56}
\]

\[
\{\lambda^*(t)\} = \frac{\delta H}{\delta \{x^*(t)\}} \mid_{*} \tag{7.57}
\]

\[
\{\lambda^*(T)\} = \frac{\delta K(\{x(T)\})}{\delta \{x(T)\}} \mid_{*} \tag{7.58}
\]

\[
\{g(\{x^*(t)\}, \{y^*(t)\}, t)\} = -\frac{\delta H}{\delta \{y^*(t + 1)\}} \mid_{*} \geq 0 \tag{7.59}
\]

\[
\{y^*(t)\} \geq 0 \tag{7.60}
\]

\[
\{g(\{x^*(t)\}, \{y^*(t)\}, t)\}'\{y^*(t + 1)\} = 0 . \tag{7.61}
\]

Condition (7.55) repeats the difference equation (7.48), and (7.59) is the constraint (7.49). The necessary conditions are essentially equivalent to the Kuhn-Tucker conditions of nonlinear programming. Equations (7.55) and (7.57) are referred to as the canonical difference equations (Athans, 1971; p. 458). Conditions (7.56) and (7.58) are the boundary conditions.

\(^{14}\)Co-state and co-constraint variables in optimal control are similar to Lagrange multipliers in function optimization. They may be interpreted as shadow prices associated with the constraints (Pindyck, 1973; pp. 35-38).
Note that the minimum principle yields only necessary conditions for optimality, which are valid locally. Global optimality also requires sufficiency conditions. These involve the convexity of the functions.

If (7.54) is solved for \( \{y(t)\} \) in terms of \( \{x(t)\}, \{\lambda(t)\} \) and \( \{\mu(t)\} \), and if the resulting expression for \( \{y(t)\} \) is then substituted into equation (7.48) and (7.57), a two-point boundary value problem results. A number of numerical methods are available for solving these problems. Methods such as steepest descent, conjugate directions, conjugate gradient, quasi-linearization, and the Newton-Raphson method are the best known. A description of these algorithms falls beyond the scope of this study. The interested reader should consult Bryson and Ho (1969, Chapter 7), Sage (1968), McReynolds (1970) or Noton (1972). Noton illustrates his exposition with simple numerical examples. Special algorithms, which fit some specific population policy models, have been developed by Evtushenko and MacKinnon (1975) and by Mehra (1975).

7.3.3. The Linear-Quadratic Control Problem

The linear-quadratic (LQ) control problem is one of many possible optimal control problems. It deserves special attention because it is the only optimal control problem for which the solution may be expressed analytically, and because it generalized Theil's idea of quadratic objective function with linear constraints. The LQ control problem fits two types of policy problems. In the first,
the policy maker desires to transform an initial state, say the actual population distribution, to a desired state at the planning horizon, while exhibiting an acceptable behavior of the control and state variables on the way. In the second situation, he tries to keep a system within an acceptable deviation from a reference condition using acceptable amounts of control. In both situations, the optimal control is described by feedback equations known as terminal controllers and as regulators, respectively (Bryson and Ho, 1969, Chapter 5).

The basic ingredients of the LQ problem are\textsuperscript{15}:

1) A linear state equation,

\[
\{x(t + 1)\} = \sum x(t) + B\{u(t)\} .
\] (6.3a)

2) The boundary condition,

\[
\{x(0)\} = \{x_0\} .
\] (7.50)

The planning horizon T is fixed.

3) A quadratic performance index,

\[
J = \frac{1}{2} \{x(T)\}' P\{x(T)\} + \frac{1}{2} \sum_{i=0}^{T-1} \{\sum x(t)\}' Q\{\sum x(t)\} + \{u(t)\}' R\{u(t)\} .\] (7.62)

\textsuperscript{15} The LQ problem has received much attention in the literature. See, for example, Bryson and Ho (1969, Chapter 5), Pindyck (1973; pp. 27-35), Noton (1972; pp. 158-165) and Bar-Ness (1975; pp. 49-56).
The rationale for the quadratic performance index is identical to the one on which the Theil model is based. To assure the convexity of the objective functional, the matrices \( \mathcal{F} \) and \( Q \) are assumed to be positive semi-definite, while \( R \) is positive definite. They may be functions of time. However, the \( t \)-index is deleted for convenience. The matrices \( \mathcal{F} \), \( Q \) and \( R \) give the weights attached to the state and the control variables. They will normally be diagonal. The matrix \( G \) is \( N \times N \) and \( B \) is \( N \times K \) with \( N \) the number of targets and \( K \) the number of instruments at each period of time.

