DATA BASES AND ACCOUNTING FRAMEWORKS FOR IIASA'S COMPARATIVE MIGRATION AND SETTLEMENT STUDY

Philip Rees
Frans Willekens

December 1981
CP-81-39

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

Philip Rees, School of Geography, University of Leeds, Leeds
LS2 9JT, United Kingdom

Frans Willekens, Netherlands Interuniversity Demographic Institute, Box 955, NL-2270, AZ Voorburg, Netherlands

FOREWORD

The evolution of human populations over time and space has been a central concern of many scholars in the Human Settlements and Services Area at IIASA during the past several years. From 1975 through 1978 some of this interest was manifested in the work of the Migration and Settlement Task, which was formally concluded in November 1978. Since then, attention has turned to disseminating the Task's results, to concluding its comparative study, and to exploring possible future activities that might apply the mathematical methodology to other research topics.

This paper is part of the Task's dissemination effort. It is a draft of a chapter that is to appear in a volume entitled *Migration and Settlement: A Comparative Study*. Other selected publications summarizing the work of the Migration and Settlement Task are listed at the back.

Andrei Rogers
Chairman
Human Settlements
and Services Area

ACKNOWLEDGMENTS

The authors wish to thank Andrei Rogers, Jacques Ledent, Dimiter Philipov and Peer Just for their valuable comments on an earlier draft of this paper. The contributions of Maria Rogers (editing) and Susie Riley (typing) are also gratefully acknowledged.

ABSTRACT

This paper reviews the accounting frameworks underpinning the data inputs to the Comparative Migration and Settlement (CMS) task now that the majority of country studies are either published as IIASA Research Reports or are in the publication pipeline. Rigorous comparisons are made of accounting, sex, age, time, and regional definitions used in the CMS set of studies, and the methods used to estimate required model inputs from available data are described. These comparisons and descriptions should serve as required reading for researchers embarking on further analysis of the CMS task outputs or on any future exercise in comparing the spatial patterns of population change and movement in different countries.

CONTENTS

DATA BASES AND ACCOUNTING FRAMEWORKS FOR IIASA'S COMPARATIVE MIGRATION AND SETTLEMENT STUDY

1. INTRODUCTION

1.1. The Comparative Migration and Settlement (CMS) Study

In 1975, a group of research scholars from IIASA's National Member Organization (NMO) countries met in Laxenburg to discuss common issues of internal migration and spatial population distribution. It was suggested that IIASA undertake a quantitative assessment of recent migration patterns and spatial population dynamics in all of its 17 NMO countries (for details see Rogers 1976a and Willekens 1978). Such a study would involve national scholars, use a common methodology, and be carried out on a comparative basis as much as possible. The methodology of multiregional demographic analysis was proposed as the common analytical framework for the study of spatial population dynamics since it would enable one to simultaneously consider migration and differential fertility and mortality among the regions. The proposals were adopted and members of the Migration and Settlement Task began this work at IIASA.

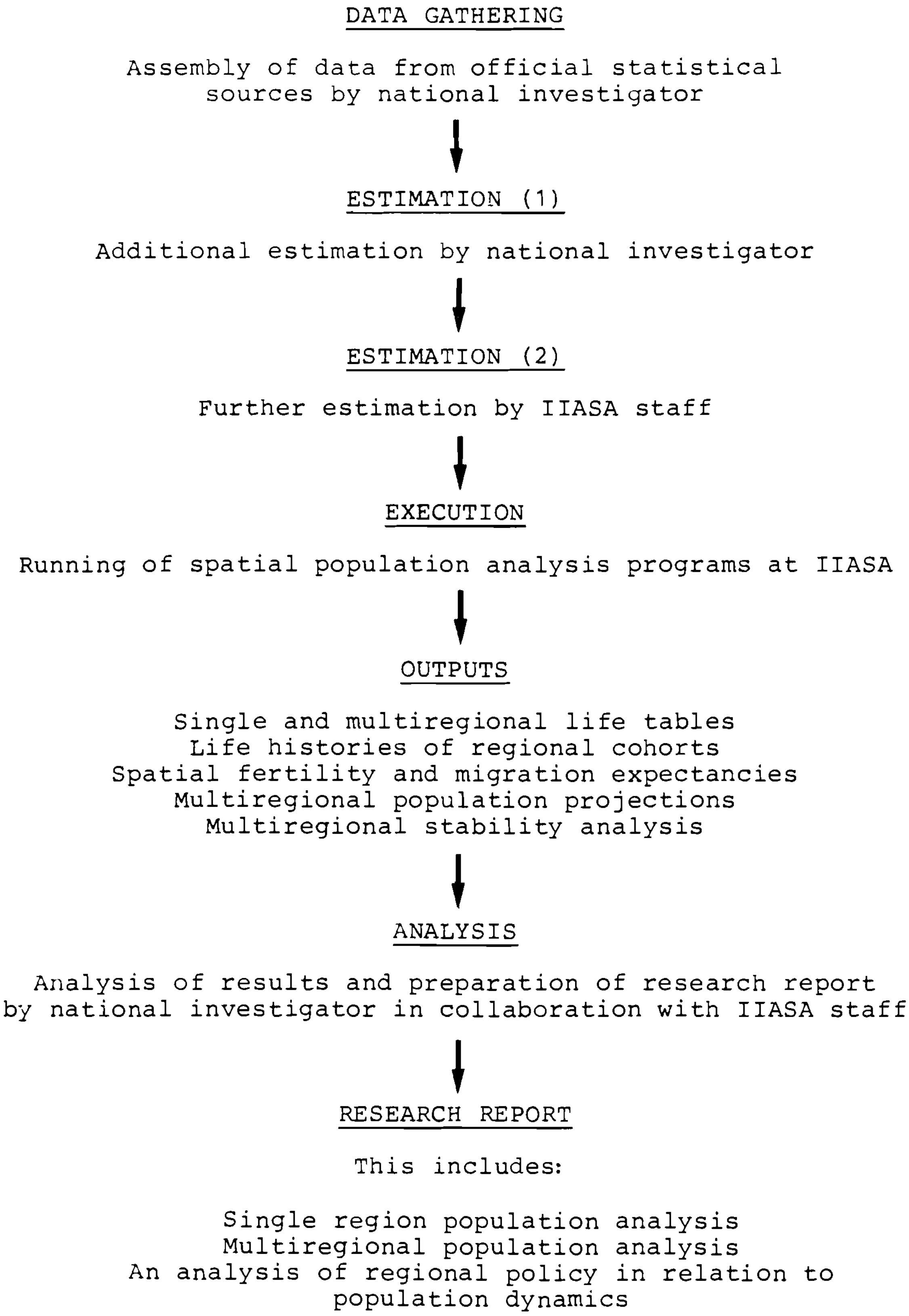
One of the expected outcomes of the CMS study was a set of 17 research reports, one for each participating country, written by a national scholar. Each report was intended to provide an overview of recent patterns of change in migration and in regional fertility and mortality, to illustrate the application of multiregional demographic techniques and the additional insights in population redistribution that can be gained from it, and to give a brief review of population distribution policy issues. A common outline for the report was given to each collaborating author. IIASA assisted extensively in data processing and in the preparation of the reports.

Each study involved a number of steps which are set out in Figure 1. Firstly, the national investigator assembled population, births, deaths, and migration data for the set of regions to be studied from official published or unpublished sources. Regions were defined by the national collaborators and were generally continguous units of territory that divided up the countries concerned. The number of regions in the studies ranged from a minimum of 4 to a maximum of 13. The second and third steps involved the estimation of model input variables from the available data either by the national investigator or by IIASA staff. The data thus estimated were then used as input to the suite of spatial population analysis programs (SPA programs) described in Willekens and Rogers (1978) and the country analysis was run on IIASA's PDP 11 computer. A variety of outputs concerning the population structure and dynamics of the country's multiregional system was thus obtained, and these were then analyzed by the national investigator and incorporated in the country's research report. Of course, in practice, there was a good deal of reiteration of steps in the CMS task, and several cycles were carried out more than once.

The purpose of this paper is to describe in general terms the kinds of data that were used, the kinds of estimation techniques that were employed and the consequences of data problems that were encountered in carrying out the CMS study.

1.2. The IIASA Set of Countries

What kind of countries were included in the CMS task as a result of IIASA membership? Table 1 provides a set of basic indicators for all countries extracted from the research reports and from the World Development Report (World Bank 1980). The set



DATA GATHERING

Assembly of data from official statistical
sources by national investigator

ESTIMATION (1)

Additional estimation by national investigator

ESTIMATION (2)

Further estimation by IIASA staff

EXECUTION

Running of spatial population analysis programs at IIASA

OUTPUTS

Single and multiregional life tables
Life histories of regional cohorts
Spatial fertility and migration expectancies
Multiregional population projections
Multiregional stability analysis

ANALYSIS

Analysis of results and preparation of research report
by national investigator in collaboration with IIASA staff

RESEARCH REPORT

This includes:

Single region population analysis
Multiregional population analysis
An analysis of regional policy in relation to
population dynamics

Figure 1. Steps in the CMS study for each member nation.

Table 1. Basic demographic and economic indicators for IIASA member nations

Country (study period)	Area[a] (1000s of sq. km.)	Population ($x 10^6$) in study year[b]	1978	Ave. annual growth of population (per 1000) 1970-78[a]	Crude birth rate (per 1000) 1978	Crude death rate (per 1000) 1978
1 United Kingdom (1970)	244	55.4[c]	55.8[k]	1	12	12
2 Finland (1974)	337	4.7	4.8	4	14	9
3 Sweden (1974)	450	8.2	8.3	4	12	11
4 GDR (1975)	108	16.8	16.7	-2	13	13
5 Netherlands (1974)	41	13.5	13.9	8	13	8
6 Canada (1966-71)	9,976	20.8[d]	23.5	12	16	8
7 Hungary (1974)	93	10.4	10.7	4	16	12
8 Soviet Union (1974)	22,402	250.9	261.0	9	18	10
9 FRG (1974)	249	62.0	61.3	1	9	12
10 Austria (1966-71)	84	7.5[e]	7.5	2	11	12
11 Poland (1975)	313	34.2[f]	35.0	9	19	9
12 Bulgaria (1975)	111	8.7	8.8	5	16	11
13 France (1968-75)	547	52.7[g]	53.3	6	14	10
14 Czechoslovakia (1975)	128	14.8	15.1	7	18	11
15 Japan (1970)	372	105.5[h]	114.9	12	15	6
16 United States (1965-70)	9,363	203.8[i]	221.9	8	15	9
17 Italy (1971)	301	54.0[j]	56.7	7	13	9

Country (study period)	Life expectancy at birth (years) 1978	Total fertility rate (per woman) 1978	GNP per capita ($) 1978	Percent of population of working age 1978	World Bank classification
1 United Kingdom (1970)	73	1.7	5,030	64	IC
2 Finland (1974)	72	1.7	6,820	68	IC
3 Sweden (1974)	75	1.7	10,210	64	IC
4 GDR (1975)	72	1.8	5,710	63	CPE
5 Netherlands (1974)	74	1.6	8,410	65	IC
6 Canada (1966-71)	74	1.9	9,180	66	IC
7 Hungary (1974)	70	2.2	3,450	66	CPE
8 Soviet Union (1974)	70	2.4	3,700	65	CPE
9 FRG (1974)	72	1.4	9,580	65	IC
10 Austria (1966-71)	72	1.7	7,030	63	IC
11 Poland (1973)	71	2.3	3,670	66	CPE
12 Bulgaria (1975)	72	2.3	3,230	66	CPE
13 France (1968-75)	73	1.9	8,260	63	IC
14 Czechoslovakia (1975)	70	2.4	4,720	64	CPE
15 Japan (1970)	76	1.8	7,280	68	IC
16 United States (1965-70)	73	1.8	9,590	65	IC
17 Italy (1971)	73	1.9	3,850	64	IC

Notes: GDR - German Democratic Republic, FRG - Federal Republic of Germany, GNP - Gross National Product, IC - Industrialized country, CPE - Centrally planned economy. See next page for notes on data sources.

of countries includes three of the most areally extensive in the world (Soviet Union, United States, and Canada) together with moderately and small-sized European countries. In population size the countries range from Finland (just under 5 millions in 1981) to the Soviet Union (261 millions in 1978). The total population of IIASA member countries in 1978 was 969.2 millions or 23 percent of the world total. In terms of level of development the IIASA set contains 17 out of the 30 countries with per capita income over $3,000 per annum in 1978 (see Figure 2). In addition, they have some of the highest life expectancies among the countries of the world (Figure 2) comprising 17 of the 39 countries with life expectancies of 70 or over in 1978. The CMS task thus involved the investigation of the regional population dynamics of a set of relatively rich, developed countries, a majority of which fall in the World Bank's "industrialized countries" category, a minority of which fall in its "centrally planned economies" class.

1.3. The Systems of Interest in the CMS Study: The Systems Studied

For methodological and for practical reasons, the scope of the studies under the CMS task was limited. This was done in order to make the execution of the task feasible. To carry out a 17-nation comparative study it was necessary that the methodology adopted be specified and fixed prior to the start of the study. To have changed the methods used mid-way through the study would have been foolish, for it is on such grounds that many projects fail.

Notes on data sources for Table 1.

[a] World Bank (1980), Table 1, pp. 110-111.

[b] The figure is taken from the research report for the country concerned. See the list given at the end of this paper.

[c] The corresponding Great Britain (the study area for the multiregional population analysis) population is 53.9 million.

[d] The figure for Canada omitting the Yukon and Northwest Territories (the study area for the multiregional population analysis) is 20.7 millions. The population is an arithmetic average of the 1966 and 1977 Census populations in both cases.

[e] This is the 1971 Census population for Austria.

[f] Further multiregional analyses were carried out for 1973 and 1974 in the Poland study when the mid-year population was 33.5 and 33.8 millions.

[g] This is the Census 1975 population

[h] An October 1st, 1971 figure.

[i] Population for July 1st, 1970.

[j] An average of January 1st, 1971 and January 1st, 1972 population estimates.

[k] The Great Britain population was 54.3 millions.

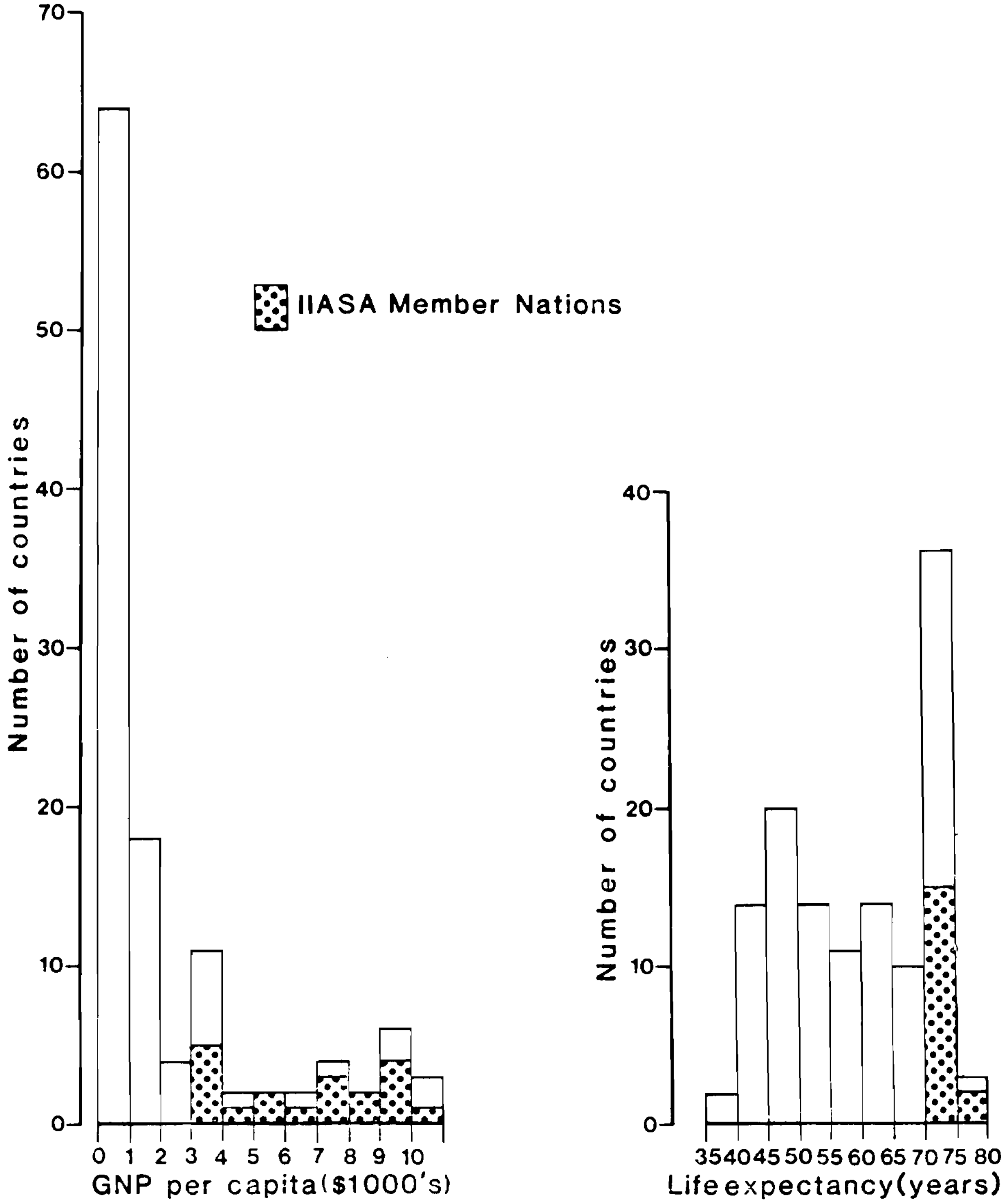


Figure 2. The distribution of countries and IIASA member nations by per capita income and by life expectancy. (Source of data: World Bank 1980).

The important limitations of the CMS task were as follows:

i) The age classification was by five-year age groups.
ii) The regions were limited in number.
iii) The temporal frame of reference was limited to one period in the recent past (end dates vary from 1970 to 1975).
iv) The system studied was limited to the country concerned and interactions with the rest of the world were not generally included in the multiregional analyses. The studies concentrated therefore on the internal redistribution of national populations.
v) Further classifications of the population—by birth place or by ethnic or socioeconomic group—were not attempted.

Detailed comments are made on each of these limitations in the course of this paper. In future studies many of them can and will be eliminated, but in the first such effort of comparing the population dynamics of regions within a set of 17 countries, the limitations, we feel, are justified.

This paper sets out a systematic presentation of the accounting frameworks within which the multiregional population analyses of the CMS task are embedded. The concepts are described in general terms in section 2. Age and time frameworks are developed in section 3. Section 4 deals with the spatial frameworks adopted in the individual country studies, and section 5 describes the estimation problems and procedures involved for the stocks, events, and flows data used as input to the multiregional analysis. An overview of some of the principal problems associated with comparing the results and the lessons to be learned for future work are drawn out in the last section.

2. ACCOUNTING FRAMEWORKS

2.1. Population Accounts in General

Multiregional demographic analysis aims at a better understanding of the dynamics directing population growth and spatial distribution. To accomplish this goal, it depicts the process of change in the population size (stock) and composition in terms of flows of people among various states of the "demographic system." The flow-perspective requires flow data. These data may conveniently be arranged in an accounting framework. Accounts are not only convenient data representation schemes, they also provide a useful framework for evaluating the completeness and accuracy of the available data and for estimating the missing data.

In this paper, the data base of the CMS study will be approached in an accounting framework. The CMS project did not involve the preparation of population accounts prior to "spatial population" analysis, since the main publications on spatial demographic accounts became available after the research strategy for the CMS study had been finalized.

Population accounts are two- or multidimensional tables of population flows. All flows are accounted for by including all possible states of origin as row classes for the account matrices and all possible states of destination as column classes. Flows may be given for the total population (aggregate accounts) or for each of several population categories (disaggregate accounts). In multiregional analysis, age-disaggregated accounts are used, i.e., all flows must be given for each of the (five year) age groups, considered in the study. A complete exposition of population accounts is beyond the scope of this paper; the reader is referred to other works for details (Stone 1971, 1975, Illingworth 1976, Rees and Wilson 1977, Rees 1977a, 1980, 1981). We limit ourselves to the presentation of the two types of accounts that are fundamental in spatial population analysis, since they relate to different ways of measuring migration flows: movement and transition accounts. This distinction has important implications for the analysis and interpretation of the results, as we will show.

2.2. Movements and Transitions

Quite distinctive instruments are used to measure migration flows. Registration systems are generally used in Europe, where each change of address (and hence each move) must be registered with the local authorities. In countries with a registration system, each move (passage from one state to another) is counted, and the statistical data that represent the number of passages are said to be *movement data*. Other countries, like the U.S., Japan, and the U.K., derive migration statistics from a retrospective question in the national census. In this question, respondents are asked to state what their address (place of residence) was some unit number of years ago. Individual moves are not recorded; only the transition a person made between the start and end of a given time interval is recorded. These data representing migration are therefore referred to as *transition data*.

The distinction between movements and transitions are visually illustrated in Figure 3. The figure shows the mobility experience of 8 individuals during the interval from t to t + τ. Each line represents a lifeline of an individual. Every time a person (and his lifeline) crosses the boundary between region i and region j, a move occurs. A person can make several moves within the interval. Persons (3) and (8) make two moves, for example, in Figure 3. Although a person can make several moves within an interval, he/she can make only one transition. Person (3) for instance, will not figure in the transition count, despite having moved to region j for a short period in the time interval. Some transitions—for example, a person alive in region i at time t and dead in region j later in the period—do carry information on the additional move that must have taken place. Note that the distinction between movements and transitions is numerically important only for migration; for birth and death totals the movement and transition counts are identical.

In the CMS study, both registration-based movement data and census-based transition data were employed. Figure 4 shows the type of migration data employed in each national study on base maps depicting the 1978 populations of the countries involved. Movement data were used in 11 out of 17 country studies and transition data were employed in 6 out of 17 countries. In no country were both kinds of data simultaneously available for interregional migration streams. The argument against collecting both types of statistics by national statistical offices has

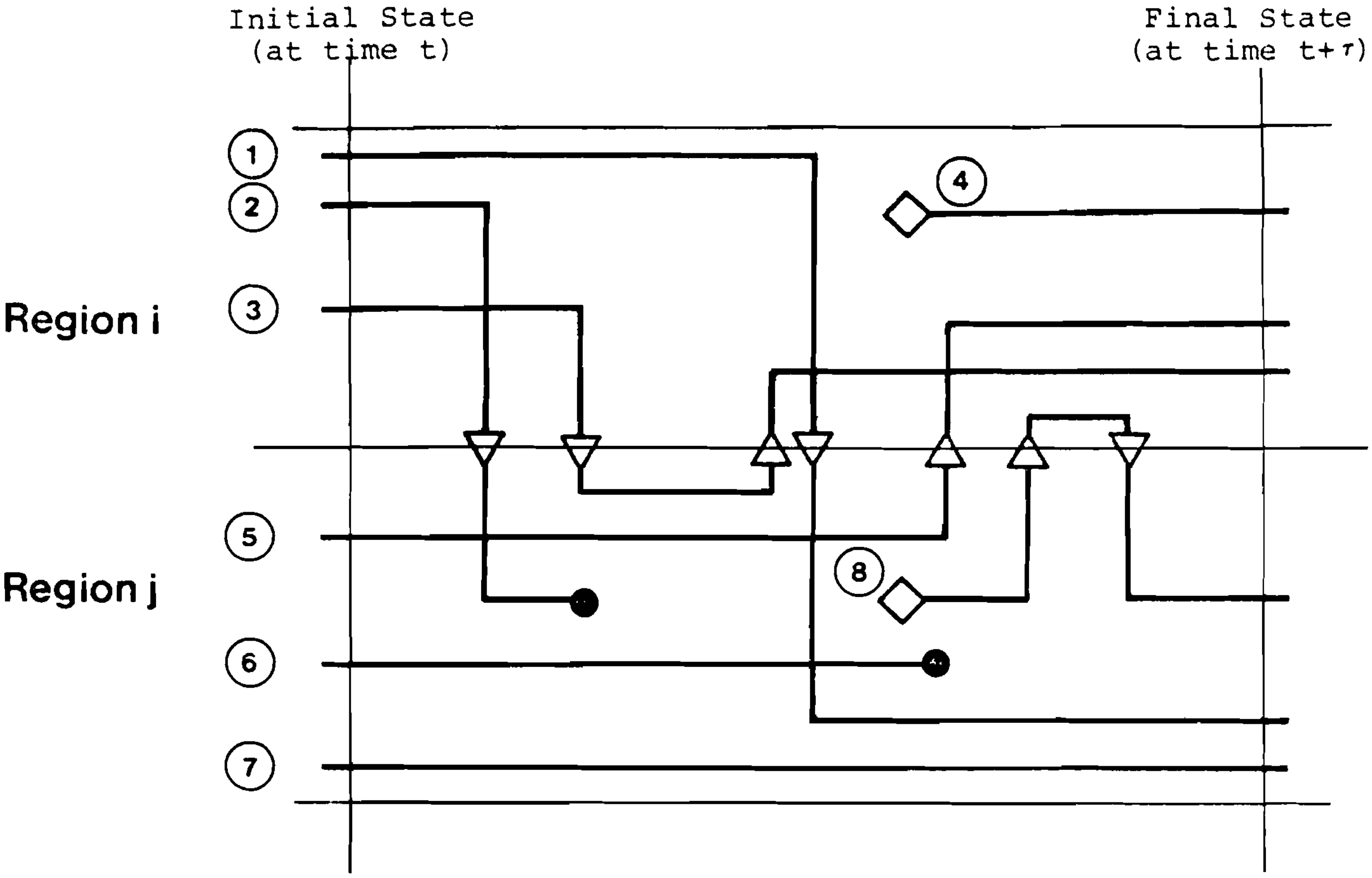


Figure 3. Transitions from the initial state to the final state and movements of individuals represented by lifelines (—), a move from region i to region j (▽), a move from region j to region i (△), a birth (◇), and a death (●).

been that duplication is thereby avoided and costs saved. However, as we have argued in Figure 3 and show in several places later in the paper, the measures of migration are sufficiently different in numerical magnitude to suggest that this argument is unfounded. Views are now beginning to shift and some countries (e.g., the United Kingdom) which have in the past relied on census measures have now developed a registration and survey-based system of data collection (Ogilvy 1980a).