The optimal control problem is to minimize (7.62) subject to (6.3a) and (7.50). How the LQ model relates to the Theil model and to other policy models has been discussed in Chapter 5. The optimal controls \( \{u^*(t)\} \), \( t = 0, \ldots, T-1 \) are found by applying the discrete minimum principle. Not all of the necessary conditions listed in the previous paragraph must be met, since there are no inequality constraints. The Hamiltonian is

\[
H = \frac{1}{2} \{x(T)\}' \mathcal{F} \{x(T)\} + \frac{1}{2} \sum_{t=0}^{T-1} \left\{ \{x(t)\}'Q\{x(t)\} + \{u(t)\}'R\{u(t)\} \right. \\
+ \left. \{\lambda(t+1)\}'[G\{x(t)\} + B\{u(t)\}] \right\} 
\]

(7.63)

where \( \{\lambda(t+1)\} \) is the co-state vector evaluated at period \( t+1 \). From (7.57), we see that \( \{\lambda(t+1)\} \) is the solution of:

\[
\{\lambda(t)\} = \frac{\delta H}{\delta \{x(t)\}} = Q\{x(t)\} + G'\{\lambda(t+1)\} 
\]

(7.64)
or
\[
\{\dot{\lambda}(t + 1)\} - \{\dot{\lambda}(t)\} = -Q\{\dot{x}(t)\} - (G - I)'\{\dot{\lambda}(t + 1)\}
\]

with the final value fixed by (7.58):
\[
\{\dot{\lambda}(T)\} = \frac{\delta}{\delta\{\dot{x}(T)\}} \left[\{\dot{x}(T)\}'F\{\dot{x}(T)\}\right] = F\{\dot{x}(T)\} . \tag{7.65}
\]

Along the optimal trajectory, \(J\) and \(H\) are minimized with respect to \(u(t)\). The necessary conditions yielding the extremum are:

a) \[\frac{\delta H}{\delta\{u(t)\}} = \{0\} = R\{u(t)\} + B'\{\dot{\lambda}(t + 1)\} . \tag{7.66}\]

b) The constraint (7.3a). This condition is formulated as
\[
\{\dot{x}(t + 1)\} = \frac{\delta H}{\delta\{\dot{\lambda}(t + 1)\}} = \{0\} = G\{x(t)\} + B\{u(t)\}
\]

with the initial condition
\[
\{x(t_0)\} = \{\dot{x}_0\} . \tag{7.68}
\]

Since \(R\) is positive definite, we derive from (7.66) the optimal trajectory of the control vector
\[
\{\dot{u}^*(t)\} = -\dot{\lambda}^{-1}B'\{\dot{\lambda}(t + 1)\} . \tag{7.69}
\]

In order for \(\{\dot{u}^*(t)\}\) to minimize \(H\), \(\dot{\lambda}\) must have an inverse
and
\[
\frac{\delta^2 H}{\delta \{u(t)\}^2} = R
\]
must be positive definite. After substituting (7.69) into (7.67), we have a system of 2N first-order difference equations to solve, together with 2N boundary conditions

\[
\{\lambda(t)\} = Q\{x(t)\} + G'\{\lambda(t + 1)\}
\]
(7.64)

\[
\{x(t + 1)\} = G\{x(t)\} + BR^{-1} B'\{\lambda(t + 1)\}
\]
(7.70)

\[
\{x(0)\} = \{x_0\}
\]
(7.68)

\[
\{\lambda(T)\} = F\{x(T)\}
\]
(7.65)

The solution to this two-point boundary-value problem is derived in the Appendix to this part. It starts out from the assumption that there exists a linear relation between \{\lambda(t)\} and \{x(t)\} at the optimum:

\[
\{\lambda^*(t)\} = K(t) \{x^*(t)\}
\]
(7.71)

The feedback matrix \(K(t)\) is the solution of the Riccati equation. Once \(K(t)\) is known for all \(t\), the trajectory of the state vector is given by

\[
\{x^*(t + 1)\} = \left[ I_N - \tilde{B} [R + \tilde{B} K(t + 1) \tilde{B}']^{-1} \tilde{B}' K(t + 1) \right] \{x^*(t)\}
\]
(7.72)
and the optimal control, or control law, is

$$\{u^*(t)\} = -R^{-1}B'K(t + 1)\{x^*(t + 1)\}$$ (7.73)

which gives the control vector in linear state feedback form. The trajectory of the co-state variables is

$$\{\lambda^*(t)\} = K(t)\{x^*(t)\}.$$ (7.74)

The optimal value of the cost functional

$$J^* = \frac{1}{2}\{x(0)\}^T\tilde{K}(0)\{x(0)\}$$ (7.75)

depends only on the initial condition \(\{x(0)\}\) and on \(\tilde{K}(0)\). The matrix \(\tilde{K}(0)\), however, depends on \(G, F, Q, R\) and \(B\) and on the feedback matrices \(K(t), t = 1,\ldots,T\).

Some useful extensions of the LQ model have been made. We present them here as illustrations of how the LQ model may fit policy problems. The first is the dual tracking problem where the policy-maker is looking for a regulator to keep the target and control variables as close as possible to predefined, most acceptable values. By way of a second illustration, we take up the horizon constrained optimal control problem again. Finally, it is shown how the LQ model may handle additional constraints. The idea is to assign penalties for the constraints which are not met.

Illustration a: The dual tracking problem.

In most policy applications of the linear quadratic problem, the objective is to minimize the deviations from desired values of the target vector and eventually also of
the control vector. Rather than having the objective to minimize a function with the arguments expressed as deviations from zero, we have

\[
\text{min } J = \frac{1}{2} \{\hat{x}(T)\}' \tilde{P}\{\hat{x}(T)\} + \\
+ \frac{1}{2} \sum_{t=0}^{T-1} \left[\{\hat{x}(t)\}' Q\{\hat{x}(t)\} + \{\hat{u}(t)\}' \tilde{R}\{\hat{u}(t)\}\right]
\]

(7.76)

where

\[
\{\hat{x}(t)\} = \{x(t)\} - \{\bar{x}(t)\}
\]

\[
\{\hat{u}(t)\} = \{u(t)\} - \{\bar{u}(t)\}
\]

with \{\hat{x}(t)\} and \{\hat{u}(t)\} the desired or most acceptable values for the trajectory of the target vector and the control vector, respectively.

The optimum may be found in the same manner as in the original problem.

Illustration b: Zero terminal error problem.

The dual tracking problem may be supplemented by the additional requirement that at the planning horizon some, say \(\bar{N}\), of the desired levels of the state or target variables must be met exactly, rather than approximately. This means that the following constraint must hold

\[
\{\hat{x}_1(T)\} = \{0\}
\]

(7.77)

where \{\hat{x}_1(T)\} is a \(\bar{N} \times 1\) vector with \(\bar{N} \leq N\).
The control problem is now

\[
\min J = \frac{1}{2} \{\hat{x}(T)\}' P\{\hat{x}(T)\} + \\
+ \sum_{t=0}^{T-1} \left[ \{\hat{x}(t)\}' Q\{\hat{x}(t)\} + \{\hat{u}(t)\}' R\{\hat{u}(t)\} \right]
\]

subject to

\[
\{\tilde{x}(t+1)\} = G\{\tilde{x}(t)\} + B\{\tilde{u}(t)\}
\]

(6.3a)

\[
\{\tilde{x}_1(T)\} = \{0\}
\]

(7.77)

and with \{\tilde{x}(0)\} = \{x_0\} being given.

This is the exact formulation of the horizon constrained optimal control problem of the previous paragraph. Therefore, policy problems where the target vector is given for the planning horizon, and where the restrictions on the state and control trajectory are not so stringent as those discussed previously, may be formulated as dual tracking problems with zero terminal error.

To form the Hamiltonian, we adjoin equation (6.3a) to \(J\) with a multiplier sequence \{\tilde{\lambda}(t)\}, \(t = 1, \ldots, T\), and, in addition, we adjoin (7.77) with a set of \(\tilde{N}\) multipliers \((\nu_1, \nu_2, \ldots, \nu_{\tilde{N}}) = \{\nu\}'\). Thus

\[
H = \frac{1}{2} \{x(T)\}' F\{x(T)\} + \sum_{t=0}^{T-1} \left[ \{x(t)\}' Q\{x(t)\} + \{u(t)\}' R\{u(t)\} \right] \\
+ \{\tilde{\lambda}(t+1)\}' [G\{\tilde{x}(t)\} + B\{\tilde{u}(t)\}] + \{\nu\}' \{x_1(T)\}
\]

(7.78)
Application of the minimum principle yields a two-point boundary-value problem. Solution algorithms have been discussed by Bryson and Ho (1969; pp. 158-164) and by Mehra (1975; pp. 12-16).

**Illustration c: Sign restriction.**

In policy making it is often desirable to restrict a target or a control variable in sign. For example, let \( \{u(t)\} \) be the net migrants of each region, and suppose that the policy maker, in addition to his quadratic objective function, would like to make sure that some regions have no net out-migration or only an "allowable" net out-migration for some or all the periods between 1 and \( T \). It implies that the value of the control variable for these regions and time periods must be positive. He also might want to impose the restriction that the total population of some regions may not fall below a predetermined level. Such constraints may be handled by the formulation of penalty functions. The procedure has been described by Mueller and Wang (1975; p. 610). Although their exposition relates to the continuous model, the application to the discrete version is straightforward. To each state and control variable is attached a number, which plays the role of a penalty or cost if the sign restriction is violated. The extended objective functional becomes:

\[
\min J = \frac{1}{2} \{\tilde{x}(T)\}' F\{\tilde{x}(T)\} + \frac{1}{2} \sum_{t=0}^{T-1} \left[ \{x(t)\}' Q\{x(t)\} + \{u(t)\}' R\{u(t)\} + 2\{x(t)\}' \{g(t)\} \right. \\
\left. + 2\{\tilde{u}(t)\}' \{\tilde{r}(t)\} \right]
\]  

(7.79)
where the elements of \( \{g(t)\} \) and \( \{r(t)\} \) are penalties. If an element \( g_i(t) \) is positive, then \( x_i(t) \) will be penalized when it is positive. A similar idea holds for \( \{r(t)\} \). The magnitudes of the elements of penalty-vectors reflect the weight that the policy-maker puts on the nonnegativity restrictions of the elements of \( \{x(t)\} \) and \( \{u(t)\} \). The objective (7.79) may also be formulated in terms of \( \{\dot{x}(t)\} \) and \( \{\dot{u}(t)\} \). The optimal control is found by applying the necessary conditions to the Hamiltonian. No special difficulties are introduced by the sign restrictions.

The use of penalty functions may be extended to include other equality and inequality constraints as well. The reader may refer to Evtushenko and MacKinnon (1975).
CHAPTER 8
CONCLUSION

The purpose of this study has been the discussion of some of the analytical problems of population distribution policy. It extends the work of Rogers (1975) on spatial population dynamics to the policy domain.