The distinction between movements and transitions is crucial in spatial population analysis. Researchers on migration (Courgeau 1973a, 1973b, 1980, Rees 1977b) have long been aware of the distinction but its implications for population models such as the multiregional life table model have only recently been realized (Ledent 1980a, Ledent and Rees 1980). A major implication

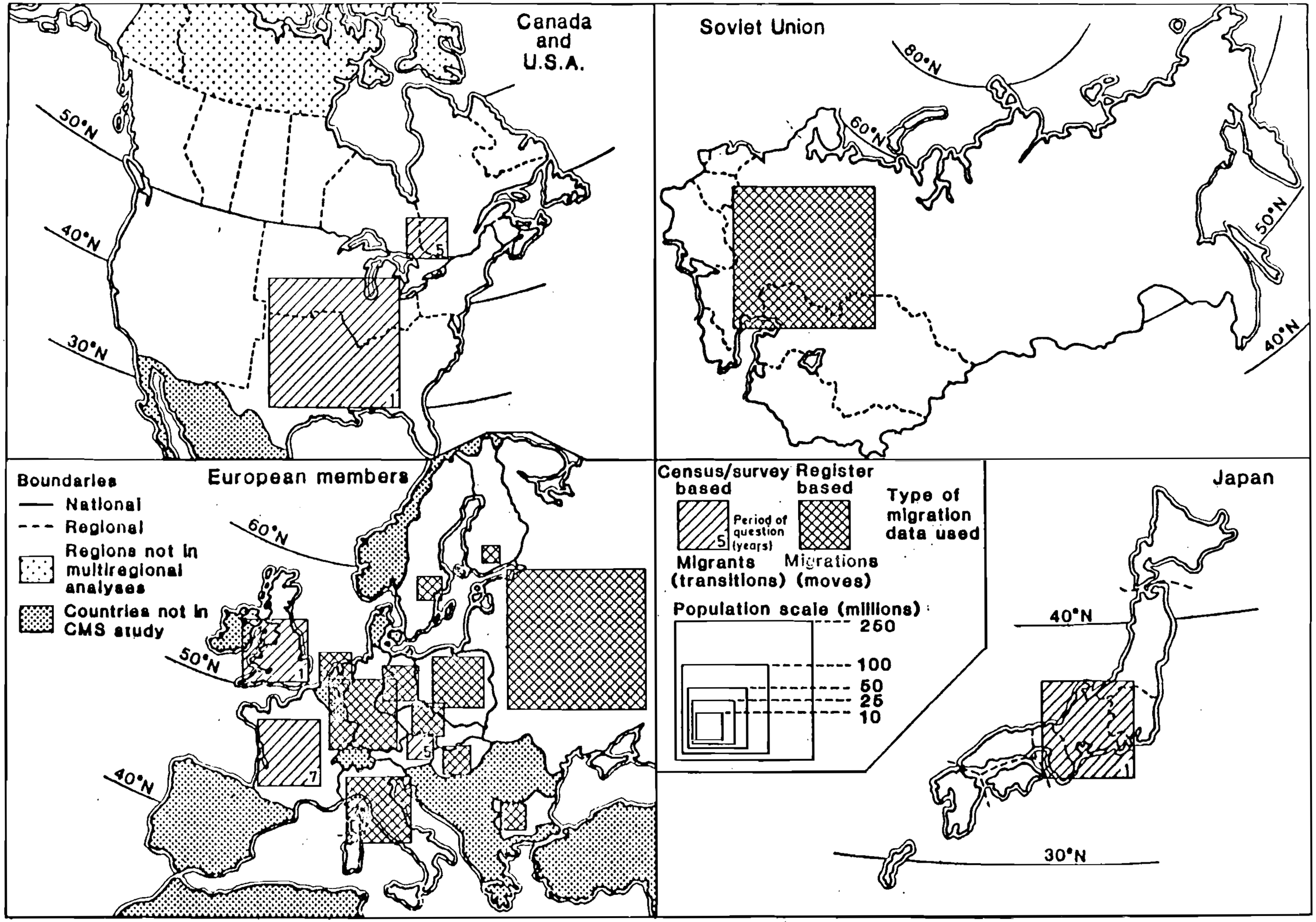


Figure 4. Types of migration data used in the CMS task.

is that the life-table probabilities, which are transition probabilities, should be calculated differently for movement and transition data. In the next section, we elaborate on this implication. First, however, we show how different types of data lead to different accounts.

2.3. Movement Accounts and Transition Accounts

Accounts provide a useful framework for distinguishing between movements and transitions. The structure of a movement account is shown in Table 2, the transition account is presented in Table 3. Each account involves a set I of regions internal to the country, there being 1,2,...,N internal regions, and an external region set, O, called for present purposes "the outside world." To these "existing" states we must add the entry point of birth and the exit point of death. In both accounts initial population stocks are linked to final population stocks via accounting equations. For the moment, no age or sex detail is included. In the movement case for typical region i, the accounting equation is as follows

$$P^{\cdot i} = P^{j\cdot} - \sum_{j \in I} M^{ij} - M^{iO} - D^{i} + \sum_{j \in I} M^{ji} + M^{Oi} + B^{i} \tag{1}$$

stock at end of period = stock at start of period
- outflows + inflows

where I refers to the set of regions internal to the system being studied (usually the country). The first four terms in which total migrations out of region i and deaths in region i are subtracted from the initial population in the region, also serve to define the diagonal entry in the movement accounts, which is designated R^{i} (for residual). This term acts as a balancing item in the accounts, has no substantive interpretation, and may on extreme occasion be negative (as in the example of a movement accounts in Table 4). This R^{i} term is to be distinguished from the term that often appears in the diagonal of movements tables: namely, the total number of movements within the region, which might be termed M^{ii}.

Table 2. Movement account for a multiregional system within a single country in a single period.

State prior to move in period		State after move in period						
		Destination regions						
		Internal regions				Outside world		
		1	2	···	N	0	Death	Totals
Origin regions	1	R^1	M^{12}	···	M^{1N}	M^{10}	D^1	$P^{1\cdot}$
Internal regions	2	M^{21}	R^2	···	M^{2N}	M^{20}	D^2	$P^{2\cdot}$
	⋮	⋮	⋮		⋮	⋮	⋮	⋮
	N	M^{N1}	M^{N2}	···	R^N	M^{N0}	D^N	$P^{N\cdot}$
Outside world		M^{01}	M^{02}	···	M^{0N}	$\emptyset$	$\emptyset$	$M^{0\cdot}$
Birth		B^1	B^2	···	B^N	$\emptyset$	$\emptyset$	$B^{\cdot}$
Totals		$P^{\cdot 1}$	$P^{\cdot 2}$	···	$P^{\cdot N}$	$M^{\cdot 0}$	$D^{\cdot}$	T

Definitions of variables

R^i — residual accounts balancing term for region i (no substantive meaning)

M^{ij} — migrations (moves) from region i to region j

M^{i0} — emigrations (moves) from region i to outside world

D^j — deaths in region j

$P^{i\cdot}$ — population at the start of the period in region i. The . refers to summation over all possible future states in the period and indicates that the population stock is defined for the start of the period

M^{0j} — immigrations (moves) from the outside world to region j

$\emptyset$ — item ignored or set to zero

$M^{0\cdot}$ — the total number of immigrations to internal regions from the outside world

B^i — births in region i

$P^{\cdot j}$ — population at the end of period in region j. The . refers to summation over all possible past states in the period and indicates that the population stock is defined for the end of the period

$M^{\cdot 0}$ — the total number of emigrations from the internal regions to the outside world

$D^{\cdot}$ — the total number of deaths in the internal regions

T — the grand total

Table 3. Transition account for a multiregional system within a single country in a single period.

			Final state in period											
			Survival, s at time t					Death, d in period t, t + T					
	Initial state in period		Internal regions 1	2		... N	Outside world 0	Internal regions 1	2		... N	Outside world 0	Total
Existence, e at time t	Internal regions	1	$K^{e1,s1}$	$K^{e1,s2}$	...	$K^{e1,sN}$	$K^{e1,s0}$	$K^{e1,d1}$	$K^{e1,d2}$	...	$K^{e1,dN}$	$K^{e1,d0}$	$K^{e1,\cdot\cdot}$
		2	$K^{e2,s1}$	$K^{e2,s2}$	...	$K^{e2,sN}$	$K^{e2,s0}$	$K^{e2,d1}$	$K^{e2,d2}$	...	$K^{e1,dN}$	$K^{e2,d0}$	$K^{e2,\cdot\cdot}$
		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
		N	$K^{eN,s1}$	$K^{eN,s2}$	...	$K^{eN,sN}$	$K^{eN,s0}$	$K^{eN,d1}$	$K^{eN,d2}$	...	$K^{eN,dN}$	$K^{eN,d0}$	$K^{eN,\cdot\cdot}$
	Outside world	0	$K^{e0,s1}$	$K^{e0,s2}$	...	$K^{e0,sN}$	$\emptyset$	$K^{e0,d1}$	$K^{e0,d2}$	...	$K^{e0,dN}$	$\emptyset$	$K^{e0,\cdot\cdot}$
Birth in t, t+T	Internal regions	1	$K^{b1,s1}$	$K^{b1,s2}$	...	$K^{b1,sN}$	$K^{b1,s0}$	$K^{b1,d1}$	$K^{b1,d2}$	...	$K^{b1,dN}$	$K^{b1,d0}$	$K^{b1,\cdot\cdot}$
		2	$K^{b2,s1}$	$K^{b2,s2}$	...	$K^{b2,sN}$	$K^{b2,s0}$	$K^{b1,d1}$	$K^{b2,d2}$	...	$K^{b2,dN}$	$K^{b2,d0}$	$K^{b2,\cdot\cdot}$
		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$			$\vdots$	$\vdots$	$\vdots$
		N	$K^{bN,s1}$	$K^{bN,s2}$	...	$K^{bN,sN}$	$K^{bN,s0}$	$K^{bN,d1}$	$K^{bN,d2}$	...	$K^{bN,dN}$	$K^{bN,d0}$	$K^{bN,\cdot\cdot}$
	Outside world	0	$K^{b0,s1}$	$K^{b0,s2}$	...	$K^{b0,sN}$	$\emptyset$	$K^{b0,d1}$	$K^{b0,d2}$	...	$K^{b0,dN}$	$\emptyset$	$K^{b0,\cdot\cdot}$
	Totals	.	$K^{\cdot\cdot,s1}$	$K^{\cdot\cdot,s2}$	...	$K^{\cdot\cdot,sN}$	$K^{\cdot\cdot,s0}$	$K^{\cdot\cdot,d1}$	$K^{\cdot\cdot,d2}$	...	$K^{\cdot\cdot,dN}$	$K^{\cdot\cdot,d0}$	$K^{\cdot\cdot,\cdot\cdot}$

Definitions of variables

T — length of period
t — time at start of period
K — count of transitions or persons making a transition
$K^{ei,sj}$ — persons in existence in region i at time t who survive in region j at time t + T
$K^{ei,dj}$ — persons in existence in region i at time t who die in region j before time t + T
$K^{bi,sj}$ — persons born in region i in period t to t + T who survive in region j at time t + T
$K^{bi,dj}$ — persons born in region i in period t to t + T who die in region j before time t + T
$K^{ei,\cdot\cdot}$ — population of region i at time t
$K^{bi,\cdot\cdot}$ — births in region i
$K^{\cdot\cdot,dj}$ — deaths in region j
$K^{\cdot\cdot,sj}$ — population of region j at time t + T
$K^{eO,\cdot\cdot}$ — immigrants from the outside world to regions within the country
$K^{bO,\cdot\cdot}$ — infant immigrants from the outside world to regions within the country
$K^{\cdot\cdot,sO}$ — surviving emigrants from regions within the country to the outside world
$K^{\cdot\cdot,dO}$ — nonsurviving immigrants from regions within the country to the outside world

Table 4. A movement and a transition account for the Figure 3 example.

A. MOVEMENT ACCOUNTS

State prior to move in period		State after move in period				
		Destination regions				
		Internal regions		Outside world		
		i	j	0	Death	Totals
Origin regions						
Internal regions	i	-1	4	0	0	3
	j	3	-2	0	2	3
Outside world	0	0	0	0	0	0
Birth		1	1	0	0	2
Totals		3	3	0	2	8

B. TRANSITION ACCOUNTS

Initial state in period		Final state in period						
		Survival			Death			
		Internal regions		Outside world	Internal regions		Outside world	
		i	j	0	i	j	0	Totals
Existence								
Internal regions	i	1	1	0	0	1	0	3
	j	1	1	0	0	1	0	3
Outside world	0	0	0	0	0	0	0	0
Birth								
Internal regions	i	1	0	0	0	0	0	1
	j	0	1	0	0	0	0	1
Outside world	0	0	0	0	0	0	0	0
Totals		3	3	0	0	2	0	8

The corresponding equation for the transition accounts that links initial and final populations is

$$K^{\cdot\cdot,si} = \left(K^{ei,\cdot\cdot} - \sum_{\substack{j\in I \\ j\neq i}} K^{ei,sj} - K^{ei,sO} - K^{ei,di} - \sum_{\substack{j\in I \\ j\neq i}} K^{ei,dj} - K^{ei,dO} \right) + \sum_{\substack{j\in I \\ j\neq i}} K^{ej,si} + K^{eO,si} + K^{bi,si} + \sum_{\substack{j\in I \\ j\neq i}} K^{bj,si} + K^{bO,si} \qquad (2)$$

where the terms in brackets define the diagonal entry in the accounts $K^{ei,si}$ which refers to persons who survive and stay in regions over the period. This has substantive meaning in contrast to R^i.

2.4. Placement of the CMS Study in the Accounting Frameworks

We can now place the multiregional analyses of the CMS study within these two accounting frameworks.

There are two aspects of this statement: (1) method of survival probability estimation and (2) degree of closure of the system.

2.4.1. Methods of Calculating the Life Table Survival Probabilities

Methods of calculating survival probabilities from different types of data have been discussed in detail by Ledent and Rees (1980). Three approaches to probability estimation are distinguished: the movement approach, the hybrid approach, and the transition approach.

In the movement approach, movement concepts are applied in the equation for estimating survival probabilities and movement data are employed to estimate the migration rates

$$\underset{\sim}{p}_x = (\underset{\sim}{I} - \frac{n}{2T} \underset{\sim}{M}_x)(\underset{\sim}{I} + \frac{n}{2T} \underset{\sim}{M}_x)^{-1} \qquad (3)$$

where $\underset{\sim}{p}_x$ is a matrix of multiregional survival probabilities for age transition x to x + n, $\underset{\sim}{M}_x$ is a special matrix of migration and mortality rates (see Rogers and Ledent 1976, Willekens and Rogers 1978 for details), n is the age interval and T the time interval involved. This equation is designated "option 3" in Willekens and Rogers (1978). In the hybrid approach, the same equation is used but transition data are used to compute the migration rates involved. Finally, in the transition approach survivorship rates, $\underset{\sim}{S}_x$, are computed directly or indirectly from transition type data (sometimes via accounts) and the survival probabilities computed by interpolation between successive $\underset{\sim}{S}_x$ matrices using

$$\underset{\sim}{p}_x = \frac{1}{2} (\underset{\sim}{S}_x + \underset{\sim}{S}_{x+n}) \qquad (4)$$

or some more sophisticated interpolation technique.

If we cross-tabulate concept (or equation type) against data type, we obtain a box within which the CMS country studies can be classified (Table 5). Since all country studies except that for France employ equation (3) to estimate survival probabilities the classification follows the data type distinction with the exception of the French study. In the French study a method involving transition concepts (Ledent and Courgeau 1980) was applied in which seven-year migration and stayer rates conditional on survival were factored down to a five-year time interval (by raising the matrix of rates to the power 5/7) and then used to compute survival probabilities using conventional mortality rates and an interpolation technique.

The transition approach was also applied by Ledent and Rees (1980) to the United Kingdom data, and a comparison was made between the results of the two approaches for a common three region system (East Anglia, South East, and the Rest of Britain). The results differed significantly: the percentage of life predicted to be spent in the region of birth differed by 4.5 percent in the case of the smallest region, East Anglia. The difference can ultimately be attributed to the way the diagonal terms were handled in these particular applications of hybrid and transition approaches. (For details, see Ledent and Rees 1980).

Table 5. The CMS country studies classified by approach to multiregional life table construction

	Data	
Concepts	Moves	Transitions
Movement	Movement Approach	Hybrid Approach
	Finland Sweden GDR Netherlands Hungary FRG Poland Bulgaria Czechoslovakia Soviet Union Italy	United Kingdom (1) Canada Japan Austria United States
Transition		Transition Approach
		France United Kingdom (2)

Notes:
GDR — German Democratic Republic
FRG — Federal Republic of Germany

Source of figure framework: Ledent and Rees (1980) Table 1.
United Kingdom (1): the Research Report analysis (Rees 1979c).
United Kingdom (2): analysis reported in Ledent and Rees (1980).

2.4.2. *Degree of Closure of the Multiregional Systems*

The importance of migration streams external to the country's system being studied was usually discussed if considered relevant in each CMS report, but external migration flows were not incorporated in the multiregional analysis. This is a common practice in life table analysis where life histories are drawn of persons born in the country considered and where a life expectancy, that is comparable to the conventional concept, can only be calculated by assuming people do not leave the multiregional system. Excluding external migration is an appropriate practice for stable population analysis, where the asymptotic behavior of a population with fixed demographic rates is investigated. The practice is, however, not suited for population forecasting, since external migration affects both population size and population distribution. At the time the CMS study was carried out, the necessary computer programs for incorporating external flows in projections were still under development and there were great difficulties in assembling comparable international migration statistics for regions within countries. It was therefore decided at the 1975 Workshop not to include external migration. It is, however, useful to make an assessment of the importance of external migration as a component of population change.

To make an assessment of the degree to which the various country studies are closed in the CMS study, a common table was drawn up (Table 6). This table is equivalent to either the movement account without a births row and deaths column (see Table 2) or the top left-hand or exist-survive quadrant of the transition account. Added to the table are rows for "total internal inmigration," "immigration flows," and "total inmigration," together with additional columns for "total internal outmigration," "emigration flows," "total outmigration," and "intraregional migration." Our best estimates of these quantities from evidence given in each research report and from other sources is given in Table 7. All figures have been converted into equivalent annual averages by dividing by T, the length of the period (not really the correct method in the case of transition data but not too inaccurate empirically), and the rates have been worked out using mid-year or average populations at risk for the study periods (again not really the correct method in the case of transition data but necessary for comparison with the movement data). In many cases fairly crude estimates had to be made of immigration and emigration flows so that only "relative orders of magnitude comparison" are really possible.

How important are external migration flows? The percentage contribution of external migration to the combined total of internal interregional migration and external migration is computed in column (1) of Table 8 for the countries for which the relevant figures are available. For some three countries—the United Kingdom, Canada, and the Federal Republic of Germany—more than half of the migration flows are external; for three countries—Sweden, the Netherlands, and France—between a quarter and a half

Table 6. A table of migration flows, including external migration.

From		To Internal regions 1 2 ••• N	Totals •	Outside world O	Totals •	Within the region
Internal regions	1	Interregional migration flows* Diagonal elements set to zero Interregional migration flows*	Total internal out-migration	Emigration flows	Total out-migration	Intra-regional migration
	2					
	•					
	•					
	•					
	N					
Totals	•	Total internal inmigration	(B)	(D)	(F)	(A)
Outside World	O	Immigration flows	(C)	0	(C)	
Totals		Total inmigration	(E)	(D)	(G)	

(A) — sum of intraregional migration flows (both origin and destination within region)
(B) — sum of internal interregional migration flows
(C) — sum of immigration flows
(D) — sum of emigration flows
(E) — sum of total inmigration flows = (B) + (C)
(F) — sum of total outmigration flows = (B) + (D)
(G) — sum of all migration flows = (B) + (C) + (D)

*Migration variables included in the CMS study multiregional analyses. The term "migration flows" is used to refer to either migration transitions or migration movements.

Table 7. Internal and external migration flows and rates, IIASA nations, annual averages.

a. Flows (1,000s)

		Number of regions	Total intra-region (A)	Total inter-region (B)	Immi-gration (C)	Emi-gration (D)	Note on source and derivation
United Kingdom	1970-71	10	5,035	834	385	475	(1)
Finland	1974	12	160	116	16	14	(2)
Sweden	1974	8		132	37	28	(3)
GDR	1975	5		100			(4)
Netherlands	1974	5	520	194	93	61	(5)
Canada	1966-71	10	1,478	213	183	90	(6)
Hungary	1974	6	373	339			(7)
Soviet Union	1968-70	15	5,530	1,397	15	15	(8)
FRG	1974	11		929	630	640	(9)
Austria	1966-71	9		170			
Poland	1975	13		298			(10)
Bulgaria	1975	7	86	38			(11)
France	1968-75	8		562	240	119	(12)
Czechoslovakia	1975	12		119			(13)
Japan	1970	8	10,385	2,175			(14)
United States	1965-70	4	38,686	74	371		(15)
Italy	1971	4		442			(16)

b. Rates (per 1,000 population per year)

United Kingdom	1970-71	10	92.8	15.4	7.1	8.8	(1)
Finland	1974	12	34.2	24.7	3.3	3.0	(2)
Sweden	1974	8		16.2	4.5	3.4	(3)
GDR	1975	5		5.9			(4)
Netherlands	1974	5	38.5	14.4	6.8	4.5	(5)
Canada	1966-71	10	71.1	10.2	8.8	4.3	(6)
Hungary	1974	6	35.7	32.4			(7)
Soviet Union	1968-70	15	23.1	5.8	0.1	0.1	(8)
FRG	1974	11		15.0	10.2	10.3	(9)
Austria	1966-71	9		23.8			
Poland	1975	13		8.6			(10)
Bulgaria	1975	7	9.8	4.3			(11)
France	1968-75	8		11.0	4.7	2.3	(12)
Czechoslovakia	1975	12		8.1			(13)
Japan	1970	8	99.2	20.8			(14)
United States	1965-70	4	189.8	0.4	1.8		(15)
Italy	1971	4		8.3			(16)

Notes and sources for Table 7

1. Columns A, B, and C: OPCS (1974), Table 1A, p. 1. Column E from Table 11, p. 27 in Rees (1979c). Emigration estimation procedure explained in Rees (1980).
2. Columns A and B: computed from Appendix A in Rikkinen (1979). Columns C and D: estimated from Table 2.1 and figures on p. 5 in Rikkinen (1979).
3. Column B: computed from Appendix A of Andersson and Holmberg (1980). Columns C and D: estimated from data in Jenkins (1976).
4. Column B: computed from Appendix A of Mohs (1980).
5. Columns A and B: computed from Appendices A and F of Drewe (1980). Columns C and D: estimated from data in Jenkins (1976).
6. All statistics refer to annual averages of five-year figures computed by dividing the five-year figures by five. Column A: computed from Statistics Canada (1977), Table 4.65, p. 217. Column B: computed from Appendix A in Termote (1980). Columns C and D: extracted from Liaw (1977), pp.90-91.
7. Columns A and B computed from Appendix A in Bies and Tekse (1980). The migration figures are for both permanent and temporary migrants.
8. Computed from Tsentral'noe Statisticheskie Upravlenie (1972), Table 1. The number of emigrants is assumed equal to the number of immigrants. (The multiregional analysis was carried out for a system of 8 regions only).
9. Column B: computed from the Appendix in Koch and Gatzweiler (1980). Columns C and D: estimated from data in Jenkins (1976).
10. Columns A and B: estimated from Dziewonski and Korcelli (1978), Tables 5 and 9.
11. Columns A and B: computed from Appendix A, Tables 1.6 and 2.2 in Philipov (1978).
12. Column B: computed from Table 27, p. 97 in Ledent and Courgeau (forthcoming) Columns C and D: Table 19, p. 77 in Ledent (1980a).
13. Column B: computed from Appendix A of Kuhnl (forthcoming).
14. Columns A, B, and C total combined: rate from Long and Boertlein (1976), Table 2, p. 6 and population from Nanjo, Kawashima, and Kuroda (forthcoming) Table 7, p. 22, Migrants from Table 11, p. 31 in Nanjo, Kawashima, and Kuroda (forthcoming). Column A is computed by subtraction assuming the column C figure is zero.
15. Totals of Columns A, B, and C from rate in Long and Boertlein (1976), Table 2, p. 6 and resident population of USA at July 1, 1970 in U.S. Bureau of the Census (1975); Column B figure from Table 3.2, p. 31 in Long and Frey (forthcoming); Column C total from U.S. Bureau of the Census (1975). Column A obtained by subtraction.
16. Column B: computed from Just and Rogers (1980), Appendix A.

General Note: a caveat

The different countries have rather varying definitions of the terms "immigrant" and "emigrant." The United Kingdom statistics (in the Census and in the International Passenger Survey) include British citizens: in fact, these are the largest citizenship group in both immigration and emigration streams. Canadian immigration statistics, on the other hand, do not appear to count returning Canadian citizens who have spent some permanent residence time abroad. United States statistics fail to mention the existence of emigrants (to the authors' knowledge). The principal difficulties stem from the distinction between an international migrant in the demographic sense (a person changing permanent residence across an international border with an intention to reside for a year at least at the destination) and an international migrant in the legal sense (a person admitted to a country under the provisions of a particular piece of legislation, either for settlement or for work). As a result the statistics in this table can only be regarded as reflecting the order of magnitude of external flows. Precise comparison of flow figures in the table is therefore unwise.

Table 8. Degrees of closure of the CMS study regional systems of the IIASA nations.

Country	Period	External migration as a percentage of total interregional $\frac{(C+D)}{(B+C+D)}$ 100^{a} (1)	Percentage closure of CMS task regional system 100+(1) (2)	Immigrants as a percentage of total inmigrants $\frac{C}{(B+C)}$ 100 (3)	External migration as a percentage of total migration $\frac{(C+D)}{(A+B+C+D)}$ 100 (4)
United Kingdom	1970-71	51	49	32	13
Finland	1974	20	80	12	10
Sweden	1974	33	67	22	
GDR	1975	$<5^{b}$			
Netherlands	1974	44	56	32	18
Canada	1966-71	56	44	46	12
Hungary	1974	$<5^{b}$			
Soviet Union	1968-70	2	98	1	0
FRG	1974	57	43	29	
Austria	1966-71	$5-24^{b}$			
Poland	1973	$<5^{b}$			
Bulgaria	1975	$<5^{b}$			
France	1968-75	39	61	30	
Czechoslovakia	1975	$<5^{b}$			
Japan	1970	<5			
United States	1965-70	$5-24^{b}$			
Italy	1971	$5-24^{b}$			

[a] Definitions in terms of Table 7 columns.

[b] Guesstimates.

of the migration flows are external. For the Soviet Union external migration makes up only 2 percent of total migration flows. Of the remaining countries we would guess that for Finland, Austria, the United States, and Italy, external migration contributes between 5 and 25 percent of total flows, whereas for Japan, the German Democratic Republic, Hungary, Poland, Bulgaria, and Czechoslovakia external flows probably make up less than 5 percent of the total. The pattern of estimates and guesses is displayed in Figure 5. The significance of international labor exchanges for Western Europe and North America and their significance for Eastern Europe, the Soviet Union, and Japan is clear. The second column of Table 8 gives the same statistics as the first but in a complementary fashion. The third column assesses the contribution of immigrants to the migrant populations entering the CMS regions—less than 1 percent of Russian inmigrants come from abroad, whereas 46 percent of Canadian inmigrants are immigrants. The final column of the table provides a statistic that is independent of the number and size of regions chosen for analysis in the country studies by expressing the number of external migrants as a percentage of all migrants irrespective of the distance of migration. This reduces the importance of international migration as a phenomenon, of course, but in certain countries its contribution to total migrant flows is substantial.

What kinds of effects are introduced by ignoring external migration flows in the 10 studies where the flows are important?

Three effects may be distinguished.

(i) The effect on computation of overall life expectancies (either single region or multiregion) will be fairly unimportant. More important in this respect may be the method of estimating mortality probability from mortality rates for the first age interval.

(ii) The effect on the computation of projected populations will be of some importance, especially if emigration and immigration are very differently distributed within the country concerned. This is certainly the case in the United Kingdom, France, and the Netherlands in which immigration flows are concentrated in the capital regions (South East, Paris region, and South Holland, respectively), and in which emigrant flows are more evently distributed.

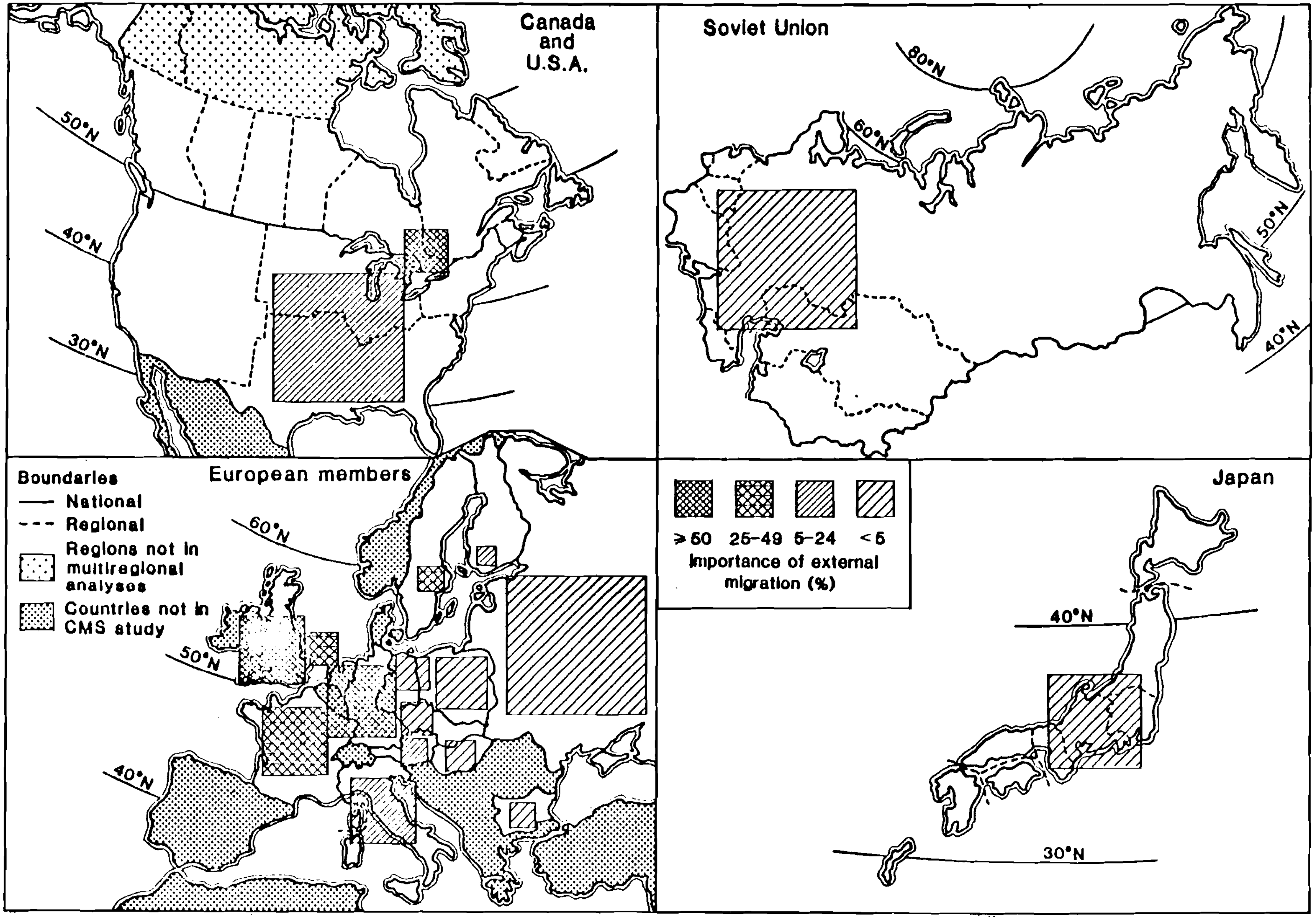


Figure 5. The contribution of external migration to total migration flows.

(iii) The major impact, however, will be on these products of multiregional population analysis involving "matrix" results—for example, the matrix of life expectancies by place of birth and place of residence. A significant proportion of the lives of people living in countries with high emigration rates are likely to be spent outside that country. This is not within the demographic system covered in the CMS study. What happens in the CMS analysis is that the years that would have been spent abroad are added wholly to those spent in the region of birth in the movement and hybrid approaches, and are added proportionally over all internal regions in the transition approach (see Ledent and Rees 1980, section 5.2 for the detailed arguments).

3. SEX, TIME, AND AGE FRAMEWORKS

3.1. Disaggregation by Sex

Population, births, deaths, and migration data were not available for all countries by sex. Where these data were available, they were aggregated to persons prior to the analysis in most studies though occasionally analyses were carried out for both sexes (as in the multiregional life tables of the Canadian and French studies). The aggregation was simply made to limit the amount of output that national authors had to digest.

For the demographic analyses, single-sex models were employed rather than the more customary female dominant models. It was assumed that the births were distributed by age group of mother at time of maternity. Although the distribution by average age of parents will differ slightly from that by female parent only, this difference is, in practice, negligible. Extension of the computer program to handle two-sex models in future exercises is planned.

3.2. Time Frameworks

Demographic data are collected continuously or periodically over time. The population analyst has to select the appropriate statistics for a base period for input to this model. The base period should ideally be the same for all component inputs (Figure 6).

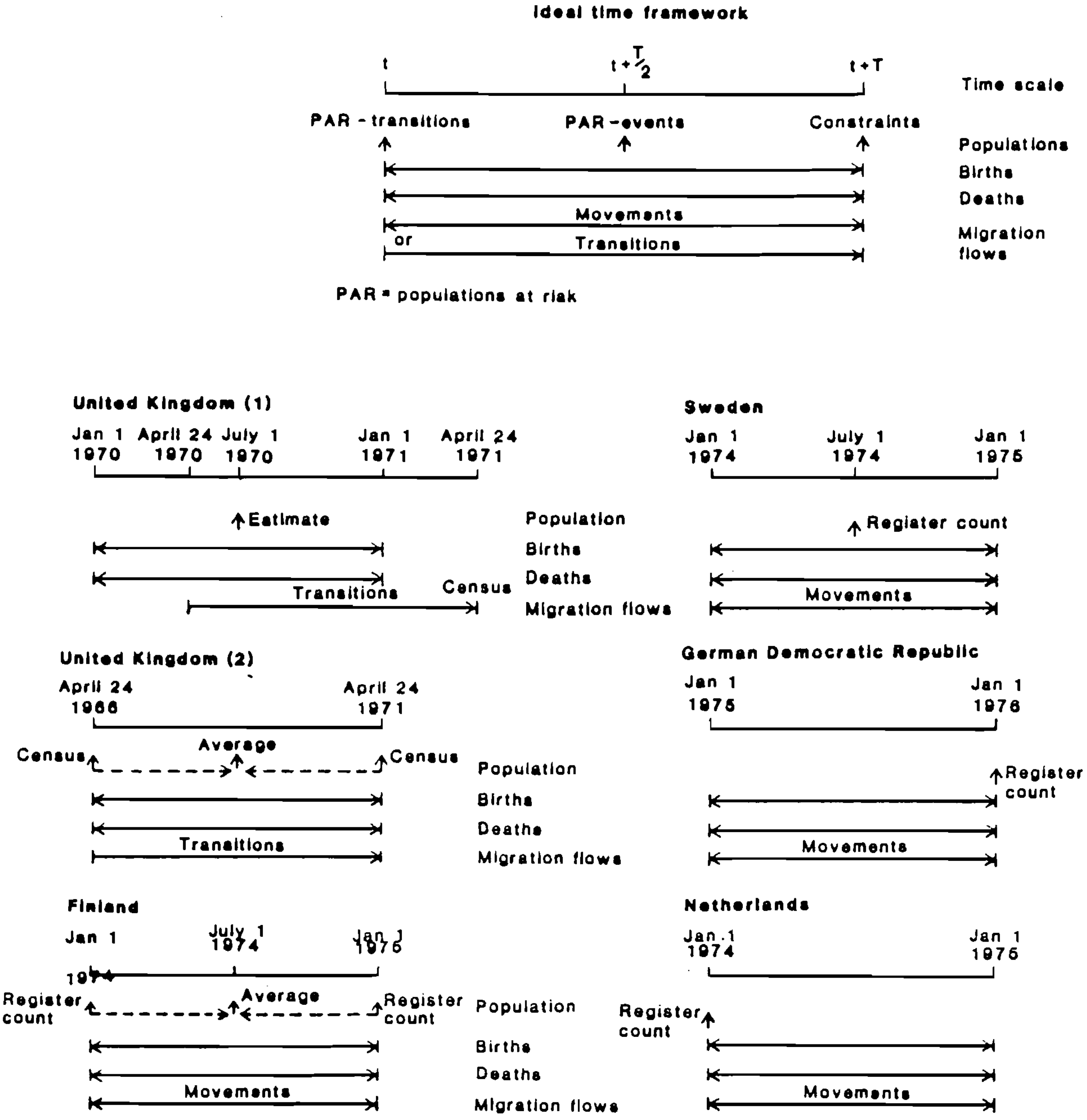


Figure 6. The ideal time framework for the CMS task.

The population stock data needed for multiregional population analysis should fulfill two functions in an ideal framework. They should serve as constraints on population change over the period and as a population at risk for events and flows during the period. Thus, if we were to build population accounts prior to our multiregional analysis, we would need data for beginning of period populations, end of period populations, and an estimate of the populations at risk for births, deaths, and movements data, which might be either the average of beginning and end of period populations or the mid-period population estimate or a multiregional population at risk computed from associated transition accounts (Rees and Wilson 1977). The populations at risk for transition data would be the beginning of period populations (Ledent and Rees 1980), since here survivorship proportions (probabilities) and not rates are calculated from the data.

However, as Figure 7 reveals the actual frameworks used are, in several cases, far from ideal, even given the absence of underpinning population accounts. In general, the studies which employ movements data derived from register data are able to match the periods used for births, deaths, and migration flow data precisely and use the mid-period or average population stocks as populations at risk in their rate calculation (Finland, Sweden, Federal Republic of Germany, Bulgaria, Czechoslovakia). Because of lack of adequate data, other movement-based studies employed beginning of period or end of period populations (German Democratic Republic, Netherlands, Hungary, Soviet Union, Poland), thereby introducing errors into their subsequent rate calculations. The errors are expected to be small, however, since the time interval of the periods of study is one year.

Transition-based studies faced the problem that the periods of measurement of transition data, being keyed to the census data, were not the same as those for the births and deaths data. The solution in the case of the Canadian study (and the United Kingdom analyses reported in Rees 1979a and Ledent and Rees 1980, hereafter referred to as United Kingdom 2) was to estimate the requisite period vital statistics from detailed time series (monthly). In other cases (United Kingdom analysis reported in

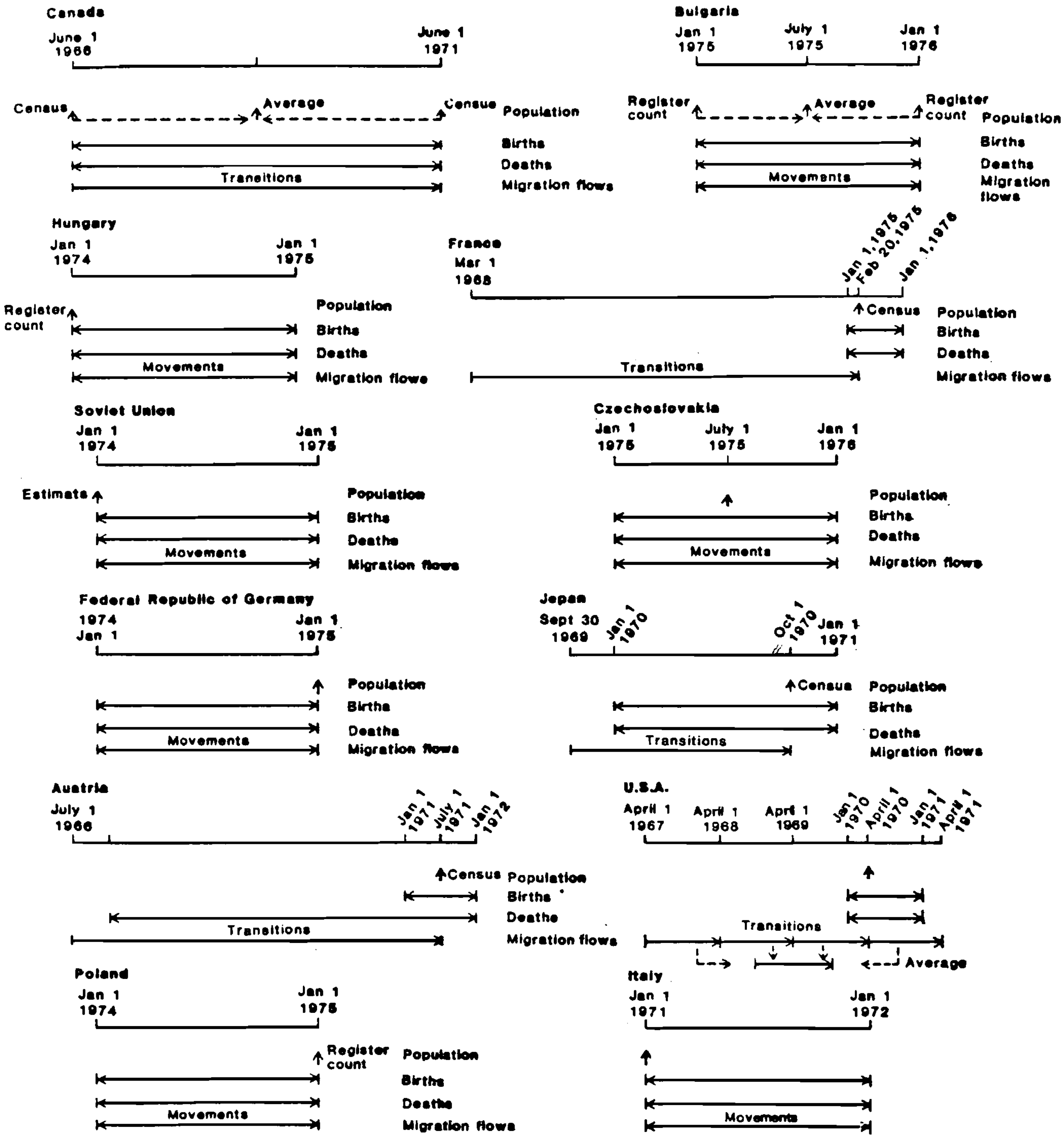


Figure 7. The time framework for the data inputs to the CMS task.

in Rees 1979c, hereafter referred to as United Kingdom 1, France, and Austria) some of the discrepancies in time reference were left unresolved or only partially adjusted for.

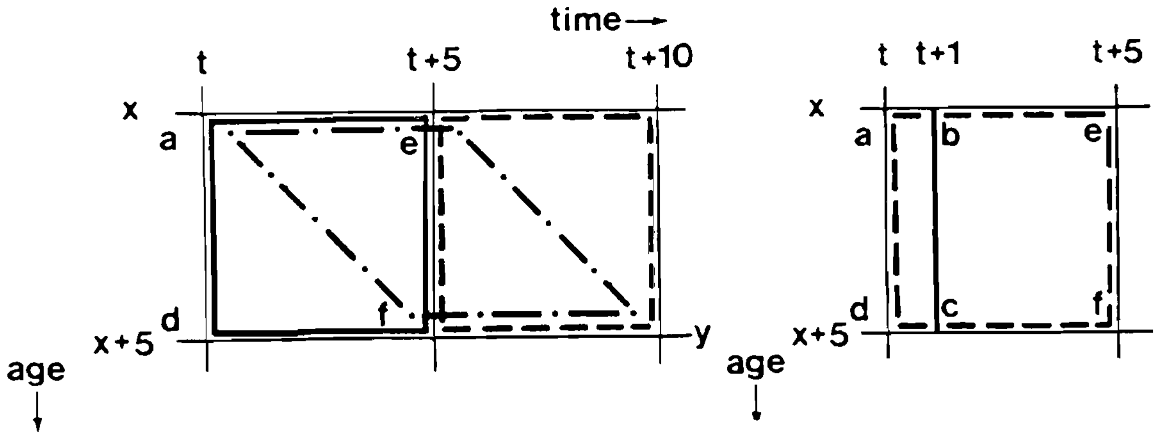
The errors that result from the kind of misspecification of time reference of the input components are probably not great for the one-year-period transition data. The degree of error depends on the degree of instability in fertility and mortality rates during the period. The procedure adopted in the Canadian study of compiling birth and death statistics for all the years in the period of transition is to be recommended.

Conversely, the discrepancies highlight the inadequacy of periodic censuses for monitoring migration behavior unless these censuses are quinquennial and linked (as in the Canadian practice). In the United Kingdom, the pressure from population practitioners has persuaded the Office of Population Censuses and Surveys to exploit partial register data (the National Health Service Central Register) to generate movement data for quarterly and annual periods (Ogilvy 1980a, 1980b). There is, however, one good argument for not being rigid in adhering to the time framework dictated by accounting principles. The flexibility in time frameworks allowed by the programs used in the CMS study makes possible analysis of population dynamics when the data are less than ideal.

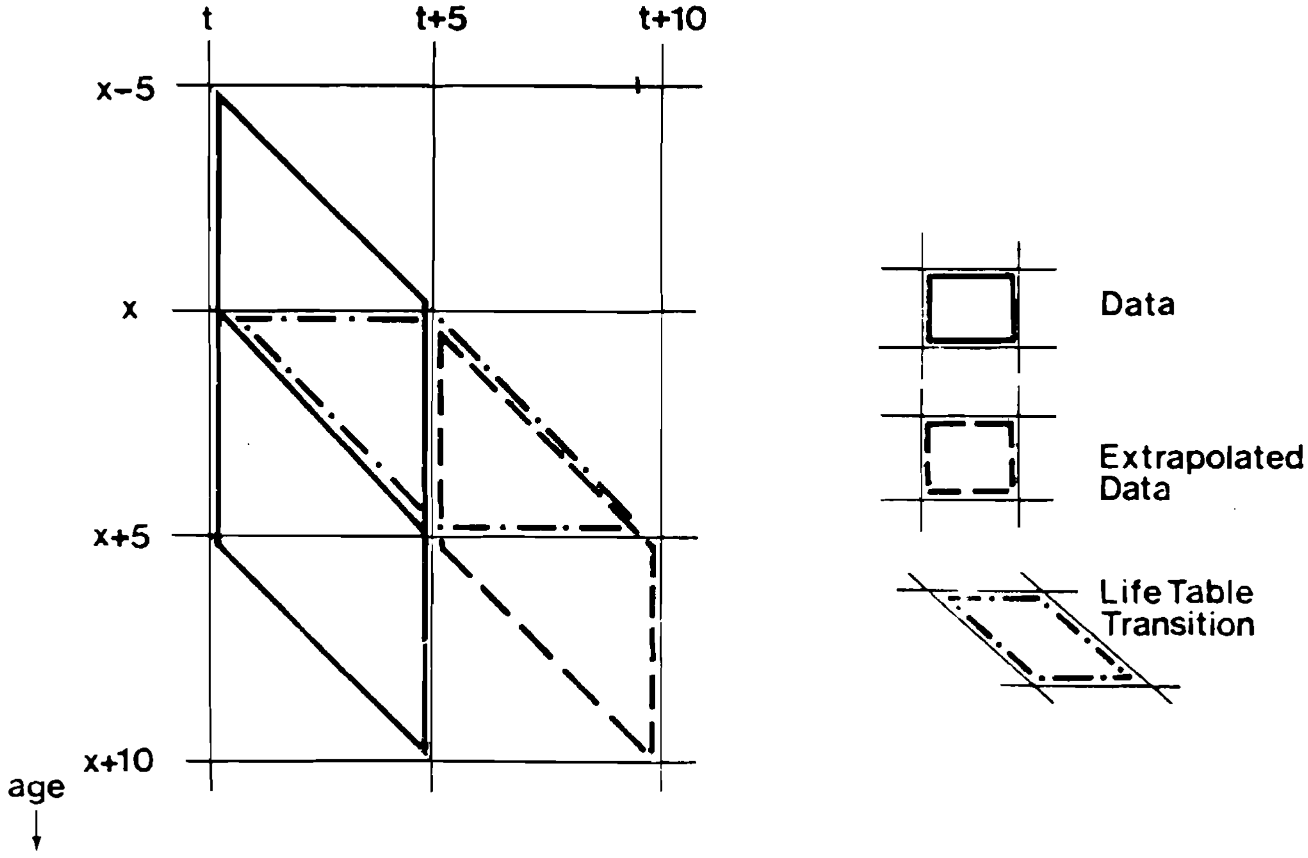
3.3. Age-Time Frameworks

In the movement approach all data are normally classified by age at the time of the event (birth/maternity, death, or migration): for example, events over one year could therefore be counted in space abcd in Figure 8A's Lexis diagram and extrapolated (by multiplication by five) to fill space aefd. In the transition approach, transitions (of migrants) are counted by cohorts moving from being aged x - 5 to x at time t to being aged x to x + 5 at time t + 5 as in Figure 8B.

It is easy to see how the all-age transition accounts of Table 3 can be converted into age-specific accounts (Table 9). All we have done is add a subscript, say x, to indicate the



A. Movement approach framework



B. Transition approach framework

Figure 8. Lexis diagrams showing the age-time frameworks of the movement and transition approaches.

Table 9. A transition accounts table for cohort x.

		Final state in period										
		Survival at time t					Death in period t to t + T					
Initial state in period		Internal regions 1	2	...	N	Outside world 0	Internal regions 1	2	...	N	Outside world 0	Totals .
Internal regions	1	$K_x^{e1,s1}$	$K_x^{e1,s2}$	...	$K_x^{e1,sN}$	$K_x^{e1,s0}$	$K_x^{e1,d1}$	$K_x^{e1,d2}$	...	$K_x^{e1,dN}$	$K_x^{e1,d0}$	$K_x^{e1,..}$
	2	$K_x^{e2,s1}$	$K_x^{e2,s2}$	...	$K_x^{e2,sN}$	$K_x^{e2,s0}$	$K_x^{e2,d1}$	$K_x^{e2,d2}$	...	$K_x^{e2,dN}$	$K_x^{e2,d0}$	$K_x^{e2,..}$
	⋮	⋮	⋮		⋮	⋮	⋮			⋮	⋮	⋮
	N	$K_x^{eN,s1}$	$K_x^{eN,s2}$	...	$K_x^{eN,sN}$	$K_x^{eN,s0}$	$K_x^{eN,d1}$	$K_x^{eN,d2}$	...	$K_x^{eN,dN}$	$K_x^{eN,d0}$	$K_x^{eN,..}$
Outside world	0	$K_x^{e0,s1}$	$K_x^{e0,s2}$	...	$K_x^{e0,sN}$	$\emptyset$	$K_x^{e0,d1}$	$K_x^{e0,d2}$	...	$K_x^{e0,dN}$	$\emptyset$	$K_x^{e0,..}$
Totals	.	$K_x^{e.,s1}$	$K_x^{e.,s2}$	...	$K_x^{e.,sN}$	$K_x^{e.,s0}$	$K_x^{e.,d1}$	$K_x^{e.,d2}$	...	$K_x^{e.,dN}$	$K_x^{e.,d0}$	$K_x^{e.,..}$

Notes:

1. Infant accounts have the same structure: a b superscript is substituted for e, and the value of x set to -n.
2. The x subscript means "aged x to x + n at time t, aged x + n to x + 2 at time t + T for survivor variables; the subscript means "aged x to x + n at time t, dying before attainment of age x + n at time t + T for non-survivor variables.

parallelogram of interest [with points (x,t), (x + 5,t), (x + 5, t + 5), and (x + 10, t + 5) in Figure 8B], following Stone (1971) and Rees (1981). The accounts are reduced in size because the births portion in the aggregate representation simply becomes the first, infant cohort. The transition rates or survivorship rates, S, for multiregional population analysis are simply defined as

$$S_x^{ij} = K_x^{ei,sj} \;/\; K_x^{ei,..} \tag{5}$$

and

$$S_{-n}^{ij} = K_{-n}^{bi,sj} \;/\; K_{-n}^{bi,..} \tag{6}$$

where S_x^{ij} is the rate at which the region i population aged x to x + n survives to be aged x + n to x + 2n in region j. The s_{-n}^{ij} is the equivalent rate for persons born (aged "-n to 0" at the start) in the period.

We can make the same disaggregation of the movement accounts (Table 10) but this involves shifting the age-time framework to that of the transition approach (Figure 8B). If we stay with the Figure 8A framework, we must define two accounts referring to triangles aef and afd involving two separate cohorts, labelled c and c + 1, respectively. In these two accounts only one of the opening and closing stocks refer to population at a point in time. The other one refers to population attaining an exact age in a period. When we put them together to form a life table parallelogram (aegf), by assuming the rates of adf apply to efg, an opening stock of persons attaining age x is matched with a closing stock of persons attaining age x + n. Information for these opening and closing stocks is simply not available (although the accounts can be constructed—see Chapters 14 and 15 in Rees and Wilson 1977).

3.4. Age Classifications

All of the CMS studies adopt five-year age groupings as the level of disaggregation for which it was feasible to gather or estimate the necessary input data streams. Care, however, is

Table 10. Movement accounts for cohorts and for age-group-cohort combinations: A. cohort accounts, B. cohort-age-group accounts 1, C. cohort-age-group accounts 2.

		State after move						
		Destinations						
		Internal regions				Outside world		Totals
State before move		1	2	...	N	O	Death	.
A. COHORT ACCOUNTS								
Origins	1	R_x^1	M_x^{12}	...	M_x^{1N}	M_x^{10}	D_x^1	$P_x^{1\cdot}$
Internal regions	2	M_x^{21}	R_x^2	...	M_x^{2N}	M_x^{20}	D_x^2	$P_x^{2\cdot}$
	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$
	N	M_x^{N1}	M_x^{N2}	...	R_x^N	M_x^{N0}	D_x^N	$D_x^{N\cdot}$
Outside world	O	M_x^{01}	M_x^{02}	...	M_x^{0N}	0	0	$M_x^{0\cdot}$
Totals		$P_x^{\cdot 1}$	$P_x^{\cdot 2}$	...	$P_x^{\cdot N}$	$M_x^{\cdot 0}$	$D_x^{\cdot}$	T_x
B. AGE-GROUP-COHORT ACCOUNTS 1								
Origins	1	${}_cR_x^1$	${}_cM_x^{12}$	...	${}_cM_x^{1N}$	${}_cM_x^{10}$	${}_cD_x^1$	${}_cQ_x^{1\cdot}$
Internal regions	2	${}_cM_x^{21}$	${}_cR_x^2$	...	${}_cM_x^{2N}$	${}_cM_x^{20}$	${}_cD_x^2$	${}_cQ_x^{2\cdot}$
	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$
	N	${}_cM_x^{N1}$	${}_cM_x^{N2}$	...	${}_cR_x^N$	${}_cM_x^{N0}$	${}_cD_x^N$	${}_cQ_x^{N\cdot}$
Outside world	O	${}_cM_x^{01}$	${}_cM_x^{02}$	...	${}_cM_x^{0N}$	0	0	${}_cM_x^{0\cdot}$
Totals	.	${}_cP_x^{\cdot 1}$	${}_cP_x^{\cdot 2}$	...	${}_cP_x^{\cdot N}$	${}_cM_x^{\cdot 0}$	${}_cD_x^{\cdot}$	${}_cT_x$

Table 10 continued.