The growth of a multiregional population may be represented by a system of linear, first-order, homogenous difference equations with constant coefficients:

\[ \tilde{x}(t+1) = G\tilde{x}(t) \]  

(3.1)

with \( \tilde{x}(t) \) the state vector representing the distribution of the population over space and/or age, and \( G \) the growth matrix. To transform (3.1) into a policy model, we add a control vector \( \tilde{u}(t) \):

\[ \tilde{x}(t+1) = G\tilde{x}(t) + B\tilde{u}(t) \]  

(6.3a)

The vector \( \tilde{u}(t) \) contains the instruments of population distribution policy. It has been argued that a fundamental feature of population distribution policy is that it does not occur in a vacuum. It is subordinate to social and economic policies. The ultimate goals are non-demographic in nature, and the instruments are socio-economic. This has been illustrated in Part I of this study. The policy models must, therefore, reflect this connection. The elements of \( \tilde{u}(t) \) are socio-economic variables representing the instruments. The relation between \( \tilde{u}(t) \) and the population
distribution \( \{x(t)\} \) is assumed to be linear and constant in time. The matrix multiplier \( B \) plays a pivotal role in our discussion of policy models. The relation between the population distribution \( \{x(t)\} \) and the vector of socio-economic policy objectives \( \{y(t)\} \) is assumed to be linear too:

\[
\{y(t)\} = C \{x(t)\} \quad .
\]

(6.3b)

Equations (6.3a) and (6.3b) constitute the policy model we have devoted our attention to. It takes the form of a state-space model. Without loss of generality, we have assumed in several instances that \( C = I \), which means that the objectives of the population distribution policy are expressed in terms of the multiregional distribution of people. The policy model becomes then (6.3a).

The state-space model is a powerful tool for policy analysis, once the behavior of the system is known and the policy objectives and the range of instruments are identified. In most of the literature on quantitative policy, it has been assumed that these conditions are satisfied. We made similar assumptions in Part III of this study. Before doing that, the validity of these assumptions was questioned in Part I. The usefulness of the state-space model for the analytical treatment of policy problems is maximal if it is time-invariant. Time invariance of the coefficients of the policy models has been assumed in Part III. This assumption has been critically evaluated in Part II. Do we really know the behavioral dynamics of population distribution? What are the goals and the means of a
population distribution policy? Who defines them, and how do they relate to social well-being?

8.1. THEORETICAL ISSUES OF POPULATION DISTRIBUTION POLICY

The major factor in shaping spatial population growth is migration. In the past decades, several scientists have attempted to identify the determinants of migration as well as its consequences. This research has resulted in a large amount of regression models, the most advanced of which are simultaneous equation models. These models are referred to as demometric models. They form the starting point of the policy models in Part III. How one goes from demometric models to the state-space model has been demonstrated in Chapter 6. Since the policy models are based on demometric models, they cannot be better than the latter. From a practical point of view, highly sophisticated policy models are not justified as long as the demometric models have no sound basis. The theoretical basis of demometric models should be founded on migration theory.

The theoretical underpinning of the goals-means relationship of population distribution policy is the theory of externalities and the theory of government intervention. The ultimate goal of any action is to maximize the quality of life or, in terms of welfare economics, to achieve Pareto optimality. In a society, where the assumptions of a perfect market economy are met, the private behaviors of the migrants assure that Pareto optimality is reached, and hence there is no need for policy intervention. However, a society does not behave this way. Private decisions have unwanted effects
on other individuals. These externalities cause the private optimum to diverge from the social optimum. The role of government is to internalize the externalities, assuring the achievement of the social optimum. The goals and means of population distribution policy should reflect this optimum. Throughout Part III, we assumed the existence of a policy maker, who defines the social optimal policy objectives and who implements the instruments.

We have looked upon a policy model as derived from a demometric model by adding a new dimension to it, namely the goals-means relationship of population distribution policy. Ideally, these policy models should be based on a theory of population distribution policy. Such a policy does not exist. The purpose of Part I was to throw some light on the three pillars of such a theory: migration theory, theory of externalities, and theory of government intervention.

8.2. DYNAMICS OF STRUCTURAL CHANGE IN DEMOGRAPHIC ANALYSIS

No matter what the goals and the means are, a population distribution policy will affect the basic characteristics of the demographic system. From mathematical demography, we know that demographic change may be traced back to a change in age-specific fertility, mortality and migration rates. A population policy, in order to be effective, influences the vital rates. But how do the changing rates affect the demographic system? This question was the subject of Part II. We derived a set of sensitivity functions relating a change in demographic statistics to a change in
the vital rates. The primary purpose was to contribute to the knowledge of spatial population dynamics by presenting a unifying technique of impact assessments. In the single-region mathematical demography, ordinary differential calculus is used to perform sensitivity analysis. In multi-regional demography, where we deal with matrix and vector functions, the application of ordinary calculus is very complicated. Instead, matrix differentiation techniques prove to be very useful. A review of these techniques has been given in the Appendix to Part II. These mathematical tools have been applied to derive analytical expressions for multiregional demographic features, such as life table statistics, population projection, and stable population characteristics, representing the impacts of changes in vital rates. The sensitivity functions reveal how each spatial demographic characteristic depends on the age-specific rates and how it reacts to changes in those rates. Matrix differentiation techniques form a powerful tool for the analysis of structural change in multiregional systems.