State before move		State after move Destinations Internal regions 1	2		N	Outside world 0	Death	Totals
C. AGE-GROUP-COHORT ACCOUNTS 2								
Origins	1	${}_{c+1}R^{1}_{x}$	${}_{c+1}M^{12}_{x}$	...	${}_{c+1}M^{1N}_{x}$	${}_{c+1}M^{10}_{x}$	${}_{c+1}D^{1}_{x}$	${}_{c+1}P^{1}_{x}$
Internal regions	2	${}_{c+1}M^{12}_{x}$	${}_{c+1}R^{2}_{x}$	...	${}_{c+1}M^{2N}_{x}$	${}_{c+1}M^{20}_{x}$	${}_{c+1}D^{2}_{x}$	${}_{c+1}P^{2}_{x}$
	⋮	⋮	⋮		⋮	⋮	⋮	⋮
	N	${}_{c+1}M^{N1}_{x}$	${}_{c+1}M^{N2}_{x}$	...	${}_{c+1}R^{N}_{x}$	${}_{c+1}M^{N0}_{x}$	${}_{c+1}D^{N}_{x}$	${}_{c+1}P^{N}_{x}$
Outside world	0	${}_{c+1}M^{01}_{x}$	${}_{c+1}M^{02}_{x}$	...	${}_{c+1}M^{0N}_{x}$	0	0	${}_{c+1}P^{0\cdot}_{x}$
Totals		${}_{c+1}Q^{\cdot 1}_{x}$	${}_{c+1}Q^{\cdot 2}_{x}$	...	${}_{c+1}Q^{\cdot N}_{x}$	${}_{c+1}M^{\cdot 0}_{x}$	${}_{c+1}D^{\cdot}_{x}$	${}_{c+1}T_{x}$

Notes:

1. Infant accounts follow the same structure except that variables $B^{1} \cdots B^{N}$ replace $P^{1\cdot}_{x} \cdots P^{1\cdot}_{x}$ and -n replaces x.
2. The meaning of x is the same as that used in the transition accounts (Figure 9).

Definitions: as in Table 2, except for

c — label for cohort born t - x to t - x + n years ago.

c+1 — label for cohort born t - x + n to t - x years ago.

${}_{c}Q^{i\cdot}_{x}$ — count of persons in cohort x attaining their xth birthday in region i in period t to t + T.

${}_{c+1}Q^{\cdot i}_{x}$ — count of persons in cohort c + 1 attaining their x + nth birthday in region i in period t to t + T.

needed in handling the age classifications. For births, deaths, and movements data the classification is straightforwardly linked to the model framework (Figure 9A). However, for transition-type migration data care is needed in interpreting and adjusting the age classification.

Censuses tabulate migration data by age group at the time of the census. It is incorrect to assume that the figures in the tables can be entered directly into the rate definition equations of either the movement or hybrid approach (Figures 9B and 9C): mobility rates should be defined as neither

$$m_x^{ij} = \frac{K_x^{ei,sj}}{K_x^i \ (T + \frac{T}{2})} \tag{9}$$

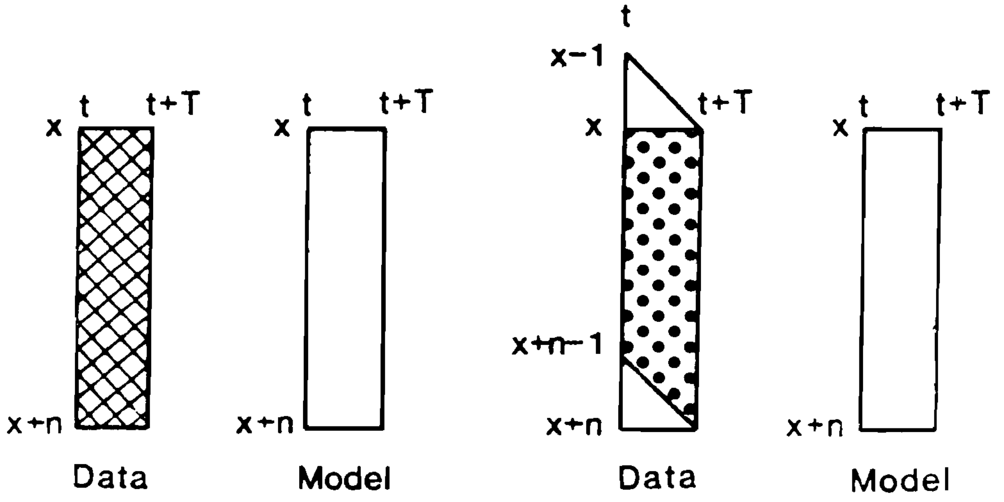
nor

$$m_x^{ij} = \frac{K_{x+n}^{ei,sj}}{K_x^i \ (T + \frac{T}{2})} \tag{10}$$

Instead, the necessary model inputs should be estimated from adjacent age groups data thus:

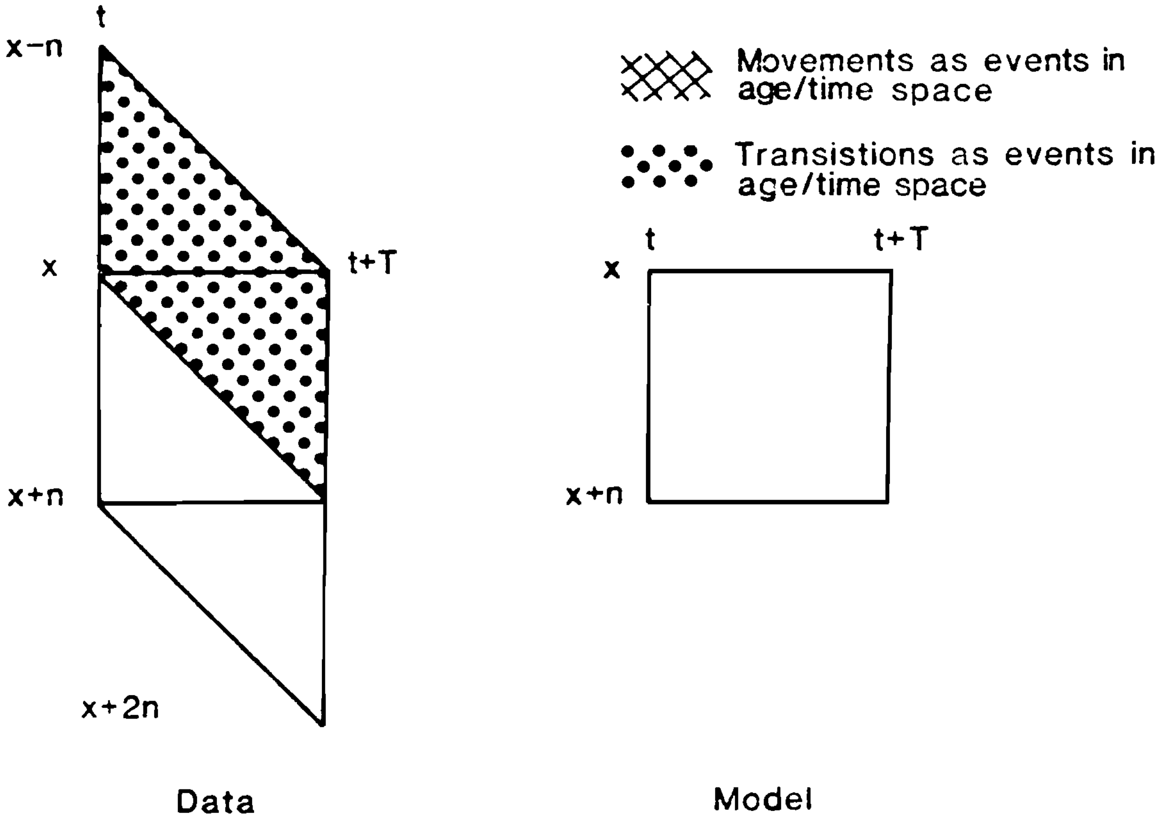
$$m_x^{ij} = \frac{(1 - \frac{T}{2n}) \ K_x^{ij} + (\frac{T}{2n}) \ K_{x+n}^{ij}}{K_x^i \ (T + \frac{T}{2})} \tag{11}$$

In the United Kingdom this equation is used to convert the data to the model age-time space (Figure 9B). In the Austrian study equation (9) is used and in the Canadian study equation (10) is used to derive the mobility rates. However, Ledent and Rees (1980, Table 25) show that the error introduced does not have very serious effects on statistics in the multiregional population analysis in which the age subscript has been summed over.



A. Movement approach:
1 year
(n - 5, T - 1)

B. Hybrid approach:
1 year
(n - 5, T - 1)



C. Hybrid approach: 5 year
(n - 5, T - 5)

Figure 9. The data and model matching problem. (Source: Ledent and Rees 1980, Figure 4.)

3.5. Time Period of Migration Measurement

A glance back at Figure 7 shows that the time period over which movement data were gathered was always one year. Figure 8A shows that this data is extrapolated to an imaginery five-year period in the multiregional population analysis [because of the $\frac{n}{T}$ terms on the right hand side of equation (3)]. If we had measured movements over five years instead, and the underlying migration pattern had not changed, we would have recorded the same number of moves; moves like births and deaths are additive over time.

In the case of studies where transition data were used, the time periods were one year (United Kingdom 1), five years (Canada, Austria, United States, United Kingdom 2), and seven years (France). The one-year data were multiplied by five in the United Kingdom 1 case and the seven-year French migration matrix was raised to the fraction power ($\frac{5}{7}$) in order to make the time interval (T) and age interval (n) equivalent. The other data matched requirements as to n and T, and so were not adjusted. A comparison in Table 11 of the United Kingdom 1 (Rees 1979c) and United Kingdom 2 life tables (Ledent and Rees 1980) reveals a profound difference between the outcomes in terms of distribution of life among regions of residence for the one-year and five-year based analyses. Although exact analysis of these differences awaits further study and more detailed information, the principal reason appears to be that the transition probabilities are not homogeneous across subgroups of the population.

What implications do these findings have for comparing the results of the CMS studies? Clearly, they mean that we are unable to compare results with precision unless the studies concerned use the same type of migration data and same period length. The differences that can be generated are displayed in Figure 10. The figure presents retention percentages, i.e., the population-weighted average percentages of life expected to be lived in the region of birth, for a three region United Kingdom system (East Anglia, South East and Rest of Britain) and a four-region United States system (averages are calculated over the regions).

Table 11. Life expectancies and percentage distributions of life for three United Kingdom regions.

Region of birth	Percentage of life spent resident in: East Anglia	South East	Rest of Britain	Life expectancy (years)
ONE-YEAR MIGRATION AND DEATHS DATA (U.K. 1)				
East Anglia	41	26	33	72.4
South East	4	65	31	72.5
Rest of Britain	2	16	81	71.5
FIVE-YEAR MIGRATION AND DEATHS DATA (U.K. 2)				
East Anglia	56	20	24	72.8
South East	4	74	23	72.5
Rest of Britain	2	12	87	71.7

Source: Ledent and Rees (1980) Appendix A.3.1., runs 7 through 15. Not all the percentages sum to 100 because of rounding error.

The figure indicates that on the basis of five-year migration data a person in the United Kingdom may on the average expect to spend 82 percent of his/her lifetime in his/her region of birth. The retention percentage is 75 if one-year migrant data are employed. The statistic labelled "one-year migrant x 1.220" needs some explanation. The one-year complements of the retention figures (100-retention percentage—that is, the percentage of life expected to be spent outside the region of birth) were multiplied by the reciprocal of the ratio of transitions to moves for interregional migration plans reported by Ogilvy (1980a). This 69 percent is thus a guess as to the results that might have been obtained if movement data had been used.

The equivalent United States figure using five-year migrant data is 69 percent. The difference of 13 percentage points between this United States figure and its United Kingdom counterpart is a real one and means that persons born in the United States are expected to spend 13 percent more of their lives outside their birth region than are persons born in the United Kingdom.

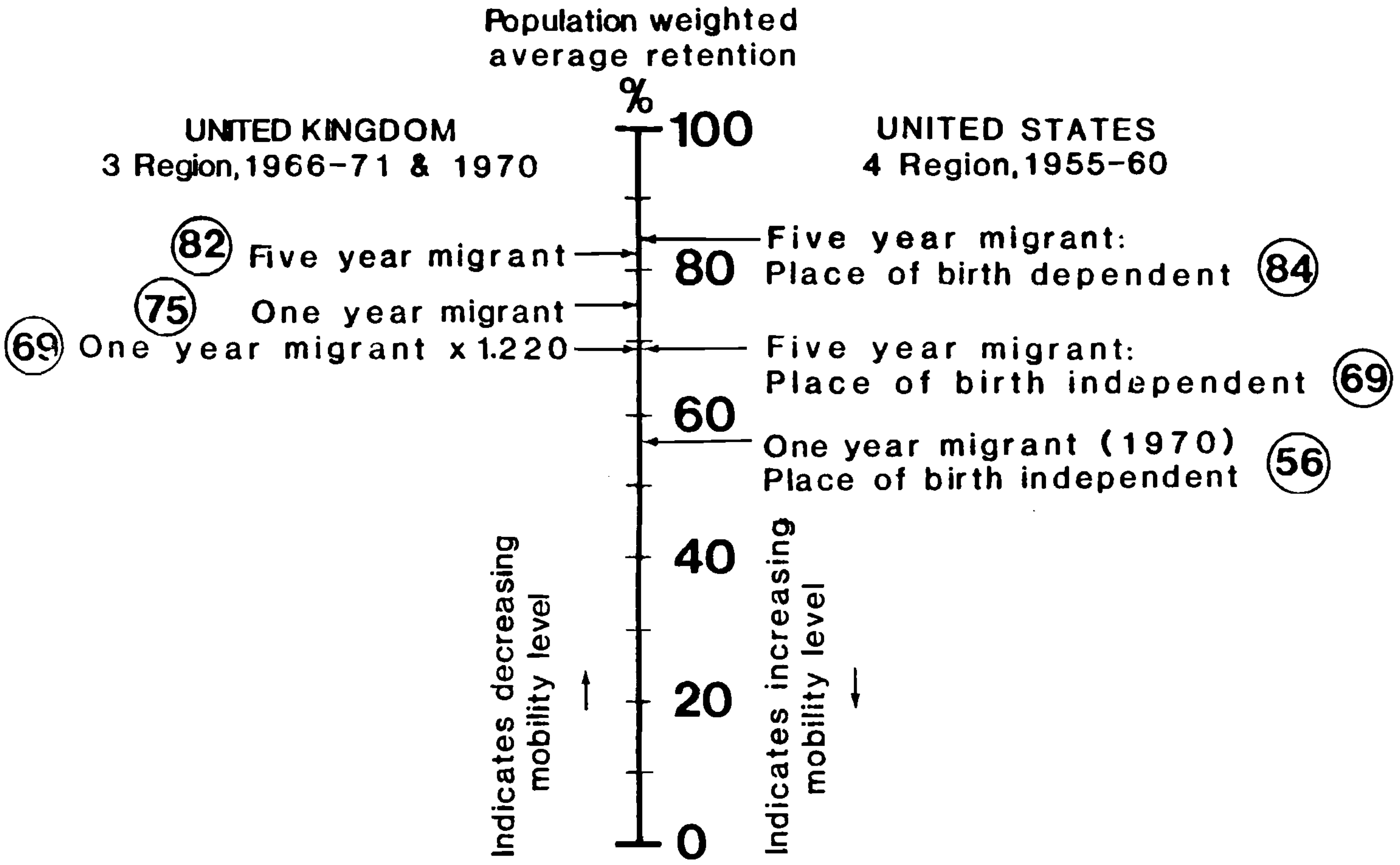


Figure 10. The effect of different migration data types on life table regional retention percentages (numbers in circles). (Source: United Kingdom computed from data in Table 11 and population figures given in Ledent and Rees 1980, Appendix A2.1. United States computed from Ledent 1981, p. 44, Table 5 using population figures in Rogers 1976b, p. 523 for 1958.

A comparison of the United Kingdom figures demonstrates the importance of having the same type of migration data and same period length if results ought to be comparable. Differences in retention percentages are made of differences due to

movement rather than transition concept
75 - 69 = 6 points

one-year rather than five-year period
82 - 75 = 7 points

The real difference of 13 percentage points between the United States and the United Kingdom migration propensity would have been masked if the United Kingdom one-year migrant data or the movement data were used.

Is there anything that could have been done in the CMS study to have avoided disturbing incomparabilities due to differences in the type of migration data and in period lengths? The answer must be "probably not" on two counts. Firstly, interregional migration statistics are expensive demographic statistics to collect and are usually collected and tabulated in only a few ways. The only countries in which there appears to have been any choice are the United Kingdom (one-year and five-year migration tables, Census 1971) and the Soviet Union (two-year migration tables, Census 1970; registration based movement tables for single years). For Hungary, movement data were used, although the 1970 census collected five-year transition data; these data have, however, not been tabulated. For Japan, 1970 one-year transition data were used; movement data were generated by a registration system, but these data were not tabulated for age categories (Nanjo 1981). In member countries of the European Communities, transition data (one- and two-year periods) are nowadays collected as part of labor force surveys; the migration data are, however, generally not tabulated. The second argument is that the nature of the differences and their effects have only become clear because the CMS study has been carried out.

We turn next to an examination in detail of the spatial frameworks adopted in the country studies, which we have so far neglected.

4. THE REGIONS USED

4.1. The Concept of "Region"

The concept of a "region" has been much argued about in the social sciences, particularly in geography (Grigg 1967; Haggett, Cliffe, and Frey 1977). Two views have been in conflict. The first sees countries being divided up into functional, "real" regions that organize human activity (commuting, trade, traffic, information flows, and control). The second view identifies regions by using classificatory principles of uniform grouping with a contiguity constraint addressed to the purpose at hand. The regions used for administrative or planning purposes may have the characteristics of either functional or formal regions or a mixture of both or of neither, being *ad hoc* products of historical evolution, usually associated with a collective regional consciousness.

Geographers have put forward strong arguments that the evolution of the population distribution within a country should be studied using functional urban regions as in Berry and Gillard (1977) and Hall and Hay (1980). Researchers build up their regions by adding to significant employment centers those areal units (operational taxonomic units) that have strong connections in terms of journey-to-work flows. The changes in population and employment in the system of functional urban regions are then studied, and in some instances (Kennett 1980) the components of population change (natural increase, net migration, gross migration flows) are studied.

There are, however, problems with adopting such regional units: some practical, some methodological. Firstly, in most countries functional urban regions are not adopted as the areal units for publication of the necessary population stock, vital, and migration statistics broken down by age and sex. To estimate the missing data items would be a major research task in itself. Secondly, there tend to be many functional urban regions within a country (70 in Sweden, 130 odd in Great Britain, 200 plus in the United States). As a result, the migration flows matrix would be extremely sparse and any analysis based on it would be rather unreliable. There would also be the problem of reworking the existing computer programs (Willekens and Rogers 1978) to cope with an order-of-magnitude increase in the number of regions to be handled.

Of course, demographic analysis of functional urban regions is perfectly possible on a single region or aggregated system basis (say, n systems of three regions consisting of the functional urban region, the rest of the country, and the rest of the world), as Long and Frey (forthcoming) have shown in their study of the United States. In general, such an analysis on a comparative basis, would pose substantial data problems.

As a result of these problems a decision was taken very early in the CMS study to leave the delineation of regions very much to the national collaborators. They were asked to decide upon a policy-relevant set of regions into which their country could be divided and for which most population data could be obtained without too much difficulty. The number of regions to be identified was constrained to be less than or equal to 12, which was felt to be a maximum for the purposes of the project (which was to a large extent a learning exercise for the participants) and to which the version of the computer program used was constrained. (For Poland, 13 regions were used and the program was executed on a larger computer in Warsaw.)

4.2. The Regions Chosen

What kinds of regions were chosen by authors of the CMS study and what problems of comparability do they pose?

Table 12 sets out the names and numbers of regions that exist and are used in each country. The regional sets are classified into coarse (less than six regions in the set), medium (six to 15 regions), and fine (more than 15 regions). Multiregional analyses were carried out at either the coarse or the medium scale or both; in some studies additional single-region analyses were carried out at the fine scale (for example, Koch and Gatzweiler 1980, section 2.2). The coarse scale consists of sets of regions which are aggregations of medium scale regions for which all the necessary data were available.

Table 12. The regions used in the CMS study.

	Scale of regions		
Country	Coarse	Medium	Fine
1. United Kingdom	2 Standard Regions & Remainder of Country[a]	10 Standard Regions[b]	18 Conurbations & Region Remainder 61 Counties & Regions
2. Finland	---	12 Läani (provinces)[b]	16 Economic Regions
3. Sweden	---	8 Regions[b]	24 Counties[c] 70 A-Regions
4. GDR	5 Regions[b]	15 Regions (districts)[cd]	219 Kreise (counties)
5. Netherlands	5 geographic Regions[b]	12 Provinces[cd]	40 COROP Regions 129 Economic Geographic Areas
6. Canada	---	10 Provinces[b]	
7. Hungary	---	6 Economic Planning Regions[b]	25 Counties & Country Towns[c]
8. Soviet Union	Urban and Rural Areas[a]	8 Units: 7 Urban Regions & 1 rural remainder[b]	15 Republics
9. FRG	---	11 Länder[b] (stages)	58 Republics
10. Austria	4 Länder aggregations[a]	9 Länder[b] (states)	95 Gemeinden
11. Poland	---	13 Regions[b]	22 Voivodships (to 1975), 49 voivodships (since 1975)[c]
12. Bulgaria	---	7 Regions[b]	28 Districts
13. France	---	8 Zeats[b]	22 Regions[c]
14. Czechoslovakia	---	12 Regions[b]	
15. Japan	---	8 Regions[b]	47 Prefectures
16. United States	4 Regions[a]	9 Census Divisions[b]	50 States
17. Italy	4 Statistical Regions[b]	---	20 Administrative Units

[a] Principal multiregional analyses carried out at this scale.

[b] Additional multiregional analyses carried out at this scale.

[c] Additional single region analyses carried out at this scale.

[d] Data provided in Research Report at this scale for multiregional analysis.

Notes for this Table are on the following page.

A majority of the regional sets used consisted of aggregations of administrative (governmental or planning) areas. In some cases (for example, Sweden, Japan) this aggregation was made by the authors; in others (for example, the United Kingdom, France) the aggregations were widely used for statistical reporting purposes by government agencies. In other studies the regions analyzed were used for planning or policy purposes. The Läani of Finland, the Provinces of Canada, the Economic Planning Regions of Hungary, the geographic regions of the Netherlands, the Länder of the Federal Republic of Germany, and the Länder of Austria all fall into this category.

The boundaries of the regions used in the principal multiregional analyses of the CMS studies are shown in Figures 11A and 11B, and the names are spelled out in full in Table 13. A few comments are in order about the nature and scale of some particular regions.

Notes for Table 12.

1. The United Kingdom region analysis covers 11 regions: the eight standard regions of England, plus Wales, Scotland, and Northern Ireland. In the multiregional analysis Northern Ireland was omitted. The three regions (coarse regionalization) are used in the United Kingdom chapter analysis and the Ledent and Rees (1980) study. The standard regions are aggregations for statistical purposes of the administrative counties.
2. The provinces are administrative units.
3. The regional units are amalgamations of counties (administrative units).
4. The multiregional analysis of the German Democratic Republic was carried out principally using five macroregions, though some analysis was done with 15 regions which were the 15 administrative districts of the German Democratic Republic (Bezirke). The macroregions were aggregations of the administrative districts.
5. The five regions are groups of the 12 administrative provinces.
6. The Canadian study omits the Yukon and North West Territories from the multiregional analysis. The provinces are administrative units.
7. The six regions are groupings of the 25 administrative districts.
8. The urban regions are not contiguous.
9. The Länder are administrative regions.
10. The four regions are groupings of the nine Austrian Länder.
11. The 13 Polish regions are groupings of the 49 (post-1975) administrative voivodships. Before 1975 there were 22 voivodships.
12. The seven Bulgarian regions are groupings of 28 administrative districts.
13. The "Zeats" are the "Zones d'Etude et d'Amenagement du Territoire," originally defined for the regionalization of the Sixth National Plan. They are groupings of the 22 programming regions.
14. Some eight of the regional units fall in the Czech Republic and four in the Slovak Republic.
15. The eight regions are aggregations of the 47 administrative prefectures.
16. The four regions are aggregations of the nine census divisions which are amalgamations of the 50 administrative states.
17. The four regions are amalgamations of the 20 administrative divisions.

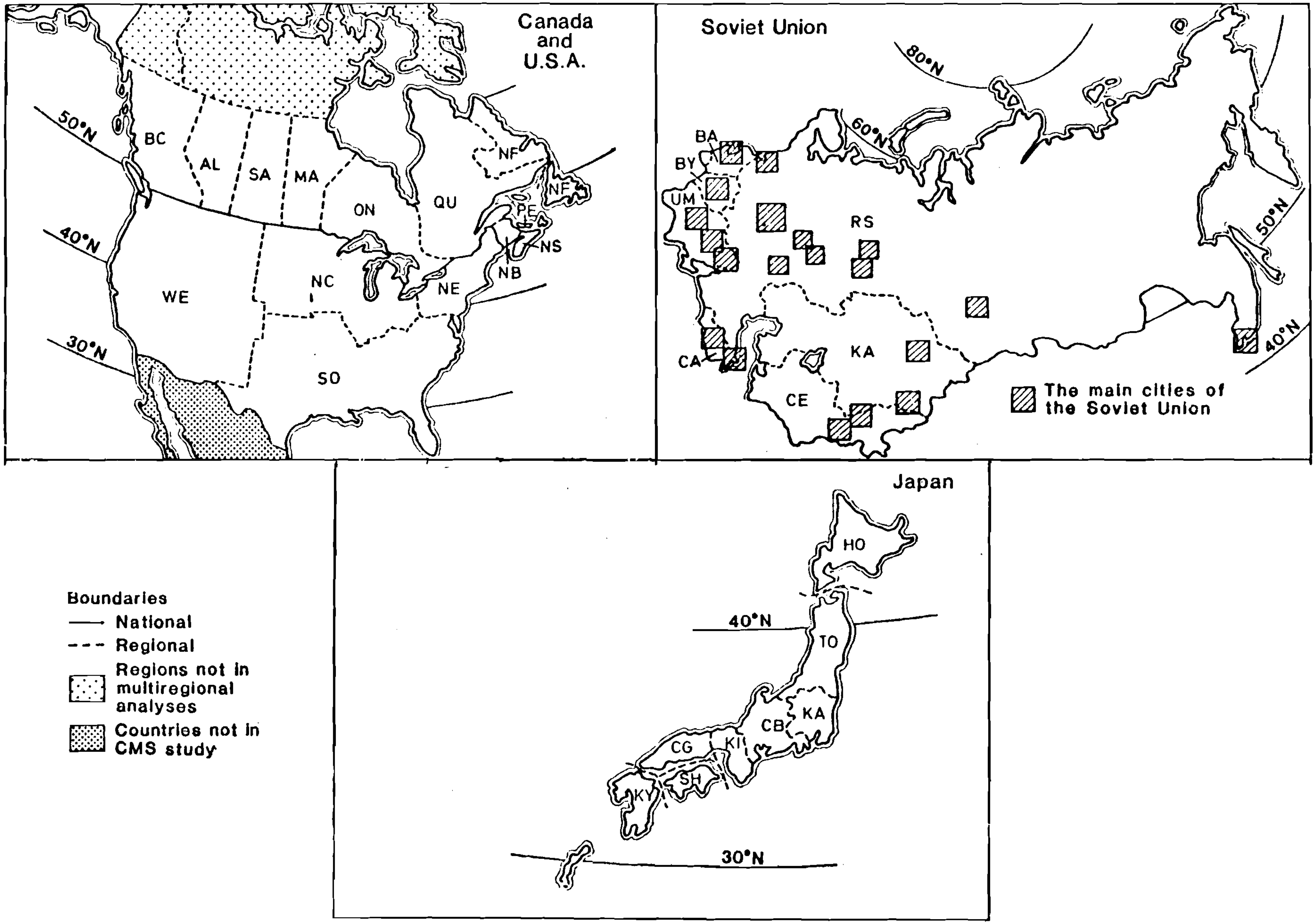


Figure 11A. The regions used in the CMS study: (A) North America, Soviet Union, and Japan.