A secondary objective of Part II was to contribute to the reconciliation of the discrete and continuous models of demographic growth. Traditionally, there has been a sharp distinction between the discrete model and the continuous model of population growth. It is our belief that the reason is mainly historical. We have attempted to show that the results derived for the continuous model, may easily be extended to the discrete model. Therefore, the discrete and continuous models of demographic growth are equivalent tools for the analysis of population dynamics.
8.3 OPTIMAL MIGRATION POLICIES

After discussing some theoretical issues of the goals-means relationship of population distribution policy, and after exploring the dynamics of the system we wish to control, we shifted our focus to policy models, in which the dominant factor is migration. Although regional differences in fertility and mortality affect the population distribution, their instrumental value to the policy maker is limited. Therefore, we have limited ourselves in Part III to migration policy models.

The models are based on the assumptions that the behavior of the system to be controlled has been described by a system of linear equations, denoted as demometric models, and that the objectives and instruments of population distribution policy have been formulated in precise terms by the policy maker. The policy dimension is introduced into the demometric model, following the Tinbergen paradigm: the policy-relevant part of the system is isolated. It has been shown that any linear descriptive or explanatory model may be converted to a policy model if and only if all the target variables of the policy model belong to the set of endogenous variables of the descriptive or explanatory model, and if at least one of the exogenous variables is controllable.

The general formulation of a policy model is (Tinbergen, 1963):

\[ \{y\} = \mathbf{R}\{z_1\} + \mathbf{S}\{z_2\} \]  \hspace{1cm} (5.3)
with \( \{y\} \) the vector of target variables, \( \{z_1\} \) the vector of instrument variables and \( \{z_2\} \) the vector of uncontrollable exogenous and lagged endogenous variables. A crucial role in policy analysis is played by the matrix multiplier \( R \). Our discussion of policy models centers around this multiplier. This is consistent with the economic literature on policy models. However, we go beyond the traditional approach in economics and draw from recent findings of mathematical system theory and the theory of optimal control. To present an overview of policy models, a classification scheme has been set up that is based on the rank and the structure of \( R \). This scheme enables us to relate seemingly unrelated models to each other. For example, it has been shown that the linear-quadratic control problem may be derived from the Tinbergen and Theil model by assuming inter-temporal separability of the objectives and unidirectional causality of the population system. The state-space model of (6.3) also may be derived from the Tinbergen model, and from the reduced form model in general.

The fundamental questions of quantitative migration policy may be expressed in terms of existence and design. In Chapters 6 and 7, we have dealt with these two topics. The discussion revolves around the matrix multiplier. Whether arbitrarily specified levels of target variables can be reached by the existing set of instruments, depends on the rank of \( R \). The conditions that must be satisfied for a population system to be controllable are formulated in a number of existence theorems. These theorems enable us to uncouple the controllable parts of a not-completely controllable system, and to compute the minimal number
of instruments that assure the achievement of the targets. It has been shown, for example, that under well-defined circumstances represented by a specific transformation of the matrix multiplier, all the desired target-values can be reached with a single instrument. This result is intriguing and totally contrary to the thinking engendered by Tinbergen's Theory of Policy.

The design procedure of optimal policies is dictated by the structure and the rank of the matrix multiplier $\mathcal{R}$. If $\mathcal{R}$ is nonsingular, then the unique solution to (5.3) for $\{z_1\}$ is found by simply inverting $\mathcal{R}$. When $\mathcal{R}$ is singular, there may be no instrument vector leading to the desired target values, or there may be an infinite number of them. To find a unique optimal solution, an objective function reflecting the policy maker's preferences is introduced, and mathematical programming techniques may be applied. There is a wide variety of algorithms available in the literature. The common characteristic of most of them is that they determine the optimal solution numerically. In this study, we have directed our attention to cases where solutions to policy problems can be found analytically.

In this regard, there is the applicability of the notion of generalized inverse. We have shown how the minimizing properties of generalized inverses may be relevant in solutions of policy models with a singular multiplier matrix. For example, no matter what the rank of the $N \times K$ matrix $\mathcal{R}$ is, a unique solution to (5.3) is given by

$$\{z_1\} = \mathcal{R}^D[\{\tilde{y}\} - S\{z_2\}]$$
where $\tilde{R}^P$ is the Moore-Penrose inverse (Ben-Israel and Greville, 1974; p. 7). If $\tilde{R}$ is nonsingular, then $\tilde{R}^P$ is the ordinary inverse; if $\tilde{R}$ is singular and of rank $N$, i.e., the number of instruments exceeds the number of targets, then $\tilde{R}^P$ defines a minimum norm solution to (5.3); and if $\tilde{R}$ is singular and of rank $K$, i.e., the targets exceed the instruments in number, then $\tilde{R}^P$ defines a solution to (5.3) that minimizes the squared deviations between the desired and the realized values of the target variables. No explicit objective function has been specified, but it is implicit in the minimizing properties of the generalized inverses. The interesting feature of generalized inverses is that they provide an analytical solution to policy models.

Another case for which the optimal policy may be found analytically is the initial period control problem, namely: the case where the target vector is given for the planning horizon and the control vector at each time period is a linear combination of the control vector of the previous time period. It then can be shown that the initial period control problem reduces to a single-period problem and the control only needs to be specified at the initial period.

A final policy problem for which a solution may be expressed analytically, is the linear-quadratic control problem. In this trajectory-optimization problem, the policy maker wants to minimize a quadratic function of target variables and instrument variables, subject to linear constraints imposed by the behavior of the system and by the initial condition.
8.4. RECOMMENDATIONS FOR FUTURE RESEARCH

We have based our treatment of population distribution policy models on three fundamental assumptions:

i) The dynamic behavior of the population system and its interaction with socio-economic conditions can be modeled adequately.

ii) This model takes the form of a system of simultaneous linear equations with constant coefficients.

iii) There exists a policy maker who expresses the goals-means relationship of population distribution policy in specific terms of a target vector or in terms of a social welfare function, who sets up a range of instruments, and who is willing and able to implement the policy.

We have already discussed some issues related to the validity of these assumptions. A lot more research is needed in this regard. The prerequisite for good population distribution models is a well developed migration theory. There is no consensus yet on the determinants of migration and on the way the population system interacts with the socio-economic system. As long as the dynamics of the population system are not fully understood, government intervention cannot have a sound basis.

Apart from the problem of identifying the determinants of population growth and distribution, there is the problem of modeling the population system once the determinants are known. Specification and estimation of population models is the subject of demometrics. This new science
ultimately should provide the necessary input information for policy analysis.

The third assumption on which our discussion of policy models has been based concerns the goals–means relationship of population policy. Not much research has been done to provide a theoretical underpinning for this relationship. The approach has instead been pragmatic. The emerging theories of externalities and of government intervention may be important building blocks for a theory of population distribution policy. We are convinced that this theory is a limiting factor for a sound analysis of population distribution policy.
APPENDIX

THE LINEAR-QUADRATIC CONTROL MODEL:
SOLUTION OF THE TWO-POINT
BOUNDARY-VALUE PROBLEM

The application of the discrete minimum principle to the LQ control problem yields a system of first-order difference equations, together with a system of equations representing the boundary conditions at the initial and at the terminal time periods, respectively. The optimal control of the LQ model is given by the solution of this two-point boundary-value problem. The system has been derived in Chapter 7, and is given by (7.64), (7.70), (7.68) and (7.65):

\[
\begin{align*}
\dot{\lambda}(t) &= Q(\dot{x}(t)) + G'\{\lambda(t + 1)\} \quad (A7.1) \\
\dot{x}(t + 1) &= G(\dot{x}(t)) + BR^{-1}B'\{\lambda(t + 1)\} \quad (A7.2) \\
\dot{x}(0) &= \{x_0\} \quad (A7.3) \\
\dot{\lambda}(T) &= F(\dot{x}(T)) \quad (A7.4)
\end{align*}
\]

where \(\{x(t)\}\) and \(\{\lambda(t)\}\) are the state vector and the co-state vector, respectively.

The solution to the two-point boundary-value problem starts out with the assumption that there exists a linear relation between the co-state vector and the state-vector at the optimum:
\{\lambda^*(t)\} = K(t) \{x^*(t)\} \quad (A7.5)

where the feedback matrix \(K(t)\) is the solution of the discrete Riccati equation. Since by (7.69)

\{u^*(t)\} = -R^{-1}B'\{\lambda(t + 1)\} \quad (A7.6)

we may write the feedback control law as

\{u^*(t)\} = -R^{-1}B'K(t + 1) \{x^*(t + 1)\} \quad (A7.7)

The closed-loop system then is

\{x^*(t + 1)\} = G\{x^*(t)\} - BR^{-1}B'K(t + 1) \{x^*(t + 1)\}

\[ I + BR^{-1}B'K(t + 1) \{x^*(t + 1)\} = G\{x^*(t)\} \quad (A7.8) \]

The matrix

\[ [I + BR^{-1}B'K(t + 1)] \]

is nonsingular, as will be shown later. Therefore

\{x^*(t + 1)\} = [I + BR^{-1}B'K(t + 1)]^{-1} G\{x^*(t)\} \quad (A7.9)

The solution to this system of homogenous difference equations is

\{x^*(t)\} = \phi(t,0) \{x(0)\} \quad (A7.10)
where \( \phi(t,0) \) is the discrete state transition matrix. The matrix

\[
\Phi(t) = \left[ I + R^{-1}B'K(t + 1) \right]
\]

is function of time. Therefore, the solution of time invariant systems

\[
\{ x^*(t) \} = \left[ \Phi^{-1}G \right]^t \{ x(0) \}
\]

is incorrect.

We must find a sequence \( K(t) \) such that (A7.9) and (A7.10) hold for any value of \( \{ x(0) \} \). Substituting (A7.5) into (A7.1) gives

\[
K(t) \{ x^*(t) \} = \Phi \{ x^*(t) \} + G'K(t + 1) \{ x^*(t + 1) \}
\]

where \( \{ x^*(t + 1) \} \) is given by (A7.9). We find then

\[
K(t) \{ x^*(t) \} = \Phi \{ x^*(t) \} + G'K(t + 1) \left[ \Phi^{-1}(t) G \right] \{ x^*(t) \}
\]

where \( \Phi(t) \) is given by (A7.11), whence

\[
K(t) \{ x^*(t) \} = \left[ Q + G'K(t + 1) \Phi^{-1}(t) G \right] \{ x^*(t) \} .
\]

(A7.12)