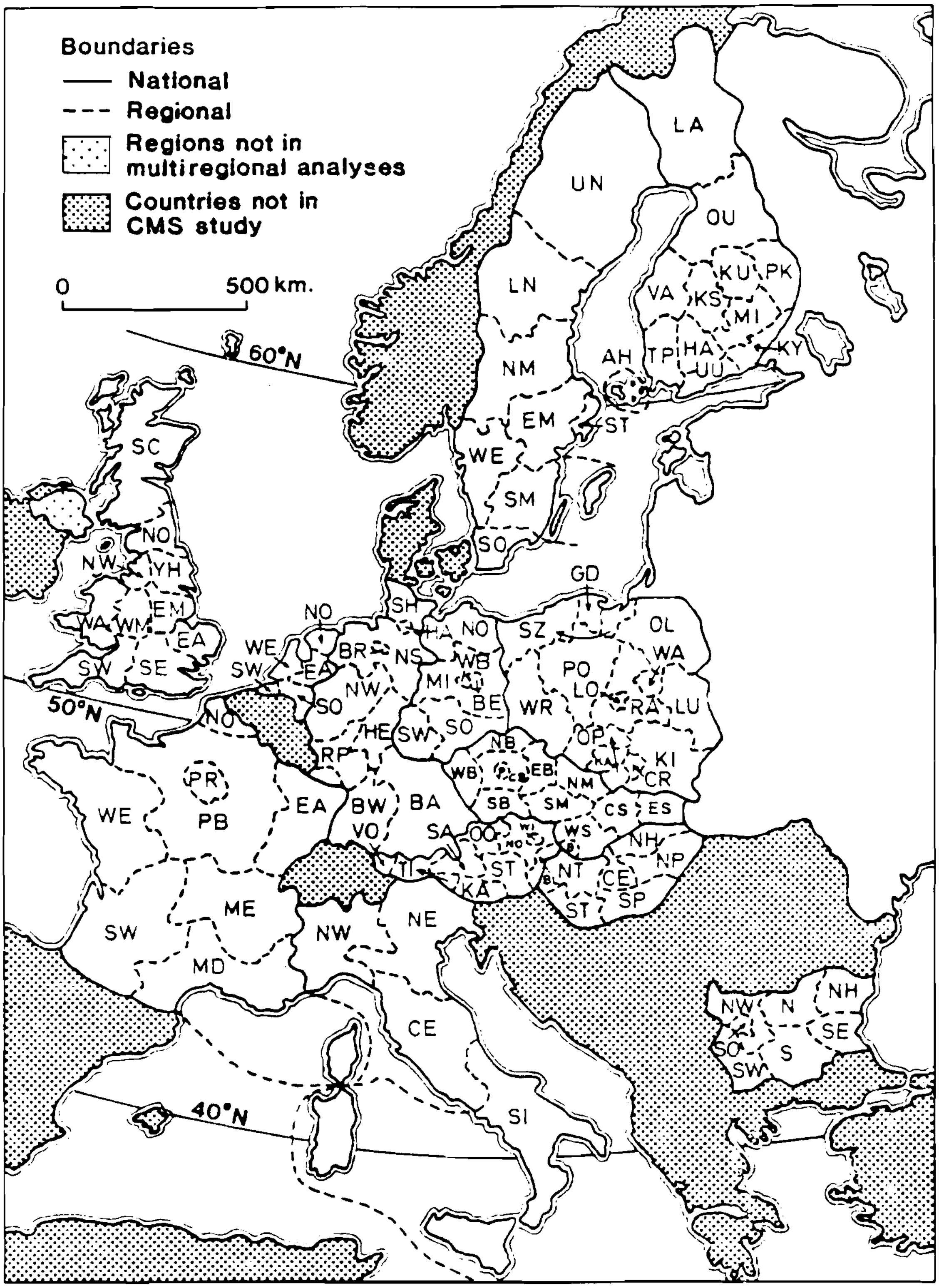


Figure 11B. The regions used in the CMS study: (B) Europe.

Table 13. Names of the regions used in the multiregional population analyses.

1.	United Kingdom
NO	North
YH	Yorkshire and Humberside
NW	North West
EM	East Midlands
WM	West Midlands
EA	East Anglia
SE	South East
SW	South West
WA	Wales
SC	Scotland

2.	Finland
UU	Uusimaa
TP	Turku and Pori
AH	Ahvenanmaa
HA	Häme
KY	Kymi
MI	Mikkeli
PK	Phojois-Karjala
KU	Kuopio
KS	Keski-Suomi
VA	Vaasa
OU	Oulu
LA	Lappi

3.	Sweden
UN	Upper North
LN	Lower North
NM	North-Middle
EM	East-Middle
ST	Stockholm
SM	South-Middle
WE	West
SO	South

4.	German Democractic Republic
NO	North
BE	Berlin
SW	South West
SO	South
MI	Middle

5.	Netherlands
NO	North
EA	East
WE	West
SW	South West
SO	South

6.	Canada
NF	Newfoundland
PE	Prince Edward Island
NS	Novia Scotia
NB	New Brunswick
QU	Quebec
ON	Ontario
MA	Manitoba
SA	Saskatchewan
AL	Alberta
BC	British Columbia

7.	Hungary
CE	Central
NH	North Hungary
NP	North Plain
SP	South Plain
NT	North Trans-Danubia
ST	South Trans-Danubia

8.	Soviet Union
RS	Urban areas of the RSFSR
UM	Urban areas of the Ukrainian and Moldavian SSRs
BY	Urban areas of the Byelorussian SSR
CE	Urban areas of the Central Asian Republics (Uzbek, Kirghiz, Tadzhik and Turkmen SSRs)
KA	Urban areas of the Kazakh SSR
CA	Urban areas of the Caucasian Republics (the Georgian, Azerbaijan and Armenian SSRs)
BA	Urban areas of the Baltic Republics (the Estonian, Latvian and Lithuanian SSRs)
RU	Rural areas of the USSR

9.	Federal Republic of Germany
SH	Schleswig-Holstein
HA	Hamburg
NS	Niedersachsen
BR	Bremen
NW	Nordrhein-Westfalen
HE	Hessen
RP	Rheinland-Pfalz
BW	Baden-Württemburg
BA	Bayern
SA	Saarland
WB	Berlin (West)

Table 13 continued.

10.	Austria
WI	Vienna
NO	Lower Austria
BU	Burgenland
KA	Carinthia
ST	Styria
OO	Upper Austria
SA	Salzburg
TI	Tyrol
VO	Vorarlberg

11.	Poland
WA	Warsaw
LO	Łódź
GD	Gdańsk
KA	Katowice
CR	Cracow
RA	Radom (East Central)
OL	Olsztyn (North East)
SZ	Szczecin (North West)
OP	Opole (South)
KI	Kielce (South East)
LU	Lublin (East)
PO	Poznan (West Central)
WR	Wroclaw (West)

12.	Bulgaria
NW	North West
NO	North
NE	North East
SW	South West
SH	South
SE	South East
SP	Sofia

13.	France
PR	Paris Region
PB	Paris Basin
NO	North
EA	East
WE	West
SW	Southwest
ME	Middle East
MD	Mediterranean

14.	Czechoslovakia
PR	Prague
CB	Central Bohemia
SB	South Bohemia
WB	West Bohemia
NB	North Bohemia
EB	East Bohemia
SM	South Moravia
NM	North Moravia
BR	Bratislava
WS	West Slovakia
CS	Central Slovakia
ES	East Slovakia

15.	Japan
HO	Hokkaido
TO	Tohoku
KA	Kanto
CB	Chubu
KI	Kinki
CG	Chugoku
SH	Shikoku
KY	Kyushu

16.	United States
NE	Northeast
NC	North Central
SO	South
WE	West

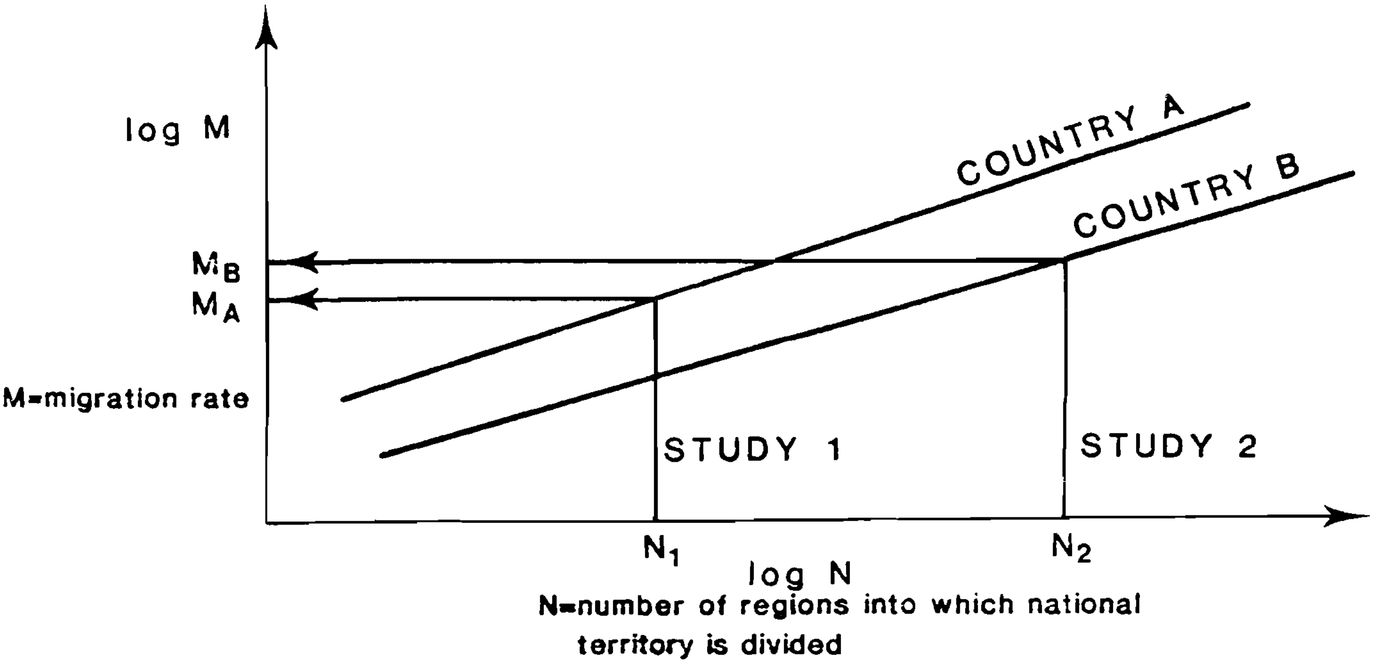
17.	Italy
NW	North West
NE	North East
CE	Central
SI	South and Islands

The acute observer will note that a couple of regionalizations fail to be exhaustive of the national territory. Northern Ireland is omitted from the United Kingdom multiregional analysis because of the lack of published data on migration to Northern Ireland from the regions of Great Britain. This deficiency could surely have been rectified. The second instance is the omission of the Yukon and North West Territories from the Canadian study on the grounds of their small population size, although Liaw (1977, 1979) does include them in his work.

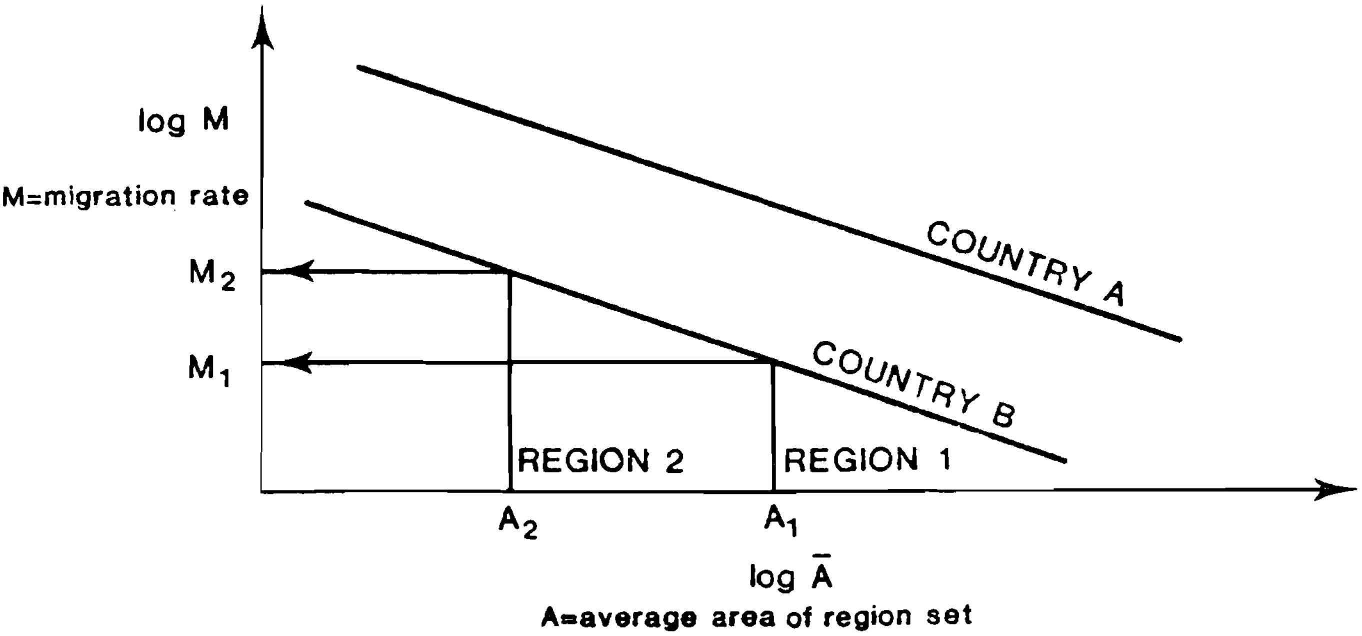
The varying scale of regions (both in area and in population size) gives rise to problems of comparability both within countries and between them. Courgeau (1973b) has studied the relationship between the level of migration (migration rate) and the number of units into which the national territory is divided (Figure 12A). He proposes a power function to express this relationship. A parallel proposition is given in Figure 12B. The migration level is an inverse power function of the average areal size. Country A's migration rates are consistently higher than those of country B when equivalent numbers of regions are taken into account. However, it might be quite feasible to observe a higher interregional migration level for country B if the number of units involved was sufficiently greater, as is the case when studies 1 and 2 are compared in Figure 12A.

A similar effect can apply when two regions are compared within a country. The smaller the region, other things being equal, the higher the observed outmigration rates. But the propensity of migrants to move over given distances may be exactly the same. Thus, the Länder of Bremen and Hamburg are much smaller than the other Länder of the Federal Republic of Germany, and the migration levels observed are much higher.

Of course, Bremen and Hamburg Länder are under bounded definitions of their respective functional urban regions (Figure 5, p. 13 in Koch and Gatzweiler 1980), and thus a major part of their higher than average migration levels is probably a result of the suburbanization process overlapping Länder boundaries.



A. Courgeau's relationship.



B. The inverse of Courgeau's relationship.

Figure 12. Migration rates as a function of the number and average size of measurement units.

However, the biggest contrast between and within sets of regions is the contrast between the regions of the Soviet Union and the other countries. The Soviet Union regions are collections of urban islands in a rural sea; only the rural remainder preserves the property of contiguity. Seven of the eight "regions" will be small in collective areal extent, and the eighth will be vast. Since the Soviet Union is currently in the process of rapid urbanization (Soboleva 1980, Rogers 1980), this particular regionalization is especially pertinent but difficult to compare with other regionalizations.

4.3. Aggregation Experiments

In several of the studies, multiregional analyses were carried out at more than one scale (United Kingdom 2, German Democratic Republic, Netherlands, Austria). These "shrinking" experiments show that aggregation before input to multiregional analysis gave very similar results to aggregation after analysis. Thus, the life expectancy in East Anglia of persons born in East Anglia was 29.80 years in the 10 region analysis and 29.72 in the three region analysis. These findings confirm those of Rogers (1976b) who showed that decomposition (tearing the regional system apart) had a much greater effect than aggregation.

4.4. Assessment

It is clear from the above discussion that there are considerable problems in comparing the regional patterns of mortality, fertility, and migration in the CMS countries. It is also clear that no uniform criteria or regionalization rules are yet available to deal with the problem of the dependence of migration levels on regional disaggregation. Regional population dynamics must be viewed as a variety or spectrum of scales (compare the conclusions of an equivalent debate on the segregation of ethnic groups in cities in Poole and Boal 1973 and Rees 1979b, pp. 297-302). Table 13 tells us that this is what many of the study's authors did intuitively: studying population and migration patterns using coarse, medium, and fine levels of resolution. However, as we show in the next section, the difficulties in preparing the data for multiregional population research generally restrict analysis to the medium scale.

5. THE ESTIMATION OF MODEL INPUTS

5.1. Principles

The first steps in the CMS study were to estimate for each country the necessary input data on population stocks, births, deaths, and migration flows for the regions chosen (see Figure 1). For one reason or another the data gathered and published by national statistical offices are never quite in the correct form for analysis. Official data must often be adjusted or disaggregated to yield variables for input to population models. Sometimes this work of estimation is needed because the data are missing; more often, estimations are carried out because the work of retabulation of the original census or register files is either prohibitively costly or prohibitively lengthy in time.

Some five principles should govern the estimation process.

(1) The data should match the concept(s) demanded of the variables input to the population model.
(2) The data should be corrected for known under- or over-enumeration.
(3) The time points or periods of data collection should match those of the model.
(4) The data should apply to regions used in the model.
(5) The age disaggregation of the input data should match that demanded in the model.

Application of these principles leads to estimation involving *concept adjustment*, *correction for under- or over-enumeration*, *temporal adjustment*, *regional admustment*, and *age disaggregation*.

5.2. The Estimation of Population Stocks

Population stocks can be measured in several ways. The population can be enumerated in a census or counted on a register, or estimated using accounting principles from previous counts, from demographic events and flows counts, and from other correlated data series such as electoral registers and housing statistics (OPCS 1980b). In fact, the classification of country studies by source of population stock data in Table 14 is the same as that for migration statistics in Figure 4 except in the case of the

Table 14. Population stock sources, temporal adjustments, and age breakdowns.

Country	Source Census Temporal adjustment Needed Carried out	 Not carried out	 Not needed	 Register Temporal adjustment Needed Carried out	 Not carried out	 Not needed	 Estimate Temporal adjustment Needed Carried out	 Not carried out	 Not needed	Last age group in raw regional data
United Kingdom 1									X	85+
United Kingdom 2	X									75+
Finland				X						75+
Sweden						X				100+
GDR				X						75+
Netherlands					X					85+
Canada	X									85+
Hungary					X					85+
Soviet Union					X					70+
FRG					X					85+
Austria		X								85+
Poland					X					80+
Bulgaria				X						60+
France		X								95+
Czechoslovakia										85+
Japan										85+
United States										85+
Italy										85+

Notes: GDR—German Democratic Republic, FRG—Federal Republic of Germany.

United Kingdom 1 study in which an estimate is used. Indirect population estimates are not needed in countries with population registers.

Concept adjustment is occasionally necessary for population stocks. Census populations are tabulated in one of two ways: as either the enumerated (*de facto*) population actually present in a household on census night or the "usually resident" (*de jure*) population counted by place of usual residence, census night visitors being transferred back to households of usual residence. In the United Kingdom censuses of 1966 and 1971 the main region-by-age-by-sex population stock tables are for enumerated populations, and the detailed group numbers must be adjusted to a "usually resident" total,

Population stocks can also be classified differently. In England and Wales, for instance, a distinction is made between "home," "total," or "civilian" population. The home population is that population actually resident in England and Wales distributed by area. The total population consists of the home population *plus* members of national armed forces serving abroad but minus the forces of other countries temporarily stationed in the country. Only the home population can be used for regional analysis but the total population may be employed in the national analysis. In that event, regional populations need not necessarily aggregate to the national.

Correction for under- or over-enumeration may be applied to census or register based population stocks data if a measure of likely error is available in the form of a post-census survey for census derived populations or in the form of a census for register derived populations. Correction for a 1.4 percent under-enumeration was needed for the 1966 census based on populations in the United Kingdom 2 study. Andersson and Holmberg (1980) include a discussion of errors associated with register counts.

The time frameworks for population stocks has been discussed already in section 3.2 (see Figure 7). Table 14 notes whether or not the necessary adjustments have been made.

Regional adjustment is rarely necessary for study periods of only one year and in statistical systems in which the building blocks (departments in France, states and counties in the United States, Länder in Austria and the Federal Republic of Germany) remain fixed over long periods of time. Where radical changes have taken place in the local government system of units and boundaries (as in Poland or the United Kingdom) the reconstruction of time series of spatially compatible time series of demographic data is an involved and error-prone process. Every time a spatial revision takes place, the timer series must be adjusted. There was some suspicion that the resort to demographic cartograms in the United Kingdom study was an attempt to gloss over this problem.

The need for further age disaggregation of population stock data is rare. In order to make reasonable life expectancy estimates, the population age breakdown (by five-year age groups) should extend to at least 85 years. Where the age breakdown of regional populations was insufficient, deconsolidation proportions were applied to disaggregate an age group such as "60+" into quinquennial age groups "60-64, 65-69,...,80-84, 85+." The deconsolidation proportions were derived either from national population data (as in the Bulgarian study), from equivalent regions within the country, or from other countries (as in the Soviet Union study where deconsolidation proportions from Poland were used).

A particular age grouping is used for the German Democratic Republic. For planning purposes, age groups 0-1, 1-3, 3-6, 6-10, 10-15, 15-18, 18-21, 21-25, followed by the five-year age groups 25-30, 30-35, and so on, are considered in statistical publications. For the CMS study, the age grouping was rearranged into five-year age groups on this basis of the age composition at the national level, which was available in single years of age.

5.3. The Estimation of Deaths and Births

Concept adjustment is not normally required for deaths data since death registrations are tabulated as a matter of course by area of usual residence. Under-enumeration in developed countries is not regarded as serious and the existence of combined birth and death registration systems ensures accuracy of the reported information.

Adjustment of the deaths data to the relevant time period is occasionally needed if an intercensal period is used to match the migration data time span. Annual counts of deaths for regions can be adjusted by applying temporal fractions based on quarterly or monthly national data. Termote (1980) describes the procedures involved in the Canadian study.

Regional adjustment was not required in any of the studies (apart from United Kingdom 1). In the absence of a revised statistical series (as in OPCS's *Population Trends* journal) based on the spatial reassignment of deaths records, knowledge of the population transfers that have resulted can be used to effect an estimate of "new" region deaths.

It is usually possible to estimate the contents of the following population matrix

$$\underset{\sim}{P} = \begin{bmatrix} p^{11} & p^{12} & \dots & p^{1N} \\ p^{21} & p^{22} & \dots & p^{2n} \\ \vdots & \vdots & & \vdots \\ p^{N1} & p^{N2} & \dots & p^{NN} \end{bmatrix} \qquad (12)$$

where p^{ij} is the population at a point in time t classified by "old" region i and "new" region j. The row totals of this matrix $p^{i\cdot}$, are the time t "old" region populations; if we divide each element of the $\underset{\sim}{P}$ matrix by its row total we obtain a matrix of redistribution coefficients, r^{ij}, arranged in a similar matrix, $\underset{\sim}{R}$. We can now postmultiply row vectors of deaths by the matrix of redistribution coefficients

$$\underset{\sim}{D}_x(\text{old})\ \underset{\sim}{R} = \underset{\sim}{D}_x(\text{new}) \qquad (13)$$

to yield an estimate of deaths distributed according to the new regional classification. This technique was occasionally employed to build consistent time series of the components of growth in the United Kingdom study (Rees 1979a, 1979c).

Occasionally, the age disaggregation of deaths data was insufficiently detailed for model input, as in the Bulgarian and United Kingdom studies. The age groups for which deaths data are available and are needed are shown in Table 15. Since a detailed age composition of deaths was available at the national level, broad age classes were disaggregated using national data. In the case of Bulgaria, disaggregation is done proportionally to the national age structure of deaths within the broad age classes. For the United Kingdom, deaths in broad age classes are disaggregated using national death *rates* for the age groups that compose the age classes. The procedure is described by Rees (1980). It implies that the *shape* of the national mortality curve within an age class is imposed onto the region, whereas the *level* of the curve is determined by the number of deaths in the age class in the region.

Table 15. The age groups for which deaths data are available and needed, United Kingdom and Bulgaria.

National Age Groups	Regional Age Groups		Model Age Groups
	United Kingdom	Bulgaria	
< 1	< 1	< 1	0 - 4
1 - 4	1 - 4	1 - 4	
5 - 9	5 - 14	5 - 9	5 - 9
10 - 14	15 - 24	10 - 14	10 - 14
15 - 19	15 - 24	15 - 19	15 - 19
20 - 24		20 - 29	20 - 24
25 - 29	25 - 34		25 - 29
30 - 34		30 - 39	30 - 34
35 - 39	35 - 44		35 - 39
40 - 44		40 - 49	40 - 44
45 - 49	45 - 54		45 - 49
50 - 54		50 - 59	50 - 54
55 - 59	55 - 64		55 - 59
60 - 64		60 - 69	60 - 64
65 - 69	65 - 74		65 - 69
70 - 74		70 +	70 - 74
75 - 79	75 +		75 - 79
80 - 84			80 - 84
85 +			85 +

Let the age group be x to x + n, and the age class, containing x to x + n, be w to w + zn. Let D_w^i be the number of deaths in region i and age class w to w + zn; P_x^i is the population in region i and age group x to x + n; d_x^n is the age-specific national death rate. The number of deaths in region i and age group x to x + n is

$$D_x^i = D_w^i \frac{d_x^n P_x^i}{\sum\limits_{x \in w} d_x^n P_x^i}$$

In one instance age classified data had to be aggregated: both the United Kingdom and the Bulgarian regional death tables distinguish between deaths under one year of age and deaths at ages one to four. This disaggregation has long been recognized as necessary in "abridged" single-region life tables although in most developed countries, it had relatively little effect on average life expectancies [Rees (1979a, p. 51) suggests that adopting the finer disaggregation lowers the life expectancies for British regions by about three weeks].

Little needs to be added to the above discussion when births data are considered. The only estimation problem concerned the meaning of the age classification in relation to the population model, and this we have discussed earlier in section 3.1

5.4. The Estimation of Migration Flows: Minor Problems

We concentrate in this section on the problems of estimating migration flows sufficiently classified by age for input to the multiregional population models, after first referring to the problems of concept adjustment, under-enumeration, and regional adjustment.

The various migration concepts used in the CMS countries have been discussed in some detail in preceding sections. The conceptual differences between movement and transition data and between transition data over short and long time intervals, have been noted. Choice between concepts was occasionally available to CMS authors (United Kingdom), but in general insufficient information was available to transform one type of migration data into another, even if a model of the process could be proposed (Kitsul and Philipov 1980).

These major conceptual differences were not the only problems faced in the CMS study, however. In the Hungarian case study, the authors were faced with two sets of migration statistics, those described as temporary and those described as permanent. Each person can have a temporary place of residence in addition to his/her permanent residence. Although a migration is normally defined as a change of permanent residence, in countries or regions where restrictions are placed on such changes in permanent residence—usually because some attempt is being made to limit the growth of the largest metropolis in a country (e.g., Budapest)—temporary migrations (visits) may take on a rather permanent character. This has happened to a major degree in Hungary and Bies and Tekse (1980) therefore add temporary and permanent migration together before using them as inputs to their multiregional population analysis. This raises regional migration levels in Hungary above those observed in other Eastern European countries in the CMS study.