This equation is a result of the necessary conditions for an optimum. Therefore, it must hold for any initial condition \( \{ x_0 \} \). Since only \( \{ x^*(t) \} \) depends on the initial condition, (A7.12) must hold for any \( \{ x^*(t) \} \), and we must have
\[ \tilde{X}(t) = Q + G' \tilde{X}(t + 1) \tilde{Y}^{-1}(t) \tilde{G}, \quad \text{for all } t \]

\[ = Q + G' \tilde{X}(t + 1) \left[ I + BR^{-1}B' \tilde{X}(t + 1) \right]^{-1} \tilde{G}. \]

(A7.13)

Equation (A7.13) is the discrete Riccati equation. The Riccati equation may be solved backwards, starting with the boundary condition

\[ \tilde{X}(T) = F. \]

(A7.14)

That this boundary condition holds follows from substituting (A7.5) into (A7.4)

\[ \tilde{X}(T) \{ \tilde{x}^*(T) \} = F \{ \tilde{x}^*(T) \} \]

or

\[ \tilde{X}(T) = F. \]

If \( \{ \tilde{x}(T) \}' F \{ \tilde{x}(T) \} = 0 \), i.e., \( \tilde{F} = 0 \), then

\[ \{ \lambda(T) \} = \{ 0 \} \text{ and } \tilde{X}(T) = 0. \]

It is easy to show that \( \tilde{X}(t) \) is positive semi-definite for all \( t \). Since \( \tilde{F} \) was assumed to be positive semi-definite, \( \tilde{X}(T) \) is positive semi-definite. The matrix \( \tilde{X}(T - 1) \) may be found by the relation

\[ \tilde{X}(T - 1) = Q + G' \tilde{X}(T) \left[ I + BR^{-1}B' \tilde{X}(T) \right]^{-1} \tilde{G} \]

(A7.15)

\[ = Q + G' \tilde{X}(T) \tilde{G} + G' \tilde{X}(T) BR^{-1}B' \tilde{X}(T) \tilde{G}'. \]
Since \( Q \) is positive semi-definite and \( R^{-1} \) is positive definite, \( \tilde{\kappa}(t - 1) \) must be positive semi-definite, and so all \( \tilde{\kappa}(t) \). Now since all of the \( \tilde{\kappa}(t) \) are positive semi-definite, the matrix

\[
\tilde{\gamma}(t) = \mathbb{I} + \tilde{B}R^{-1}\tilde{B}'K(t + 1)
\]

is nonsingular, and (A7.9) has a unique solution. It is also clear from (A7.15) that \( \tilde{\kappa}(t) \) is symmetric if \( Q \), \( \tilde{R} \) and \( \tilde{B} \) are symmetric. Since \( \tilde{R} \) is nonsingular \( \tilde{\kappa}(t) \) is also unique.

The solution of the Riccati equation requires the inversion of \( \tilde{\gamma}(t) \) at each time period. \( \tilde{\gamma}(t) \) is an \( N \times N \) matrix, \( N \) being the number of elements in the state vector \( \{x(t)\} \), i.e., the number of targets. However, it is possible to reduce the dimension of the matrix to be inverted. It is known from matrix algebra that

\[
[I_N + WZ']^{-1} = I_N - W[I_{K} + Z'W]^{-1}Z' \quad (A7.16)
\]

where \( W \) and \( Z' \) are \( N \times K \) and \( K \times N \) respectively, with \( K \leq N \). Let

\[
W = \tilde{W}
\]

\[
Z' = \tilde{R}^{-1}\tilde{B}'\tilde{K}(t + 1)
\]

then

\[
\tilde{\gamma}^{-1}(t) = I_N - B[I_{K} + \tilde{R}^{-1}\tilde{B}'\tilde{K}(t + 1) B]^{-1}\tilde{R}^{-1}\tilde{B}'\tilde{K}(t + 1)
\]
\begin{equation}
\begin{align*}
= \mathbf{I}_N - \mathbf{B}[\mathbf{I}_K + \mathbf{R}^{-1}\mathbf{B}'\mathbf{K}(t + 1) \mathbf{B}]^{-1} \mathbf{B}'\mathbf{K}(t + 1) \\
= \mathbf{I}_N - \mathbf{B}[\mathbf{R} + \mathbf{B}'\mathbf{K}(t + 1) \mathbf{B}]^{-1} \mathbf{B}'\mathbf{K}(t + 1)
\end{align*}
\end{equation}

(A7.17)

where the matrix to be inverted

$$[\mathbf{R} + \mathbf{B}'\mathbf{K}(t + 1) \mathbf{B}]$$

is \(K \times K\). \(K\) is the number of instruments, and is generally much smaller than \(N\).

Once \(\mathbf{K}(t)\) is known for all \(t\), the sequence of the state vector is computed from (A7.3) and (A7.18):

\begin{equation}
\begin{align*}
\{\mathbf{x}^*(t + 1)\} &= \left[\mathbf{I}_N - \mathbf{B}[\mathbf{R} + \mathbf{B}'\mathbf{K}(t + 1) \mathbf{B}]^{-1} \mathbf{B}'\mathbf{K}(t + 1)\right] \{\mathbf{x}^*(t)\}
\end{align*}
\end{equation}

(A7.18)

and the sequence of the control vector follows from (A7.19)

\begin{equation}
\begin{align*}
\{\mathbf{u}^*(t)\} &= -\mathbf{R}^{-1}\mathbf{B}'\mathbf{K}(t + 1) \{\mathbf{x}^*(t + 1)\}
\end{align*}
\end{equation}