Under-enumeration and misreporting are always a potential problem in the collection of migration statistics. In censuses the most mobile of the population are the most likely to be missed. Retrospective census questions upon which the migration statistics rely depend on accurate recall by the respondent and accurate classification by census clerical staff. The usual assumption that researchers make is that the errors will tend to cancel out. Registration systems avoid most of these difficulties but not entirely as Andersson and Holmberg's (1980) discussion of the Swedish system reveals. An additional problem associated with many censuses is that the migration question is only asked or tabulated for a sample (10, 25, 50 percent) of the population. It might be interesting to run the analyses in such cases using the upper and lower confidence limit migration flow values to assess whether results based on mean values were the average of a narrow or a wide band of alternatives.

Regional adjustment of migration flows, when required by changes in regional boundaries, is peculiarly difficult to effect unless the statistical offices reanalyze the individual migration flow data. Equations (12) and (13) do not give very good results for migration flow tables. Instead migration model based techniques (Stillwell 1978) must be used.

5.5. The Estimation of Migration Flows: The Problem of Age

Estimation of missing migration flows by quinquennial age groups proved to be the greatest problem in preparing the national data sets for input to the Willekens and Rogers (1978) programs. This involved substantial research into the application and further development of techniques, developed in regional science and transportation science to infer spatial interaction flows from incomplete data (Willekens 1977, Willekens, Pór, and Raquillet 1979). Recently, the estimation methods have been simplified and extended to allow for the combination of various sources of prior information in order to come up with the best estimates possible (Willekens 1981). The review presented in this section, draws on this research; the numerical results shown are, however, identical to those obtained by the entropy maximization technique proposed in Willekens, Pór, and Raquillet (1979).

A major feature of the estimation method is its focus on the structure of the whole data set to predict values of missing elements. The structural representation of migration data is provided by accounts. The strategy to predict (or estimate) missing cell values in migration tables consists of five stages (Willekens 1981, p. 2).

(1) Set up the accounting frames. (The account is a multidimensional contingency table.)

(2) Develop a model of the data in the accounts. (Although various models of data structures are available, the parametric log-linear model has been found most appropriate for our purpose since its parameters denote particular effects of interaction between the cross-classified variables.)

(3) Enter the available data into the account. (Fill the account as far as possible and list other prior information separately.)

(4) Determine the parameter-values of the parametric model on the basis of the different types of prior information, supplemented by hypotheses about certain structural relationships in the data to be estimated.

(5) Apply the model to infer the values of the missing elements.

Step 4 may be skipped, i.e., the missing elements may be predicted directly from the available data without explicitly estimating the model parameters. This shortcut will be followed in this section. The steps are now discussed in some detail.

5.5.1. Accounting Framework

Multiregional population analysis requires migration flows by age and by region of origin and of destination. The required data may be arranged to constitute a three-dimensional account: region of origin, region of destination, and age. Age-specific accounts are shown in Tables 9 and 10. Only the first quadrants of the movement and transition accounts are of interest since they contain all the required data on internal migration. Consider the movement account. The information on age-specific migrations may be arranged in layers of two-way tables (Table 16). Note that the diagonal elements represent the number of movements within the region and therefore differ from the R_x^i elements in Table 10.

5.5.2. Model of Data in the Account

The investigation of large data sets becomes relatively simple by fitting models to the data. During the past decade, analytical techniques have been developed for structural analysis of multidimensional contingency tables (see, for example, Bishop, Fienberg, and Holland 1975). These techniques, which were originally designed to identify patterns of association among several cross-classified categorical variables, may fruitfully be applied for estimating missing cell values in the

Table 16. Account of movements for age group x to x + n.

	Destinations				
Origins	1	2	·······	N	Total
1	M_x^{11}	M_x^{12}	·······	M_x^{1N}	$M_x^{1\cdot}$
2	M^{21}	M_x^{22}	·······	M_x^{2N}	$M_x^{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
N	M_x^{N2}	M_x^{N2}	·······	M_x^{NN}	$M_x^{N\cdot}$
Total	$M_x^{\cdot 1}$	$M_x^{\cdot 2}$	·······	$M_x^{\cdot N}$	$M_x^{\cdot\cdot}$

contingency table or account. In fact, the problem of estimating cell values in a multidimensional account is equivalent to the problem of quantifying appropriate interaction effects (hypothesis testing)(Willekens 1981). This can easily be seen by appropriately modeling the data. The model is given in the first equation in Figure 13. It is a multiplicative or log-linear model. It has eight terms, the number of terms depending on the dimension of the account or the number of cross-classified variables (in this case three: A denotes region of origin, B region of destination, and C age group). Each term represents a particular structural effect on the cell values M_x^{ij}.

According to the model, the expected cell count is the product of various effects. The overall effect, w, is a size effect; it is the geometric mean of all expected cell counts (Figure 13). The main effects denote the effects on M_x^{ij} of relative size differences between the various univariate marginals. For instance, w_x^C is the effect of the average age composition of the migrants on the number of M_x^{ij}. When all else is equal, large age groups result in large migration flows. The age-effect is the ratio

---

Model $M^{ij}_{x} = w\, w^{A}_{i}\, w^{B}_{j}\, w^{C}_{x}\, w^{AB}_{ij}\, w^{AC}_{ix}\, w^{BC}_{jx}\, w^{ABC}_{ijx}$

Overall mean effect $w = \left[\prod_{i,j,x} M^{ij}_{x}\right]^{1/RCL}$

Main effects $w^{A}_{i} = \frac{1}{w}\left[\prod_{j,x} M^{ij}_{x}\right]^{1/CL}$

w^{B}_{j} and w^{C}_{x} : analogous

First-order interaction effects (two-way or pairwise interaction)

$$w^{AB}_{ij} = \frac{1}{z}\left[\prod_{x} M^{ij}_{x}\right]^{1/L}$$

with $z = w\, w^{A}_{i}\, w^{B}_{j}\, w^{C}_{x}$

w^{AC}_{ix} and w^{BC}_{jx} : analogous

Second-order interaction effect (three-way interaction)

$$w^{ABC}_{ijx} = \frac{1}{z'} M^{ij}_{x}$$

with $z' = w\, w^{A}_{i}\, w^{B}_{j}\, w^{C}_{x}\, w^{AB}_{ij}\, w^{AC}_{ix}\, w^{BC}_{jx}$

Constraints $\prod_{i} w^{A}_{i} = \prod_{j} w^{B}_{j} = \prod_{x} w^{C}_{x} = 1$

$$\prod_{i} w^{AB}_{ij} = \prod_{j} w^{AB}_{ij} = \prod_{i} w^{AC}_{ix} = \prod_{x} w^{AC}_{ix} = \prod_{j} w^{BC}_{jx} = \prod_{x} w^{BC}_{jx} = 1$$

$$\prod_{i} w^{ABC}_{ijx} = \prod_{j} w^{ABC}_{ijx} = \prod_{x} w^{ABC}_{ijx} = 1$$

R : number of rows
C : number of columns
L : number of layers

Figure 13. Multiplicative formulation of log-linear model.

between the geometric mean of the xth layer and the overall geometric mean. The term w_{ix}^{AC} represents the interaction effect between age and origin. Note that the pattern of interaction denoted by this term is the average interaction over all AC-tables (i.e., for all possible destinations or j-values of B). The pattern may differ for each level of B, which results in values of w_{ijx}^{ABC} different from unity.

By introducing this multiplicative model, we have transformed the problem of predicting the M_x^{ij} values into a problem of estimating the parameters w_x^{ij}, i.e., of quantifying the interaction effects. The parameters may be derived, and hence the migration flows may be determined from the available data.

5.5.3. Enter the Available Data into the Accounts

In countries where migration information is not abundantly available, the existing data are generally limited to aggregate information about the flows. We might, for example, know the total migration flows between origins and destinations (that is, aggregated over age); the total outmigration flows from origins by age group (that is, aggregated over destinations); or the total inmigration flows to destinations by age group (that is, aggregated over origins). Each of the known items is a two-dimensional array or matrix, which is a marginal (bivariate) total of the three-dimensional account, and may be entered at the appropriate place $(M_{\bullet}^{ij}, M_x^{i\cdot}, M_x^{\cdot j})$. The estimation problem, for which three bivariate marginal totals are given, is referred to by Willekens, Pór, and Raquillet (1979) as the "three faces (3F)" problem, since the available data can be imagined as constituting the three faces of a cube (Figure 14).

Although in the CMS study the 3F-problem is common, migration flows may be predicted from less data. In a two face (2F) problem, only two of the three faces (bivariate marginal totals) are given. There are three varieties of the 2F-problem. In a one face and one edge (1FE) problem, we would only know the contents of some face together with an edge. The edges represent univariate marginal totals $(M_x^{\cdot\cdot}, M_{\bullet}^{i\cdot}, M_{\bullet}^{\cdot j})$; hence there are three variates of the 1FE-problem. The data limitation is most severe

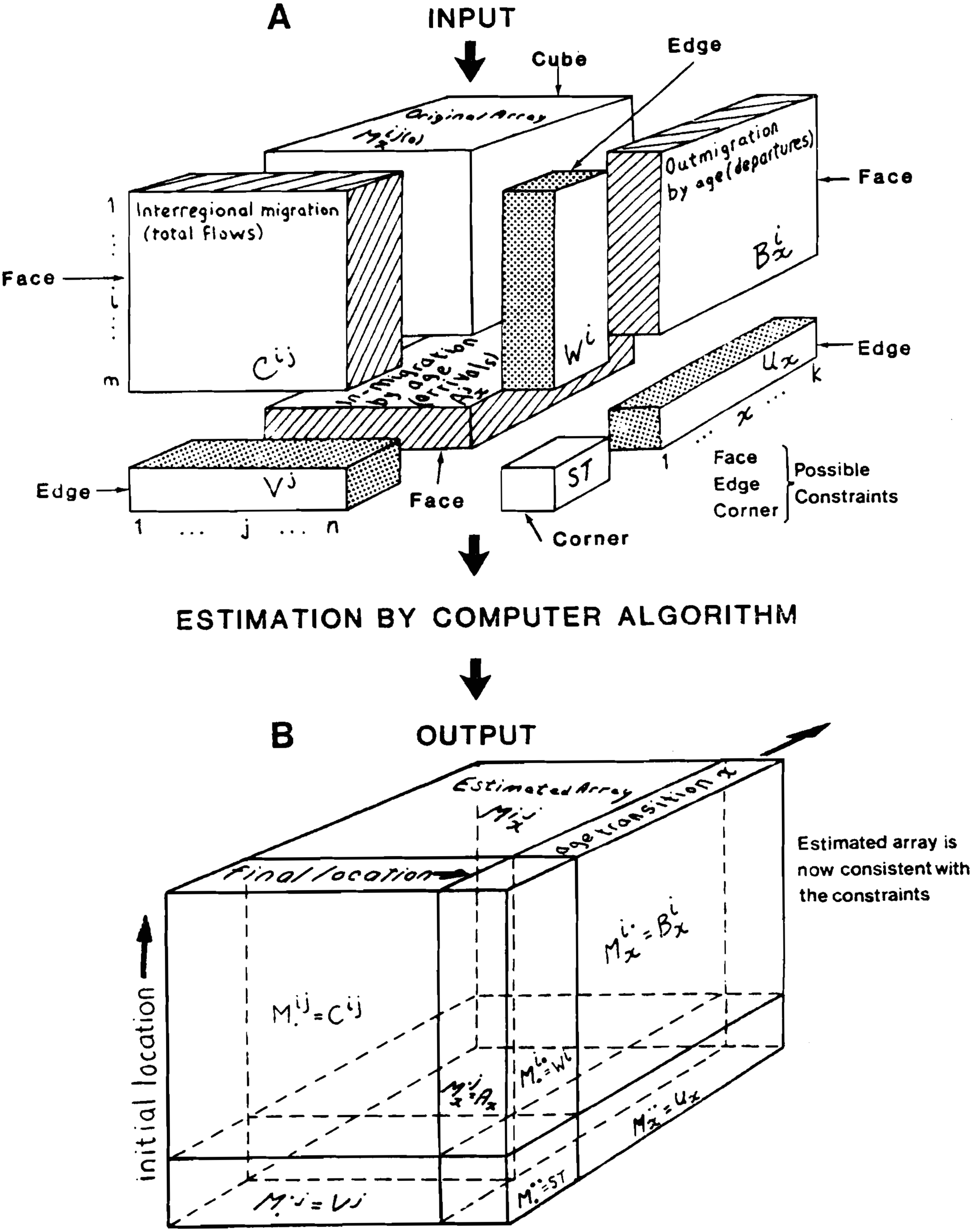


Figure 14. A diagramatic representation of the migration estimation problem. (Source: Adapted from Willekens 1977, Figure 3.)

when only the three edges are known (three edge or 3E-problem). In the 3E-problem, the available data are limited to a single age composition of the migrants, a vector of departures by region, and a vector of arrivals by region.

The 3F-, 2F-, 1FE-, and 3E-problems represent various situations of data availability. Data limitations that are experienced in a few countries participating in the CMS study, can be related to the four types. Table 17 notes what kind of problems had to be solved in each country's study. A full 3F-problem had to be solved for Bulgaria and the Netherlands. The estimation procedure for the Netherlands is described by Drewe and Willekens (1980). In the Finland, Sweden, Federal Republic of Germany, and Poland studies, the necessary migration flows data disaggregated by five-year age groups were available and no further estimation was required.

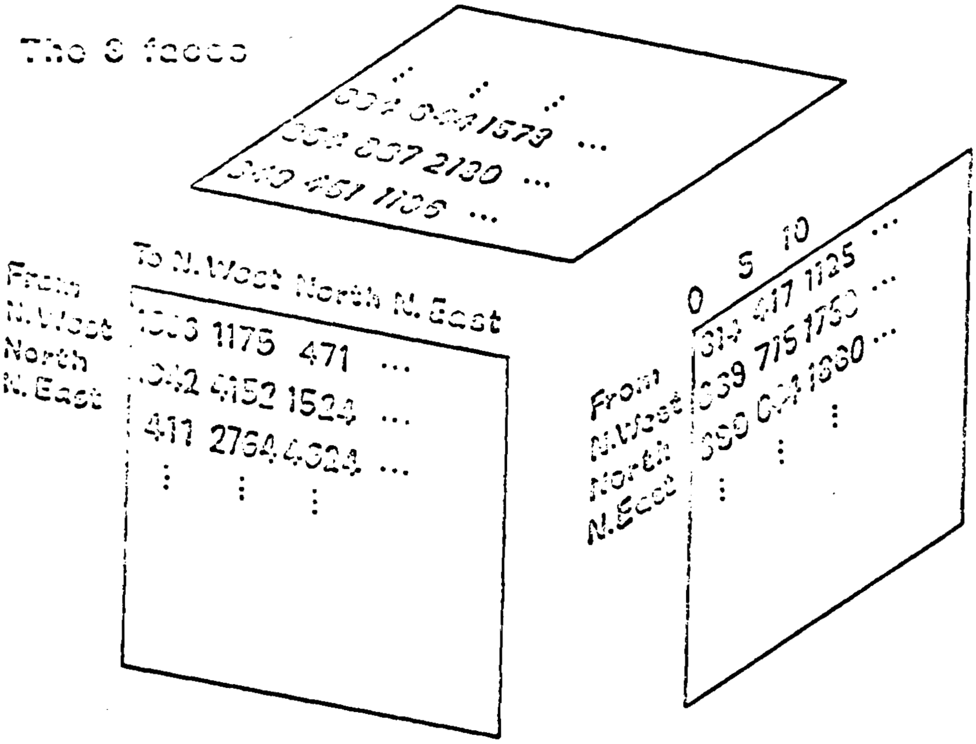
In the Canada, Austria, and France studies, migration flows were classified by five-year age groups, but conversion of the classification to the appropriate age-time framework had to be carried out (as discussed in section 3). In the Netherlands and Bulgaria studies full 3F-problems had to be solved. Figure 15 shows a portion of the Bulgarian problem and its solution. Between these two extremes were situations in which the age breakdown was by 10 or 15-year age groups for interregional migration flows (United Kingdom 1 and 2). In other cases the only age classified data available were not for quinquennial ages (German Democratic Republic) or were not sufficiently disaggregated at the older ages (Japan, United Kingdom, Soviet Union, Bulgaria). In these cases model migration schedules derived from the work of Rogers, Raquillet, and Castro (1978) or other interpolative or extrapolative techniques had to be used.

5.5.4. Predict Missing Migration Flows

The migration flows may be predicted directly from the marginal totals without first estimating the model parameters. The estimation procedure starts out with a set of preliminary estimates (guesses) of the unknown migration flows. To design a set of appropriate preliminary estimates, prior information on

Table 17. Nature of the migration age estimation problems solved.

Country	Description of the problem solved
United Kingdom 1	Reduced 1FE-problem for age groups 5-14, 35-44, 45-59, 65+ (ages are age at last birthday). Age conversion for transition data.
United Kingdom 2	Reduced 3F-problem for age groups 5-14, 35-44, 45-59, 65+ (ages are ages at last birthday). Age conversion of transition data.
Finland	No problems to be solved.
Sweden	No problems to be solved.
German Democratic Republic	Model migration schedules used to redistribute migrants from age groups 1, 1-3, 3-6, 6-10, 15-18, 18-21, 21-25, to 0-5, 5-10 and 15-20, 20-25 (ages are end points of intervals).
Netherlands	Full 3F-problem.
Canada	Age conversion for transition data. Infant migrants from birth place tables.
Hungary	Reduced 1FE-problem for age groups 40-49, 50-59, and 60+ (ages at last birthday) and a full 3F-problem.
Soviet Union	Further disaggregation of 70+ age group using Polish schedules.
Federal Repbulic of Germany	No problems to be solved.
Austria	Age conversion for transition data. Missing infant migrants.
Poland	No problems to be solved.
Bulgaria	Reduced 1FE-problem for 70+ age group. Full 3F-problem.
France	Age conversion for transition data.
Czechoslovakia	3F-problem but with some additional constraints.
Japan	Interpolation from age groups 35-44, 45-54, 55-64, and 65+ to five-year age groups 35-39 to 85+ (1FE-problem).
United States	Age conversion for transition data.
Italy	Extended 3E-problem (work in progress).



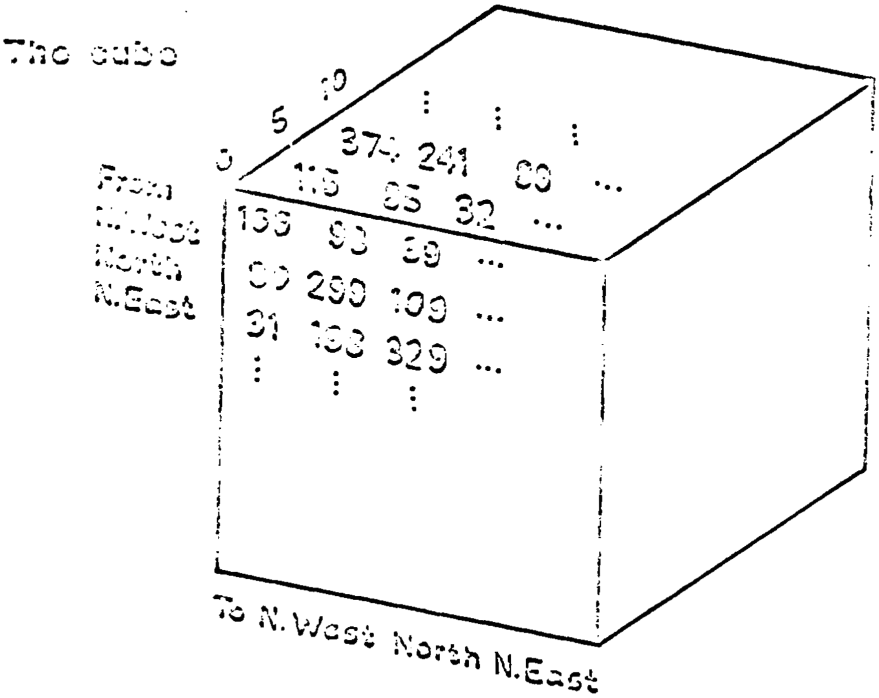


Figure 15. The three faces problem for Bulgaria illustrated.

migration, other than the marginal totals, may be used. For instance, we may disregard intraregional migration and therefore force the diagonal elements of the migration matrices by age to be zero. Information on locational preferences may be introduced in the set of off-diagonal preliminary estimates. How this may be done is illustrated in Willekens (1981). In the CMS study we did not have additional information on migration that would warrant special attention given to the design of a set of preliminary estimates. Therefore, the estimation procedure was started from a uniform distribution:

$$^{o}M_{x}^{ij} = 1 \qquad\qquad \text{for all } i,j,x$$

where $^{o}M_{x}^{ij}$ denotes the preliminary estimate of the $^{o}M_{x}^{ij}$ element. In most cases, we could deduct the intraregional migration flows from the given marginal totals and hence limit the estimation problem to interregional migration. (Diagonal elements were forced to be zero: $^{o}M_{x}^{ii} = 0$ for all i,x.)

The estimation technique that was applied is an extension of the entropy method, widely used in transportation science and geography to infer spatial interaction flows from aggregate data. The method is fully described and illustrated in Willekens, Pór, and Raquillet (1979), where also the computer program (written in FORTRAN) is given. The algorithm is an iterative procedure. However, it can be shown that in three of the four problem types, the iteration takes only one step, i.e., the estimation problem has a closed-form solution. Table 18 shows the solutions to the various types of estimation problems in three-dimensional accounts. Also given is the associated log-linear model and its statistical interpretation. The model assumes a uniform distribution of preliminary estimates and the final estimates are maximum-likelihood estimates. Apart from the trival case (CI) when all M_{x}^{ij}-flows are known, higher order interaction terms are equal to unity. The explanation is simple. Since in the uniform distribution (set of preliminary estimates) all interaction affects are absent, the only interaction effects contained in the final estimates of migration flows are those exhibited by the given marginal totals.

Table 18. Maximum-likelihood estimates in three-dimensional accounts.

Case	Available Data	Log-Linear Model	Closed-form Cell Estimates	Interpretation of Log-Linear Model
3E	{A, B, C}	$M_x^{ij} = w\, w_i^A\, w_j^B\, w_x^C$	$\hat{M}_x^{ij} = \frac{1}{N^2} M_\bullet^{i\bullet}\, M_\bullet^{\bullet j}\, M_x^{\bullet\bullet}$	-Mutual independence (variables A, B, and C are independent)
IFE	{AB, C}	$M_x^{ij} = w\, w_i^A\, w_j^B\, w_x^C\, w_{ij}^{AB}$	$\hat{M}_x^{ij} = \frac{1}{N} M_\bullet^{ij}\, M_x^{\bullet\bullet}$	-Multiple independence (joint variable AB is independent of C)
2F	{AB, BC}	$M_x^{ij} = w\, w_i^A\, w_j^B\, w_x^C\, w_{ij}^{AB}\, w_{jx}^{BC}$	$M_x^{ij} = M_\bullet^{ij}\, M_x^{\bullet j} / M_\bullet^{\bullet j}$	-Conditional independence (A independent of C, given B)
3F	{AB, AC, BC}	$M_x^{ij} = w\, w_i^A\, w_j^B\, w_x^C\, w_{ij}^{AB}\, w_{jx}^{BC}\, w_{ix}^{AC}$	no closed-form solution	-Pairwise association (each two-way interaction is independent of level of third variable)
CI	{ABC}	$M_x^{ij} = w\, w_i^A\, w_j^B\, w_x^C\, w_{ij}^{AB}\, w_{jx}^{BC}\, w_{ix}^{AC}\, w_{ijx}^{ABC}$	M_x^{ij} given	-Three-way interaction (association between every pair of variables varies with level of third variable)

For instance, in the 1FE-problem, only one bivariate marginal total is given and hence only one set of first-order interaction effects are different from unity. If a set of preliminary estimates are used that deviate from the uniform distribution, then the higher-order interaction effects of the final estimates are those exhibited by the preliminary estimates.

In the 3F-problem, no closed-form solution exists, and the estimates must be determined by an iterative procedure. Willekens, Pór, and Raquillet (1979) propose an iterative multiproportional adjustment algorithm, which gives a solution of the desired characteristics and which is equivalent to iterative proportional fitting algorithms (Bishop, Fienberg, and Holland 1975, pp. 83-97).

In this paper, a slightly different solution strategy is presented which is better suited for a statistical interpretation. The final estimates may be written as follows:

$$M_x^{ij} = {}^oM_x^{ij}\ a_x^i\ b_x^j\ c^{ij} \tag{14}$$

where

$$a_x^i = M_x^{i\cdot} / \sum_j {}^oM_x^{ij}\ b_x^j\ c^{ij} \tag{15}$$

$$b_x^j = M_x^{\cdot j} / \sum_i {}^oM_x^{ij}\ a_x^i\ c^{ij} \tag{16}$$

$$c^{ij} = M_\cdot^{ij} / \sum_x {}^oM_x^{ij}\ a_x^i\ b_x^j \tag{17}$$

where a_x^i, b_x^j, and c^{ij} are multiplicative balancing factors, associated with the face constraints $M_x^{i\cdot}$, $M_x^{\cdot j}$, and $M_\cdot^{ij}$. The non-linear equations (15) through (17) are solved iteratively. Expression (14) closely resembles the log-linear model of the 3F-problem (Table 18). It shows the absence of a second order effect unless the ${}^oM_x^{ij}$ array exhibits such an effect. The procedure presented above is equivalent to the algorithm proposed by Chilton and Poet (1973) and was also used by Willekens (1977).

6. FINAL REMARKS AND SUGGESTIONS FOR FUTURE STUDIES

6.1. Warning: Do Not Compare the Incomparable

The main objective of the CMS study was a quantitative assessment of recent migration patterns and spatial population dynamics in all of IIASA's 17 National Member Organization countries. By involving national scholars and by using a methodology which was considered the best one available and which was being developed at IIASA, substantial research capacities were built up in the different countries. The application of a common methodology to different data sets provided a first step to producing comparative results. An illustrative overview of the results is given in the Appendix. Comparability is, however, severely handicapped by the considerable degree of incomparability of the input data, in particular the migration data. We have already noted in section 3 the deviations that exist between movement and transition data and the limitations on the time period for which transition data are available. A consequence is that we can perhaps compare Finland, Sweden, the German Democratic Republic, the Netherlands, Hungary, Soviet Union, the Federal Republic of Germany, Poland, and Bulgaria in one set (movement data over one year) and Canada, France, and Austria in another (transition data over five or seven years). Perhaps the United Kingdom (transition data over one year) can be included in the first set.

Within the former set the high levels of interregional migration propensity stand out in Finland, Hungary, and the Soviet Union. In the case of Hungary this may be associated with the inclusion of temporary migrations in the migration flows input data; in the Soviet Union the high levels are undoubtedly associated with the urban-rural region definitions adopted,* in Finland, however, migration levels between provinces are "genuinely" high.

Poland, the German Democratic Republic, and Bulgaria show the highest retention levels and lowest migration propensities in the first set of countries. In Poland's case (Dziewonski and Korcelli 1981) there has been a strong secular decline in postwar migration rates. The Netherlands, Sweden, the Federal Republic of Germany, and the United Kingdom fall in the middle of the range of observations.

---

*The interrepublic migration rates quoted in Table 7 are much lower.

In the second set, Canada shows substantial regional contrasts in retention level with the regions of net outmigration (Saskatchewan, Manitoba, the Maritimes, and Newfoundland) retaining the least fraction of their birth cohorts. Mobility levels in France and Austria appear to be lower and more even over the regions (although the French levels are influenced by the longer period of measurement of the original data - seven years).

The comparability of the results of the CMS study is further handicapped because of varying sizes of regions considered in the study. We refer to section 4, where the regions used were discussed.

To these problems may be added differences in base period. Although the year of observation for all countries are situated in the early seventies, there are variations (see Figure 7). The reason for this lack of uniform base year is that not all participating countries had the necessary data available for the same year. In countries with registration systems, migration data became available annually and hence a recent year could be selected as the basis for the study. Countries which do not have a registration system, must rely on a census for their migration data. Since all basic data on the components of demographic change should relate to the same period, fertility and mortality data were collected for the period (or mid-period) to which the migration data refer.

Another strategy could have been to assemble as much data as possible for a given year (1975, say) and to assemble other data for the nearest available period. This procedure may be illustrated with reference to the United Kingdom study. The base period chosen was the calender year 1970, since it was the year closest to the one-year period for which the latest migration data were available (Rees 1979a, p. 74). If a more recent year, say 1974, would have been selected, the steps involved in assembling the input data for the multiregional analysis would be the following:

(1) Assemble for that year, births, deaths, and population at risk (mid-year), and compute fertility and mortality rates.

(2) Assemble for the nearest available period, migration flow statistics and associated populations at risk. For the United Kingdom such a period would be the year prior to Census 1971 (April 24/25, 1971). The populations at risk would be for the mid-point of that "census" year, and would be interpolated between the 1970 mid-year estimates (June 30) and the census populations (April 24/25) or the June 30, 1971 estimates (which incorporate the census information). Migration rates would then be computed.

(3) A synthetic multiregional life table would then be computed from the 1974 mortality rates and the 1970-71 migration rates using the hybrid approach equations.

(4) If the transition approach were used, and accounts could be constructed for 1970-71, then it would be possible to improve comparability further by using migation rates conditional on survival (see, the French study—Ledent and Courgeau, forthcoming) for 1970-71 and so removing any influence of 1970-71 mortality patterns on the migration rates. The migration rates would then be computed by multiplying 1974 survivorship rates by 1970-71 conditional migration rates.

In terms of programming requirements, it is easy to alter the appropriate subroutine in the Willekens and Rogers (1978) suite so that for each component input and associated and separate population at risk is also available if required. Thus, although we criticized the French study (Ledent and Courgeau, forthcoming) for mismatching periods of data collection, this flexibility, in fact, turns out to be a feature that could be employed in future analyses.

This section is a warning against comparing the incomparable. The CMS study did not produce results that were directly comparable. A uniform methodology was used throughout in the hopes of obtaining truly comparable results for all participating countries. Severe data limitations, however, as well as the focus on individual country reports and the building up of research capabilities in the NMO countries prevented the generation of fully comparable input data within the time and budget constraints imposed on the study. The findings of the CMS study must be assessed very carefully in relation to the nature of the inputs used.

6.2. A Framework for Future Studies

The principle purpose of this paper has been to provide the reader with a framework within which the CMS study can be set and to provide the user of the programs with a guide for preparation of input data sets in new situations.

In this review of accounting frameworks and data inputs we have often been critical of what has been done. Only in that way can future analyses be improved and our understanding of spatial population behavior be increased. Yet despite the severe problems attending comparison of separate national studies, we feel such an approach was justified. Only by confronting the theory of multiregional population analysis with the problems of implementation in a wide variety of contexts could its content and applicability be improved. We therefore conclude the paper by discussing in what ways we would recommend the research specification be improved should anyone wish to repeat and update the exercise later in the 1980s.

(1) The framework for data preparation should be elaborated from a model involving only the input of data to one involving the input of *rates*, the rates being derived from accounts in the base period and through other procedures in other and future periods.

(2) External (international) migration should be included in the population projection exercises (as is done in the extended version of the projection model and computer program).

(3) To avoid the effect of age groupings, a single-year of age data should be used; to reduce the huge data set to manageable proportions, model schedules of age-specific migration rates could be used (Rogers, Raquillet, and Castro 1978).

(4) If migration and settlement studies ought to generate directly comparable results, attention should be given to the

harmonization of the available migration statistics. Harmonization of migration statistics is a priority for comparative analyses.

(5) Given a choice between different migration data types and period lengths, we would recommend employment of transition data rather than movement data, for a time interval equal to the age interval (Ledent and Rees 1980), preferably classified by region of birth as well (Ledent 1980b, Rogers and Philipov 1981) (in most countries, this choice does not exist, however).

(6) The problem of adopting sets of regions more attuned to the detailed settlement structure of developed countries (a system of many city regions) might be solved by dropping the requirement that they be studied as a single multiregional system (an impossible task)(Long and Frey, forthcoming). Instead they could be studied in a set of smaller (three region) systems—city region, rest of country, rest of world. Most of the important statistical outputs of multiregional population analysis would be generated for such systems, and for many more and meaningful regions. The migration data estimation problems would be fairly straightforward.

(7) Careful attention should be paid to some of the detailed problems of data estimation analysis such as age classification of migration, proper time specification of populations at risk, comparable treatment of intraregional migration: these can all have a nonnegligible effect on results.

We hope to live long enough to see some of these recommendations come to pass.

APPENDIX: AN OVERVIEW OF FINDINGS

A.1. Introduction

The principal discussions of the findings of the CMS study is carried out elsewhere. However, it is useful in this Appendix to review selected aspects of the multiregional population dynamics of the IIASA countries in order to show how a knowledge of the accounting framework and data inputs must be used in interpreting the results.

A.2. Life Expectancies

One of the principal products of the CMS task is a set of estimates of life expectancy at the regional level within countries. Two measures of life expectancy were generated: the conventional abridged (five-year age interval) life table expectancy, which we call the single-region life expectancy, and the multiregional life expectancy, in which the expectation of life of a regional birth cohort allowed to migrate is calculated. These measures are set out for all the IIASA countries in Table A1. Also noted in the table is the "retention level" or the proportion of its expected life that a regional birth cohort can expect to spend in its region of birth. The complement to the retention level (1 - retention level) provides a measure of the

Table A1. Single and multiregional life expectancies (LE) at birth and retention levels (RL)

	Life expectancy at birth		Rentention level
Region	SR	MR	RL
United Kingdom 1970: T1			
NO	71.1	71.7	.539
YH	71.2	71.6	.513
NW	70.5	71.4	.593
EM	72.0	72.1	.465
WM	71.6	71.9	.558
EA	73.5	72.6	.411
SE	73.1	72.6	.653
SW	72.8	72.5	.434
WA	71.1	71.1	.540
SC	70.2	71.2	.612
Finland 1974: M1			
UU		71.9	.528
TP		72.1	.524
AH		72.7	.592
HA		71.9	.439
KY		71.3	.438
MI		71.6	.310
PK		71.2	.336
KU		71.4	.359
KS		71.6	.359
VA		72.0	.489
OU		71.5	.433
LA		71.4	.391
Sweden 1974: M1			
UN	74.9	75.0	.533
LN	74.4	74.8	.464
NM	74.5	74.9	.506
EM	75.1	75.2	.487
ST	75.0	75.2	.499
SM	75.4	75.2	.479
WE	75.6	75.4	.641
SO	75.9	75.6	.622
German Democratic Republic 1974: M1			
NO	71.3	71.3	.725
BE	71.1	71.1	.743
SW		71.1	.788
SO		72.0	.800
MI		71.4	.745

	Life expectancy at birth		Retention level
Region	SR	MR	RL
Netherlands 1974: M1			
NO	74.7	74.6	.600
EA	74.4	74.5	.565
WE	75.1	74.8	.649
SW	75.7	74.8	.461
SO	74.0	74.3	.689
Canada 1966-71: T5			
NF	72.4	72.4	.575
PE	72.8	72.5	.428
NS	72.2	72.5	.505
NB	72.5	72.5	.500
QU	71.7	72.1	.787
ON	72.9	73.0	.790
MA	73.5	73.2	.448
SA	74.2	73.2	.373
AL	74.0	73.4	.574
BC	73.3	73.2	.731
Hungary 1974: M1			
CE		68.4	.488
NH		69.1	.424
NP		69.1	.372
SP		69.1	.460
NT		69.7	.506
ST		68.8	.471
Soviet Union 1974: M1			
RS	69.4	69.5	.666
UM	71.5	70.8	.493
BY	73.5	71.4	.394
CE	68.3	67.8	.436
KA	68.6	69.1	.330
CA	71.5	70.5	.602
BA	71.7	71.2	.472
RU	68.2	69.1	.386

Notes:
SR—single region, MR—multiregion, RL—retention level
T—transition data, M—movement data, 1,5,7—year for data
See Table 13 for the full region names.

Table A1 continued

Region	Life expectancy at birth		Retention level
	SR	MR	RL
Federal Republic of Germany 1974: M1			
SH		72.0	.460
HA		72.0	.331
NS		71.7	.559
BR		71.9	.271
NW		71.9	.689
HE		72.1	.535
RP		71.8	.451
BW		72.3	.614
BA		72.0	.694
SA		71.4	.475
WB		71.6	.418
Austria 1966-71: T5			
WI		70.7	.809
NO		70.3	.766
BU		69.9	.732
KA		70.3	.817
ST		70.2	.853
OO		70.6	.881
SA		71.1	.819
TI		71.6	.882
VO		71.4	.863
Poland 1975: M1			
WA		71.5	.839
LO		70.2	.764
GD		71.1	.761
KA		70.1	.820
CR		71.2	.771
RA		70.4	.584
OL		71.1	.614
SZ		70.2	.592
OP		70.5	.691
KI		70.7	.711
LU		71.0	.662
PO		70.7	.744
WR		70.6	.693
Bulgaria 1975: M1			
NW		71.4	.742
NO		71.2	.823
NE		71.1	.847
SW		70.9	.775
SH		70.6	.867
SE		70.5	.758
SO		70.6	.842

Region	Life expectancy at birth		Retention level
	SR	MR	RL
France 1975: T7			
PR	74.6	74.0	.572
PB	73.3	73.3	.641
NO	70.2	71.6	.688
EA	72.3	72.9	.682
WE	73.0	73.4	.699
SW	74.4	74.1	.658
ME	73.6	73.8	.705
MD	74.7	74.2	.656
Czechoslovakia 1975: M1			
P	70.2	70.1	.703
CB	69.9	70.2	.606
SB	70.9	70.7	.687
WB	69.8	70.2	.640
NB	68.7	69.3	.668
EB	71.2	70.9	.714
SM	71.5	71.2	.796
NM	70.3	70.4	.794
B	70.6	70.3	.730
WS	70.3	70.4	.768
CS	70.4	70.4	.775
ES	69.9	69.8	.817
Japan 1970: T1			
HO	71.8	72.1	.420
TO	71.1	72.0	.413
KA	72.5	72.3	.734
CB	72.4	72.3	.589
KI	72.6	72.5	.611
CG	72.5	72.3	.431
SH	71.6	72.0	.359
KY	71.6	72.1	.352
United States 1970: T1			
NE	71.0	71.0	.586
NC	71.3	71.1	.561
SO	69.9	70.5	.560
WE	71.8	71.1	.530
Italy 1971: M1			
NW		71.7	.931
NE		72.0	.950
CE		73.2	.943
SI		71.7	.895

Notes to Table Al:

Sources: The following sources refer to the Migration and Settlement Reports, which are listed in numerical order at the end of the references. The final reports for France, Czechoslovakia, Japan, the United States, and Italy have not been published, therefore working drafts were used as references.

1. United Kingdom, Rees (1979): LE, SR—Table 17, p. 50; LE, MR—Table 28, p. 91; RL—Table 29, p. 94.
2. Finland, Rikkinen (1979): LE, SR—average of male and female values, 1961-1965, given in Table 2.7, p. 20; LE, MR—Table 3.5a, p. 51; RL—Table 3.5b, p. 51.
3. Sweden, Andersson and Holmberg (1980): LE, SR—Table 3.4, p. 37; LE, MR—Table 3.4, p. 37; RL—computed from Table 3.2, p. 36.
4. German Democratic Republic, Mohs (1980): LR, MR—Table 12, p. 28; RL—computed from Table 12, p. 28.
5. Netherlands, Drewe (1980): LE, SR—Table 3, p. 8; LE, MR—Table 6, p. 18; RL—computed from Table 6, p. 18.
6. Canada, Termote (1980): LE, SR—average of male and female values in Table 5, p. 20; LE, MR—average of male and female values in Table 12, p. 34; RL—computed from Table 12, p. 34.
7. Hungary, Bies and Tekse (1980); LE, SR—average of male and female county values in Table 12, p. 19 for 1969-1970; LE, MR—Table 25, p. 42; RL—Table 26, p. 42.
8. Soviet Union, Soboleva (1980): LE, SR—Table 26, p. 48; LE, MR—Table 31, p. 55; RL—computed from Table 31, p. 55.
9. Federal Republic of Germany, Koch and Gatzweiler (1980): LE, MR—Table 7, p. 31; RL—computed from Table 7, p. 31.
10. Austria, Sauberer (1981): LE, MR—Table 16, p. 34; RL—computed from Table 16, p. 34.
11. Poland, Dziewonski and Korcelli (1981): LE, MR—Table 10, p. 45; RL—Table 11, p. 48.
12. Bulgaria, Philipov (1981): LE, MR—Table 16, p. 27; RL—Table 17, p. 28.
13. France, Ledent and Courgeau (forthcoming): LE, SR—Table 51, p. 157; LE, MR—Table 51, p. 157; RL—Table 49, p. 154 for males. The life expectancy values are labelled "1975" because this is the period to which the deaths data apply. The migration flows data refer to the period 1968-1975, however. See Figure 12.
14. Czechoslovakia, Kuhnl (forthcoming): LE, SR, LE, MR, and RL recomputed by Rees from original data.
15. Japan, Nanjo, Kawashima, and Kuroda (forthcoming): LE, SR—average of male and female values in Table 9, p. 26; LE, MR—Table 13, p. 35; RL—Table 13, p. 35.
16. United States, Long and Frey (forthcoming): LE,SR—Table 3.1, p. 29; LE, MR, and RL—Table 3.4, p. 35.
17. Italy, Just and Rogers (1980 IIASA monograph): LE, MR, and RL—Table 14, p. 21.

migration propensity of the regional birth cohort (in a stationary life table model population). Also noted in the table are the year or longer period to which the data apply, and above each country's retention level statistics there is a code denoting the nature of the migration flow data used and the period over which it was observed. The notes to the table give details of the sources used to compile the statistics, which were the CMS study Research Reports, to which the reader can turn for further details.

The first property of the life expectancies which we examine is their range and variation. The type of question we have in mind is the extent to which the life expectancies of the regions vary among the countries or within them. A qualitative answer to this question is provided if we graph the minima, medians, and maxima of the countries' regional life expectancy set in rank order of medium multiregional life expectancy (Figure A1). The difference between highest and lowest median for countries is 6.1 years; the widest range within a country is 3.6 years in the Soviet Union, and most countries have much narrower ranges.

This impression of wider international variation and narrower intranational variation is reinforced when we map the life expectancies (Figures A2 and A3). The picture for Europe is a familiar one of the highest life expectancies to be found in the countries closest to the Northwest Europe-Scandinavia axis running from, say, Paris to Stockholm. Away from this axis, life expectancies decline.

However, there are problems in interpreting the results. In particular, the year of observation varies quite a bit, and since life expectancies are improving continuously in most countries (perhaps by one year per decade in the CMS set), countries with more recent periods of study compare favorably with countries with less recent periods of study. So Austria, Canada, and the United Kingdom have lower life expectancies than they would have if statistics had been available for 1974 (the model year in the studies), and France and Bulgaria have higher expectancies.

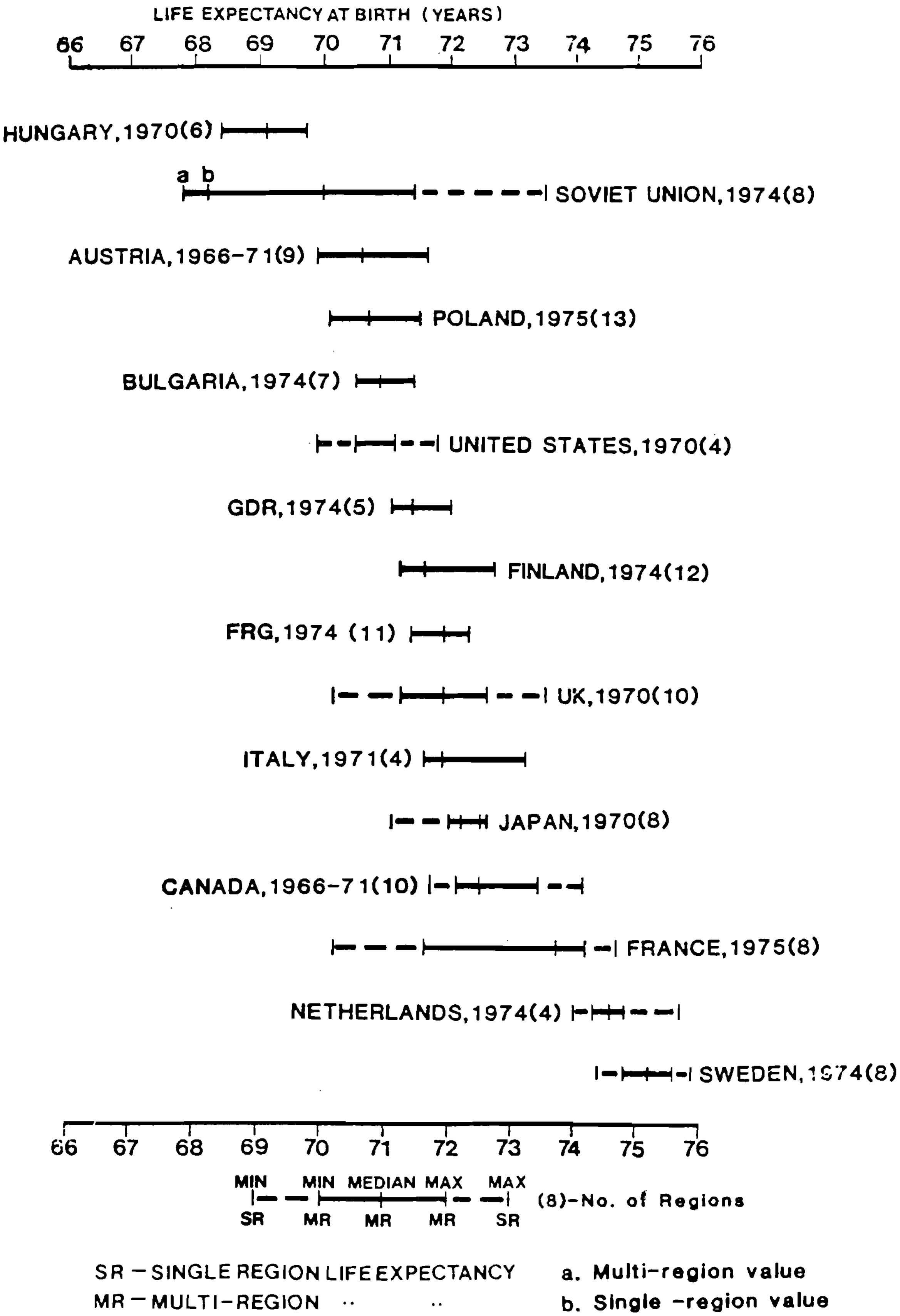


Figure A1. The range of life expectancies in the CMS country studies.

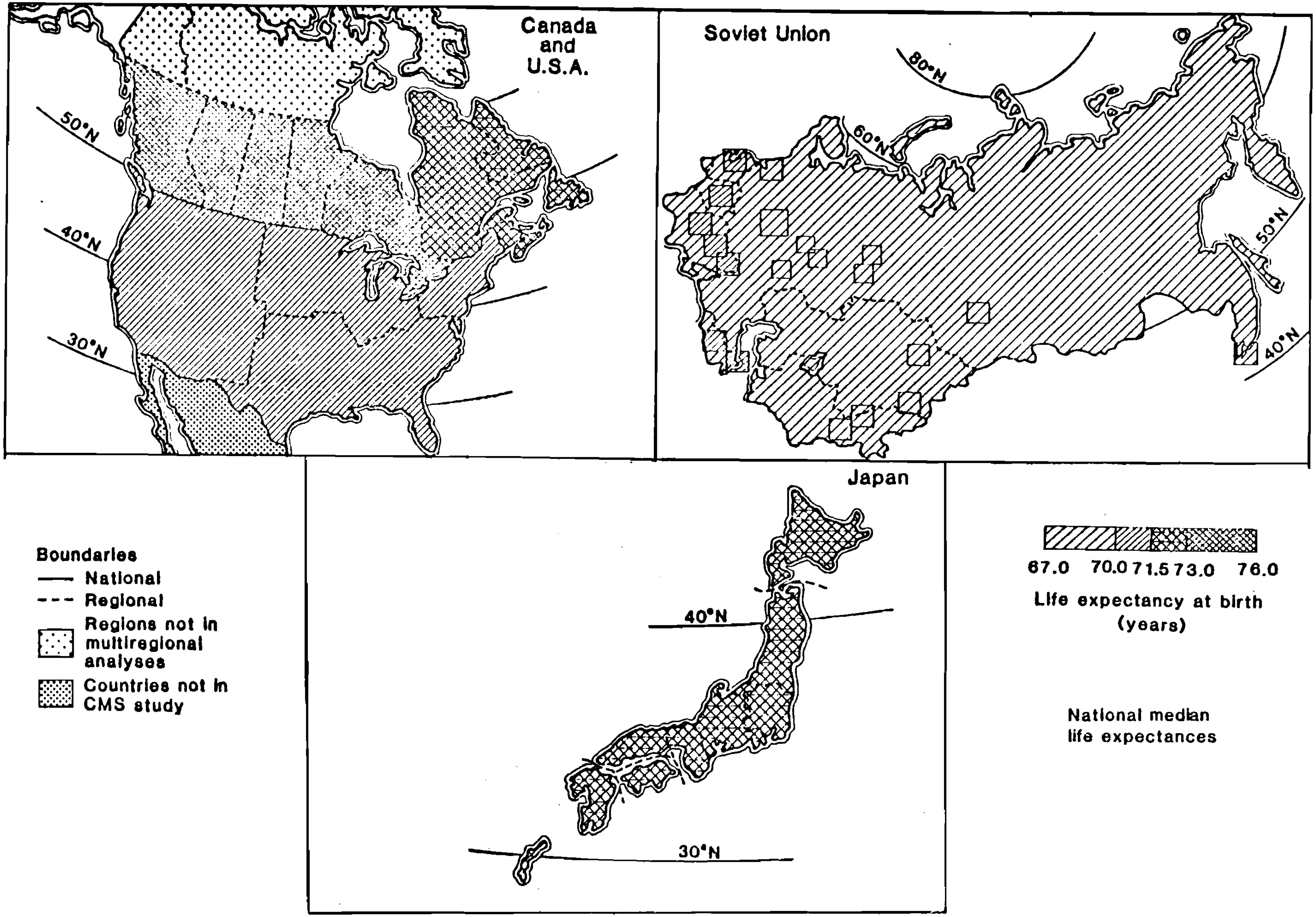


Figure A2. Life expectancies (multiregional) in the CMS regions: North America, Soviet Union, and Japan.

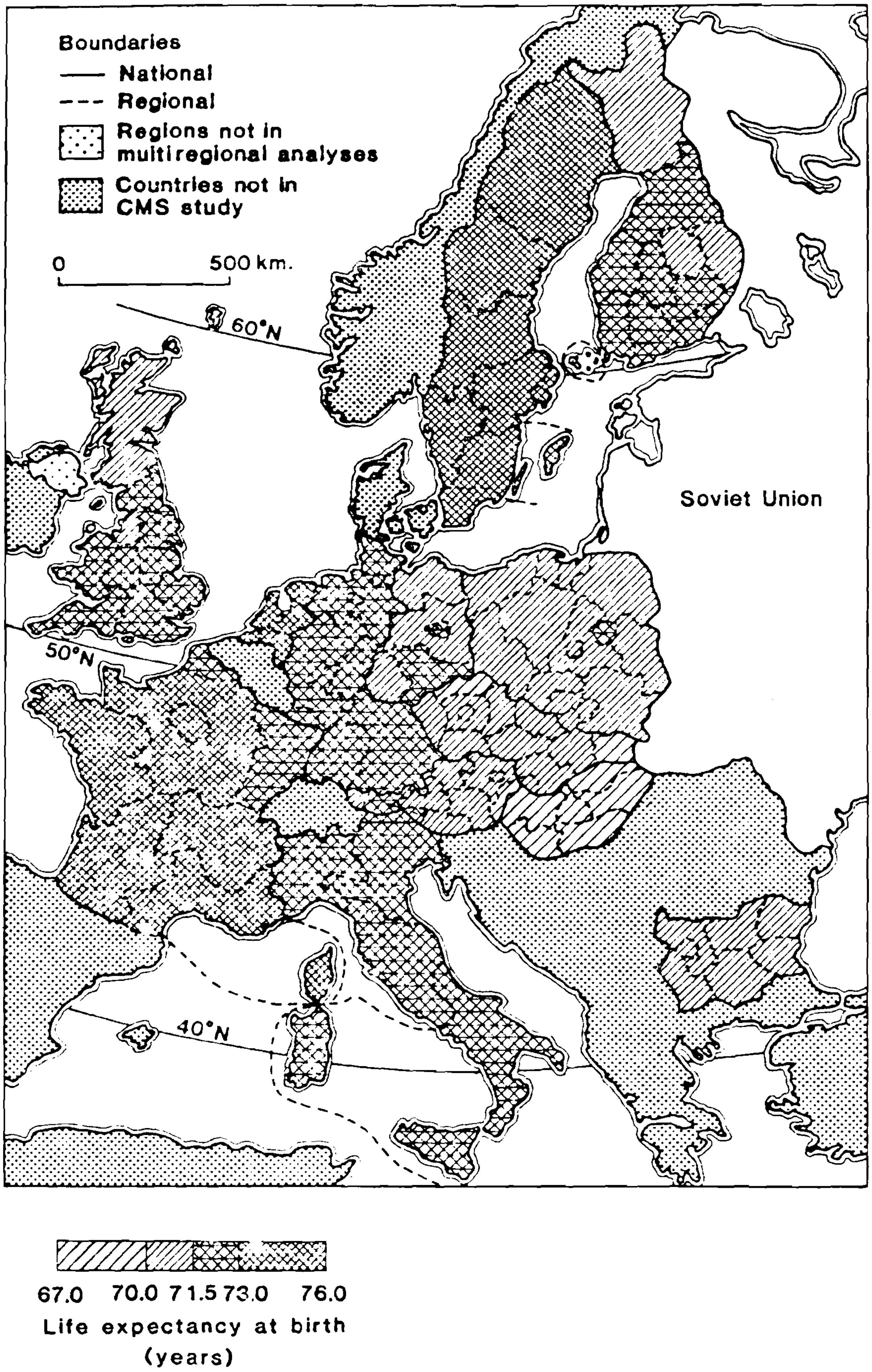


Figure A3. Life expectancies (multiregional) in the CMS regions: Europe.

The reason for the incomparability is, as we have noted earlier in section 3, the limitations on the time period for which migration flows data (particularly, transition type data) are available.

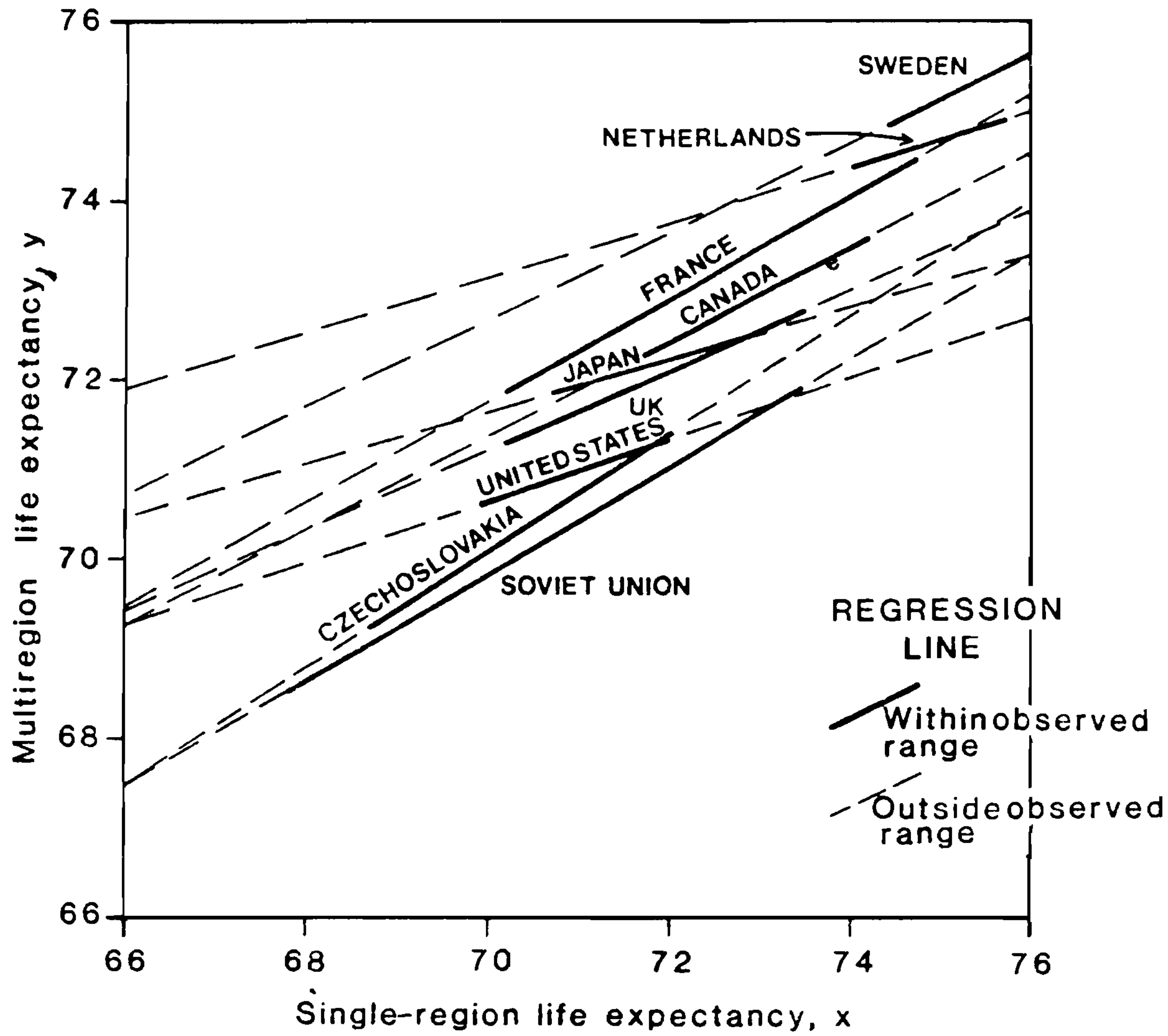
A.3. The Relationship between Single and Multiregional Life Expectancies

In the United Kingdom and Sweden studies, much attention was devoted to the relationship between life expectancies calculated in a single region and in a multiregional fashion. As Figure A1 shows the range of single-region life expectancies is greater (with the exception of the Soviet Union regional minima, something that in theory should not happen) than the multiregional range, and the variance is higher. It was suggested in the United Kingdom study that the multiregional life expectancies represented a regression of the single-region values to the mean. Migration was then interpreted as a process that smoothed out or equalized life changes within a nation, although it was recognized that this interpretation depended critically on the assumption that migrants did not carry their past history with them.

This regression analysis can be repeated for a number of the CMS countries and the results are shown in Figure A4. In all analyses the slope of the regression line is substantially less than unity (when there would be no effect). In fact, the United Kingdom turns out not to be the country with the highest degree of regression to the mean: that falls to the Netherlands. The weakest regression effect occurs in the Soviet Union, where although migration levels are relatively high, their nature and direction do not favor equalization of life changes as much.

A.4. Retention Levels and Migration Propensities

The values from Table A1 are plotted on Figures A5 and A6. Unfortunately, it is only possible to compare within sets of countries with approximately similar migration data inputs: we can perhaps compare Finland, Sweden, the German Democratic Republic, the Netherlands, Hungary, the Soviet Union, the Federal Republic of Germany, Poland, and Bulgaria in one set (movement data over



Country Study	Correl-ation coeff.	Regression coeff.	Regression intcpt.	LE,MR mean	LE,SR mean	LE,MR stan. dev.	LE,SR stan. dev.	MR/SR stan. dev %
United Kingdom	0.99	0.44	40.25	71.94	71.71	0.50	1.11	45
Sweden	0.97	0.49	38.41	75.16	75.10	0.26	0.52	50
Netherlands	0.94	0.30	51.83	74.60	74.78	0.21	0.65	32
Canada	0.93	0.52	34.96	72.80	72.95	0.45	0.80	56
Soviet Union	0.93	0.59	28.49	69.93	70.34	1.25	1.97	63
France	0.99	0.57	31.83	73.44	73.26	0.86	1.49	58
Czechoslovakia	0.95	0.66	24.11	70.32	70.31	0.49	0.72	69
Japan	0.92	0.29	51.14	72.20	72.01	0.47	1.48	32
United States	0.91	0.34	46.80	70.95	71.00	0.25	0.70	36

LE - life expectancy at birth   MR - multiregion   SR - single region

Figure A4. Multiregional versus single-region life expectancies: regression analysis.

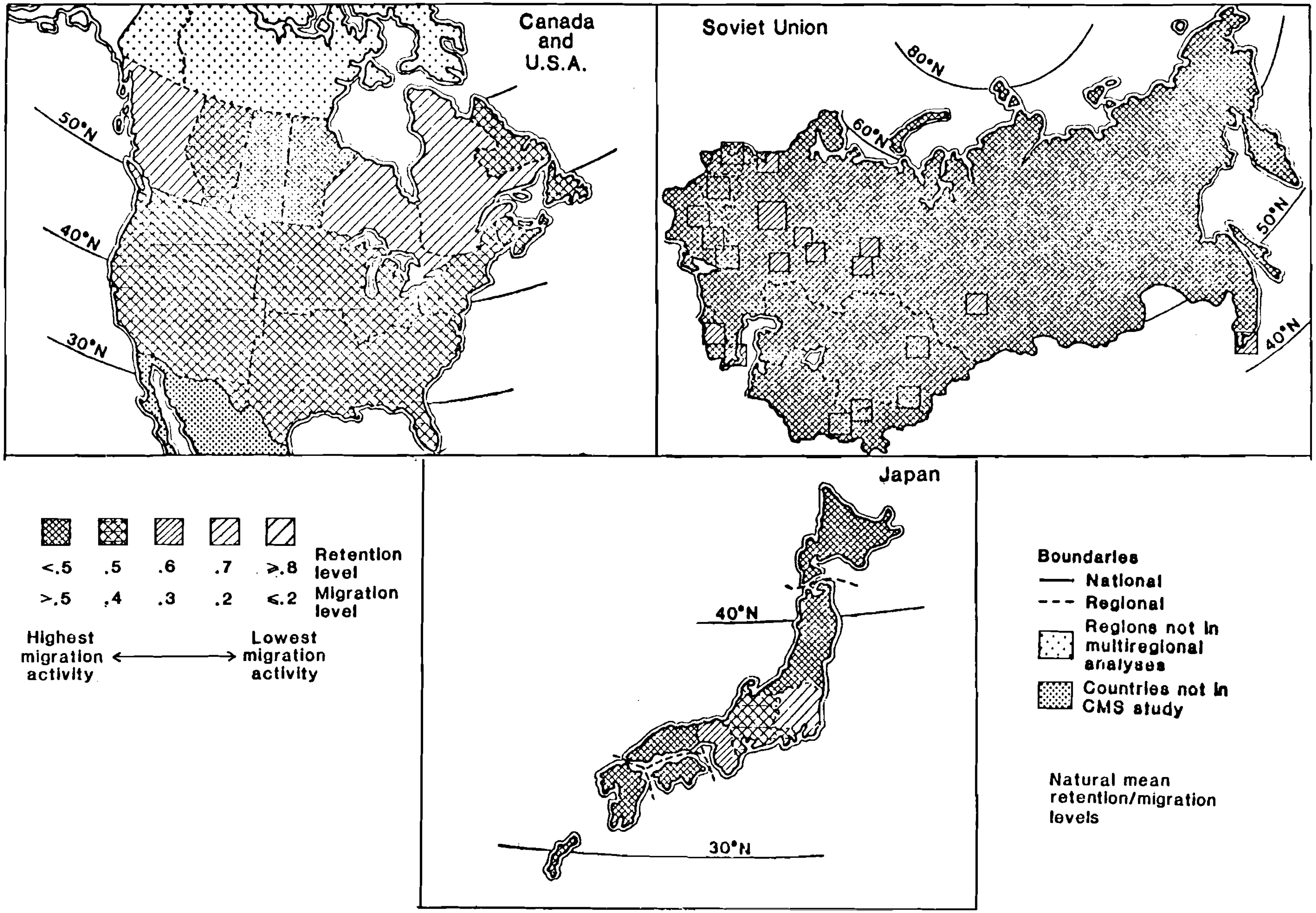


Figure A5. Retention/migration levels in the CMS regions: North America, Soviet Union and Japan.

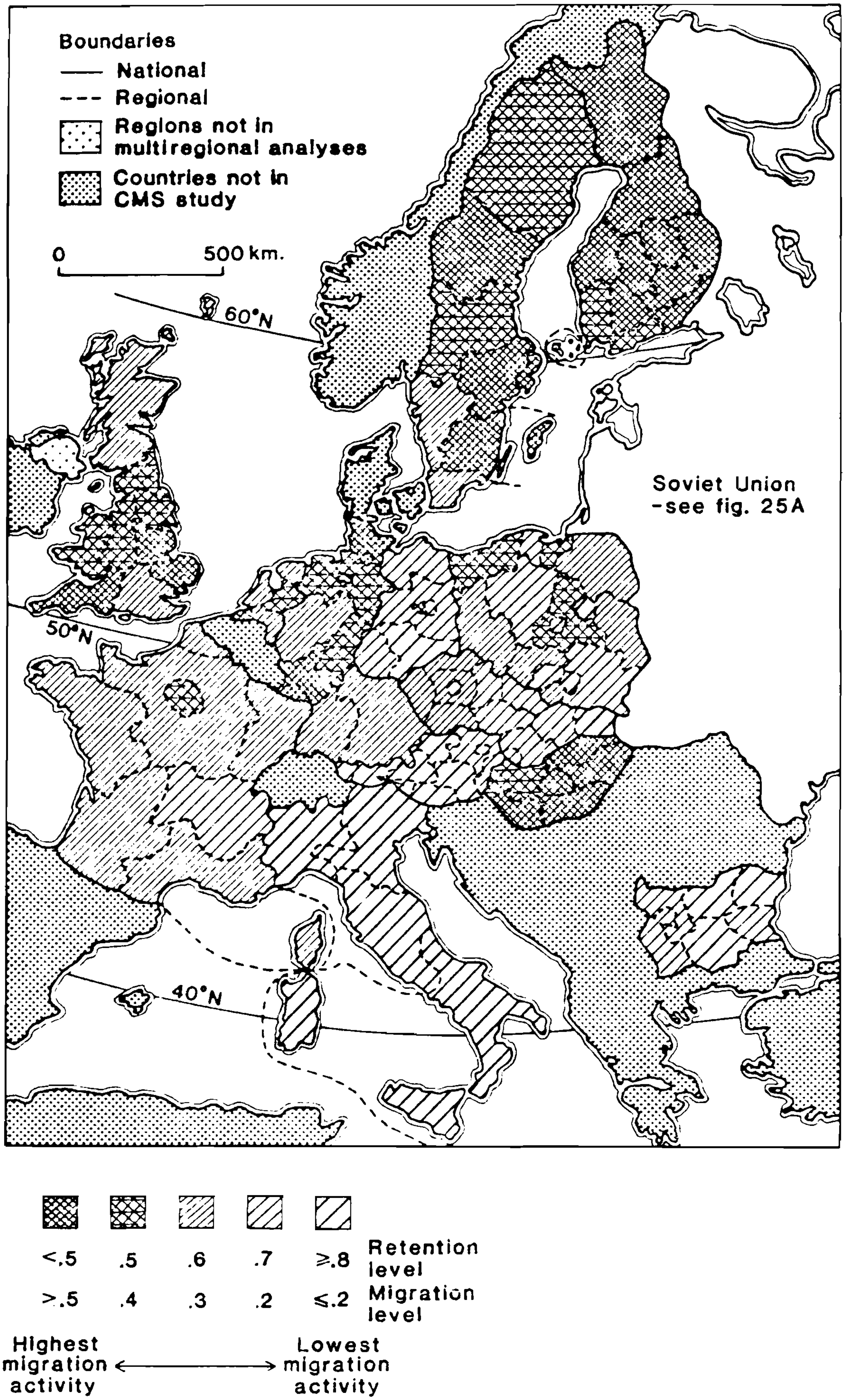


Figure A6. Retention/migration levels in the CMS regions: Europe.

one year) and Canada, France, and Austria in another (transition data over five or seven years). Perhaps the United Kingdom (transition data over one year) can be included in the first set.

Within the former set the high levels of interregional migration propensity stand out in Finland, Hungary, and the Soviet Union. In the case of Hungary this may be associated with the inclusion of temporary migrations in the migration flows input data; in the Soviet Union the high levels are undoubtedly associated with the urban-rural region definitions adopted,* in Finland, however, migration levels between provinces are "genuinely" high.

Poland, the German Democratic Republic, and Bulgaria show the highest retention levels and lowest migration propensities in the first set of countries. In Poland's case (Dziewonski and Korcelli 1981) has been a strong secular decline in postwar migration rates. The Netherlands, Sweden, the Federal Republic of Germany, and the United Kingdom fall in the middle of the range of observations.

In the second set, Canada shows substantial regional contrasts in retention level with the regions of net outmigration (Saskatchewan, Manitoba, the Maritimes, and Newfoundland) retaining the least fraction of their birth cohorts. Mobility levels in France and Austria appear to be lower and more even over the regions (though the French levels are influenced by the longer period of measure of the original data—seven years).

However, the map of retention levels/migration propensities clearly needs further detailed and careful analysis, and supplementation with an analysis of migraproduction levels. Unfortunately, in the migraproduction analyses yet another element of variability was added—in some country studies, interdistrict/intercounty migrations with regions were included in the calculation of migraproduction statistics, while in others they were not. The findings of the CMS study must be assessed very carefully in relation to the nature of the inputs used.

---

*The interrepublic migration rates quoted in Table 7 are much lower.

REFERENCES

Andersson, A., and I. Holmberg (1980) *Migration and Settlement 3: Sweden.* RR-80-5. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Berry, B.J.L., and Q. Gillard (1977) *The Changing Shape of Metropolitan America: Commuting Patterns, Urban Fields, and Decentralization Processes 1960-1970.* Cambridge, Mass.: Ballinger.

Bies, K., and K. Tekse (1980) *Migration and Settlement 7: Hungary.* RR-80-34. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Bishop, Y.M., S.E. Fienberg, and P.W. Holland (1975) *Discrete Multivariate Analysis: Theory and Practice.* Cambridge, Mass.: Massachusetts Institute of Technology Press.

Chilton, R., and R. Poet (1973) An entropy maximizing approach to the recovery of detailed migration patterns from aggregate census data. *Environment and Planning A*, 5:135-146.

Courgeau, D. (1973a) Migrants et migrations (Migrants and migrations). *Population*, 28(1):95-129.

Courgeau, D. (1973b) Migrations et decoupages du territoire (Migrations and spatial divisions). *Population*, 28(3):511-537.

Courgeau, D. (1980) *L'analyse Quantitative des Migrations Humaines* (The quantitative analysis of human migration). Paris: Masson.

Drewe, P. (1980) *Migration and Settlement 5: The Netherlands.* RR-80-13. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Drewe, P., and F. Willekens (1980) Maximum likelihood estimation of age-specific migration flows in the Netherlands. *DELFT Progress Report*, 5:92-111.

Dziewonski, K., and P. Korcelli (1978) *Migration and Settlement in Poland.* Paper presented at the IIASA Conference on the Analysis of Multiregional Population Systems: Techniques and Applications, 19-22 September, 1978, Laxenburg, Austria.

Dziewonski, K., and P. Korcelli (1981) *Migration and Settlement 11: Poland.* RR-81-20. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Grigg, D. (1967) Regions, models, and classes, in *Models in Geography*, edited by R.J. Chorley and P. Haggett. London: Methuen.

Haggett, P., A. Cliffe, and A. Frey (1977) *Locational Methods.* London: Edward Arnold.

Hall, P., and D. Hay (1980) *Growth Centres in the European Urban System.* London: Hinemann Educational Books.

Illingworth, D. (1976) Testing some new concepts and estimation methods in population accounting and related fields. Ph.D. dissertation. Leeds: University of Leeds, School of Geography.

Jenkins, J.C. (1976) *Management Models, Demographic Accounts, and Associated Policy Studies.* Research Committee Paper. Leeds, University of Leeds, School of Geography.

Just, P., and A. Rogers (1980) *Migration and Population Redistribution in Italy.* IIASA Monograph. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Kennett, S. (1980) Migration within and between the metropolitan economic labour areas of Britain, 1966-1971. In *Regional Demographic Development*, edited by J.N. Hobcraft and P.H. Rees. London: Croom Helm.

Kitsul, P., and D. Philipov (1980) The one-year five-year migration problem. Pages 1-33 in *Advances in Multiregional Demography*, edited by A. Rogers. RR-80-6. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Koch, R., and H.P. Gatzweiler (1980) *Migration and Settlement 9: Federal Republic of Germany.* RR-80-37. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Ledent, J. (1980a) *An Improved Methodology for Constructing Increment-Decrement Life Tables from the Transition Perspective.* WP-80-104. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Ledent, J. (1980b) *Constructing Multiregional Life Tables Using Place-of-Birth-Specific Migration Data.* WP-80-96. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Ledent, J., and D. Courgeau (forthcoming) *Migration and Settlement in France.* Research Report draft. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Ledent, J., and P.H. Rees (1980) *Choices in the Construction of Multiregional Life Tables.* Working Paper 289. Leeds: University of Leeds, School of Geography. Also WP-80-173. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Liaw, K.L. (1977) *Discrete-Time, Sex-, and Age-Disaggregated Spatial Population Model: Theoretical Analysis and Canadian Case Study.* Research Report to the Canadian Council. Hamilton, Ontario: McMaster University, Department of Geography.

Liaw, K.L. (1979) *Analysis and Projections of the Canadian Interregional Population System.* Research Report to the Social Science and Humanities Research Council of Canada. Hamilton, Ontario: McMaster University, Department of Geography.

Long, J.F., and C.B. Boertlein (1977) *Comparison of Migration Measures for Different Intervals.* Mimeograph. Washington, D.C.: U.S. Census Bureau.

Long, L., and W. Frey (forthcoming) *Migration and Settlement 16: United States.* Research Report draft. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Mohs, G. (1980) *Migration and Settlement 4: German Democratic Republic.* RR-80-6. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Nanjo, Z. (1981) *The Projection Method of Future Regional Population in Japan Using Migration Data Derived from the Basic Resident Registers.* Paper distributed at the IIASA Conference on Urbanization and Development, 1-4 June, 1981, Laxenburg, Austria.

Nanjo, Z., T. Kawashima, and T. Kuroda (forthcoming) *Migration and Settlement 13: Japan.* Research Report draft. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Ogilvy, A. (1980a) *Interregional Migration Since 1971: An Appraisal of Data from the National Health Service Central Register and Labour Force Surveys.* Occasional Paper 16. London: Office of Population Censuses and Surveys.

Ogilvy, A. (1980b) Migration—the influence of economic change. *Futures,* 11(5):383-394.

Office of Population Censuses and Surveys (OPCS)(1974) *Census 1971, Great Britain, Migration Tables.* London: Her Majesty's Stationary Office.

Office of Population Censuses and Surveys (OPCS)(1980a) *Population Trends 21.* London: Her Majesty's Stationary Office.

Office of Population Censusus and Surveys (OPCS)(1980b) *Local Authority Population Estimates Methodology.* Occasional Paper 18. London: Population Statistics Division.

Philipov, D. (1978) Migration and settlement in Bulgaria. *Environment and Planning A*, 10:593-617.

Poole, M.A., and F.W. Boal (1973) Religious residential segregation in Belfast in mid-1969: a multilevel analysis. In *Social Patterns in Cities*, edited by B.D. Clark and M.B. Gleave. London: Institute of British Geographers.

Rees, P.H. (1977a) Une famille de comptes demographiques (A family of demographic accounts). In Colloques Nationaux due Centre National de la Recherche Scientifique No. 934. *L'analyse Demographiques et Ses Applications.* Paris: CNRS.

Rees, P.H. (1977b) The measurement of migration from census data and other sources. *Environment and Planning A*, 9:247-272.

Rees, P.H. (1979a) Migration and Settlement in the United Kingdom. Unpublished paper. Leeds: University of Leeds, School of Geography.

Rees, P.H. (1979b) *Residential Patterns in American Cities: 1960.* Research Paper 189. Chicago: University of Chicago, Department of Geography.

Rees, P.H. (1979c) *Migration and Settlement 1: United Kingdom.* RR-79-3. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rees, P.H. (1980) Multistate demographic accounts: measurement and estimation procedures. *Environment and Planning A*, 12:499-531.

Rees, P.H. (1981) *Accounts Based Models for Multiregional Population Analysis: Methods, Program, and Users' Manual.* Working Paper 295. Leeds: University of Leeds, School of Geography.

Rees, P.H., and A.G. Wilson (1977) *Spatial Population Analysis.* London: Edward Arnold.

Rikkinen, K. (1979) *Migration and Settlement 2: Finland.* RR-79-9. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A. (1976a) *The Comparative Migration and Settlement Study: A Summary of Workshop Proceedings and Conclusions.* RM-76-1. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A. (1976b) Shrinking large-scale population projection models. *Environment and Planning A*, 8:515-541.

Rogers, A. (1980) Sources of Urban Population Growth and Urbanization, 1950-2000: A Demographic Accounting. Unpublished paper. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A., and J. Ledent (1976) Increment-decrement life tables: a comment. *Demography*, 13:287-290.

Rogers, A., and D. Philipov (1979) *Multiregional Methods for Subnational Population Projections.* WP-79-40. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A., R. Raquillet, and L. Castro (1978) Modeling migration schedules and their applications. *Environment and Planning A*, 10:475-502.

Sauberer, M. (1981) *Migration and Settlement 10: Austria.* RR-81-16. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Soboleva, S. (1980) *Migration and Settlement 8: Soviet Union.* RR-80-36. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Statistics Canada (1977) *Canada Year Book 1976-1977.* Ottawa, Canada: Special Edition Statistics Canada.

Stillwell, J.C.H. (1978) Interzonal migration: some historical tests of spatial-interaction models. *Environment and Planning A*, 10:1187-1200.

Stone, R. (1971) *Demographic Accounting and Model-Building.* Paris: Organization for Economic Cooperation and Development.

Stone, R. (1975) *Towards a System of Social and Demographic Statistics.* New York: United Nations.

Termote, M. (1980) *Migration and Settlement 6: Canada.* RR-80-29. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Tsentral'noe Statisticheskie Upravlenie (1972) *Itogi perepisi naseleniya 1970 goda, volume VII* (Results of the 1970 Population Census, volume VII). Moscow: Statistika.

U.S. Bureau of the Census (1975) *Statistical Abstract of the United States: 1975.* Washington, D.C.: Government Printing Office.

U.S. Bureau of the Census (1979) *Statistical Abstract of the United States: 1979.* Washington, D.C.: Government Printing Office.

Willekens, F. (1977) *The Recovery of Detailed Migration Patterns from Aggregate Data: An Entropy Maximizing Approach.* RR-77-58. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Willekens, F. (1978) *The Comparative Migration and Settlement Study: Background Paper for the September 1978 Conference.* WP-78-10. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Willekens, F. (1981) *Multidimensional Population Analysis with Incomplete Data.* Working Paper No. 19. Voorburg, Netherlands: Netherlands Interuniversity Demographic Institute.

Willekens, F., A. Pór, and R. Raquillet (1979) *Entropy, Multiproportional, and Quadratic Techniques for Inferring Detailed Migration Patterns from Aggregate Data. Mathematical Theories, Algorithms, Applications, and Computer Programs.* WP-79-88. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Willekens, F., and A. Rogers (1978) *Spatial Population Analysis: Methods for Computer Programs.* RR-78-18. Laxenburg, Austria: International Institute for Applied Systems Analysis.

World Bank (1980) *World Bank Development Report, 1980.* Washington, D.C.: The World Bank.

COMPARATIVE MIGRATION AND SETTLEMENT RESEARCH REPORTS

Migration and Settlement 1: United Kingdom
P.H. Rees (1979) RR-79-3

Migration and Settlement 2: Finland
K. Rikkinen (1979) RR-79-9

Migration and Settlement 3: Sweden
A.E. Andersson and I. Holmberg (1980) RR-80-5

Migration and Settlement 4: German Democratic Republic
G. Mohs (1980) RR-80-6

Migration and Settlement 5: Netherlands
P. Drewe (1980) RR-80-13

Migration and Settlement 6: Canada
M. Termote (1980) RR-80-29

Migration and Settlement 7: Hungary
K. Bies and K. Tekse (1980) RR-80-34

Migration and Settlement 8: Soviet Union
S. Soboleva (1980) RR-80-36

Migration and Settlement 9: Federal Republic of Germany
R. Koch and H.P. Gatzweiler (1980) RR-80-37

Migration and Settlement 10: Austria
M. Sauberer (1981) RR-81-16

Migration and Settlement 11: Poland
K. Dziewonski and P. Korcelli (1981) RR-81-20

Migration and Settlement 12: Bulgaria
D. Philipov (1981) RR-81-21

Migration and Settlement 13: France
J. Ledent (forthcoming)

Migration and Settlement 14: Czechoslovakia
K. Kuhnl (forthcoming)

Migration and Settlement 15: Japan
N. Nanjo, T. Kawashima, and T. Kuroda (forthcoming)

Migration and Settlement 16: United States
L.H. Long and A. Frey (forthcoming)

Migration and Settlement 17: Italy
A. La Bella (forthcoming)

RELATED IIASA PUBLICATIONS

Andrei Rogers (editor), *Migration and Settlement: Selected Essays*. RR-78-6. Reprinted from a special issue of *Environment and Planning A*.

Andrei Rogers and Frans Willekens, *Migration and Settlement: Measurement and Analysis*. RR-78-13.

Frans Willekens and Andrei Rogers, *Spatial Population Analysis: Methods and Computer Programs*. RR-78-18.

Andrei Rogers, *Migration Patterns and Population Redistribution*. RR-80-7. Reprinted from *Regional Science and Urban Economics*.

Andrei Rogers (editor), *Essays in Multistate Mathematical Demography*. RR-80-10. Reprinted from a special issue of *Environment and Planning A*.

Nathan Keyfitz, *Multidimensionality in Population Analysis*. RR-80-33. Reprinted from *Sociological Methodology 1980*.

Andrei Rogers (editor), *Advances in Multiregional Demography*. RR-81-6.

Kao-Lee Liaw, *Multiregional Population Projection: An Analytic Approach*. WP-81-81.