(A7.19)

The co-state variables are computed by (A7.20)

\begin{equation}
\begin{align*}
\{\mathbf{\lambda}^*(t)\} &= \mathbf{K}(t) \{\mathbf{x}^*(t)\}
\end{align*}
\end{equation}

(A7.20)

The optimal cost functional is

\begin{equation}
\begin{align*}
J &= \frac{1}{2} \{\mathbf{x}^*(T)\}' \mathbf{F} \{\mathbf{x}^*(T)\} + \frac{1}{2} \left[ \sum_{i=0}^{T-1} \{\mathbf{x}^*(t)\}' \mathbf{Q} \{\mathbf{x}^*(t)\} \right] \\
&\quad + \{\mathbf{u}^*(t)\}' \mathbf{R} \{\mathbf{u}^*(t)\}
\end{align*}
\end{equation}

(A7.21)
Substituting for \( \{\hat{y}^*(t)\} \) yields

\[
J = \frac{1}{2} \{\hat{x}^*(T)\}' P\{\hat{x}^*(T)\} + \frac{1}{2} \left[ \sum_{i=0}^{T-1} \{\hat{x}^*(t)\}' \mathcal{Q}\{\hat{x}^*(t)\} \right. \\
+ \left. \{\hat{x}^*(t+1)\}' K(t+1)' \mathcal{BR}^{-1} B' \mathcal{K}(t+1) \{\hat{x}^*(t+1)\} \right]
\]

Substituting \( \mathcal{Q}\{\hat{x}^*(t)\} \) using (A7.12) yields

\[
J = \frac{1}{2} \{\hat{x}^*(T)\}' P\{\hat{x}^*(T)\} + \frac{1}{2} \left[ \sum_{i=0}^{T-1} \{\hat{x}^*(t)\}' K(t) \{\hat{x}^*(t)\} \right. \\
- \left. \{\hat{x}^*(t)\}' G' K(t+1) \mathcal{V}^{-1}(t) G\{\hat{x}^*(t)\} \right. \\
+ \left. \{\hat{x}^*(t+1)\}' K(t+1)' \mathcal{BR}^{-1} B' \mathcal{K}(t+1) \{\hat{x}^*(t+1)\} \right]
\]

But by (A7.9)

\[
\mathcal{V}^{-1}(t) G\{\hat{x}^*(t)\} = \{\hat{x}^*(t+1)\}
\]

and, therefore,

\[
J = \frac{1}{2} \{\hat{x}^*(T)\}' P\{\hat{x}^*(T)\} + \frac{1}{2} \left[ \sum_{i=0}^{T-1} \{\hat{x}^*(t)\}' K(t) \{\hat{x}^*(t)\} \right. \\
+ \left. \{\hat{x}^*(t+1)\}' K(t+1)' \mathcal{BR}^{-1} B' - \{\hat{x}^*(t)\}' G' \right] \\
\mathcal{K}(t+1) \{\hat{x}^*(t+1)\}
\]

(A7.22)

By (A7.9)

\[
\{\hat{x}^*(t)\}' G' = \{\hat{x}^*(t+1)\}' [I + \mathcal{BR}^{-1} B' \mathcal{K}(t+1)]'.
\]
And

\[
\{x^*(t + 1)\}' \ K(t + 1)' \ BR^{-1}B' - \{x^*(t)\}' \ G' \]

\[= \{x^*(t + 1)\}' \ K(t + 1)' \ BR^{-1}B' - \{x^*(t + 1)\}
- \{x^*(t + 1)\}' \ K(t + 1)' \ BR^{-1}B' = [-x^*(t + 1)] \ .
\]

The objective function (A7.22) becomes

\[
J = \frac{1}{2} \{x^*(T)\}' \ F\{x^*(T)\} + \frac{1}{2} \left[ \sum_{i=0}^{T-1} \{x^*(t)\}' \ K(t) \ {x^*(t)} \right]
- \{x^*(t + 1)\}' \ K(t + 1) \ {x^*(t + 1)} \]
\[
J = \frac{1}{2} \{x^*(T)\}' \ F\{x^*(T)\} + \frac{1}{2} \{x(0)\}' \ K(0) \ {x(0)}
- \frac{1}{2} \{x^*(T)\}' \ K(T) \ {x^*(T)} \ .
\]

But since \(K(T) = F\), we have

\[
J^* = \frac{1}{2} \{x(0)\}' \ K(0) \ {x(0)} \ . \quad (A7.23)
\]

The optimal value of the objective functional depends on the initial condition \(\{x(0)\}\) and on \(K(0)\). The matrix \(K(0)\) depends on the matrices \(G, F, Q, R, B\) and on the feedback matrices \(K(t), t = 1, \ldots, T\).
References


Vita

Frans Willekens was born in Mol (Belgium) on March 5, 1946.

He has been educated in agricultural engineering, economics and sociology at the University of Leuven (Belgium). He holds a masters degree in agricultural engineering.

After graduating, Frans Willekens worked with a consulting firm on the evaluation of agricultural investment projects in developing countries. From 1971 to 1973, he was assistant professor in agricultural economics at the National University of Zaire.

While a graduate student at Northwestern University, he was a research assistant of Dr. A. Rogers. Since June 1975, he has been his research assistant at the International Institute for Applied Systems Analysis at Laxenburg, Austria.

Publications:
